Training, Productivity and Wages:
Direct Evidence from a Temporary Help Agency

By

Xinwei Dong
(The University of Tokyo)
Dean R. Hyslop
(Motu Economic and Public Policy Research)
Daiji Kawaguchi
(The University of Tokyo)

May 2022

CREPE DISCUSSION PAPER NO. 122

CENTER FOR RESEARCH AND EDUCATION FOR POLICY EVALUATION (CREPE)
THE UNIVERSITY OF TOKYO
http://www.crepe.e.u-tokyo.ac.jp/

Training, Productivity and Wages: Direct Evidence from a Temporary Help Agency[1]

Xinwei Dong[2] Dean R. Hyslop[3] Daiji Kawaguchi[4]

May 15, 2022

[1]We thank Marco Francesconi, Drew Griffen, Takao Kato, Peter Kuhn, Dave Maré, Katsuya Takii, Till von Wachter, Atsushi Yamagishi, and seminar participants at the University of Essex, Keio University, the University of Tokyo, Transpacific Labor Seminar, Waseda University, and Victoria University of Wellington for helpful comments and discussion. We are grateful to the temporary help agency we analyzed for providing us with data. Confidentiality restrictions prevent us from revealing the name of the company or other identifying information. This study is supported by JSPS grant-in-aid for scientific research (19H00589, 18H03632).

[2]University of Tokyo, Email: xwdong@g.ecc.u-tokyo.ac.jp

[3]Motu Economic and Public Policy Research, Email: dean.hyslop@motu.org.nz

[4]University of Toyko, Email: kawaguchi@e.u-tokyo.ac.jp, Address: Hongo 7-3-1, Bunkyo-ku, Tokyo, Japan, Telephone: +81-3-5841-5508; REITI; IZA

Abstract

Firms frequently provide general skill training to workers at the firm's cost. Theories proposed that labor market frictions entails *wage compression*, larger productivity gain than wage growth to skill acquisition, and motivates a firm to offer opportunities for skill acquisition, but few studies directly test the hypothesis. We use unusually rich data from a temporary help service firm that records both workers' wages and their productivity as measured by the fees charged to client firms. We first document that the firm provides upfront training, and show that both workers' tenure and the initial fee charged to clients are positively related to the length of training, but the initial wage paid to workers is not. We then demonstrate that the fees charged to clients grow faster over workers' tenure than the wages paid to workers. Finally, we find that about one-quarter of the fee growth is associated with client quality upgrading, but that workers receive none of this growth. Each of these results are consistent with wage compression that skills acquired through training and learning-by-doing increases productivity more than wages.

JEL classification codes: J24, J42
Key words: Training, General skill, Temporary help service agency, Productivity, Wages

1 Introduction

Employers frequently provide free training upfront for their workers to acquire apparently general skill. This observation poses a long-standing puzzle in labor economics because, in a perfectly competitive labor market, the wage offer for general skill is bid up to the value of the marginal product of labor (MPL), and the return to human capital is perfectly accrued to the workers. Thus, employers have an incentive to invest in general skill only when they can shift the cost of training to the workers by paying them lower wages than their productivity (Becker, 1964).

However, in reality, employers provide training opportunities that endow general skill to the trainees whose productivity is apparently not high enough to cover the training cost. Examples include the German apprenticeship system, long-term training offered by large Japanese firms, and general skill training offered by temporary help service firms before assignment to clients (Acemoglu and Pischke, 1998; Holzhausen, 2000; Krueger, 1993). Previous studies have attempted to resolve the puzzle by arguing that the labor market frictions enable firms to reap higher productivity than wage returns to general skill investment, creating *wage compression*, which motivates employers to invest in general skill training of workers (Acemoglu and Pischke, 1998, 1999a,b). However, to our knowledge, no study has directly observed wage compression because measuring productivity growth due to human capital investment is fundamentally difficult.

We use unique worker level data from a temporary help service (THS) firm in Japan to directly observe the MPL of individual workers along with their wages. The business of the THS firm is to procure labor services from workers in the labor market and sell these services to client firms. In each transaction, we observe both the wage paid to a worker and the fee charged to a client, which represents the MPL of the worker to the THS firm.[1] If the markup of the fee over the wage increases with skill formation through formal training or learning-by-doing, the THS firm has an incentive to offer formal training or assign a worker to a client with the opportunity of learning-by-doing even if the skill is technically transferable across firms.

This firm's main service is to assign information communication technology (ICT) engineers to the client firms. The firm employs workers on permanent contracts and pays each worker a monthly salary regardless of whether or not they are assigned to a client. This in contrast to the standard practice of THS firms that hire workers only during the periods they are assigned to their clients. At the start of employment, the firm provides workers training opportunities to acquire or update their ICT skills. The data set covers the period between 2015 and 2020 for around 2,000 employees. Our analysis sample contains information on the monthly fees, wages, billable hours recorded for each worker, and the client the worker is assigned to. It also contains workers' background information such as gender, educational attainment, age, date (year-month) of the entry to the firm, and the branch location the worker is registered with. We infer the training period from the initial non-placement periods in the record. Drawing on this panel data, we are able to track the dynamic paths of the fee, wage, and billable hours of each worker.

The dataset of this particular THS firm has three attractive features for testing the hypothesis on the general skill training done at the cost of the employers. First, we can infer the length of the initial training period from the time

[1] The fee is the lower bound of the MPL of a worker at the client firm because the client firm will not hire a worker if the fee is higher than MPL. More specifically, if the product market (i.e. the THS service market) is perfectly competitive, the client firm will hire workers until the MPL is equal to the fee. In this case, the fee corresponds to the MPL of the marginal worker, while the MPL is greater than the fee among non-marginal workers. Alternatively, if the client firm has market power over the THS firm, they reduce the service purchase to suppress the fee; while if the THS firm has the market power over the client firm, the THS firm will reduce the service supply to increase the fee. In either case, the MPL at the client firm exceeds the fee.

between the start of a worker's employment with the firm and their assignment to a client. During this training period, workers are paid their full monthly salary and are not involved in the production activity. Thus the workers do not pay the training cost in the form of receiving lower wages than productivity. Second, the IT related skills acquired through the training is transferable across employers because workers are assigned to various clients.[2] Third, and most importantly, we directly observe the fee charged to the clients, that is the each worker's MPL, together with their wage. These features suggest that data provide an ideal opportunity to test the wage compression hypothesis.

We first document the career paths of workers in the firm by analyzing the length of the initial training period and their tenure at the firm after the training period. This confirms that the THS firm typically provides training prior to the initial assignment to a client: the average training period is 2.2 months and varies across employees. Using survival analysis incorporating the right censoring of tenure length we find that workers with longer training periods, on average, have lower hazard rates and longer tenure with the firm. We also find that workers with high service fees charged to clients at initial assignments have substantially higher hazard rates, and that university graduates have lower hazard rates than non-university graduates. These results imply that the composition of workers changes with length of tenure at the firm, and so controlling for such changes is important when estimating tenure-fee and wage profiles, as emphasized in literature on the return to tenure (e.g. Altonji and Shakotko, 1987; Abraham and Farber, 1987; Topel, 1991).

We next examine how the hourly fees charged to clients and wages paid to workers are determined after workers are assigned to clients. Once a worker is assigned to a client, the initial fee is about 38% higher than their initial wage

[2]As further evidence, the workers typically receive the network engineer certificate issued by CISCO that is widely recognized in the industry.

on average. The THS firm charges higher initial fees for workers who receive more training but does not pay corresponding higher wages, thus the initial markup rate is higher for those with more training. Although the algorithm determining the length of training is complicated and likely endogenous, this result is consistent with the wage compression hypothesis, and suggests the firm may partially recover the cost of longer training by increasing the gap between the fee and the wage.

We then track the evolution of fees and wages over the course of workers' tenure with the THS firm, with a motivation to observe the MPL and wage returns to skill acquired through learning-by-doing on assignments to client firms. Controlling for observed worker characteristics and worker fixed effects, over the first 15 months of tenure, workers' wages are essentially constant, while the fees charged to clients increase linearly at an annual rate of about 6%, so the markup increases similarly. After 15 months, fees increase at nearly 8% annually while wages increase at about 5.3%, resulting in a continuing annual increase in the firm's markup of about 2.5%. Controlling additionally for client fixed effects reduces the estimated annual fee growth by about 1% over the first 15 months and 2% after that, but has almost no effect on the estimated wage growth. This implies the firm is able to increase the fee charged by assigning the workers to clients with higher skill requirement, and workers do not share any of these gains in terms of higher wages. The difference between fee and wage growth associated with client switches suggests a further source of labor market friction. Each of these results suggest that the firm captures a premium on the return to skill acquired through learning-by-doing.

Finally, we consider the value to the THS firm of hiring and training workers. We do this by estimating the internal rate of return associated with the skill investment to the firm over a 10-year horizon. We infer the cost of training from the training period length and initial monthly salary. Based on the analysis of the training period, tenure length, and the evolution of the markup, we calculate the expected return as the product of the expected probability of staying with the THS firm and the expected markup. To calculate the true rate of return, we need to know the indirect cost associated with hiring an additional worker such as the cost of hiring, administrative cost of dispatching the employees to the clients, and the employer's contribution to the social security insurance. Using available information for the industry, we assume the firm's fixed costs of operation account for 21.6% of its wage costs, and that the cost of hiring a worker is approximately 200,000 JPY. From this, we estimate the IRR across all workers is 25.5%, with higher IRR among university graduates. This suggests that the firm can improve by modifying the recruitment policy toward university graduates, which is consistent with the actual policy change of the firm.

Our study contributes to the literature by directly testing the wage compression hypothesis first proposed by Stevens (1994) and further developed by Acemoglu and Pischke (1998, 1999a,b). Some of these studies provide evidence that is consistent with the theoretical prediction, but do not directly show the productivity return to skill investment is larger than the wage return at individual worker level. Using firm level data, research generally finds positive effects of training on productivity, that is often larger than the effects on wages, which implies that firms earn some of the returns to training and so have incentives to pay for it. For example, Dearden et al. (2006) estimated that a 1 percentage point (pp) increase in the fraction of workers receiving training increased value-added per worker by about 0.6% and average wages by 0.3% for firms in the UK, and Konings and Vanormelingen (2015) estimated a 1pp increase in the fraction of workers trained increased productivity by 0.17–0.32%, and average wages by 0.1–0.17% for Belgium firms. In contrast, recent evidence by Morikawa

(2021) for Japan finds training has low but similar effects on both productivity and wages, with elasticities of about 0.02. Our study adds to the literature by showing the gap between MPL and wages based on individual employee level data.

Our study also contributes to the understanding of the operation of THS firms. In the context of upfront training provided by the firm, Krueger (1993) reports that about 60 percent of THS firms that provide secretarial services to the client firms offer computer training to its workers before assigning them to the clients and almost all the firms do so at the cost of the THS firms. Autor (2001) develops a specific model of THS firms to explain the upfront training offered to the workers. Autor demonstrated that THS workers who received training from firms receive lower wages, which is consistent with the theoretical prediction. However, as his worker data does not contain the information on the fees charged to clients, a complete test of the theory was not possible. Our study fills this gap in the literature.

The rest of the paper is organized as follows. In the next section we begin by describing the THS firm's data used in the analysis. In section 3 we present a simple model based on Acemoglu and Pischke (1999b) to help motivate our analysis. We then document the patterns of the initial training provided to workers, and their subsequent tenure in section 4. In section 5 we present and discuss the main results of our analysis of the dynamics of workers' fees and wages, and the implications for the rate of return to training provision in section 6. The paper then concludes with a summary discussion.

2 Data description

The main data used in this study is obtained from a THS company, focusing on the Information Communication Technology (ICT) industries. The company is based in the Kantō region, and has several branches located around Japan. The firm employs workers to provide a variety of temporary placements with clients to perform ICT-related tasks for varying period lengths. Throughout our discussion, we will refer to the temporary-help company as the *firm*, the workers it employs as *workers*, and the client firms they are placed in as *clients*.

The THS firm hires its workers on permanent contracts including the periods the workers are not assigned to clients are used for training. This is in sharp contrast to the typical THS firms that hire workers on a contingent temporary contract basis to cover the service period provided to the client firms. The THS firm employs both non-college and college graduates, as well as workers with and without prior ICT-industry experience. The firm gives intensive training to its new employees before placing them to clients. As we analyze in detail based on administrative records, the new employees receive intensive training in training rooms in the corporate head quarter (Panel A of Figure 1). The training program emphasize the hands-on instruction, and trainees are assigned problems and solve the problems as a team (Panel B of Figure 1). The training curriculum includes the recovery of the server: the instructor intentionally sets the problem on the server and the trainees are supposed to diagnose and fix the problems (Panel C of Figure 1). At the end of training period, the trainees are encouraged to obtain Cisco's CCNA certificate, which is the entry level certificate for the network and program engineer.

Workers who complete the initial training program are then assigned to client firms. The clients are typically large firms that attempt to absorb the demand fluctuation by procuring IT related services through the THS firm. The workers who are assigned to clients onsite are involved in network and server maintenance or software development. While client firms may poach workers to avoid paying the fee-wage margin, from discussions with the management this appears to occur infrequently for two reasons. First, the large client firms tend to have high skill requirements corresponding to their high wage scale, and many workers cannot clear the bar. Second, the THS and client firms generally have an ongoing relationship, and the THS firm has some bargaining power over client firms by assigning a group of workers: if a client porches workers, the THS firm can abruptly stop assigning workers to retaliate. In addition, because of stringent employment protection laws in Japan,[3] large firms tend to commit to long-term employment and use THS workers for short-term assignments and to absorb demand fluctuations. Although being poached by client firms is not common, workers may quit the firm presumably because they accumulate the skill through the initial training and clients' onsite learning-by-doing and receive better wage offers from outside firms.

The main dataset consists of a single pay record of each worker-client pair in each month, covering the period April 2015 to February 2020.[4] Each record includes the worker and client identifiers, the monthly fee the client is charged for the worker, the worker's monthly wage, and their hours worked for (i.e. charged to) the client in the month. In addition, workers' background characteristics, such as gender, age, education level, the date of entering the temporary help firm, and the branch location the worker is registered at, are collected and merged to the main dataset based on the worker's ID.

[3]Japanese employment contract law requires firms to prove 1) the need for termination of the employment contract, 2) the possibility of reallocation of the worker within a firm is exhausted, 3) the selection of the terminated worker is fair, and 4) the procedure for the termination is according to formal procedure. If the firm fails to prove these conditions are satisfied, the dismissal is judged as unjust and the judges request the reinstatement of the worker.

[4]There are a small number of cases with multiple records recorded in a single month for a worker-client pair. According to a manager of the firm, this may occur due to billing additional charges, correcting for mistakes, or duplication of a record. For the first two cases, we need to sum multiple records to obtain the monthly amount. For the last case, the duplicated record should be dropped. To address these cases, we keep one record with an imputed fee and wage. To do this, we calculate both the sum of monthly fee (wage and hours worked), and the average of each from all duplicated records for a worker-client pair in a month. Imputation is then made by choosing either the sum or the average that is closest to the client mean level, calculated over all worker-months. The hourly fee and hourly wage are then calculated based on the imputed data. Around 2.7% of the records are dropped in this adjustment.

We first restrict the sample to observations on workers who joined the temporary-help firm in or after April 2015, when the earliest pay record is available. As workers only appear in the data once they are assigned to clients, new employees being observed implies that they start working at the clients right after joining the temporary-help firm. Shorter than average training periods are possibly associated with higher skill or longer prior experience, leading to a positive selection concern. For this reason, workers entering after November 2019 are also excluded from the sample to ensure a minimum of three-month-stay at the firm over the observation period that ends in February 2020. Workers at a branch are excluded due to the small sample size. Because the wage-tenure relationship appears to become relatively unstable over long tenure range, we also drop observations with tenure greater than 48 months (the 99th percentile). The process above gives us the full sample consisting of 35,414 observations from 1,908 workers and 412 clients.

We calculate the hourly fee and wage by dividing the monthly fee and wage by the hours worked, and calculate the relative markup by dividing the fee by wage. The worker's tenure with the THS firm is defined by the total number of months from entering the firm to the current month of record. We define the initial training period (discussed in detail later) as the number of months from when a worker joins the firm until they are assigned to their first client. We estimate the worker's potential work experience in years as *(age - years of education - 6)*.

Because workers commonly start or end a placement *during* a month, the fee charged for the first and last placement months is typically lower than the intervening months reflecting the shorter actual service hours. In contrast, the worker is paid their regular full-month wage regardless of the shorter hours worked at the client. Thus, calculating the hourly wage for these months by dividing the monthly salary by the service hours provided to the client is misleading because the calculated hourly wage does not accurately correspond to the compensation for the labor service provided to the client. Instead, we replace the hourly wage in the first and last month of each worker working at each client with the second and second-to-last month values respectively. A consequence of this is that we require worker-client spells to last at least three months, and those less than 3 months are excluded, since the 'regular' monthly hours worked and thus wage rate are not available. This restriction results in the exclusion of about 2% of monthly observations: our main analysis sample consists of 34,729 observations from 1,784 workers and 376 clients.

Table 1 provides summary statistics of the original sample and the analysis sample. The comparison of the means for the full sample reported in Column (1) and the analysis sample in Column (2) indicate how dropping the worker-client pair that lasts less than three months affects the sample characteristics. Except for hourly wage, the means of the variables of the two samples are almost identical. As for hourly wage, the mean wage of the analysis sample is about 10 percent lower than the mean wage of the original sample. This lower average wage is largely due to the adjustment to the hourly wage in the first and last month for each worker-client pair.

Focusing on the analysis sample reported in Column (2), female workers make up about one-third of the sample. The firm mainly employs of younger workers, with their average age being 27 years, and average (post-education) experience of 6 years. About two-thirds of workers at least hold a bachelor's degree. Workers' average initial training period is 2.2 months (9-10 weeks), and average tenure is 16 months suggesting high turnover.[5] The average hourly fee

[5]Note that this average tenure is measured across all monthly observations. The average maximum tenure across all workers is 22.3 months, and the average completed tenure across workers with completed spells (30.7% of all workers) is 20 months. Right censoring will be discussed further in section 4.2.

charged to the clients is 2,900 yen, while the average hourly wage is 2,000 yen, and the average fee/wage ratio is 1.43. The average hourly wage is slightly lower than the national average of 2,300 yen and substantially lower than the internet related service industry average of 2,860 yen.[6] On the other hand, the average fee/wage ratio is slightly lower than the national average of 1.53.[7] The average hourly wage The monthly average billable hours worked is 156 hours, which is about the hours worked by a full-time workers (8 hours per day for 20 days). The initial hourly fee is slightly lower than the average hourly fee, consistent with there being fee growth. In contrast, the initial hourly wage is slightly higher than the average hourly wage: as we discuss in detail below, this occurs because of the negative selection of workers over tenure. The initial fee/wage ratio is 1.38, which is also lower than the average fee/wage ratio, suggesting markup grows with tenure.

The gender differences in the analysis sample, shown in columns (3) and Column (4), are rather minor. Males are about 1 year older than females, with correspondingly more potential experience when they join the firm. Males also receive on average 0.2 months less initial training, and have about 1 month longer tenure, than females. Males are less educated than female: 61 percent of males have a University qualification compared to 67 percent of females. Fees and wages are quite comparable between genders. The relatively minor gender differences in fees and wages suggest that we can pool both male and female in the analysis.

[6] According to the Basic Survey of Wage Structure of 2017, the average monthly regular cash compensation was 333,800, the average bonus compensation in the previous year was 905,900, the average scheduled monthly hours was 165 and the average overtime was 13 hours. The average hourly wage is calculated as (333,800+905,900/12)/(165+13) = 2,300. The average hourly wage among employees in internet related service industry was 2,860 yen based on the same method.

[7] Mean hourly fee of THS service was 2,644 yen and hourly wage was 1,729 yen in 2017 according to Annual Report of THS Service by Ministry of Health, Labor and Welfare. Implied relative mark up is 1.53.

3 Theoretical Background

In this section, we present a simple two period model for the firm's decisions regarding training, wage and fee settings. Our model captures the essence of Acemoglu and Pischke (1999b), and aims to motivate our empirical analysis.[8] We assume constant returns to scale in order to abstract from the determination of the number of workers employed.

In the first period, the firm hires a worker at wage w_1 and trains her with the intensity τ with the cost of training $c(\tau)$. Note that the opportunities for skill formation τ is provided through either formal training or assignment to client firms that enables learning-by-doing in our context. The cost function is a strictly convex function and satisfies Inada conditions $c' \geq 0$, $c'' > 0$, $c'(0) = 0$ and $c'(\infty) = \infty$. No production takes place in the first period. The training amount τ is public information and outside firms observe it. We assume that a fraction $(p : 0 < p < 1)$ of workers quit between the first and the second periods for exogenous reasons.

In the second period, the THS firm produces a service flow by assigning the worker to a client firm, which is represented by the fee charged to the client, $f(\tau)$. Note that the fee $f(\tau)$ is the marginal product of labor from the view point of the THS firm, while it is not necessarily so from the view point of the client firm (i.e. the fee provides a productivity lower bound for the client). Given the outside option of the worker $v(\tau)$, the firm pays a wage w_2: $w_2 = v(\tau) + \beta(f(\tau) - v(\tau))$, where β $(0 < \beta < 1)$ is the Nash bargaining power of the worker.

[8] Although Autor (2001) explicitly models the operation of THS, we do not adopt his modelling because the source of the labor market friction is the information asymmetry between incumbent THS and outside firms. Modelling information asymmetry as a source of labor market imperfection is similar to Acemoglu and Pischke (1998). These models predict a positive selection of workers over workers' tenure because incumbent firms terminate the contracts with low ability workers, but our empirical results show the opposite. Thus, we do not employ these models.

The firm's problem is to maximize the following profit expression:

$$\pi(\tau) = (1 - p)(1 - \beta)(f(\tau) - v(\tau)) - (c(\tau) + w_1), \tag{1}$$

assuming a zero discount rate. The first order condition is

$$(1 - p)(1 - \beta)(f'(\tau^*) - v'(\tau^*)) = c'(\tau^*). \tag{2}$$

With the above assumptions on the cost function, β and p, $\tau^* > 0$ if and only if $f'(0) - v'(0) > 0$. This condition requires that the marginal return to training in terms of the service fee must be higher than that in terms of the wage for the training investment takes place. This condition is known to be *wage compression* in the literature and we test if this condition holds in terms of skills acquired through upfront training and learning by doing.

4 Upfront training and workers' tenure

The theories of firm provided general training argue that the employer provides the general training upfront and recoups the investment cost over time from the retained workers. In this section, we examine how much training the THS firm provides, and how the firm succeeds in retaining its workers.

4.1 Length of training period

The THS firm provides IT skill training upfront. How intensive is the training? While we do not have direct record of training participation, all the workers including trainees are employed on a full time basis, thus we can infer the training period from their date of the entry to the firm and the first month placed with a client.

To describe the training period inferred from the dataset, Figure 2 presents the distribution of the length of the initial training period, measured as the number of months between when a worker is hired by the firm and first assigned to a client. This figure implies that the training period typically lasts for 1-3 months for most of the workers. That is, about 3% of workers are placed with a client in their first month of employment, while about three quarters (74%) of workers have 1-2 months of training before placement, 13% have 3 months, and the remaining 10% have 4 or more months of training before being placed. The median and modal training period is 2 months, and the average is about two and a quarter (2.3) months.[9]

The length of training varies across workers for several reasons. In theory, both positive and negative self-selection occurs. If a worker who is identified as eligible receives extended training so that the firm can assign them to a project (client) with high skill requirement, then the ability of the worker and the length of training is positively associated. On the other hand, the firm may extend the training period for slow learners. In this case, the ability of a worker and the length of training is negatively associated. A corporate executive claims that both cases occur, but the positive self-selection is more probable because the firm often trains eligible workers for a longer period to assign them to projects with high skill requirements.

To examine whether there is significant heterogeneity in the training period across workers' demographic characteristics, we regress the number of months of training on workers' observed characteristics, and present results in Table 2. In column 1, we tabulate the OLS estimates without controlling for the entry cohort fixed effect. These results confirm there is a statistically significant

[9] The mean tenure is slightly different from the mean tenure reported in Table 1, which is 2.2 months, because Table 1 reports the descriptive statistics calculated over worker $\times$ month observations whereas the descriptive statistics reported here is calculated based on worker observations.

gender difference in training, with women and workers with university education receiving about one-quarter of a month (1 week) more training than men and those with less education on average. But differences across other dimensions are not statistically significant. The finding that university graduates receive longer training is consistent with the finding in the literature that educated workers are more likely to participate in training programs (Brunello, 2004; Ikenaga and Kawaguchi, 2013). Since the data set covers various entry cohorts, we estimate the same model with cohort fixed effects. The results reported in column 2 are not substantially different from those in column 1.

To summarize the findings from the training period analysis, we find that workers in this THS firm receive around two and a quarter months of upfront training on average. The length of training period varies across gender and qualification groups, with females and university qualified workers receiving about 10% (1 week) longer training periods on average. The positive correlation between academic credentials and the training length suggests a presence of positive self-selection at least on average.

4.2 Job tenure with the THS firm

The THS firm potentially recoups the cost of training from retained workers, through the surplus (markup) between the fee charged to clients and the wage paid to workers. Thus the length of tenure of its workers critically determine the return from the upfront general skill investment.

One feature of the tenure length variable is right censoring associated with ongoing tenure at the end of the sample period. The fraction of the workers with right-censored spells (i.e. are still employed at the sample end) is 69.3%. Figure 3 draws the Kaplan-Meier survival estimate that indicates the probability of staying with the THS firm by the month of tenure, addressing the right-censoring issue. The figure shows there is little separation during the first six months of tenure and separation then occurs at a fairly constant rate after that. The probability of workers staying with the THS firm after 48 months is slightly less than half.[10]

Our goal is to estimate the growth rates of fees and wages, but workers' composition changes over tenure if the workers' complete tenure are different across workers' observed and unobserved characteristics. As the literature on the return to tenure shows, the systematic change of the workers composition poses a challenge in the estimation of the growth rates of fees and wages along with tenure.[11]

As a simple way to illustrate the selection over tenure, Figure 4 shows the means of hourly initial fees and hourly initial wages by the length of tenure. Both initial fees and wages are individual specific and if the attrition occurs at random, the means of these variables should be constant over tenure. In contrast, the mean of the initial fee decreases over tenure length, suggesting that the employees with high initial fees are more likely to quit. On the other hand, we do not observe a systematic change in mean wages by employees' tenure. The high initial fee arguably captures the high skill of the employees and thus decreasing mean initial fees over tenure implies that the employees are negatively selected over tenure.

Some workers leave the THS firm early and others have long tenures. To examine the determinants of the tenure, we attempt to characterize the determinants of tenure. To handle the right censoring of the tenure variable, we estimate the duration model. Among the parametric duration models, we choose

[10]To give further idea of length of tenure with the firm, among workers who started with the firm in 2015 (the first year of observation), 56% have right-censored spells, the average maximum tenure of workers is 39 months, and the average completed tenure (i.e. among those who are not right-censored) is 32 months.

[11]Altonji and Shakotko (1987), Abraham and Farber (1987), and Topel (1991) are the representative works in the field.

the log-normal as the baseline hazard function among alternative baselines, such as Exponential, Log-logistic, Weibull and Generalized Gamma, using the Akaike information criterion.[12] Figure 3 shows that the survival rate predicted with log-normal model is similar to the Kaplan-Meier estimates.

We also attempt to characterize the composition changes of workers' quality over tenure. To capture the unobserved workers' quality, we include the initial fee charged to the client and the initial wage paid to the worker as explanatory variables in the log-normal hazard model. We also control for a quadratic in potential years of labor market experience, and indicator variables for female, 4-year university graduates, and the firm's branch location.

Table 3 reports the estimates of the log-normal hazard model. The estimated coefficients show the effects on the hazard relative to the baseline hazard rate: a coefficient larger than 1 implies higher hazard rate than baseline, and smaller than 1 implies lower hazard rate than baseline. We use the initial fee charged to the client, the initial wage paid to the worker, and the initial markup rate (defined as the ratio of the fee to the wage), to proxy for unobserved workers' characteristics. Since these three variables are highly co-linear, we include each variable separately in turn.[13] We also consider alternative specifications to handle the heterogeneity in monthly fees (wages or markup rates), and hours worked: first including the monthly fees or wages together with the hours worked as a separate explanatory variable; and second including *hourly* measures of fees, wages and markup rates.

The first three columns of Table 3 report the regression estimates of the specification using the initial *monthly* fees, wages and the markup, along with the hours worked as the explanatory variables. First, we find that the length

[12] The estimates based on Cox proportional hazard model, where the shape of the base line hazard function is not specified, render almost identical estimates.

[13] We have also estimated specifications that include both the initial fee and wage. This results in the respective estimated coefficients becoming extenuated relative to those presented in Table 3, but otherwise the results are largely consistent.

of a worker's initial training period is (positively) associated with lower hazard rates, hence longer tenures with the firm.[14] For example, the estimates imply that an extra month of training is predicted to lower the hazard rate by about 2.5 percent. Interpreting this effect is complicated by possible endogeneity of the training offered by the firm. For instance, the firm may provide more intensive training for the workers that are expected to have longer tenures. For this reason, we do not interpret it causally; nonetheless, it does suggest training may have positive effects on workers' tenure at the firm.

Column 1 of Table 3 shows that workers with high initial monthly fees are (statistically significantly) more likely to separate: a 10 percent higher fee is predicted to increase the hazard rate by about 4.9 percent ((1.486-1) $\times$ 0.1). This large coefficient implies workers are dynamically *negatively* selected over tenure. While workers with the high initial fee are attractive to the THS firm, this result suggests the firm struggles to retain them under the current wage scheme. The column 2 of Table 3 shows the converse that, workers with high initial monthly wages have statistically significantly lower hazard rates. Consistent with these results for fees and wages, the results in Column 3 show that a high initial fee-wage margin significantly increases the hazard rate. Initial hours worked do not appear to have systematic effects on the hazard rate.

The last three columns of Table 3 replicate the results using the initial *hourly* fees, wages and the markup. The estimates are broadly similar to those in the previous columns, although with more muted effects: workers with higher initial hourly fees or markups are more likely to separate, while those with higher wages are less likely to (but not statistically significantly so). Consequently, the workers with low outside options stay with the firm and as such workers will be

[14]This is consistent with Royalty (1996), who finds that training is associated with lower turnover. The effects are statistically significant in the specification including either the initial fee or markup, but not with initial wage (columns 2 and 5); however, the coefficients are similarly sized across the specifications.

negatively selected over tenure.

In addition, the results in Table 3 imply there are important composition change of workers in terms of observed characteristics. Across the specifications, we robustly find that the workers with 4-year university degree are about 9-13 percent less likely to separate at any moment of the tenure. In this regard, workers are *positively* selected over tenure. Furthermore, the gender differences in the hazard rate suggest female workers are 8-10 percent more likely to separate than the male workers, though they are not statistically significant.

To summarize the findings from the survival analysis of the tenure length, we find that the workers with longer training periods are less likely to separate, while those with high initial fees are more likely to separate. Thus, the average initial fee decreases as the tenure increases because of the composition change of workers. On the other hand, 4-year university graduates are systematically less likely to separate, thus the fraction of workers with 4-year university degrees increases as tenure deepens. In the end, the workers' selection over tenure is nuanced and complicated. Thus, the estimation of fee, wage and markup growth without correcting for composition changes may suffer from either the upward or downward biases. A main take away for the fee and wage growth analysis is the importance of controlling for the composition change of workers both in terms of unobserved and observed characteristics.

A few comments on the relevance of the survival analysis results and the theoretical predictions. According to the models that generate wage compression because of the information asymmetry in the labor market (Acemoglu and Pischke, 1998; Autor, 2001), the gap between MPL and wages originates from the information rent. That is, the incumbent firm selects only high ability workers based on their private information. Thus, the canonical model predicts dynamic positive selection of workers. In contrast, we find the evidence of dynamic negative selection of worker over tenure, most probably because high skilled workers receive better outside offers. Thus, our empirical findings are not consistent with the prediction of wage compression due to the information asymmetry.

5 Fees and wages

The THS firms presumably attempts to recoup the upfront cost of general skill investment from the gap between the fees charged to clients and wages paid to retained workers. In this section, we analyze first how the initial fees and wages are determined, and then how these variables evolve over workers' tenure with the firm.

5.1 Initial fees, wages and markup

After the initial general skill training period, each worker is assigned to a client firm. In this first assignment, how are fees and wages determined? To address this question, we examine how the initial fees, wages and consequent markup are determined based on workers' characteristics.

First, in Figure 5, we plot the average initial assignment fee and wage by the length of training period in month. The left panel shows that the length of training and the average initial fee are positively correlated, with the average fee increasing almost monotonically with length of training. While the average fee among workers who receive at least 6 months of training is relatively high, as shown in Figure 2 few workers receive this amount of training. In contrast, the right panel plots the average initial assignment wage by the length of initial training, which shows no obvious relationship between the length of training and the average hourly wage. This figure suggests that MPL increases with the length of training but wages do not.

Next, we examine the relationship between the length of training and the fee and wages, conditional on observed characteristics of workers. Table 4 tabulates the regression results of initial fees, wages and markups on the workers' characteristics in the initial month and the length of their initial training period. The explanatory variables include the length of training period, a quadratic in potential years of labor market experience, and indicator variables for female, university graduates, and the firm's branch location. First, consistent with Figure 5, we find that workers' initial training has a positive association with the initial fee charged to clients (each month of training is associated with 1.5% higher fee), but has zero correlation with the wages paid, and so is also positively associated with the initial markup. Establishing the causal impact of training period on fee, wage and the mark up is difficult because the length of training is a choice variable of the firm and is likely to be endogenous. For instance, the firm may prolong the training period of those workers who exhibit high ability during the training period and dispatch such workers to the clients charging high fees. However, these patterns are consistent with the theoretical prediction that the workers' skill increases the productivity of worker at incumbent firm but does not increase the outside option of the worker. Thus, this finding suggests that the firm offering the general training is able to recoup the cost of investment due to the rent created by the friction in the labor market (Acemoglu and Pischke, 1999b).

We also find statistically significant effects of potential experience on fees, wages and markups. Column 1 reports the estimates for the initial fee regression: these show there is a roughly linear relationship with potential years of experience: a worker with one year longer potential experience receives 1.4 percent higher fees. In contrast, the relationship between initial wages and potential experience, reported in column 2, is convex (i.e. positive second derivative). The combination of linear fee and convex wage generates a concave relationship between the initial markup rate and potential experience, as seen in column 3. The markup rate is increasing with respect to the potential labor market experience until 15.6 years of potential experience (i.e. $15.6 = 0.014/(2 \times 0.00045)$). Given the average potential experience in the analysis sample is 6.4 years, this implies the THS firm gains a higher margin by hiring workers with more potential experience. Again we find evidence that the skill of workers, approximated by the potential years of experience, is positively associated with the fee but not with the wage. While the potential years of experience is public information equally observed by the incumbent firm and the outside firms, the THS firm seems to capture the rent from the labor market friction.

The estimated coefficients on the Female indicator variable imply there are no significant gender effects on initial fees, wages or markups. We estimate no significant effects of University graduates on initial fees, wages or markups.

5.2 Growth rates of fees, wages and markups

Thus far we have analyzed the effect of the initial training on the initial assignment fees and wages to analyze the returns to skill accumulation through formal training. The workers' skill are also formed through workers' experience at client sites via learning-by-doing or on-the-job-training.[15] We now examine how workers' skill acquired on the job affects their fees and wages. For this purpose, we analyze the growth rates of fees, wages and markups with tenure to shed light on the division of the return of skill upgrading between the firm and the workers.

Our analysis begins with a linear returns to tenure model, which captures the main results. But, based on the empirical pattern of wage growth, we then extend this baseline model to consider a linear spline model.

[15]Distinction between learning-by-doing and on-the-job-training is conceptually clear as articulated by Heckman et al. (2002), but empirical distinction is difficult with our data.

5.2.1 Baseline linear model

For our baseline analysis, we estimate alternative specifications of the linear tenure model:

$$\ln(Y_{ijt}) = \beta_1 Ten_{it} + \beta_2 Train_i + X_{it}\gamma + c_i + d_j + u_{ijt}, \tag{3}$$

where Y_{ijt} is either the hourly fee, wage or markup of worker-i at the client firm-j in month-t; Ten_{it} is the worker's current tenure in months (measured in years); $Train_i$ is the length of training period; X_{it} is a vector of control variables; c_i and d_j are worker and client fixed effects respectively; and u_{ijt} is an idiosyncratic error term. We estimate the model by the weighted least squares using the service hour of each month as the weight. We calculate the standard errors robust against the clustering within an individual employee. In contrast to the literature (e.g. Abraham and Farber, 1987; Altonji and Shakotko, 1987) that emphasizes the importance of job matching in estimating the returns to tenure in wages, we do not require controls for the worker-firm match effects, because our data comes from a single firm and the worker fixed effects fully captures the worker-firm match effects. In the specification with the worker fixed effects, the training period is absorbed in the worker fixed effects. Furthermore, we are able to control for client fixed effects to examine the contribution of changing clients on the evolution of the firm's fees and a worker's wages.

Due to the standard identification problem associated with co-linearity of cohort, age, and time effects, we cannot include both year-month and individual fixed effects along with the tenure length. That is, conditioning on individual worker fixes the starting date and thus adding the tenure length exactly matches a specific year and month. Instead, we control for regional time varying labor market effects using the quarterly unemployment rate measured for nine regions, and regional inflation using the consumer price index (CPI).[16] Additional control variables include the length of the worker's initial training period, gender, education, a quadratic in initial potential experience, the firm-branch.

As expressed, equation (3) assumes that the natural logarithm of hourly fees, wages and the markup depend linearly on tenure length. In order to check this, we have estimated this equation with separate dummy variables for each tenure month to allow non-parametric tenure profiles. More specifically, we estimate the model:

$$ln(Y_{ijt}) = \sum_{s=1, s\neq 4}^{48} \beta_s \mathbb{1}[Ten_{it} = s] + X_{it}\gamma + c_i + u_{ijt}. \quad (4)$$

The model includes the individual fixed effects because the previous analysis points to the importance of the selection.

In Figure 6 we plot each of the estimated hourly fee, wage and markup tenure profiles (together with their 95 percent confidence intervals) from a specification that also includes observable controls and worker fixed effects.[17] The pattern of fee growth appears remarkably linear, with fees increasing at about 5% annually. In contrast, wages appear roughly flat over the first 18 month or so, before rising approximately linearly and in parallel to fees after that. These patterns imply the markup increases at approximately the same rate as fees over the first couple of years, and then much slower after that as wages increase. Given these patterns, we will estimate both simple linear specifications for each of the (fee, wage and markup) outcomes, as well as linear-spline versions allowing for

[16] The unemployment rate is based on monthly Labor Force Survey. The finest unemployment rate published is at nine regions and quarterly periods to assure the precision of the estimates. The 2015-base monthly CPI for ten metropolitan areas is published by the Statistics Bureau of Japan.

[17] We have also estimated specifications controlling only for observable characteristics, and also including client or worker-client fixed effects (analogous to models 1–5 described below). The profiles are similar in terms of the linearity of fee growth and non-linearity in wage growth to those in Figure 6 when client fixed effects are also included, but generally steeper when worker and client fixed effects are excluded. A similar exercise for monthly hours worked shows average hours decline somewhat with tenure.

a break in trend after a certain threshold.

We begin by summarizing the results from models with linear tenure profiles. Table 5 tabulates the tenure coefficients from five alternative regression specifications for equation (3) for hourly fee in column (1), hourly wage in column (2), and hourly markup in column (3). In the first model, we include only the vector of control variables in addition to tenure, and estimate statistically significant positive effects of tenure on each outcome, of 1.0% per year for the hourly fee charged, 0.4% for the hourly wage, and the difference between these (0.6%) for the hourly markup. When we include worker fixed effects (model 2), consistent with the tenure patterns described in footnote 17, the annual tenure effects are substantially higher than those for model 1. In particular, we estimate fees increase 7.2% annually, while wages increase 3.4%, and the markup wedge increases 3.7%. The substantial downward bias of the tenure profiles of the OLS estimates reflects the negative selection of employees over tenure. Thus, we treat the model estimates with employee fixed effects as preferred estimates.

In the subsequent models presented in Table 5, we also include various controls for the clients that workers are assigned to: client fixed effects (model 3), additionally controls for the client order (model 4), or worker-client effects (model 5). The estimated tenure effects are comparatively stable across these three models. The annual growth in workers' hourly wages in these models (3.2-3.4%) is very similar to that in model 2, implying wages paid by the firm are independent of client effects. In contrast, the estimated growth in the hourly fee charged by the firm for workers is substantially lower in model 3 (5.5%) than that estimated in model 2 (7.2%); and as a result there is also variation in the estimated effect on the hourly markup across these models. Adding client order (model 4) or worker $\times$ client fixed effects (model 5) has little effect on the estimated growth rate (5.0%), implying that the worker-client match does not play important role.

Comparing the estimates of fee growth from models 2 and 3 implies nearly one quarter (1.7%) of the 7.2% growth in model 2 is associated with the firm improving the assignment of workers to clients paying high fees. However, the wage growth estimates imply none of this improved client quality effect is passed on to the workers in terms of higher wages.[18] The finding that THS firm assigns its experienced employees to high-fee clients over time but does not increase their wages at the timing of client change is consistent with the presence of the labor market friction. Thus the THS firm fully captures the rent due to the accumulated skill through learning-by-doing.

5.2.2 Linear-spline model

Close examination of Figure 6 suggests that the hourly wages are essentially constant until around month 18 and then grow linearly. Given this, we now relax the linearity assumption on the relationship between tenure in month and natural logarithm of fees and wages. We capture this kink in the wage profile by adopting a linear spline function with a single knot.

The linear spline model is:

$$\ln(Y_{ijt}) = \beta_{11} Ten_{it} + \beta_{12} Ten_{it} \mathbb{1}[T_{it} \geq \bar{T}] + \beta_2 Train_i + X_{it}\gamma + c_i + d_j + u_{ijt}, \quad (5)$$

where Y_{ijt} is outcomes of worker-i at the client firm-j in month-t; Ten_{it} is the worker's current tenure in months (measured in years); $Train_i$ is the length of training period; X_{it} is a vector of control variables; c_i and d_j are worker and client fixed effects respectively; and u_{ijt} is an idiosyncratic error term. As

[18] To understand the client effects on fees and wages, for workers who are assigned to at least two clients, we have also conducted an event study for fees and wages around the start date with the second client. From this, we observe steady growth in fees of about 5% annually, both before and after the client change, and a discrete jump in fees of about 10% at the time of client change. In contrast, wages show much weaker growth and no jump associated with the change in client. These patterns are consistent with the results in Table 5.

before, we weight observations by the service hours of employee i in year-month t and calculate the standard errors robust against clustering within an employee. The threshold $\bar{T}$ is the knot that determines the kink of the linear functions. We estimated the model with $\bar{T} = \{12, ..., 24\}$ and calculated the R^2 for each model. We find $\bar{T} = 15$ maximizes the R^2 and choose this as the knot point.

Table 6 summarizes results from the linear-spline tenure profiles with the knot point at 15. The estimated tenure effects for the fee models (columns (1) and (2)) are generally similar to those in Table 5, with small and statistically insignificant changes after 15 months. The estimates confirm there is essentially no wage growth over the first 15 months controlling for worker and client fixed effects, after which wages grow relatively strongly (about 5.7% in model 3). As a result of the roughly linear fee growth and linear-spline wage growth, we estimate stronger growth in markup over the first 15 months (5.8%), followed by much weaker growth (0.6%) than from the linear models in Table 5.

The difference in the estimated fee growth in the models with and without client effects in Table 6 are broadly consistent with those in Table 5. The results imply that client quality effects become more important with tenure, accounting for 0.9% of the 6.2% annual growth over the first 15 months, and 2.1% of the 7.9% growth after that. Again, we find that workers' wage growth is independent of such client quality improvement, implying the firm does not pass on any of these benefits to the workers in terms of higher wages.

Based on the results in Tables 5 and 6, and consistent with the non-parametric profiles in Figure 6, we conclude that the hourly fee-tenure profile is adequately characterized by a simple linear specification, while the wage and markup profiles are better characterized by linear-spline profiles. However, for consistency in specifications across the outcomes, we will continue to report both linear and linear-spline model results for each outcome.

5.2.3 Heterogeneous returns to tenure across employees

The estimates of the fee-tenure and wage-tenure profiles suggest that the THS firm has monopsony power among the retained workers probably because of the labor market friction. The degree of labor market friction can well be different across workers depending on their background characteristics. For example, the literature points to the difference in the labor supply elasticities between male and female explains the gender wage gap (Manning, 2013; Barth and Dale-Olsen, 2009; Webber, 2016). Motivated by this prediction, we next consider whether the tenure effects are constant across workers, or whether these vary systematically across some identifiable dimensions. To do this, we extend the equation (3) model to allow the tenure profile to vary with workers' observed characteristics:[19]

$$\ln(Y_{ijt}) = \beta_1 T_{it} + T_{it} H_i \beta_2 + X_{it}\gamma + c_i + u_{ijt}, , \tag{6}$$

where H_i is a set of demographic characteristics and other variables specific to worker-i. The vector H_i includes their initial training period, quadratic in initial experience, and dummy variables for female, 4-year university graduate and the branch fixed effects.

Table 7 summarizes the linear-tenure specification results for the hourly fee, wage and markup outcomes, based on three model specifications with various combination of fixed effects: extensions to models 2, 3 and 5 in Table 5. The estimated main tenure effects for fees are relatively similar to those in Table 5, while the main effects are lower for wages and consequently higher for markup. The estimated interaction effects are relatively consistent across the three models. Despite the length of initial training being positively correlated with the

[19] We similarly interact each of the tenure variables with worker characteristics in the spline versions of the model.

initial fee, we find no evidence that training differentially affects either the fee, wage or markup growth over tenure. We only find statistically significant tenure effects for gender wage growth (female annual wage growth is about 1% stronger than for males), although there is also some evidence of stronger wage growth for University educated workers; and fee growth by education (annual fee growth for workers with University degrees is 1-1.5% stronger than for those with less than University). F-tests for the joint hypothesis of no tenure-interactions is rejected for all models except model 5 markup.

The estimates for the linear-spline specifications of models 2 and 3 are presented in Table 8. The main tenure coefficients imply strong and essentially linear annual fee growth (about 6% in model 3), small and insignificant wage growth over the first 15 months followed by strong growth thereafter (about 4.0%), and strong markup growth over the first 15 months (about 8% in model 3) and weakly positive growth after that point. The tenure interaction effects are more complicated than in the linear models, although we again estimate that annual wage growth is around 1% faster for females than males. The experience interactions are generally statistically significant, but difficult to interpret. We again find relatively little evidence of initial training effects on fee, wage and markup growth (other than slightly negative effects on fee growth after two years).

The results in this subsection suggests the absence of any substantial heterogeneity in the fee-tenure and wage-tenure profiles. As far as growth rates are concerned, we do not find evidence for the heterogeneous labor market frictions across types of workers.

6 Internal rate of return

Thus far we have documented the wedge between fees and wages, and show that this wedge grows with the worker's tenure. This finding suggests that the THS firm potentially has an incentive to provide training opportunities to acquire general skill upfront at the cost of the THS firm. On the other hand, the survival analysis in Figure 3 showed that less than half of employees stay with the THS firm for 48 months. Considering the attrition of workers, does it pay for the THS firm to invest in the workers on average? To answer this question, we estimate the internal rate of return of the training. We approximate the cost of training by the wages paid to the workers during their initial training period, and the expected return from the training investment by the product of expected survival rate times the fee-wage gap.

In this section, we discuss the cost of training to the THS firm, and the discounted value of the accrued return to the training. In the data available to us, we are only able to observe the direct labor costs the firm must pay workers during the training periods, and not any other general fixed operational costs or training-related costs that the firm may incur. However, based on information provided by Japan Staffing Services Association (JASSA) we estimate that the firm's fixed costs of operation account for about 21.6% of its total wage costs.[20]

Given this, we calculate the cost of training to the firm as the initial wages paid to a worker (W_{i0}) multiplied by the estimated duration their initial training period (T_{i0}, the period prior to being placed with a client), and scale this up by 21.6%. Furthermore, according to the firm's management, the hiring cost is

[20]The JASSA is the industry organization of the government-approved temporary work agencies. According to JASSA data, on average the fee charged to clients consists of the direct wages paid to workers (70.0%), the employer's contribution to the social security account (10.9%), and leave payments (4.2%). Based on this, we assume the ratio of operation cost to wage payment is $(4.2 + 10.9)/70 \approx 0.216$. The remaining parts are administrative expenses (including the costs on training, customer services, internal workers, office rent, recruitment, etc.) taking up 13.7% and THS firm's profit taking up 1.2%.

approximately 200,000 JPY per worker. That is, the total cost of hiring and training a worker to the THS firm is:

$$Cost_i = T_{i0} \times 1.216 \times W_{i0} + 200,000.$$

Next, we calculate the firm's monthly flow return to the training provided to the worker as the surplus of the monthly fee the firm receives from a client (F_{it}) over the adjusted monthly cost – i.e. the scaled-up wage paid to the worker ($1.216 * W_{it}$). The expected value of the return to the training is defined as:

$$E(Return_i t) = \hat{P}_{it}(\hat{F}_{it} - 1.216 * \hat{W}_{it}),$$

where $\hat{P}_{it}$ is the estimated survival rate for worker-i in month-t, estimated using the model in column 6 of Table 3; and $\hat{F}_{it} - 1.216 * \hat{W}_{it}$ is the estimated (absolute) markup for worker-i in month-t, based on the worker fixed effect linear spline model (Model 2 in Table 6), allowing for a constant fixed cost component of 21.6% of wages. Specifically, we first estimate $log(\hat{F}_{it})$ and $log(\hat{W}_{it})$ from their respective regressions, then exponentiate each to levels and form $(\hat{F}_{it} - 1.216 * \hat{W}_{it})$.

Finally, we define the internal rate of return as the discount rate which equates the average expected discounted value of the return across workers in the main sample over a 10-year period to the average cost of hiring and training a worker. That is, the monthly internal rate of return (MIRR) to the firm is calculated as:

$$E(Cost_i) = E(\sum_{t=T_{i0}+1}^{120} (\frac{1}{1 + MIRR})^{t-1} E(Return_{it}));$$

and the annual internal rate of return (IRR) is defined as:

$$IRR = 1 - (1 - MIRR)^{12}.$$

Table 9 summarizes the estimated internal rates of return. We find that the average expected internal rate of return is 25.5% across all workers. This suggests that the firm's average rate of return associated with providing training to workers is substantial. Among the existing studies on the internal rate of return to human capital investment, Altonji (1993) for instance estimates the internal rate of return to the first year college attendance when the return to education is uncertain. He reports that the internal rate of return ranges from five to ten percent based on US data. Compared with these estimates, the estimated return to training here is substantially larger.

We examine the heterogeneity of the internal rate of returns by demographic characteristics. The expected internal rate of returns are almost identical for female (26.4%) and male (24.8%) workers. As for educational attainment, the estimated internal rate of return is substantially higher for university graduates (28.9%) than workers without University degrees (17.9%). This finding is consistent with the firm's change in the policy to hire more college graduates according to the firm's management. As for previous potential labor market experience, the internal rate of return is 25.3% for workers with 0-5 years and 26.7% for workers with 6+ years of experience, thus the internal rates of return are not much different depending on the potential years of experience.

The results here suggest that the firm expects to earn substantial returns from employing and training its workers. The provision of training increases the productivity of workers, but the firm does not have to compensate the workers by increasing the wages accordingly because of labor market frictions. There are three possible issues driving this effect of training on higher productivity.

The first is that training has a causal effect on productivity, meaning that training increases a worker's skill and enables the firm to charge higher fees to clients. The second possibility is that provision of training opportunity attracts better workers. Third, related to each of these issues is the possible non-random provision of the firm's training across workers. How much of the effect of training can be directly attributed to the effects of training depends on the extent to which the firm's training provision is attractive to prospective workers. For example, if the provision of training is a primary motivation for a worker joining the firm (even though their wages at the firm do not change), the estimates here may be attributed to the causal effect of training and the self-selection of eligible workers to the firm. However, if training is incidental to the worker's decision to join the firm, these estimates represent the marginal value of providing training. Disentangling these effects is beyond what is possible with the current data.

As a caveat, we note that the estimated internal rate of return depends on the assumption on the total cost of putting an additional worker on the payroll. As mentioned before, we inflate the wage rate by 1.216 to include the employer's contribution to the social security account and the reserve for leave payments; and add $200,000$ as the sunk cost of recruitment. Since these numbers are not definitive, we examine the sensitivity of the estimated internal rate of return by changing these parameter values.

We first examine the effect of the choice in the sunk cost of recruitment by halving and doubling it. The second row of Table 9 reports the estimated internal rate of return when the sunk cost of training is 100,000 Yen instead of the baseline case of 200,000 Yen. The estimated internal rate of return increases by 3.3 percentage points reflecting the decreased recruitment cost. On the other hand, as reported in 3rd row, increasing the recruitment cost up to 300,000 Yen decreases the estimated internal rate of return by 2.8 percentage points. Overall, the change in the recruitment cost does not change the estimated internal rate of returns much reflecting the fact that the recruitment cost is incurred only once at the initiation of the employment.

We next change the the ratio of operation cost to wage payment, which was 1.216 in the baseline case. As explained before, the firm pay 21.6% more in addition to the wage payments to cover the employer's contribution to the social security account and the reservation for leave payments. Since Japanese government regulates the social security tax and the mandatory length of the paid leave, additional labor cost of 21.6% is arguably a reasonable approximation to the additional labor cost, we examine the robustness of the calculation results by adding and subtracting 5 percentage points to/from the baseline figure. The 4th row reports the estimated internal rate of return when we reduce the inflation factor by 5 percentage points down to 1.166. The estimated internal rate of return becomes 38.1%, a increases by 12.6 percentage points from the base line model. On the other hand, the 5th row reports the estimated internal rate of return when we increase the inflation factor down to up to 26.6%, a 5 percentage points increase. The resulting internal rate of return substantially decreases down to 12.3%, a 13.3 percentage points drop. In sum, this exercise demonstrates that the estimated internal rate of return is sensitive to the assumption on the additional labor cost over the wages.

Overall, this section shows that the estimated internal rate of return is on average substantial, around 25.5%, with a caveat that the estimates are sensitive to the assumption imposed on the costs related to the mandatory social security tax and paid leave. We need to take this high internal rate of return with a grain of salt because our calculation does not take other operation costs, such as pecuniary costs of training, costs related to managing workers dispatched to clients, sales cost. Thus, the actual internal rate of return could be substantially lower than the internal rate to return reported here. However, these ball park numbers arguably assure that the THS firm has sufficient room to collect the upfront investment cost.

7 Conclusion

We use unusually rich data from a temporary help services firm to test whether skill investment brings higher productivity return than wage return, so called *wage compression*. Our data on the fees charged to clients and the wages paid to those workers allows us to directly test the hypothesis because the fees represent workers' productivity. Drawing on this unique data set, we find three pieces of evidence that are consistent with wage compression.

First, we document that the firm provides general skill training to workers at the start of their employment spell for about 2.2 months on average. Importantly, the length of a worker's training period is positively correlated with the initial fee charged on their first client placement, but is uncorrelated with their initial wage. This is consistent with training increasing workers' productivity, as reflected by the fee charged to clients, but the higher productivity is fully captured by the firm.

Second, we test whether skill acquired through learning-by-doing induces higher fee growth than wage growth over a worker's tenure. We find the hourly fee charged by the firm increases linearly with tenure at 6-8% annually while, in our preferred (linear-spline) specification, wages are roughly constant over the first 15 months before increasing at about 5.3%. Thus, the relative markup increases strongly over the first 15 months and continues to increase at about 2.5% after that.

Third, we document the importance of client upgrading as a source of productivity growth. By comparing the estimated returns to tenure from models with and without controls for client fixed effects, we estimate that about one-quarter of the annual growth in the firm's fee charged to clients is associated with client quality upgrading. In contrast, workers' wages are independent of the clients that they are placed with, implying they do not share any of the productivity benefits associated with client quality upgrading.

Each of these three findings are consistent with the wage compression hypothesis that skill accumulation, either through formal training or learning-by-doing, increases productivity more than wages. Our empirical findings corroborate the theory that explains the investment in general human capital at the firm's cost by Stevens (1994), Acemoglu and Pischke (1998), and Autor (2001). While our analysis is from a single temporary help agency operating in Japan, the findings provide clear and consistent evidence of wage compression.

Finally, our findings also shed light on the function of THS agents in the labor market. As pointed out by previous studies (Krueger, 1993; Autor, 2001), THS firms provide training opportunities to workers and place trained workers with clients. Thus, THS agencies function as the combination of a school and an employment agency, and have direct incentives to design the curriculum in response to the skills demanded by clients. For this reason, THS providers arguably have advantages in training provision over schools in response to fluctuating demand for skills. Although policy makers may criticize THS agents for exploiting their workers via an apparently high margin of the service fee over the wages, they should also pay attention to the function of such agents as promoters of skill accumulation when they design the regulation of the industry.

References

Abraham, Katharine G and Henry S Farber, "Job Duration, Seniority, and Earnings," *American Economic Review*, June 1987, *77* (3), 278–297.

Acemoglu, Daron and Jörn-Steffen Pischke, "Why do firms train? Theory and evidence," *Quarterly Journal of Economics*, 1998, *113* (1), 79–119.

_ **and Jorn-Steffen Pischke**, "Beyond Becker: Training in imperfect labour markets," *The Economic Journal*, 1999, *109* (453), 112–142.

_ **and Jörn-Steffen Pischke**, "The structure of wages and investment in general training," *Journal of Political Economy*, 1999, *107* (3), 539–572.

Altonji, Joseph G, "The Demand for and Return to Education When Education Outcomes are Uncertain," *Journal of Labor Economics*, 1993, *11* (1), 48–83.

_ **and Robert A Shakotko**, "Do wages rise with job seniority?," *Review of Economic Studies*, 1987, *54* (3), 437–459.

Autor, David H, "Why do temporary help firms provide free general skills training?," *Quarterly Journal of Economics*, 2001, *116* (4), 1409–1448.

Barth, Erling and Harald Dale-Olsen, "Monopsonistic discrimination, worker turnover, and the gender wage gap," *Labour Economics*, 2009, *16* (5), 589–597.

Becker, Gary S, *Human capital*, University of Chicago Press, Chicago, 1964.

Brunello, Giorgio, "Education, Training and Labour Market Outcomes," in Daniele Checchi and Claudio Lucifora, eds., *Education, Training and Labour Market Outcomes in Europe*, Springer, 2004, pp. 3–7.

Dearden, Lorraine, Howard Reed, and John Van Reenen, "The impact of training on productivity and wages: Evidence from British panel data," *Oxford Bulletin of Economics and Statistics*, 2006, *68* (4), 397–421.

Heckman, James J, Lance Lochner, and Ricardo Cossa, "Learning-by-doing vs. on-the-job training: Using variation induced by the EITC to distinguish between models of skill formation," in Edmund S. Phelps, ed., *Designing Inclusion: Tools to Raise Low-end Pay and Employment in Private Enterprise*, National Bureau of Economic Research Cambridge, Mass., USA, 2002.

Holzhausen, Arne, "Japanese Employment Practices in Transition: Promotion Policy and Compensation Systems in the 1990s," *Social Science Japan Journal*, 2000, *3* (2), 221–235.

Ikenaga, Toshie and Daiji Kawaguchi, "Labour-Market Attachment and Training Participation," *The Japanese Economic Review*, 2013, *64* (1), 73–97.

Konings, Jozef and Stijn Vanormelingen, "The impact of training on productivity and wages: firm-level evidence," *Review of Economics and Statistics*, 2015, *97* (2), 485–497.

Krueger, Alan B., "How computers have changed the wage structure: Evidence from microdata, 1984–1989," *Quarterly Journal of Economics*, 1993, *108* (1), 33–60.

Manning, Alan, *Monopsony in motion*, Princeton University Press, 2013.

Morikawa, Masayuki, "Employer-provided training and productivity: Evidence from a panel of Japanese Firms," *Journal of the Japanese and International Economies*, 2021, p. 101150.

Royalty, Anne Beeson, "The effects of job turnover on the training of men and women," *Industrial and Labor Relations Review*, 1996, *49* (3), 506–521.

Stevens, Margaret, "A Theoretical Model of On-the-Job Training with Imperfect Competition," *Oxford Economic Papers*, 1994, *46* (4), 537–562.

Topel, Robert H, "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," *Journal of Political Economy*, 1991, *99* (1), 145–176.

Webber, Douglas A, "Firm-level monopsony and the gender pay gap," *Industrial Relations: A Journal of Economy and Society*, 2016, *55* (2), 323–345.

Table 1: Summary statistics

Mean (SD)	(1) Full sample	(2) Main sample	(3) Main sample: Males	(4) Main sample: Females
Female	0.324	0.325	—	—
Age (years)	27.3 (3.9)	27.3 (3.9)	27.6 (4.2)	26.7 (3.1)
Education: University+	0.63	0.63	0.61	0.67
Potential work experience (years)	6.4 (4.1)	6.4 (4.1)	6.7 (4.4)	5.6 (3.4)
Training period (months)	2.2 (2.2)	2.2 (2.2)	2.1 (2.0)	2.3 (2.4)
Tenure (months)	16.3 (10.6)	16.4 (10.6)	16.8 (10.8)	15.7 (10.2)
Hourly fee	2,889 (785)	2,885 (784)	2,866 (829)	2,924 (679)
Hourly wage	2,228 (2,561)	2,043 (551)	2,040 (590)	2,050 (459)
Relative markup (fee/wage)	1.39 (0.28)	1.43 (0.35)	1.42 (0.38)	1.44 (0.28)
Hours worked (monthly)	155.2 (27.7)	155.7 (26.8)	156.6 (26.4)	154.0 (27.7)
Initial hourly fee	—	2,763 (995)	2,763 (995)	2,763 (997)
Initial hourly wage	—	2,058 (529)	2,056 (501)	2,063 (585)
Initial relative markup	—	1.38 (0.53)	1.38 (0.54)	1.39 (0.53)
No. observations	35,414	34,729	23,453	11,276
No. Workers	1,908	1,784	1,164	620
No. Clients	412	376	288	240

Notes: Column 1 represents all the observations. The main sample used in the analysis is shown in Column 2, with hourly wage adjusted and worker-client pair lasting shorter than 3 months excluded. By this restriction, 124 workers together with 36 clients are dropped. This attrition potentially results from the workers who have only been working for clients temporarily, and the clients which have never set up long-term relationship with any workers. The mean of unadjusted hourly wage in the main sample in Column 2 is 2,187 (SD=2,256), which is close to that in the full sample in Column 1, implying that the gap in hourly wage comes systematically from the downward adjustment, not sample selection.

Table 2: Determinants of initial training period

	(1)	(2)
Potential experience at entry (years)	0.002	-0.028
	(0.040)	(0.043)
Potential experience^2/100	0.213	0.267
	(0.222)	(0.246)
Female	0.233*	0.285**
	(0.123)	(0.124)
Education: University+	0.269*	0.259*
	(0.139)	(0.138)
Entry cohort FE	No	Yes
No. observations	1,784	1,784
R^2	0.012	0.046
Sample mean	2.271	2.271

Notes: Standard errors are reported in parentheses. Entry cohort is defined by the fiscal year of entry. Each model includes branch office fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Survival estimate based on Log-normal model: Tenure

	(1)	(2)	(3)	(4)	(5)	(6)
Training period	0.974^{**}	0.979^{*}	0.973^{**}	0.975^{**}	0.978^{**}	0.973^{**}
(months)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Initial monthly fee	1.486^{***}					
(in log)	(0.212)					
Initial monthly wage		0.490^{*}				
(in log)		(0.180)				
Initial monthly markup			1.794^{***}			
(in log)			(0.276)			
Initial hours worked	0.777^{*}	0.943	0.762^{*}			
(in log)	(0.118)	(0.144)	(0.114)			
Initial hourly fee				1.392^{***}		
(in log)				(0.165)		
Initial hourly wage					1.015	
(in log)					(0.154)	
Initial markup						1.647^{***}
(in log)						(0.249)
Potential experience	1.015	1.024	1.014	1.016	1.022	1.015
at entry (years)	(0.017)	(0.017)	(0.016)	(0.016)	(0.017)	(0.017)
Potential experience^2/100	0.932	0.970	0.964	0.933	0.928	0.956
	(0.076)	(0.083)	(0.079)	(0.076)	(0.077)	(0.079)
Female	1.088	1.096	1.076	1.089	1.103^{*}	1.081
	(0.061)	(0.062)	(0.060)	(0.061)	(0.062)	(0.061)
University+	0.871^{**}	0.909	0.881^{**}	0.876^{**}	0.884^{**}	0.876^{**}
	(0.052)	(0.055)	(0.052)	(0.051)	(0.053)	(0.052)
Log lik.	-1,197.75	-1,199.69	-1,194.34	-1,198.09	-1,201.88	-1,196.42
Chi-squared	31.568	27.679	38.372	30.873	23.296	34.217
No. observations	1,784	1,784	1,784	1,784	1,784	1,784

Notes: Exponentiated coefficients are reported. Standard errors calculated as $SE(coef.) \times \exp(coef.)$ are reported in parentheses. The asterisks indicate the p-value for the null hypothesis that the coefficient is 1. Test of hypothesis is in terms of original metric. For example, for the initial monthly fee in Column 1, $t-stat = (\log(1.486))/(0.212/1.486) \approx 2.78$. The initial fee, wage and hours are measured in each worker's second month of placement; the fee, wage, markup, and hours worked variables are in logarithms. Each model also includes branch office controls but the estimated coefficients are not reported. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Initial hourly fee, wage and markup (at first pay)

	(1) log(fee)	(2) log(wage)	(3) log(markup)
Training period	0.015***	-0.000	0.015***
	(0.003)	(0.002)	(0.003)
Potential experience at entry (year)	0.014***	0.001	0.014***
	(0.005)	(0.003)	(0.005)
Potential experience^2/100	0.011	0.055***	-0.045**
	(0.021)	(0.012)	(0.022)
Female	0.012	0.002	0.010
	(0.017)	(0.009)	(0.019)
University+	0.023	0.011	0.013
	(0.016)	(0.010)	(0.018)
No. observations	1,784	1,784	1,784
R^2	0.050	0.055	0.030

Note: Standard errors are reported in parentheses. Each model includes branch fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Annual growth rates – Homogeneous linear models

	(1) Fee	(2) Wage	(3) Markup
Model 1	0.010***	0.004	0.006*
(Controls only)	(0.004)	(0.003)	(0.003)
	[0.137]	[0.159]	[0.037]
Model 2	0.072***	0.034***	0.037***
(Worker FE)	(0.004)	(0.003)	(0.004)
	[0.452]	[0.485]	[0.348]
Model 3	0.055***	0.032***	0.023***
(Worker & Client FE)	(0.003)	(0.003)	(0.004)
	[0.501]	[0.522]	[0.396]
Model 4	0.050***	0.032***	0.017***
(+ Client order)	(0.003)	(0.003)	(0.004)
	[0.504]	[0.522]	[0.399]
Model 5	0.050***	0.032***	0.018***
(Worker × Client FE)	(0.003)	(0.003)	(0.004)
	[0.536]	[0.551]	[0.437]
No. observations	34,729	34,729	34,729
Workers	1,784	1,784	1,784
Clients	376	376	376

Notes: Standard errors are in parentheses (clustered at the worker level), R^2 are reported in brackets. Observations are weighted by the hours worked of each month and each worker. All models, control variables also include gender, education level, a quadratic in initial potential experience at entry, branch, and regional CPI unemployment rate and CPI. Model 1 includes no fixed effects; Model 2 includes worker fixed effects; Model 3 includes worker and client fixed effects; Model 4 includes worker and client fixed effects, and the order of client; and Model 5 includes worker × client fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Annual growth rates – homogeneous linear spline models

	Fee		Wage		Markup	
	(1) Tenure	(2) Tenure ×After 15m.	(3) Tenure	(4) Tenure ×After 15m.	(5) Tenure	(6) Tenure ×After 15m.
Model 1	0.014**	-0.001	-0.037***	0.061***	0.051***	-0.062***
(Controls only)	(0.007)	(0.009)	(0.004)	(0.006)	(0.007)	(0.008)
	[0.137]		[0.167]		[0.040]	
Model 2	0.062***	0.017**	-0.005	0.058***	0.067***	-0.041***
(Worker FE)	(0.007)	(0.008)	(0.005)	(0.005)	(0.008)	(0.008)
	[0.453]		[0.491]		[0.349]	
Model 3	0.053***	0.005	-0.006	0.057***	0.058***	-0.052***
(Worker & Client FE)	(0.007)	(0.007)	(0.005)	(0.005)	(0.007)	(0.008)
	[0.501]		[0.527]		[0.398]	
Model 4	0.045***	0.008	-0.006	0.057***	0.051***	-0.049***
(+ Client order)	(0.007)	(0.007)	(0.005)	(0.005)	(0.007)	(0.008)
	[0.504]		[0.528]		[0.401]	
Model 5	0.050***	0.002	-0.004	0.056***	0.054***	-0.054***
(Worker × Client FE)	(0.007)	(0.008)	(0.005)	(0.006)	(0.007)	(0.008)
	[0.536]		[0.555]		[0.439]	

Notes: Standard errors in parentheses (clustered at individual level). Estimation is weighted by hours worked. See notes to Table 5 for details of the model specifications. All models are based on 34,729 observations on 1,784 workers and 376 firms. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Annual growth rates – heterogeneous linear models

	Model 2: Worker FE			Model 3: Worker FE + client FE			Model 5: Worker × client FE		
	(1) Fee	(2) Wage	(3) Markup	(4) Fee	(5) Wage	(6) Markup	(7) Fee	(8) Wage	(9) Markup
Tenure	0.078***	0.018***	0.060***	0.053***	0.016***	0.037***	0.043***	0.015***	0.027***
(years)	(0.009)	(0.006)	(0.009)	(0.008)	(0.006)	(0.009)	(0.009)	(0.006)	(0.010)
×Potential experience	-0.001	0.001	-0.002	0.001	0.001	-0.000	0.001	0.002	-0.000
at entry (years)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
×Potential $\exp^2/100$	-0.003	-0.004	0.002	-0.008*	-0.004	-0.004	-0.008	-0.005	-0.003
	(0.006)	(0.004)	(0.006)	(0.004)	(0.004)	(0.006)	(0.005)	(0.004)	(0.007)
×Training period	-0.000	0.001	-0.001	-0.001	-0.000	-0.001	-0.001	-0.000	-0.000
(months)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
×Female	0.003	0.011***	-0.007	0.005	0.011***	-0.006	0.010	0.008**	0.002
	(0.006)	(0.004)	(0.006)	(0.006)	(0.004)	(0.005)	(0.006)	(0.004)	(0.006)
×University+	0.011*	0.006	0.006	0.015***	0.007*	0.008	0.015**	0.009**	0.007
	(0.006)	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.007)	(0.004)	(0.007)
No. observations	34,729	34,729	34,729	34,729	34,729	34,729	34,729	34,729	34,729
Workers	1,784	1,784	1,784	1,784	1,784	1,784	1,784	1,784	1,784
Clients	376	376	376	376	376	376	376	376	376
R^2	0.454	0.486	0.351	0.502	0.523	0.398	0.537	0.552	0.438
F-statistics	2.15*	3.16***	2.52**	4.90***	3.03***	2.10*	4.30***	2.20*	1.16

Notes: Standard errors in parentheses (clustered at individual level). Estimation is weighted by hours worked. F-statistics are for the joint hypothesis: 5 tenure interaction terms (potential experience, potential experience^2, training period, female, university+) are equal to 0. See notes to Table 5 for details of the model specifications. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Annual growth rates – heterogeneous linear-spline models

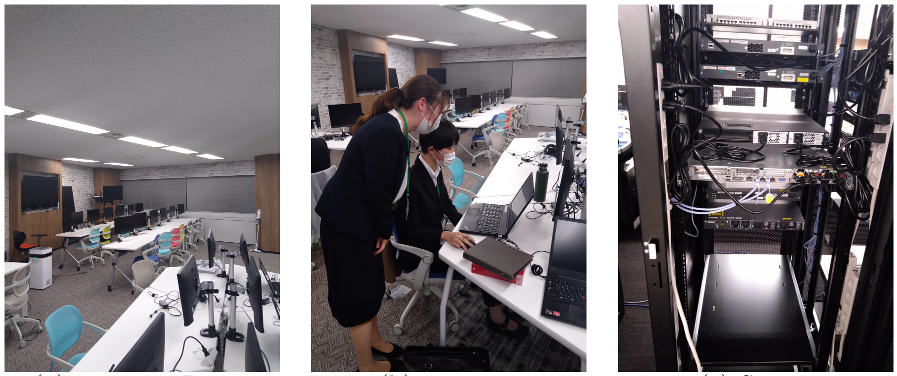
	Model 2: Worker FE			Model 3: Worker & client FE		
	(1) Fee	(2) Wage	(3) Markup	(4) Fee	(5) Wage	(6) Markup
Tenure	0.078***	-0.019	0.097***	0.059***	-0.023**	0.082***
	(0.016)	(0.012)	(0.015)	(0.015)	(0.011)	(0.016)
×After 15m.	0.012	0.061***	-0.049***	-0.000	0.063***	-0.063***
	(0.012)	(0.009)	(0.012)	(0.011)	(0.008)	(0.011)
×Pot-Exp at entry	-0.004	-0.000	-0.004	0.000	0.001	-0.001
	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)
×Pot-Exp×After 15m.	0.002	0.001	0.001	0.001	0.000	0.000
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
×Pot-Exp^2/100	0.009	-0.001	0.010	-0.006	-0.007	0.000
	(0.016)	(0.011)	(0.012)	(0.012)	(0.010)	(0.011)
×Pot-Exp^2/100×After 15m.	-0.009	-0.003	-0.006	-0.001	0.002	-0.003
	(0.013)	(0.007)	(0.010)	(0.010)	(0.006)	(0.008)
×Training period	-0.003	0.000	-0.003	-0.005**	0.000	-0.005**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)
×Training×After 15m.	0.001	-0.001	0.003	0.003*	-0.002*	0.005***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
×Female	0.014	0.011	0.004	0.008	0.006	0.002
	(0.011)	(0.007)	(0.011)	(0.010)	(0.007)	(0.010)
×Female×After 15m.	-0.008	0.001	-0.010	-0.002	0.005	-0.007
	(0.008)	(0.005)	(0.008)	(0.007)	(0.005)	(0.007)
×University+	0.021**	0.015**	0.006	0.024**	0.016**	0.008
	(0.010)	(0.007)	(0.010)	(0.010)	(0.007)	(0.010)
×University+×After 15m.	-0.008	-0.007	-0.000	-0.007	-0.007	0.000
	(0.007)	(0.005)	(0.007)	(0.006)	(0.005)	(0.006)
After 15m.	Yes	Yes	Yes	Yes	Yes	Yes
Tenure×Branch	Yes	Yes	Yes	Yes	Yes	Yes
Regional CPI/Unemployment	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.455	0.492	0.352	0.502	0.529	0.401
F-statistics	1.83**	2.41***	1.52	3.12***	3.30***	1.60

Notes: Standard errors in parentheses (clustered at individual level). Estimation is weighted by hours worked. F-statistics are for the joint hypothesis: 10 tenure interaction terms (potential experience, potential experience^2, training period, female, university+) are equal to 0. See notes to Table 5 for details of the model specifications.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Estimated annual internal rate of return

		Gender		Education		Experience at entry	
	(1) Full sample	(2) Males	(3) Females	(4) Non-university	(5) University	(6) 0-5 years	(7) 6+ years
Multiplier = 1.216 + FC = 200,000	0.255	0.248	0.264	0.179	0.289	0.253	0.267
Multiplier = 1.216 + FC = 100,000	0.288	0.281	0.297	0.217	0.320	0.284	0.304
Multiplier = 1.216 + FC = 300,000	0.227	0.220	0.234	0.146	0.263	0.227	0.235
Multiplier = 1.166 + FC = 200,000	0.381	0.376	0.386	0.338	0.402	0.372	0.401
Multiplier = 1.266 + FC = 200,000	0.123	0.114	0.132	-0.006	0.174	0.132	0.116
N	1,784	1,164	620	668	1,116	1,133	651

Notes: Internal rate of return estimates based on 10-year (120 month) horizon, and using estimated survival rates based on Table 3, and wage and fee growth from model 2 in Table 6. The multiplier is determined by the operating costs, which are assumed to account for 21.6% of wages in the first three rows, and 16.6% and 26.6% in rows four and five respectively. The fixed cost (FC) of hiring is assumed to be either 100,000 JPY, 200,000 JPY, or 300,000 JPY. See text discussion for details.



(a) Training Room (b) Trainees (c) Server

Figure 1: Training Program of the THS firm

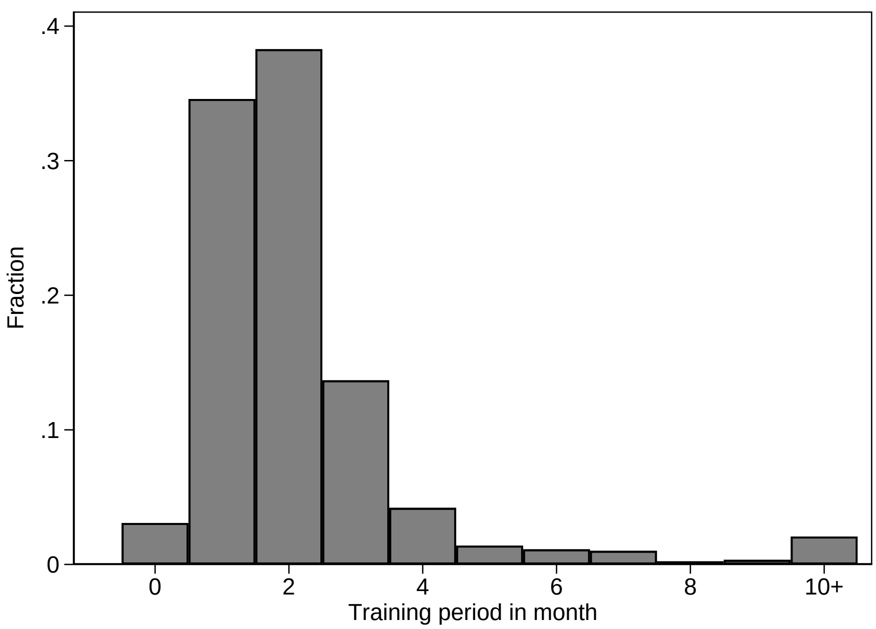


Figure 2: Histogram of pre-placement training period

Notes: The training period is defined as the number of months from a worker's entry to the firm until they are first placed with a client; 76% of workers have training period of 0–2 months.

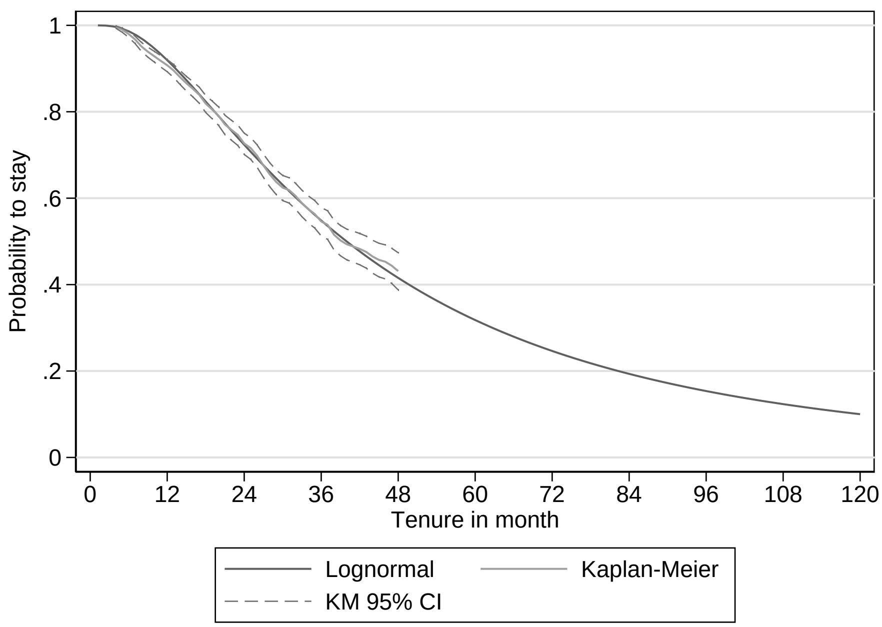


Figure 3: Estimated survival probability: Log-normal model and Kaplan-Meier estimator

Notes: The observations with 3 months and less are excluded from the sample. Therefore, the flat survival rate of the first three months is 1.

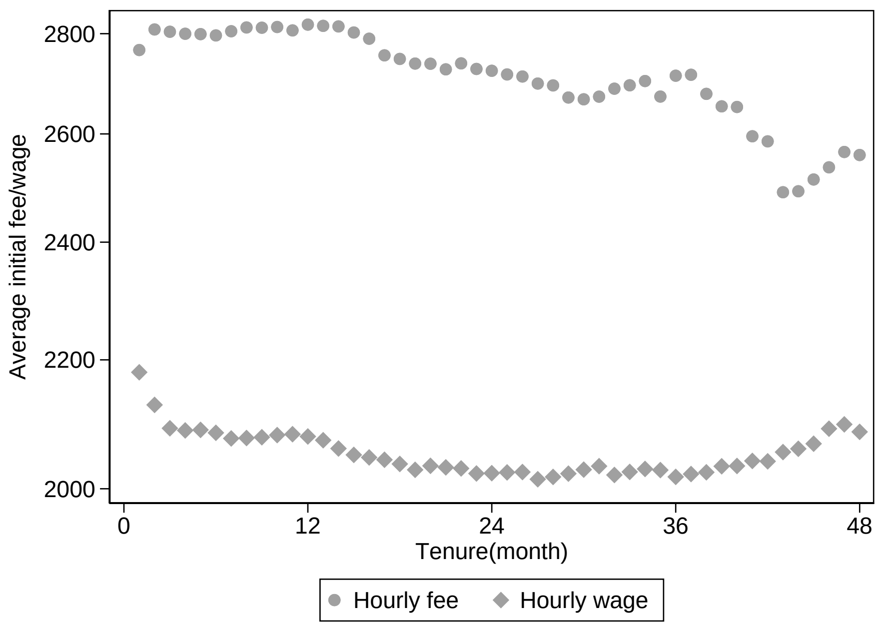


Figure 4: Average initial fee & wage by tenure

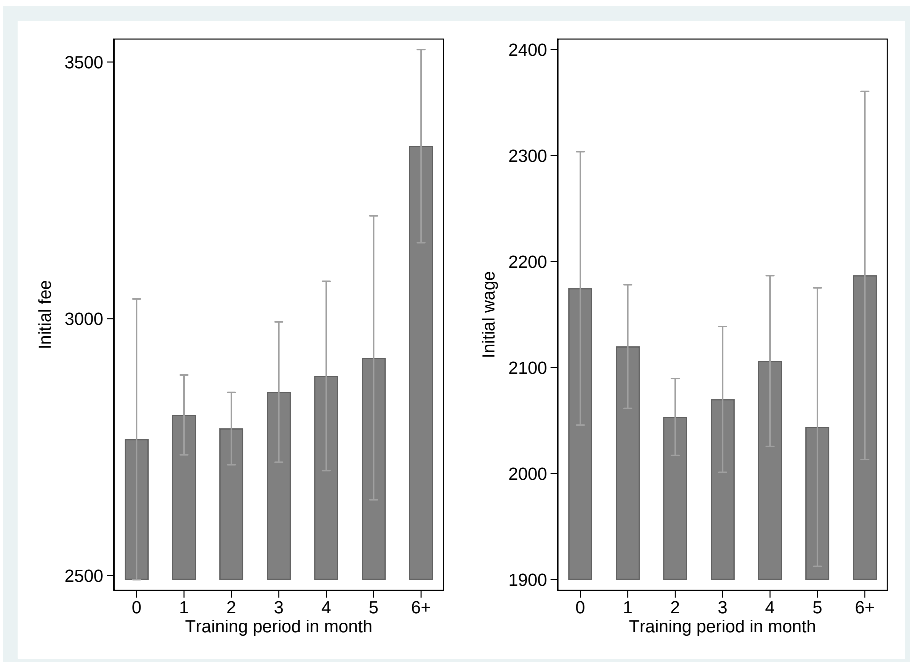


Figure 5: Average initial fees and wages by training length in month

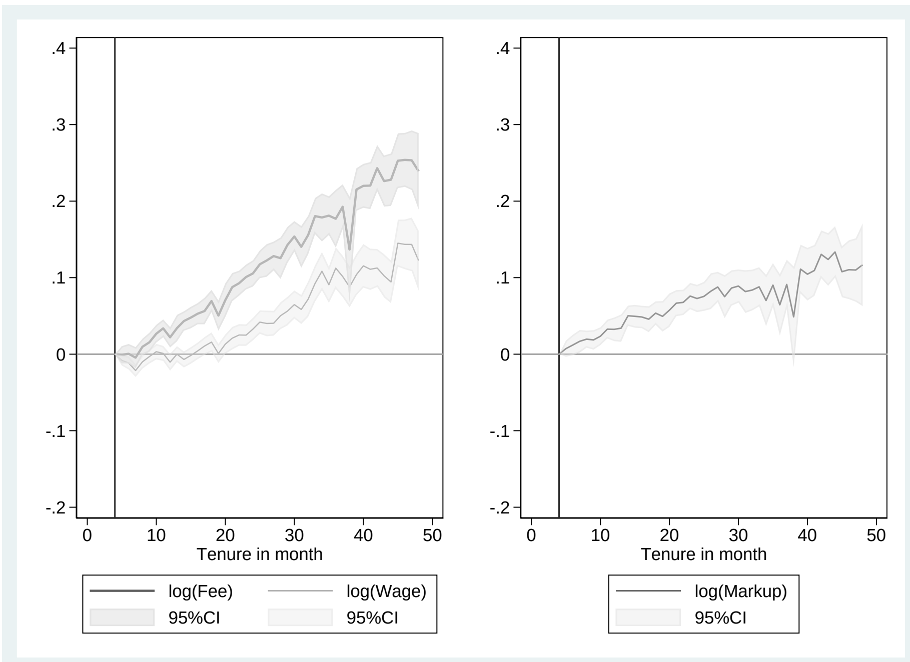


Figure 6: Non-parametric tenure profiles for hourly fee, wage and markup

Notes: Hourly fee, wage and markup (in logs) are regressed on monthly tenure dummies, controlling for worker fixed effects, regional CPI and unemployment rate (i.e the non-parametric version of Model 2). The estimated coefficients for 0-3 months are suppressed. The vertical line is at tenure = 4, which is used as the base. The estimation is based on 34,729 observations from 1,784 workers and 376 clients.