TCER Working Paper Series

Does Managerial Experience in a Target Firm Matter for the Retention of Managers after M&As?

Kenjiro Hirata
Ayako Suzuki
Katsuya Takii

May 2016

Working Paper E-108
http://tcer.or.jp/wp/pdf/e108.pdf

TOKYO CENTER FOR ECONOMIC RESEARCH
1-7-10-703 Iidabashi, Chiyoda-ku, Tokyo 102-0072, Japan

Abstract

This paper examines how managers' tenures in target firms influence their probability of retention as board members after mergers or acquisitions in Japanese firms. It develops a model that distinguishes several hypotheses about the effect of tenure on separation. Our results suggest that experience as an employee increases firm-specific skills, but at the expense of the ability to learn new skills. However, experience as a board member does not have this effect in Japanese firms, the structure of which is known to encourage specific skills. Further, we provide a novel method to correct for selection biases when using data on managers.

Kenjiro Hirata
Kobe International University
Faculty of Economics
9-1-6 Koyocho-naka, Higashinada-ku, Kobe, Hyogo 658-0032, Japan
khirata@kobe-kiu.ac.jp

Ayako Suzuki
TCER
and
Waseda University
School of International Liberal Studies
1-6-1 Nishi-waseda, Shinjuku, Tokyo 169-8050, Japan
a-suzuki@waseda.jp

Katsuya Takii
Osaka University
Osaka School of International Public Policy
1-31 Machikameyama, Toyonaka, Osaka 560-0043, Japan
takii@osipp.osaka-u.ac.jp

Does Managerial Experience in a Target Firm Matter for the Retention of Managers after M&As?*

Kenjiro Hirata[†], Ayako Suzuki[‡] and Katsuya Takii[§]

May 20, 2016

Abstract

This paper examines how managers' tenures in target firms influence their probability of retention as board members after mergers or acquisitions in Japanese firms. It develops a model that distinguishes several hypotheses about the effect of tenure on separation. Our results suggest that experience as an employee increases firm-specific skills, but at the expense of the ability to learn new skills. However, experience as a board member does not have this effect in Japanese firms, the structure of which is known to encourage specific skills. Further, we provide a novel method to correct for selection biases when using data on managers.

1 Introduction

Although several theories of takeovers presume that top managers in a target firm must be replaced after a takeover[1], there is increasing evidence that the retention of a management

*We would like to thank Boyan Jovanovic, Kei Kawai, Ryosuke Okazawa, David Yermack, and seminar participants at the APL Seminar, the Asset Pricing Workshop, the Kansai Labor Workshop, and the Monetary Economics Workshop for their helpful comments. In addition, we thank Yoko Kitaga for her assistance in the construction of the data. Financial support from the Japanese Ministry of Education, Culture, Sports, Science and Technology (a Grant-in-Aid for Scientific Research) is gratefully acknowledged.

†Faculty of Economics, Kobe International University, 9-1-6 Koyocho-naka, Higashinada-ku, Kobe, Hyogo 658-0032, Japan. khirata@kobe-kiu.ac.jp

‡School of International Liberal Studies, Waseda University, 1-6-1, Nishi Waseda, Shinjyuku, Tokyo, 169-8050, Japan. a-suzuki@waseda.jp

§Osaka School of International Public Policy, Osaka University, 1-31, Machikameyama, Toyonaka, Osaka, 560-0043, Japan. takii@osipp.osaka-u.ac.jp

1Some research considers takeovers as a disciplinary device (e.g., Martin and McConnell, 1991). Alternatively, Shleifer and Summers (1988) argue that a takeover causes a breach of trust with stakeholders and transfers rent from them to shareholders. Jovanovic and Rousseau (2002) emphasize that a takeover can reallocate capital to better manage a firm. All of these theories presume that top managers in a target company must be replaced after a takeover. On the other hand, the synergy view of takeovers, which is supported by McGuckin and Nguyen (1995) and Matsusaka (1993), does not predict the replacement of top managers.

group is important to the new firm. Matsusaka (1993) finds that the retention of managers in a firm targeted for takeover increases the bidder's return, and Cannella and Hambrick (1993) and Zollo and Singh (2004) find that the departure of executives from acquired firms is harmful to postacquisition performance. This evidence suggests that there are important skills held by executives of a target firm that a new management group in an acquiring company cannot easily replace.

Interestingly, the existing literature finds that the tenure of a CEO does not have any significant effect on his/her probability of retention after takeover (e.g., Buchholtz et al., 2003, and Wulf and Singh, 2011). Given that experience in a firm is assumed to result in the development of firm-specific skills, the lack of such an effect places doubt on the hypothesis that firm-specific skills are required to manage a newly merged firm after a merger and acquisition (M&A).

It appears that managers' tenures may have a negative impact on their retention rate after an M&A. If it is expected that managers with long tenures will have difficulty in adapting to the new environment, a newly merged firm may not want to employ a manager with a long tenure (e.g., Buchholtz et al., 2003). Therefore, this could offset the positive effect of tenure on the retention rate that occurs because of the need for the firm-specific skills developed through long tenure. Therefore, these opposing effects on retention could make it difficult to interpret the effects of tenure on retention rates.

This paper examines how the tenure of managers influences the retention rate of a management group after M&As in Japanese companies. It explicitly models the retention of managers in newly merged firms in order to separate out the two different hypotheses regarding the effect of tenure on separation probability. We show that the timing of separation provides useful information to distinguish the two effects of tenure, i.e., the importance of target-specific human capital and the ability to learn in a new environment.

Intuitively, the model is built on the following insights. On the one hand, because managers in an acquiring firm do not have much knowledge of the target firm's organization, the acquirer regards the specific skills held by the target-firm managers as highly valuable initially in managing and restructuring the target firm. On the other hand, because target managers accumulate knowledge of the new firm over time, the difference between the amount of knowledge accumulated by a manager who learns quickly and one who is a slow learner becomes larger as time goes by[2]. Therefore, if their long tenure makes it difficult for managers to adapt

[2] This intuition may not be true if the amount of total knowledge that must be learned is so little that a quick

to the environment of the new firm, and the acquiring firm does not need the specific skills developed through long tenure in the target firm, then the theory predicts that the rate of appointment of managers with long tenure to the board of a new firm after an M&A would be low. Alternatively, if the managers with long tenures do not have any difficulty in learning new management skills, then the theory predicts that, when they are appointed initially to manage the new firm, there should not be a high separation rate, even in the long run. Utilizing these theoretical predictions, we can distinguish two different hypotheses about the effect of long tenure on management retention rates by examining how the timing of separation is affected by tenure.

It is important to note that the coefficient of tenure can be contaminated by another effect. If a talented person is promoted to a management position faster than is usually the case, the short tenures observed in our data may simply indicate that a manager has a high unobserved ability. In order to deal with these selection biases, we propose a novel method to extract information about the unobserved heterogeneity of workers from the data on managers. Intuitively, if a talented person is promoted to a management position faster, the length of tenure prior to becoming a manager must contain useful information about their unobserved ability. We utilize this information to identify and control for unobserved ability to estimate the causal effects of tenure.

Using a theoretical model that is shown to be approximately estimated using a stratified Cox proportional hazards model, we conduct survival analyses of Japanese titled directors, who are considered to be top executives in Japanese companies[3]. We focus on Japanese companies that underwent M&As during the period 1990–2006. Our focus on Japanese managers provides a unique advantage for our empirical investigation of tenure. Because the Japanese promotion

learner can immediately understand everything and if an acquiring firm can wait until a slow learner catches up. However, this case is less likely to be important for our empirical studies. First, the literature on the learning curve (e.g., Jovanovic and Nyarko, 1995) suggests that a rise in productivity continues for a long time, especially if tasks are complex, which is likely to be the case for management jobs. Second, as most firms are subject to seasonal events, it is less likely that every operation can be understood without spending at least one year in the firm. Therefore, we focus on one year as the period over which extensive learning occurs in our empirical study. Third, because, as Kaplan (1994) argues, the decisions in many Japanese firms (the focus of our empirical studies) are made on a consensus basis, it is important to learn what people think about a particular strategy. As suggested by the literature on higher-order beliefs, learning about the beliefs of other people on random objects is more difficult than learning about the random objects themselves (see VeldKamp, 2011). Hence, it is likely that a target-firm manager needs time to learn the consensus views held by the managers in the new company. Finally, there is no economic reason for an acquiring firm to keep a slow learner until they start to catch up.

[3] Comparing U.S. CEOs and Japanese presidents, Kaplan (1994) argues that, because decisions in a Japanese firm are made on a consensus basis, it is important, for the sake of accurate comparison, to include other directors, not only presidents. Saito and Odagiri (2008) argue that, because there is heterogeneity among directors, not all directors are important decision makers and, therefore, they choose to focus on directors with titles as the important decision makers in Japanese firms. We follow Saito and Odagiri (2008) and identify directors with titles as the relevant executives in the sample target firms.

system is known to encourage investment in firm-specific human capital, we expect that the advantages and disadvantages of long tenure can be more accurately measured and investigated in Japanese companies than is the case for other countries.

Our empirical analyses show that an increase in a manager's tenure as an employee in a target company increases both the probability of being appointed to the board in a merged firm and the subsequent separation probability after this appointment. However, a longer tenure as a board member within the target firm does not have any significant impacts on separation probability. In addition, we find that the interaction of the initial year dummy after an M&A with the manager's tenure as an employee in the target firm reduces the separation probability, whereas the interaction of the dummy with the tenure as a board member in a target firm does not. This indicates that managers with long tenures as employees are initially more important to the new firm, whereas managers with long tenures as board members are not.

Through the lens of our theory, we can interpret our findings as follows. 1) Acquirers obtain benefits from firm-specific skills from the target firm and benefits from the skills associated with the newly merged firm that target-firm managers must learn. 2) Experience as an employee in a target firm increases firm-specific skills but lowers a manager's ability to learn new skills in a newly merged firm. 3) We cannot reject hypotheses that experience as a board member does not contain any firm-specific components and does not change the learning ability of managers. Thus, our evidence suggests that experience as an employee increases firm-specific skills, but at the expense of the ability to learn new skills, whereas experience as a board member seems to involve general skills that are unrelated to the manager's learning capability.

There have been many attempts to understand the retention of managers after M&As[4]. However, to the best of our knowledge, this is the first paper that constructs an operational model of managers' retention after M&As. Of course, constructing an explicit model requires some abstraction from reality. However, this also aids us in distinguishing several effects of tenure on separation probability. Our empirical study demonstrates that these conditions are useful for determining the importance of two different views in the literature about the role of tenure on management retention after M&As. Our results support both the specific human capital hypothesis and the low adaptability hypothesis when we use the manager's tenure as

[4] Walsh (1988), Walsh (1989), and Walsh and Ellwood (1991) investigate several factors that influence the turnover of top managers after M&As. More recently, Wulf (2004), Hartzell, Ofek, and Yermack (2004), and Bargeron Schlingemann, Stulz, and Zutter (2009) focus on the power of target CEOs to negotiate benefits, which includes a position in a new firm, as an expense borne by shareholders. Finally, Mateos de Cabo, Hagendorff, and Gimeno (2014) investigate how individual characteristics, including gender or membership of a minority group, influence appointments to directorships after M&As.

an employee as an explanatory variable, but reject both hypotheses when we use the tenure as a board member as an explanatory variable. We hope that our approach adds to and complements the existing research on management retention after M&As.

Evidence about the transferability of managerial skills between Japanese firms can provide a useful basis and comparison for the discussion of sources of high compensation for U.S. CEOs. As Bertrand (2009) and Frydman and Jenter (2010) show, some researchers consider the high compensation of U.S. CEOs to be the result of powerful managers setting their own pay rates, but others consider it to be the result of a competitive market for managerial talent. One of the important presumptions behind the market-based view (e.g., Gabaix and Landier, 2008, and Terviö, 2008) is that important managerial skills are transferable across firms. In fact, Murphy and Zábojník (2004) and Frydman (2005) argue that a rise in the importance of a CEO's general skills relative to his/her firm-specific skills can explain not only the increase in CEO compensation but also the increase in the turnover rate that has occurred since 1970. Kaplan, Klebanov, and Sorensen (2012) investigate a proprietary dataset for executives and provide evidence that general skills are important managerial skills. Our results, focusing on the Japanese context, provide evidence that complements this U.S. literature: we find that experience as a board member in a target firm may be useful for managing or monitoring new firms in Japan, even though the promotion structure of Japanese firms is considered to facilitate the accumulation of firm-specific skills to a particularly large extent (e.g., Mincer and Higuchi, 1988).

It should be noted that Japanese managers receive far less cash compensation than do their U.S. counterparts[5] and that Japanese CEOs have very long tenures at their firms relative to CEOs in U.S. firms[6]. This suggests that, even if experience as a board member in Japan is transferable across firms, such an appointment does not immediately imply high compensation and a high turnover rate. Accounting for the differences in the market for managers between the U.S. and Japan is beyond the scope of this paper. However, together with our evidence that experience gained as an employee is firm specific and is required to manage a newly merged firm, a plausible conjecture is that Japanese managers must conduct both tasks that need managerial

[5] Kaplan (1994) shows that U.S. officers earned 13.5 times the average compensation of other employed males, whereas the ratio for Japanese executives is only 4.8. Moriguchi and Saez (2008) show that, whereas the top 1% wage income share in the U.S. rose exponentially from 5% to 12% between 1970 and 2005, the share in Japan has remained nearly constant at around 5% during this period.

[6] Kaplan (1994) shows that the average tenure of a president in a Japanese firm is 34.3 years, whereas that of a U.S. CEO is 26 years.

skills and tasks that require an understanding of employees' specific skills[7], and, therefore, there are barriers to the division of labor and the creation of a market for managerial skills in Japan.

Our results also contribute to the discussion on the sources of entrepreneurial ability. Schultz (1975) defines entrepreneurial ability as the ability to interpret new information and adapt to a new environment. Takii (2003) shows that entrepreneurial ability, as defined by Schultz (1975), can explain why top business school graduates obtain jobs in risky industries. Although some of the literature argues that entrepreneurial ability can be improved by appropriate prior knowledge, which can be the result of previous experience (e.g., Fiet, 1996, and Shane, 2000), we could not find any evidence that experience as a board member in Japan improves learning capability in a new environment. This suggests that, although experience as a board member can be useful in managing other firms, it may not be a source of entrepreneurial ability, as defined by Schultz (1975).

This paper not only presents interesting empirical results, but also makes a methodological contribution to the literature. As a by-product of estimating a model using datasets on board members, we develop a novel method to correct selection biases using only the selected sample. Although it is recognized that the characteristics of a CEO change the strategies of firms (e.g., Bertrand and Schoar, 2003), researchers face difficulties in examining the causal role of CEO characteristics in a firm because publicly available data on top managers typically involve selected samples. We propose a method to deal with this problem. We argue that, even if we cannot access general population data, which is needed to estimate a selection equation in a standard two-stage estimator, such as that of Heckman (1979), we can utilize the timing of selection in a selected sample to correct for selection bias. Our methodology can be applied not only to datasets for managers and directors, but also to data on internal promotions. We hope that this novel approach can help researchers struggling with data limitations to evaluate the role of leaders.

Finally, this is the first paper that examines the retention of managers after M&As using a Japanese dataset. The number of M&As in Japan has dramatically increased since the late 1990s. Although there are several attempts to understand Japanese M&A waves (e.g., Fukao, Ito, and Kwon, 2005), few papers investigate the separation of workers after M&As. Notable exceptions are Kubo (2004) and Kubo and Saito (2012), which analyze the effect of mergers on employment and wages. In particular, Kubo (2004) analyzes the personal characteristics

[7] In fact, Jacoby (2005) compares Japan's internal labor market with the U.S. internal labor market and highlights the stronger role of personnel departments in Japan.

of employees who separated from firms before and after mergers using firm personnel data. However, neither work discusses the retention of top managers. Although the impacts of mergers on employees are interesting in their own right, the reasons for the separation of managers from the boards of newly merged firms would be different from the reasons for the separation of employees. Hence, our evidence can provide valuable new information to understand how Japanese firms adapt to merger waves.

The paper is organized as follows. The next section presents a model of managers' retention after M&As. It provides several testable conditions that distinguish the different impacts of tenure on separation probability. Section 3 introduces our empirical model and discusses how to correct a selection bias using our selected dataset. Section 4 describes our Japanese dataset and Section 5 presents our estimation model and our results, interpreted through the lens of our theory. Finally, Section 6 concludes the paper.

2 Model

In order to propose testable hypotheses about the relationship between the length of tenure and separation probability, we specify a simple model. For this purpose, we model an appointment decision and a separation decision after an M&A and abstract from the negotiation that occurs during a merger. We will discuss how we deal with a possible moral hazard problem that might occur through a negotiation during a merger in the empirical sections of our paper.

We first provide a micro foundation for the hazard function at the time of and after the appointment to boards in a merged firm. Later, we examine the impacts of tenure on the hazard rate. The proposed model provides the set of testable hypotheses that are examined in our empirical study.

Separation Probability: When a firm is merged with another firm, managers from the target firm must learn new skills and/or new routines in order to manage the merged firm in its new environment. We assume that all learning takes place at time $t = 0$ and that standard operation starts after time $t = 1$, where the analysis time t is the length of time of operation after the year in which the new firm was created by the M&A.

Let us first consider the new firm's behavior at $t \geq 1$. Assume that the marginal productivity of a manager, $mp + \zeta\left(t : \mathbf{X}_Z\right)$, where t is the time that has passed since the new firm started and the vector $\mathbf{X}_Z$ includes any variables that can be controlled for in our empirical study. The present value of the expected profit sequences from employing a particular manager, J_t at $t \geq 1$, can be described as follows:

$$J_t = mp + \zeta(t : \mathbf{X}_Z) - w_t + \beta E_t I(W_{t+1} \geq R_{t+1} + \varepsilon_{t+1}) \max\{J_{t+1}, 0\},$$

where w_t is the manager's compensation at time t, $\beta < 1$ is a discount factor, ε_t is the sum of any idiosyncratic random benefits that the manager can obtain if he/she leaves the firm at time t, and W_t and R_t are the present values of the sum of the expected income flows when a manager stays and when he/she leaves the firm at time t, respectively.

We assume that a firm reviews its managers every period. Assuming that the manager prefers to stay, which means that $I(W_{t+1} \geq R_{t+1} + \varepsilon_{t+1}) = 1$, the firm has to decide whether to retain the manager as a board member. If it retains the manager, it expects to obtain J_{t+1} in future from the manager. However, if it fires the manager, the firm must find a new manager. As assumed in the standard search model (e.g., Pissarides, 2000), the competition to find a new manager results in the net benefits from the search being zero under the assumption of free entries. Because a firm is assumed to maximize its profit, the expected profit would be $\max\{J_{t+1}, 0\}$. If $W_{t+1} < R_{t+1} + \varepsilon_{t+1}$, the manager leaves, irrespective of the firm's decision.

Similarly, the value of managers in the firm and outside the firm at the analysis time t, W_t and R_t, respectively, can be described as follows:

$$W_t = w_t + \beta E_t \left[I(J_{t+1} \geq 0) \max\{W_{t+1}, R_{t+1} + \varepsilon_{t+1}\} + (1 - I(J_{t+1} \geq 0))(R_{t+1} + \varepsilon_{t+1}) \right],$$

and

$$R_t = \hat{\mathbf{w}}'\mathbf{h}^G + \zeta^r(t : \mathbf{X}_Z) + \beta R_{t+1},$$

where $\hat{\mathbf{w}}$ is the price vector for general human capital in the external market, $\mathbf{h}^G$ is the vector of the manager's general human capital. We allow for multidimensional general human capital. The value $\zeta^r(t : \mathbf{X}_Z)$ is the sum of any other outside benefits that the manager can expect. The outside benefits may include better job possibilities, pensions, and/or the benefits of leisure.

When a manager stays in a firm, in every following period, he/she reassesses and decides whether to stay with or leave the firm. If the manager stays, he/she expects to obtain W_{t+1}. If the manager leaves, he/she expects to obtain $R_{t+1} + \varepsilon_{t+1}$. Hence, if a firm decides to continue to employ the manager, $I(J_{t+1} \geq 0) = 1$, the expected value must be $\max\{W_{t+1},\ R_{t+1} + \varepsilon_{t+1}\}$. If the firm decides to fire the manager, $I(J_t \geq 0) = 0$, the manager must leave and receive the outside benefits of $R_{t+1} + \varepsilon_{t+1}$, irrespective of his/her own decision.

Note that, if both the firm and the manager prefer to maintain the relationship, ($J_{t+1} \geq 0$ and $W_{t+1} \geq R_{t+1} + \varepsilon_{t+1}$), they can do so. However, if one party prefers to walk away, $J_t < 0$ or $W_t < R_t + \varepsilon_{t+1}$, then they must separate from each other. Because they can always negotiate w_t to avoid separation, it can be shown that they will separate at time t if and only if $S_t < \varepsilon_t$, where $S_t \equiv J_t + W_t - R_t$ and the dynamics of S_t can be described by:

$$S_t = mp - \hat{\mathbf{w}}'\mathbf{h}^G + z(t : \mathbf{X}_Z) + \beta\left[S_{t+1} + E_t\left[I(\varepsilon_{t+1} > S_{t+1})(\varepsilon_{t+1} - S_{t+1})\right]\right], \tag{1}$$

where $z(t : \mathbf{X}_Z) = \zeta(t : \mathbf{X}_Z) - \zeta^r(t : \mathbf{X}_Z)$. For any t, maintaining the relationship brings a joint instantaneous surplus of $mp - \hat{\mathbf{w}}'\mathbf{h}^G + z(t : \mathbf{X}_Z)$. If they maintain their relationship, the manager and firm can enjoy S_{t+1} in the next period. However, if the instantaneous random benefits from the outside market happen to be greater than the joint surplus, $\varepsilon_{t+1} > S_{t+1}$, then they will decide to separate and obtain the additional value of $\varepsilon_{t+1} - S_{t+1}$. Because they can separate only when the random benefits are large enough to exceed the joint surplus, the possibility of separation simply provides each party with the additional value arising from separation. This effect is captured by $E_t\left[I(\varepsilon_{t+1} > S_{t+1})(\varepsilon_{t+1} - S_{t+1})\right]$.

Assume that ε_t is exponentially distributed with a parameter λ. The separation probability, q_t, is shown to be a function of S_t:

$$q_t = E_t\left[I(\varepsilon_t > S_t)\right] = e^{-\lambda S_t}.$$

This suggests that increases in the joint surplus, S_t, reduce the probability of separation.

Given this distributional assumption, the expected gain from separation is shown to be proportional to the separation probability:

$$E_t\left[I(\varepsilon_{t+1} > S_{t+1})(\varepsilon_{t+1} - S_{t+1})\right] = \frac{q_t}{\lambda}.$$

Inserting the expected gain from separation into the dynamics of the joint surplus, we can rewrite equation (1) as the following nonlinear dynamic equation of S_t:

$$S_t = mp - \hat{\mathbf{w}}'\mathbf{h}^G + z(t : \mathbf{X}_Z) + \beta\left[S_{t+1} + \frac{e^{-\lambda S_{t+1}}}{\lambda}\right].$$

Taking a first-order approximation of $\frac{e^{-\lambda S_{t+1}}}{\lambda}$ around $S_{t+1} = v$ for any arbitrary v, we can derive an approximate solution of S_t[8]:

$$\begin{aligned} S_t &\approx \pi + Z\left(t : \mathbf{X}_Z\right) \\ \pi &= \frac{mp - \hat{\mathbf{w}}'\mathbf{h}^G}{1 - \tilde{\beta}\left(v\right)} \\ Z\left(t : \mathbf{X}_Z\right) &= \sum_{\tau=0}^{\infty} \tilde{\beta}\left(v\right)^{\tau} z\left(t + \tau : \mathbf{X}_Z\right) + \frac{\hat{\beta}\left(v\right)}{1 - \tilde{\beta}\left(v\right)}, \end{aligned}$$

where $\tilde{\beta}\left(v\right) = \beta\left(1 - e^{-\lambda v}\right)$ and $\hat{\beta}\left(v\right) = \beta\frac{(1+\lambda v)e^{-\lambda v}}{\lambda}$.

Before standard operations commence, we assume that managers must spend one period learning new skills after an M&A. The present value of the stream of instantaneous joint surpluses at time 0, S_0, can be expressed as follows:

$$S_0 = mp_0 - \hat{\mathbf{w}}'\mathbf{h}^G + z\left(0 : \mathbf{X}_Z\right) + \beta\left[S_1 + \frac{e^{-\lambda S_1}}{\lambda}\right].$$

Here, we assume that the marginal productivity of a manager, mp_0, at time 0 can differ from that at time $t \geq 1$, mp. When managers in the target firm are first appointed to the board in the merged firm, they do not have sufficient knowledge to deal with the new operation. Therefore, mp_0 can be much lower than mp. On the other hand, because managers in the acquiring firm do not have enough specific knowledge of the target firm, the target firm's managers and their specific knowledge might be more valuable in this stage. In this case, mp_0 can be greater than mp.

Using a similar approximation, we can show that:

$$\begin{aligned} S_0 &\approx \Delta y + \pi + Z\left(0 : \mathbf{X}_Z\right) \\ \Delta y &= \Delta mp, \end{aligned}$$

[8] Alternatively, we can assume that $z\left(t : \mathbf{X}_Z\right) = 0$ for all t. In this case, $S_t = S^*$, where S^* can be solved by the following equation:

$$S^* = mp - \hat{w}h^G + \beta\left[S^* + \frac{e^{-\lambda S^*}}{\lambda}\right].$$

Therefore, it can be shown that the joint surplus is an increasing function of the instantaneous surplus:

$$S^* = \Sigma\left(mp - \hat{w}h^G\right),\ \Sigma'\left(\cdot\right) > 0.$$

In this case, the separation probability can be expressed by the constant hazard model:

$$q^* = e^{-\lambda\Sigma\left(y^a - \hat{w}h^G\right)}.$$

Once we take a Taylor approximation of this expression, we can derive similar theoretical predictions.

where $\Delta mp = mp_0 - mp$.

Using the derived surplus, our separation probability is expressed by the following proportional hazards model:

$$\begin{aligned} q_t &= B(t : \mathbf{X}_Z) e^{-\lambda[\Delta y I(t=0)+\pi]} \\ \pi &= \frac{mp - \hat{\mathbf{w}}'\mathbf{h}^G}{1-\tilde{\beta}(v)} \\ \Delta y &= \Delta mp, \end{aligned}$$

where $B(t : \mathbf{X}_Z) = e^{-\lambda Z(t:\mathbf{X}_Z)}$ and $I(t=0) = 1$ when $t = 0$ and $I(t=0) = 0$ otherwise. We assume that the acquiring firm appoints the manager in the target firm to the board of the new firm at the beginning of $t = 0$. Because the probability of being appointed a manager is equivalent to $1 - q_0$, investigating the separation probability in the initial period provides information about the appointment probability as well.

Tenure and Separation Probability: We wish to derive testable hypotheses about the relationship between tenure and separation probability based on a managerial human capital theory. For this purpose, we first assume that the marginal productivity of a manager is a function of the different types of human capital possessed by managers:

$$\begin{aligned} mp_0 &= \alpha_G^0 h_M^G + \alpha_T^0 h^T + \alpha_A^0 h_0^A + \alpha_X^{0\prime}\mathbf{X} \\ mp &= \alpha_G h_M^G + \alpha_T h^T + \alpha_A h^N + \alpha_X'\mathbf{X} \\ h^N &= h_0^A + \gamma h_L^G,\ \hat{\mathbf{w}} = (\hat{w}_M, \hat{w}_L)',\ \mathbf{h}^G = \left(h_M^G, h_L^G\right)', \end{aligned}$$

where h_M^G is a general management skill, h_L^G is a general learning skill, h^T is the target firm's firm-specific human capital, h^N is the newly merged firm's firm-specific human capital that a manager in the target firm has to acquire, and $\mathbf{X}$ is the vector of other firm-level observable variables that we control for in our empirical study. The new firm's firm-specific human capital is the sum of the initial knowledge about the acquirer, h_0^A, and the amount of new knowledge accumulated during the learning period, γh_L^G. We consider h^N to consist of both an acquirer-specific skill and an entirely new skill required as a result of the establishment of the new firm. We assume that the parameters $\alpha_i \geq 0$ and $\alpha_i^0 \geq 0$, where $i = G$, T and A.

We assume that not only the productivity of managers, h_M^G, but also their ability to learn specific skills, h_L^G, involves general skills, which are transferable across firms. Nelson and

Phelps (1966) consider education to increase the speed of technology adoption because educated people are able to understand new technology. Schultz (1975) argues that human intelligence can be used to interpret new information and adapt to a new environment, and refers to such intelligence as entrepreneurial ability. Our assumptions on learning ability capture these ideas.

For a simple analysis, we assume a competitive labor market for h_M^G and for h_L^G. Under a competitive managerial labor market, the price for h_M^G, $\hat{w}_M$, is adjusted to equalize the marginal productivities of h_M^G across firms. Therefore, it is natural to assume that:

$$\hat{w}_M = \alpha_G = \alpha_G^0.$$

On the other hand, employing managers with a large h_L^G is considered to be an investment. Standard investment theory suggests that the present value of the expected profits from an investment must be equal to the cost of the investment under a competitive market. Because the new knowledge accumulated during the learning period is assumed to be specific to the new firm, the surplus from the investment must be shared by the firm and the manager. Assume that a firm can receive a fraction of the surplus, ψ. Then, the following condition must be satisfied under a competitive market:

$$\begin{aligned} \tilde{\beta}(v)\,\psi S_L^G &= \hat{w}_L h_L^G, \ where \\ S_L^G &= \frac{\alpha_A h^N - \hat{w}_L h_L^G}{1-\tilde{\beta}(v)}, \ h^N = \gamma h_L^G, \end{aligned}$$

and S_L^G is the surplus from employing a manager with h_L^G, which is the total sum of the present value of instantaneous benefits, $\frac{\alpha_A h^N - \hat{w}_L h_L^G}{1-\tilde{\beta}(v)}$. Because the firm can obtain a portion $\psi \in (0,1)$ of the surplus, the firm expects to obtain ψS_L^G one period later. Discounting ψS_L^G by $\tilde{\beta}(v)$, $\tilde{\beta}(v)\,\psi S_L^G$ is the present value of the expected profits from the investment. This must be equal to the cost of investment, $\hat{w}_L h_L^G$. This equation implies that $\hat{w}_L$ must be adjusted to satisfy:

$$\hat{w}_L = \frac{\tilde{\beta}(v)\,\psi\gamma\alpha_A}{1-(1-\psi)\,\tilde{\beta}(v)}.$$

Note that, as long as $\psi < 1$, the additional surplus from employing h_L^G is greater than the additional costs: $\frac{\tilde{\beta}(v)\gamma\alpha_A}{1-\tilde{\beta}(v)} > \frac{\hat{w}_L}{1-\tilde{\beta}(v)}$. This is a result of the holdup problem. The manager and the firm must share the benefits from employing a manager with a high h_L^G in order to maintain their relationship. Because a firm cannot obtain all the benefits from employing managers with h_L^G, the firms in a competitive economy create jobs for the managers that are less than optimal. Hence, the demand for h_G and, therefore, the price of h_G, $\hat{w}_L$, are lower than optimal. Therefore, the additional surplus from investing in h_L^G is larger than the cost, $\hat{w}_L$, in a competitive economy[9].

Imposing these theoretical restrictions, it is shown that:

$$\begin{aligned} \pi &= \tilde{\alpha}_T h^T + \tilde{\alpha}_A h_0^A + \tilde{\alpha}_N h_L^G + \tilde{\alpha}_X' \mathbf{X}, \\ \Delta y &= \bar{\alpha}_T h^T + \bar{\alpha}_A h_0^A - \alpha_N h_L^G + \bar{\alpha}_X' \mathbf{X}, \end{aligned}$$

where $\tilde{\alpha}_T = \frac{\alpha_T}{1-\tilde{\beta}(v)} \geq 0$, $\tilde{\alpha}_A = \frac{\alpha_A}{1-\tilde{\beta}(v)}$, $\tilde{\alpha}_N = \frac{\gamma\alpha_A}{1-(1-\psi)\tilde{\beta}(v)}$, $\tilde{\alpha}_X = \frac{\alpha_X}{1-\tilde{\beta}(v)}$, $\bar{\alpha}_T = \alpha_T^0 - \alpha_T \geq 0$, $\bar{\alpha}_A = \alpha_A^0 - \alpha_A$, $\alpha_N = \gamma\alpha_A$, and $\bar{\alpha}_X = \alpha_X^0 - \alpha_X$. We assume that $\bar{\alpha}_T \geq 0$. When a merger occurs, managers in the acquiring firm do not have enough knowledge about the target firm. Initially, the target firm's firm-specific knowledge must be more valuable. The assumption of $\bar{\alpha}_T \geq 0$ captures this effect.

Note that h_M^G does not influence the surplus at all. Therefore, it should not influence the separation rate. Employing a manager with high-level general skills increases output, but it is costly. Under the competitive market for h_M^G, the benefits from employing talented managers are perfectly reflective of the price. Therefore, a firm cannot obtain an additional surplus from employing a manager with high h_M^G. This result immediately implies the following proposition.

Proposition 1 *Any factors that influence a general management skill, h_M^G, including the length of tenure, should not affect the separation probability in the proportional hazards model.*

Note that our theory predicts that a rise in h_L^G increases π, but reduces Δy. If managers have a greater ability to learn in the new environment, this skill increases the firm's future profits. However, learning does not increase the initial output, when newly merged firms have just started operations. Because Δy simply reflects the initial instantaneous output, net of the future instantaneous output yielded by the manager, the impact of learning ability on Δy is negative. This unique prediction helps us to separate learning ability from other skills.

Interestingly, $\tilde{\alpha}_N > \alpha_N$ as long as $\psi < 1$. That is, $\frac{\partial S_0}{\partial h_L^G} = \frac{\partial(\pi+\Delta y)}{\partial h_L^G} > 0$ as long as $\psi < 1$. As we discussed earlier, because the additional surplus from employing a manager with h_L^G is greater than the additional costs in a competitive economy, $\frac{\tilde{\beta}(v)\gamma\alpha_A}{1-\tilde{\beta}(v)} > \frac{\hat{w}_L}{1-\tilde{\beta}(v)}$, there is scope to increase the surplus by employing a manager with a higher h_L^G.

[9]The general equilibrium consequences of the holdup problem have been investigated by Caballero and Hammour (1998) in a macroeconomic context.

Let τ_b and τ_e denote the manager's tenure as a board member and as an employee in the target firm, respectively. Following the tradition of labor economics, we assume that the target firm's firm-specific human capital is an increasing function of these tenures:

$$h^T = \eta_b \tau_b + \eta_e \tau_e + \eta_T' \mathbf{D}_T,$$

where $\mathbf{D}_T$ is the vector of any variables that influence the accumulation of a target firm's firm-specific human capital. We assume that $\eta_x \geq 0$ for $x = b$ or e, which means that increases in both types of tenures assist in the accumulation of the firm-specific human capital.

On the other hand, as h_L^G is a general skill, labor economists typically assume that h_L^G increases not only as a result of experience in the target firm, i.e., as a result of both τ_b and τ_e, but also as a result of the duration of other experiences, τ_o. Therefore, we have:

$$h_L^G = \omega_b \tau_b + \omega_e \tau_e + \mu \tau_o + \Gamma_G' \mathbf{D}_G + \Gamma_\chi \chi,$$

where $\omega_x \geq 0$ $(x = b, e)$ are the parameters that capture the importance of tenure for learning, μ is the parameter that captures the importance of other experiences, $\mathbf{D}_G$ is the vector of any other variables that influence learning ability, and χ is the unobserved individual ability. Estimating τ_o by $\tau_o = a - \tau_b - \tau_e$, where a is the manager's age when the M&A takes place, we can rewrite h_L^G as a function of tenures and age:

$$h_L^G = \omega_b \tau_b + \omega_e \tau_e + \mu \left(a - \tau_b - \tau_e\right) + \Gamma_G' \mathbf{D}_G + \Gamma_\chi \chi.$$

Buchholtz, Ribbens, and Houle (2003) argue that longer tenure may hamper adaptation to a new environment. This implies that $\mu > \omega_x$. Although this might be a reasonable assumption, we also allow for the possibility of $\omega_x \geq \mu$: i.e., that experience in a target firm can assist in learning new skills within the new firm. Several studies on spin-off effects suggest that previous experience in incumbent firms can be an important source of experience in establishing new firms (e.g., Klepper, 2001). It would be possible to apply a similar reasoning in this context. This possibility is captured by $\omega_x > \mu$.

Finally, we assume that $h_0^A = \eta_A' \mathbf{D}_A$, where $\mathbf{D}_A$ is the vector of variables that influence the initial amount of the acquirer's firm-specific knowledge. The following proportional hazards model can be derived:

$$\begin{aligned}
q_t &= B\left(t:\mathbf{X}_Z\right)e^{-\lambda[\Delta y I(t=0)+\pi]}, \\
\pi &= \tilde{\alpha}_T h^T + \tilde{\alpha}_N h_L^G + \tilde{\alpha}_A \eta_A' \mathbf{D}_A + \tilde{\alpha}_X' \mathbf{X} \\
\Delta y &= \bar{\alpha}_T h^T - \alpha_N h_L^G + \bar{\alpha}_A \eta_A' \mathbf{D}_A + \bar{\alpha}_X' \mathbf{X} \\
h^T &= \eta_b \tau_b + \eta_e \tau_e + \eta_T' \mathbf{D}_T, \\
h_L^G &= \omega_b \tau_b + \omega_e \tau_e + \mu\left(a - \tau_b - \tau_e\right) + \Gamma_G' \mathbf{D}_G + \Gamma_\chi \chi.
\end{aligned}$$

Using this derived hazard function, the predicted coefficients of the effect of tenure on separation probability are summarized as follows:

$$\frac{d\left(-\lambda\pi\right)}{d\tau_x} = -\lambda\left[\tilde{\alpha}_T \eta_x + \tilde{\alpha}_N\left(\omega_x - \mu\right)\right], \tag{2}$$

$$\frac{d\left(-\lambda\Delta y\right)}{d\tau_x} = -\lambda\left[\bar{\alpha}_T \eta_x - \alpha_N\left(\omega_x - \mu\right)\right], \tag{3}$$

$$\frac{\partial\left[-\lambda\left(\Delta y + \pi\right)\right]}{\partial\tau_x} = -\lambda\left[\left(\tilde{\alpha}_T + \bar{\alpha}_T\right)\eta_x + \left(\tilde{\alpha}_N - \alpha_N\right)\left(\omega_x - \mu\right)\right], \tag{4}$$

where $x = b$ or e.

Note that the separation probability q_t for $t \geq 1$ depends on π, but the separation probability q_0 depends on $\Delta y + \pi$. Hence, the interaction term with the $t = 0$ dummy, Δy, captures the additional benefits realized in the transition process, which enables us to explain the differences in the probability of the separation at the time of appointment decision and the probability of separation after appointment. Therefore, equation (2) provides information about the impacts of tenure on separation probability after the initial appointment to the board of a merged firm; equation (4) captures the impacts of tenure on the probability of being appointed to the board in a merged firm; and equation (3) captures the impact of tenure on the additional benefits in the transition process.

Investigating equations (2), (3), and (4) provides the following propositions that we wish to reject in our empirical study. The first proposition below shows the set of conditions under which a newly merged firm does not appreciate any specific skills of employees in the target firm and/or any ability to learn in a new environment.

Proposition 2 *Consider equations (2), (3), and (4).*

1. *Suppose that* $\tilde{\alpha}_T = \bar{\alpha}_T = 0$*, which means that the skills specific to the target firm are not*

valued in the new firm at all. Then, for both $x = b$ and e, we have:

$$sign\left[\frac{d(-\lambda\pi)}{d\tau_x}\right] = -sign\left[\frac{d(-\lambda\Delta y)}{d\tau_x}\right] = sign\left[\frac{\partial\left[-\lambda(\Delta y+\pi)\right]}{\partial\tau_x}\right]. \tag{5}$$

2. *Suppose that $\tilde{\alpha}_N = 0$ (or, equivalently, that $\alpha_N = 0$), which means that there is no important new skill that a target-firm manager must learn in the newly merged firm. Then, for both $x = b$ and e, we have:*

$$\frac{d(-\lambda\pi)}{d\tau_x} \leq 0,\ \frac{d(-\lambda\Delta y)}{d\tau_x} \leq 0,\ \frac{\partial\left[-\lambda(\Delta y+\pi)\right]}{\partial\tau_x} \leq 0. \tag{6}$$

The intuition behind equation (5) in proposition 2 is explained as follows. If $\alpha_T^* = \bar{\alpha}_T = 0$, an increase in tenure influences the separation probability only because it can change the learning capability of managers. If $\omega_x \geq \mu$, where $x = b$ or e, as an increase in tenure raises learning capability, it increases the marginal productivity of the manager in the newly merged firm and, therefore, lowers the separation probability, at all times. However, because learning takes time, the benefits from the learning capability are not initially realized. Hence, the initial marginal products of tenure, net of future marginal products, are negative and the coefficient of the interaction with the $t = 0$ dummy must be positive. However, if $\omega_x \leq \mu$, the opposite mechanism occurs. Because an increase in tenure reduces learning capability in this case, it simply lowers the marginal productivity of the manager in a merged company and, therefore, increases the separation probability at all times. However, because the initial marginal productivity of the manager after the M&A is not subject to this negative effect, the initial marginal products of tenure, net of future marginal products, are positive. Therefore, the coefficient of the interaction with the $t = 0$ dummy must be negative. Hence, for both cases, equation (5) must be satisfied.

Similarly, we can explain the intuition behind equation (6) in proposition 2 as follows. If $\tilde{\alpha}_N = 0$, an increase in tenure changes the separation probability only because it can raise the human capital specific to the target firm. As long as an increase in human capital specific to the target firm is valued in the merged firms, it would reduce the separation probability at all times. In addition, because the human capital specific to the target firm is more important before the managers in the acquiring firm learn the same skills, it would be more highly valued at $t = 0$. Therefore, the coefficient of the interaction with the $t = 0$ dummy must be negative. Equation (6) summarizes these conditions.

Note that proposition 2 requires that equations (5) and (6) must be jointly satisfied by the coefficients for both the tenure as a board member, $x = b$, and the tenure as an employee, $x = e$. Because an acquiring firm does not care about how a target manager obtains a skill, if the skills are not useful within the merged firm, the coefficients of tenures as a board member and as an employee will be influenced in a similar way. In other words, if the derived conditions are not satisfied for either $x = b$ or $x = e$, we can reject the hypotheses in proposition 2.

On the other hand, the difference between the coefficients for the two types of tenure, $x = b$ and $x = e$, provides us with information about the differences in skills obtained as a result of experience gained as a board member or as an employee. The following proposition summarizes the hypotheses about specific human capital and experience that we wish to reject.

Proposition 3 *Suppose that $\tilde{\alpha}_T \neq 0$ and $\bar{\alpha}_T \neq 0$. If $\eta_x = 0$, which means that experience gained during x, where $x = b$ or e, does not raise human capital specific to the target firm, then:*

$$sign\left[\frac{d(-\lambda\pi)}{d\tau_x}\right] = -sign\left[\frac{d(-\lambda\Delta y)}{d\tau_x}\right] = sign\left[\frac{\partial\left[-\lambda(\Delta y + \pi)\right]}{\partial\tau_x}\right]. \tag{7}$$

Equation (7) looks very similar to equation (5). The only difference is that proposition 2 requires that equation (5) must be jointly satisfied for both $x = e$ and $x = b$, whereas proposition 3 states that we can separately use equation (7) for $x = e$ and $x = b$. The intuition behind equation (7) is the same as that behind equation (5). However, by separately applying the same logic to the coefficients for the tenure as a board member and the tenure as an employee, we can determine differences in skills obtained from experience as a board member compared with experience as an employee.

Finally, the following proposition summarizes the hypotheses about learning ability and experience that we wish to reject.

Proposition 4 *Suppose that $\tilde{\alpha}_N \neq 0$.*

1. *If $\omega_x \geq \mu$, which means that a long tenure during x, where $x = b$ or e, does not hamper the ability to learn a new skill in a newly merged firm, then:*

$$\frac{d(-\lambda\pi)}{d\tau_x} \leq 0,\ \frac{\partial\left[-\lambda(\Delta y + \pi)\right]}{\partial\tau_x} \leq 0.$$

2. *If $\omega_x \leq \mu$, which means that a long tenure during x, where $x = b$ or e, does not assist in*

learning a new skill in a newly merged firm, then:

$$\frac{d\left(-\lambda\Delta y\right)}{d\tau_x} \leq 0.$$

If $\omega_x \geq \mu$, an increase in tenure would help in understanding the new environment in the merged firm and raise the speed of learning. Because an increase in tenure also increases valuable human capital specific to the target firm, a long tenure is valued by the newly merged firm and, therefore, it lowers the separation probability at all times.

However, if $\omega_x \leq \mu$, the opposite effect occurs. Because the human capital specific to the target firm is initially more highly valued, and because limited learning abilities do not initially cause any problem, the value of tenure is initially larger than it subsequently is. Therefore, a definite prediction of our theory is that the coefficient of tenure interacted with the $t = 0$ dummy must be negative. Equipped with the theoretical predictions outlined in this section, we now conduct our empirical study.

3 Empirical model

To analyze the determinants of the retention rate of target managers in a newly merged firm, we use the following stratified Cox proportional hazards model, which summarizes our theory:

$$\begin{aligned} q_{fjt} &= B\left(t : \mathbf{X}_{Z,f}\right) e^{-\lambda\left[\Delta y_{fj} I(t=0) + \pi_{fj}\right]} \\ \pi_{fj} &= \tilde{\alpha}_b \tau_{b,fj} + \tilde{\alpha}_e \tau_{e,fj} + \tilde{\alpha}_D' \mathbf{D}_{fj} + \tilde{\alpha}_X' \mathbf{X}_f + \tilde{\alpha}_\chi \chi_{fj} \\ \Delta y_{fj} &= \bar{\alpha}_b \tau_{b,fj} + \bar{\alpha}_e \tau_{bf,j} + \bar{\alpha}_D' \mathbf{D}_{fj} + \bar{\alpha}_X' \mathbf{X}_f + \bar{\alpha}_\chi \chi_{fj}, \end{aligned}$$

where $\tau_{b,fj}$ is the jth person's tenure as a board member at firm f, $\tau_{e,fj}$ is the jth person's tenure as an employee at firm f, the vector $\mathbf{D}_{fj} = [\mathbf{D}_{T,fj}, \mathbf{D}_{A,fj}, \mathbf{D}_{G,fj}, a_{fj}]$ contains any other target-firm manager characteristics of the jth person in firm f, including the jth person's age when firm f was taken over, and the vector $\mathbf{X}_f$ contains any characteristics of the fth firm that can influence the marginal productivity of managers. Finally, the vector $\mathbf{X}_{Z,f}$ contains any firm-level variables that can influence baseline hazard.

This equation depends on the individual unobserved heterogeneity of a manager j at firm f, χ_{fj}. Without controlling for χ_{fj}, our estimates might be biased. As explained in footnote 3, we focus on titled directors in Japanese companies. Because talented workers are likely to be appointed to titled director positions with less experience than less talented workers, χ_{fj} and

$\tau_{b,fj}$ can be correlated. As we can only observe data on titled directors, a simple instrumental variable regression cannot solve the problem of bias. If an exogenous event increases $\tau_{b,fj}$ for a treated group, the treated group can be in the sample, but the control group cannot. Therefore, it is likely that the average unobserved ability of the treated group in the sample is lower than that of the control group in the sample. This causes a systematic bias.

In order to deal with this bias, we estimate unobserved abilities from our data and control for them in our survival analysis. The following subsection explains how to estimate χ_{fj}.

3.1 Dealing with unobserved heterogeneity

This subsection explains the procedure to estimate χ_{fj}. Suppose that a person is appointed to a management position (titled director) in Japan if and only if:

$$\hat{h}_{fj}\left(\hat{t}\right) \geq \mathbf{H}'\mathbf{N}_h\left(f, \hat{t}\right), \tag{8}$$

where $\hat{h}_{fj}\left(\hat{t}\right)$ is the overall human capital of a person j at the fth firm in calendar year $\hat{t}$, and $\mathbf{H}'\mathbf{N}_h\left(f, \hat{t}\right)$ is the level of human capital required for appointment to a managerial position, where $\mathbf{H}$ is a parameter vector and a vector $\mathbf{N}_h\left(f, \hat{t}\right)$ summarizes the conditions that influence the required level of human capital of firm f in calendar year $\hat{t}$.

We assume the following human capital accumulation function:

$$\hat{h}_{fj}\left(\hat{t}\right) = \phi_0 + \phi_b \tau_{b,fj}\left(\hat{t}\right) + \phi_e \tau_{e,fj} + \phi_o \tau_{o,fj} + \chi_{fj}, \tag{9}$$

where $\tau_{b,fj}\left(\hat{t}\right)$ is the tenure as a board member of person j at firm f in calendar year $\hat{t}$ and $\tau_{o,fj}$ is the jth person's length of *experience before joining firm* f. Combining equations (8) and (9), when an employee j is appointed to an executive position, the following equality must be satisfied:

$$\tau_{b,fj}\left(\hat{t}_j^*\right) = \tilde{\mathbf{H}}'\mathbf{N}_h\left(f, \hat{t}_j^*\right) - \tilde{\phi}_0 - \tilde{\phi}_e \tau_{e,fj} - \tilde{\phi}_o \tau_{o,fj} - \tilde{\chi}_{fj}, \tag{10}$$

where $\tilde{\mathbf{H}} = \frac{\mathbf{H}}{\phi_b}$, $\tilde{\phi}_0 = \frac{\phi_0}{\phi_b}$, $\tilde{\phi}_e = \frac{\phi_e}{\phi_b}$, $\tilde{\phi}_o = \frac{\phi_o}{\phi_b}$, $\tilde{\chi}_{fj} = \frac{\chi_{fj}}{\phi_b}$, and $\hat{t}_j^*$ is the calendar year when employee j is appointed to an executive officer position. This equation implies that, after controlling for several observable variables, $\tau_{b,fj}\left(\hat{t}_j^*\right)$ is negatively correlated with $\tilde{\chi}_{fj}$. This means that $\tau_{b,fj}\left(\hat{t}_j^*\right)$ potentially contains useful information about $\tilde{\chi}_{fj}$.

Let us define a function R such that:

$$R\left(f,j,\hat{t}\right) \equiv \tilde{\mathbf{H}}'\mathbf{N}_h\left(f,\hat{t}\right) - \tilde{\phi}_0 - \tilde{\phi}_e\tau_{e,fj} - \tilde{\phi}_o\tau_{o,fj} - \tau_{b,fj}\left(\hat{t}\right).$$

Given the estimates of $\tilde{\mathbf{H}}$, $\tilde{\phi}_0$, $\tilde{\phi}_e$, and $\tilde{\phi}_o$, we can estimate $\tilde{\chi}_j$ from the following:

$$\tilde{\chi}_{fj} = R\left(f,j,\hat{t}_j^*\right).$$

Hence, we seek to obtain unbiased estimators of $\tilde{\mathbf{H}}$, $\tilde{\phi}_0$, $\tilde{\phi}_e$, and $\tilde{\phi}_o$. However, this is not easy because we can observe only a selected sample.

We define the deviation of the unobserved ability from the firm-level average in year $\hat{t}$ for the observed sample, $\varepsilon_{fj}\left(\hat{t}\right)$, as follows:

$$\varepsilon_{fj}\left(\hat{t}\right) \equiv \tilde{\chi}_{fj} - E\left[\tilde{\chi}_{fj}|I_b = 1, f, \hat{t}\right] = R\left(f,j,\hat{t}_j^*\right) - E\left[\tilde{\chi}_{fj}|I_b = 1, f, \hat{t}\right],$$

where I_b is the dummy variable that equals one if and only if a person is a titled director and is included in our sample and, therefore, $E\left[\tilde{\chi}_{fj}|I_b = 1, f, \hat{t}\right]$ is the conditional expectation of $\tilde{\chi}_{fj}$, given all relevant information at the fth firm in year $\hat{t}$ for the observed sample. Then, equation (10) can be rewritten as follows:

$$\tau_{b,fj}\left(\hat{t}_j^*\right) = \tilde{\mathbf{H}}'\mathbf{N}_h\left(f,\hat{t}_j^*\right) - \tilde{\phi}_0 - \tilde{\phi}_e\tau_{e,fj} - \tilde{\phi}_o\tau_{o,fj} - E\left[\tilde{\chi}_{fj}|I_b = 1, f, \hat{t}\right] - \varepsilon_{fj}\left(\hat{t}\right), \forall\hat{t}. \tag{11}$$

Note that $E\left[\varepsilon_{fj}\left(\hat{t}\right)|I_b = 1\right] = 0$ by construction. Note also that the deviation of the unobserved ability from the conditional expectation of $\tilde{\chi}_{fj}$ at the fth firm in year $\hat{t}$ for the observed sample, $\varepsilon_{fj}\left(\hat{t}\right)$, is uncorrelated with firm-level variables. Hence, we expect to satisfy an orthogonality condition that $E\left[\varepsilon_{fj}\left(\hat{t}\right)\mathbf{N}_f|I_b = 1\right] = 0$ for the vector of firm-level variables $\mathbf{N}_f$ for any $\hat{t}$. Therefore, if we know a functional form of $E\left[\tilde{\chi}_{fj}|I_b = 1, f, \hat{t}\right]$, we can obtain unbiased estimates of $\tilde{\mathbf{H}}$, $\tilde{\phi}_0$, $\tilde{\phi}_e$, and $\tilde{\phi}_o$ from a generalized method of moments (GMM) estimation, using firm-level variables from the sample of titled directors as its instruments.

In order to implement this idea, we need to estimate $E\left[\tilde{\chi}_{fj}|I_b = 1, f, \hat{t}\right]$. Because we use the sample of titled directors at target firms in the year when an M&A occurs, equation (8) implies that $I_b = 1$ if and only if $\hat{h}_{fj}\left(\hat{t}_f^m\right) \geq \mathbf{H}'\mathbf{N}_h\left(f,\hat{t}_f^m\right)$, where $\hat{t}_f^m$ is the calendar year in which the M&A occurs for target firm f. Using equation (9) and a function $R\left(f,j,\hat{t}\right)$,

$\hat{h}_{fj}\left(\hat{t}_f^m\right) \geq \mathbf{H}'\mathbf{N}_h\left(f, \hat{t}_f^m\right)$ can be rewritten as a condition for $\tilde{\chi}_{fj}$ as follows:

$$I_b = 1 \; iff \; \tilde{\chi}_{fj} \geq R\left(f, j, \hat{t}_f^m\right),$$

for any f and j. This means that the observed sample is considered to be drawn from an upper-tailed distribution of $\tilde{\chi}_{fj}$ and that:

$$E\left[\tilde{\chi}_{fj}|I_b = 1, f, \hat{t}_f^m\right] = E\left[\tilde{\chi}_{fj}|\tilde{\chi}_{fj} \geq R\left(f, j, \hat{t}_f^m\right), f, \hat{t}_f^m\right].$$

We need to estimate $E\left[\tilde{\chi}_{fj}|\tilde{\chi}_{fj} \geq R\left(f, j, \hat{t}_f^m\right), f, \hat{t}_f^m\right]$ to obtain unbiased estimators of the parameters. The standard two-stage estimator in Heckman (1979) uses a random sample from the population to estimate selection equations, the parameters of which are used to construct an inverse Mills ratio to estimate $E\left[\tilde{\chi}_{fj}|\tilde{\chi}_{fj} \geq R\left(f, j, \hat{t}_f^m\right), f, \hat{t}_f^m\right]$, by assuming that $\tilde{\chi}_{fj}$ is drawn from a normal distribution. Typically, exclusive variables are needed for the estimation of selection equations to obtain reliable estimates.

Unfortunately, a random sample from the population is not available in our case. However, note that our conditional expectation of unobserved ability, $E\left[\tilde{\chi}_{fj}|\tilde{\chi}_{fj} \geq R\left(f, j, \hat{t}_f^m\right), f, \hat{t}_f^m\right]$, is a function of $R\left(f, j, \hat{t}_f^m\right)$, which is determined by the parameters of our structural equation (11), $\tilde{\mathbf{H}}$, $\tilde{\phi}_0$, $\tilde{\phi}_e$, and $\tilde{\phi}_o$. Therefore, it might be possible to jointly estimate the parameters on the conditional expectation of unobserved ability with equation (11) using the GMM. This is possible because, although we have only a selected sample, the timing of selection is different across individuals, which gives us information about selection equations. In addition, the timing of observations, $\hat{t}_f^m$, is generally different from the timing of promotions, $\hat{t}_j^*$. Therefore, we can find several exclusive variables that are needed to identify parameters and to obtain reliable estimates.

The remaining concern is determining what is a plausible distribution of $\tilde{\chi}_{fj}$. Gabaix and Landier (2008) argue that an extreme value theory can provide a nice approximation for the upper tail of a large class of continuous distributions, including uniform, Gaussian, exponential, lognormal, Weibull, Gumbel, Fréchet, Pareto, stretched exponential, and log-gamma distributions. More specifically, suppose that $\tilde{\chi}$ is drawn from a distribution function $F(\tilde{\chi})$, where $F'(\tilde{\chi})$ is differentiable in a neighborhood of the upper bound of its support $\bar{\chi} \in R \cup \{+\infty\}$, and there exists $\xi = \lim_{\tilde{\chi} \to \bar{\chi}} \frac{d}{d\tilde{\chi}} \frac{1-F(\tilde{\chi})}{F'(\tilde{\chi})}$ and $\xi < \infty$. We define $\bar{F}(\tilde{\chi}) = 1 - F(\tilde{\chi})$ and $Q(x) = \bar{F}^{-1}(x)$. Gabaix and Landier (2008) apply an extreme value theorem and show that there exists a χ_0 and ζ such that:

$$Q'(x) \approx -\chi_0 x^{\zeta-1}.$$

Our appendix proves the following lemma.

Lemma 5 *Suppose that $Q'(x) = -\chi_0 x^{\zeta-1}$, where $\zeta \geq 0$, and that there exists a $Q(0) = \lim_{\varepsilon\to 0}\left[Q(\varepsilon) + \frac{\chi_0\varepsilon^\zeta}{\zeta}\right]$ and $Q(0) < \infty$. Then, for any small x_p, we have:*

$$E[\chi | x \leq x_p] = \frac{\zeta Q(0) + Q(x_p)}{\zeta+1}.$$

The assumption in Lemma 5 requires that the ability distribution has a finite upper bound, $Q(0)$. Gabaix and Landier (2008) provide empirical evidence that supports this assumption using data on U.S. compensation of CEOs.

We assume that a function Q may differ across firms f and across years $\hat{t}$, $Q\left(x : f, \hat{t}\right)$, and that $Q\left(0 : f, \hat{t}_f^m\right) = \mathbf{\Psi}'\mathbf{N}_\chi\left(f, \hat{t}_f^m\right)$, where $\mathbf{N}_\chi\left(f, \hat{t}\right)$ is a vector of variables that influence a potential upper bound of talent in firm f in calendar year $\hat{t}$. Let us choose x_p so that $x_p = \bar{F}\left(R\left(f, j, \hat{t}_f^m\right) : f, \hat{t}_f^m\right)$. Then, $\{x \leq x_p\} = \left\{\tilde{\chi}_{fj} \geq R\left(f, j, \hat{t}_f^m\right)\right\}$ and $Q\left(x_p : f, \hat{t}_f^m\right) = R\left(f, j, \hat{t}_f^m\right)$. Therefore, we can estimate the conditional expectation of ability in each firm $E\left[\tilde{\chi}_{fj} | \tilde{\chi}_{fj} \geq R\left(f, j, \hat{t}_f^m\right) : f, \hat{t}_f^m\right]$ by:

$$E\left[\tilde{\chi}_{fj} | \tilde{\chi}_{fj} \geq R\left(f, j, \hat{t}_f^m\right) : f, \hat{t}_f^m\right] = \frac{\zeta\mathbf{\Psi}'\mathbf{N}_\chi\left(f, \hat{t}_f^m\right) + R\left(f, j, \hat{t}_f^m\right)}{\zeta+1}.$$

In sum, we can obtain unbiased estimates of $\tilde{\mathbf{H}}$, $\tilde{\phi}_0$, $\tilde{\phi}_e$, and $\tilde{\phi}_o$ using the GMM based on the following orthogonality condition:

$$\begin{aligned} 0 &= E\left[\varepsilon_{fj}\left(\hat{t}_f^m\right)\mathbf{N}_f | I_b = 1\right], & (12)\\ \varepsilon_{fj}\left(\hat{t}_f^m\right) &= R\left(f, j, \hat{t}_j^*\right) - \frac{\zeta\mathbf{\Psi}'\mathbf{N}_\chi\left(f, \hat{t}_f^m\right) + R\left(f, j, \hat{t}_f^m\right)}{\zeta+1}, & (13)\\ R\left(f, j, \hat{t}\right) &\equiv \tilde{\mathbf{H}}'\mathbf{N}_h\left(f, \hat{t}\right) - \tilde{\phi}_0 - \tilde{\phi}_e\tau_{b,j} - \tilde{\phi}_o\tau_{o,j} - \tau_{d,j}\left(\hat{t}\right), & (14)\end{aligned}$$

where $\mathbf{N}_f = \left(\mathbf{N}_h\left(f, \hat{t}_j^*\right), \mathbf{N}_h\left(f, \hat{t}_f^m\right), \mathbf{N}_\chi\left(f, \hat{t}_f^m\right), \boldsymbol{\tau}_f\left(\hat{t}_f^m\right), \boldsymbol{\tau}_f\left(\hat{t}_j^*\right)\right)$, where $\boldsymbol{\tau}_f\left(\hat{t}\right)$ is the vector of the firm-level averages of $\tau_{b,fj}\left(\hat{t}\right)$, $\tau_{e,fj}$, and $\tau_{o,fj}$ in year $\hat{t}$. Using the estimated parameters, $\tilde{\mathbf{H}}$, $\tilde{\phi}_0$, ϕ_e, and $\tilde{\phi}_o$, we can estimate $\tilde{\chi}_{fj}$ by:

$$\tilde{\chi}_{fj} = R\left(f, j, \hat{t}_j^*\right). \tag{15}$$

Using $\tilde{\chi}_{fj}$, we can estimate:

$$\begin{aligned} q_{fjt} &= B\left(t : \mathbf{X}_{Z,f}\right) e^{-\lambda\left[\Delta y_{fj} I(t=0) + \pi_{fj}\right]}, && (16) \\ \pi_{fj} &= \tilde{\alpha}_b \tau_{b,fj}\left(\hat{t}_f^m\right) + \tilde{\alpha}_e \tau_{e,fj} + \tilde{\alpha}_a a_{fj} + \tilde{\alpha}_D' \mathbf{D}_{fj} + \tilde{\alpha}_X' \mathbf{X}_f + \alpha_\chi^+ \hat{\chi}_{fj}, && (17) \\ \Delta y_{fj} &= \bar{\alpha}_b \tau_{b,fj}\left(\hat{t}_f^m\right) + \bar{\alpha}_e \tau_{e,fj} + \bar{\alpha}_a a_{fj} + \bar{\alpha}_D' \mathbf{D}_{fj} + \bar{\alpha}_X' \mathbf{X}_f + \bar{\alpha}_\chi^+ \chi_{fj}, && (18) \end{aligned}$$

where $\tilde{\alpha}_\chi^+ = \phi_b \tilde{\alpha}_\chi$ and $\bar{\alpha}_\chi^+ = \phi_b \bar{\alpha}_\chi$.

4 Data

4.1 Firm-level data

Below, we test the propositions outlined in the previous sections, focusing on M&As during the period 1990–2006 in Japan. We identify Japanese M&As in this period from the *Delisting dataset*, which is manually constructed from various sources of information, including *Kaisha Nenkan* 1969–2006, *Kaisha Shikiho* 2000–2006, and *Tosho Yoran* 1972–1973 and 1975–2007. This dataset contains information on delisted firms in exchange markets throughout Japan from 1968 to 2007. The information includes the stock code and name of the delisted firm, the listed market, dates of listing and delisting, the reason for delisting, and the name and the stock code of the new firm, when the reason given for delisting was a merger or an acquisition.

From the database, we selected those firms that had been delisted because of a "merger" or a "full-ownership acquisition" during the sample period[10]. When the firms that were delisted as the result of M&As involve a consolidation of assets and liabilities under a company with a stock code, we can observe who is retained in the new firm after the M&A as a board member. We select these types of delisted firms, and refer to them as "target firms". In most M&As, one of the merging firms survives as the same firm with the same stock code. In other cases, mergers result in a new company with a new stock code. Our dataset includes both cases, and we refer to the postmerger firms in both cases as "new firms" in this paper. When we discuss surviving firms with the same stock codes after M&As, we refer to them as "acquirers" and we refer to all premerger firms that transacted with target firms during M&As as as "opponent firms".

Our sample contains 344 M&A cases, comprising 123 mergers and 221 full-ownership ac-

[10] Other reasons given for delisting include "stock transfer", "bankruptcy", "insolvency", "window-dressing settlement", and "business suspension". There was a single case of stock transfer in our sample, but we excluded it because the target firm was in the JASDAQ market.

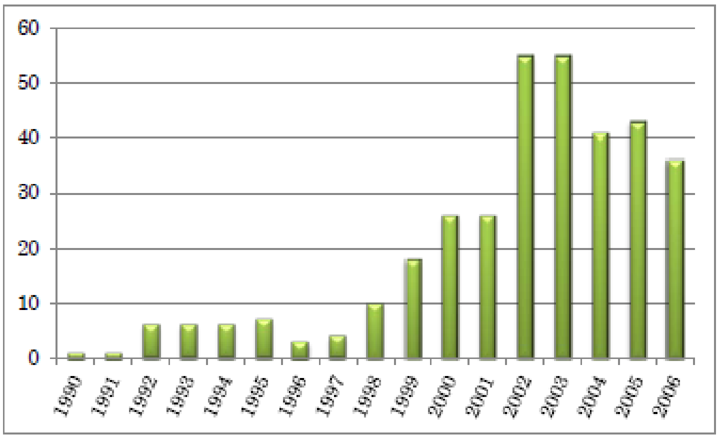

Figure 1: Number of M&As each year during the period 1990–2006

quisitions. Only mergers between listed firms are included. Figure 1 depicts the number of sample M&As over the study period. We can see that the number of M&As rapidly increased after 1998.

The exchanges on which the target firms are listed include the Tokyo, Osaka, Nagoya, and other local stock exchanges, as well as emerging and over-the-counter markets. However, we exclude cases in the JASDAQ market because director-level data for these firms are available only after 2000.

We merge data on firm characteristics for the target firms and the new firms with the above M&A data. The data on firm characteristics are from the *Nikkei NEEDS* database. We first merge the three-digit Nikkei industry code for each target and new firm.

Table 1 presents the numbers of target and new firms in our sample, based on a two-digit industry classification[11]. The second column shows the number of target firms by industry, whereas the third column shows the number of new firms by industry. The fourth column shows the number of M&A cases where the target and new firms are from the same industry. We can see that M&As within the same industry occurred most often in the electronics industry.

Next, we merge the financial characteristics in the *Nikkei NEEDS* database with our sample. More specifically, we merge operating income, sales, number of full-time employees, personnel expenses, total assets, and the stock share of the top ten stockholders for each target firm. These variables are explained in detail below. We also merge operating income, total assets, and the stock share of the top ten stockholders of the acquirer of the target firm in the year

[11] We exclude financial institutions such as banks, stock companies, and insurance companies because of the difference in accounting systems.

Industry	No. of target firms	No. of new firms	Both
Food	15	16	14
Textile	7	8	5
Pulp	13	13	10
Chemical	19	21	14
Pharmaceutical	6	3	2
Petroleum	3	6	3
Rubber	2	1	1
Ceramic	14	13	10
Iron ore	7	13	7
Nonferrous metal and metal	15	18	10
Machinery	24	24	16
Electronics	27	46	24
Shipbuilding	0	1	0
Automobile	9	9	4
Transport machinery	6	0	0
Precision instruments	2	2	1
Other manufacturers	10	9	6
Marine products	1	0	1
Mining	4	2	2
Construction	26	26	16
Trade	43	32	22
Retail	27	31	21
Other financial businesses	2	4	2
Real estate	11	1	1
Rail and bus	2	12	2
Land transportation	2	4	2
Marine transportation	7	6	6
Air transportation	1	3	1
Telecommunications	4	2	2
Electricity	0	2	0
Service	35	15	13
Total	344	344	218

Table 1: Numbers of mergers and acquisitions by industry

prior to the M&A if there is a surviving firm with the same stock code. For cases of M&As without any surviving firm, we use the average value of these variables for all opponent firms instead to represent those of the acquirer.

4.2 Personnel data

Next, we merge the firm-level data with the director-level data. The data for board members are taken from the *Directors data* published by *Toyo Keizai*. This database contains information on the directors of all listed firms in Japan from 1990 to 2007. The information includes the title, the date the person entered the firm, and a personal history of the board member, along with the name and date of birth.

Because we are interested in the retention of Japanese executives after M&As, we focus our analysis on directors with titles in the sample target firms[12]. Therefore, our sample comprises all directors with titles at target firms in the year of the M&A[13]. As explained in footnote 3, following the arguments by Kaplan (1994) and Saito and Odagiri (2008), we consider titled directors to be the top executives in Japanese companies. In the following sections, we analyze whether executives in target firms are retained as board members in the new firms after M&As.

We have 1520 observations (titled directors) for 343 targets, with an average of 4.43 titled directors per target. We specify the date of birth, title, the date a person entered the firm, and her/his history as a board member. We also identify the number of other firms (including the target firm itself) in which target-firm managers served as board members in the year of the M&A.

From the *Directors data*, we determine whether a target manager was retained as a board member (with or without a title) in the new firm after the M&A and, if so, how many years he/she stayed on the board after the M&A[14]. In addition, we identify whether a target-firm manager served as a board member of the acquirer before the M&A.

Table 2 provides summary statistics of the managers' retention characteristics. The first variable (*retained*) takes a value of one if the target manager became a board member of the new firm after the M&A and is zero otherwise. The second variable (*years of survival if retained*)

[12] Auditors are excluded.

[13] More precisely, titled directors who served in their positions until a year before the M&A are also included in our sample because of a gap between the M&A date and the recording date for the *Directors data*. We assume that titled target directors who resigned their positions more than a year before the M&A did so for reasons other than the M&A, and we exclude them from our sample.

[14] We consider the target manager to have been retained even if he/she becomes a director with a different title in the new firm or if the manager loses his/her title in the new firm. In addition, in the cases where an opponent firm of an M&A survives as the new firm and the manager of the target firm was also a director in this opponent firm and continues to be a director in the new firm, we consider this manager to have been retained.

is the number of years that the retained target manager served as a board member of the new firm after the M&A. The table shows that only 38.7% of the target managers were retained as board members after M&As and that the retained managers kept their positions for an average of less than five years.

Variable	No. of observations	Mean	S.D.	Min.	Max.
Retained	1520	0.387	0.487	0	1
Years of survival if retained	588	4.117	2.190	1	16

Table 2: Characteristics of target directors

4.3 Empirical specification

4.3.1 GMM estimation

Using the dataset explained above, we first conduct a GMM estimation using equations (12), (13), and (14) to obtain the unobserved individual ability defined in equation (15).

The summary statistics of the variables used in the GMM estimation can be found in Table 3. The first four variables are those of the manager's tenure history, which determine his/her human capital at $\hat{t}_j^*$ and $\hat{t}_f^m$ as in equation (9). *Tenure as a board member*, $\tau_{b,fj}\left(\hat{t}\right)$, is the manager's length of tenure after he/she became a board member in the target firm. We calculate tenure as a board member in the years $\hat{t}_j^*$ and $\hat{t}_f^m$ from the target manager's history as a board member in the target firm. *Tenure as an employee*, $\tau_{e,fj}$, is the manager's length of tenure as an employee in the target firm, which is calculated by determining the tenure before he/she became a board member, and *outside experience*, $\tau_{o,fj}$, is calculated from the manager's age (calculated from the date of birth) minus his/her entire period of tenure in the target firm.

For the variables that determine the human capital level required for appointment to a managerial position (a titled director position), $\mathbf{N}_h\left(f,\hat{t}\right)$, we use *Negative operating income*, *Direct*, *No. of employees*, *Wage*, *Median tenure as a titled director*, and *Median age at* $\hat{t}_f^m$. The summary statistics of these variables in the years $\hat{t}_j^*$ and $\hat{t}_f^m$ are shown in rows 7 to 16 in the table[15]. Negative operating income is an indicator variable that takes a value of one if the operating income of the target firm is negative and is zero otherwise. Kaplan (1994) shows that, if the operating income is negative, then the current managers are likely to be dismissed.

[15] More specifically, we calculate the values of these variables immediately before the years $\hat{t}_j^*$ and $\hat{t}_f^m$, except for *Direct*, which takes the same value for these two years. For median tenure as a titled director, we calculate the median tenure of all the titled directors of firm f, excluding that of the target manager. The tenure of each target manager is calculated on a monthly basis and then expressed in years. When we use the variables from Nikkei NEEDS, we use the published data for the latest fiscal year to $\hat{t}_j^*$. In some cases, especially for variables at $\hat{t}_j^*$, the financial data are not available. In these cases, we use the data for the year nearest to $\hat{t}_j^*$.

Because a firm must replace managers with new appointees for an exogenous reason in this case, the required level of human capital would be lower than is usually the case for appointments. Therefore, we expect a negative impact from negative operating income on the required level of human capital.

Direct is an indicator variable that takes a value of one if the manager is hired from outside the firm. We control for this variable because the required human capital level may differ depending on whether the manager is selected from among the employees of the firm or is hired from outside. We control for the number of employees to estimate the required human capital level because more employees may mean there is more intense competition to become a manager on the board. Therefore, we expect that the number of employees will have a positive effect on the required human capital level. Similarly, a higher average wage may imply that employees of the firm have a higher average human capital level. Therefore, a greater level of human capital may be required for appointment to a titled director position. The average wage is calculated using the number of employees and the personnel expenses of the target firm.

We include the median length of tenure as a titled director within a firm and consider the turnover rate of titled directors to be low if the tenure length as a titled director is high, on average. In such cases, it is more difficult to become a titled director. We also control the median value of age in the merger year to control for the customs of the firm. If the median age tends to be high, it may take time to become a titled director even once one has a sufficient level of human capital. Therefore, we expect its coefficient to be positive.

For a variable in a vector $N_{\chi}\left(f, \hat{t}_f^m\right)$, we assume that the potential highest level of ability in a target firm depends on the firm's size. We use the sales volume of the target firm as a proxy for the target firm's size. The summary statistics for the sales volume are shown in row 18 of the table. The next three variables in Table 3 represent the instrumental variables for the endogenous variables in the model, that is, the instrumental variables for tenure as a board member at $\hat{t}_f^m$, tenure as an employee, and outside experience. We use the target firm's median value at $\hat{t}_f^m$ for these variables as instruments.

4.3.2 Survival analysis

Once we obtain the unobserved individual ability, $\hat{\chi}_{fj}$, from the GMM estimation, we analyze the determinants of the retention of target-firm managers after M&As. We utilize the stratified Cox proportional hazards model in equations (16), (17), and (18), which requires us to specify

Variable	No.	Mean	S.D.	Min.	Max.
Tenure as a board member at $\hat{t}_j^*$ (years)	1520	2.70	2.91	0.00	20.92
Tenure as a board member at $\hat{t}_f^m$ (years)	1520	8.50	7.41	0.25	56.42
Tenure as an employee (years)	1520	11.82	13.59	0.00	43.17
Outside experience (years)	1520	40.35	15.62	15.17	71.17
Age (years)	1520	60.67	5.54	35.08	90.92
Variables in $\mathbf{N}_h\left(f,\hat{t}\right)$, for $\hat{t}=\hat{t}_j^*,\hat{t}_f^m$					
Negative operating income at $\hat{t}_j^*$	1520	0.15	0.36	0.00	1.00
Negative operating income at $\hat{t}_f^m$	1520	0.21	0.41	0.00	1.00
Direct	1520	0.20	0.40	0.00	1.00
No. of employees at $\hat{t}_j^*$ (1000)	1520	1.70	2.65	0.006	23.87
No. of employees at $\hat{t}_f^m$ (1000)	1520	1.53	2.24	0.006	16.35
Wage at $\hat{t}_j^*$ (million yen)	1520	4.25	3.37	0.08	28.39
Wage at $\hat{t}_f^m$ (million yen)	1519	5.03	3.89	0.21	25.57
Median tenure as a titled director at $\hat{t}_j^*$ (years)	1520	4.89	4.00	0	33.17
Median tenure as a titled director at $\hat{t}_f^m$ (years)	1520	4.61	3.19	0.50	24.67
Median age of titled directors at $\hat{t}_f^m$ (years old)	1520	55.35	4.34	27.58	69.17
Variables in $\mathbf{N}_\chi\left(f,\hat{t}_f^m\right)$					
Sales (1000 million yen)	1520	146.78	256.95	0.68	2877.40
$\boldsymbol{\tau}_f\left(\hat{t}_f^m\right)$, $\boldsymbol{\tau}_f\left(\hat{t}_j^*\right)$ (instrumental variables)					
Median value of tenure as a board member at $\hat{t}_f^m$ (years)	1520	7.48	4.17	0.25	32.42
Median value of tenure as an employee (years)	1520	10.84	12.17	0.00	35.83
Median value of outside experience (years)	1520	40.20	13.66	20.42	68.42
Manager characteristics					
No. of firms as a board member	1520	1.14	0.51	1.00	10.00
Board member in the acquirer before the M&A	1520	0.15	0.35	0.00	1.00
Upper	1520	0.31	0.46	0.00	1.00
Target ROA	1520	0.00	0.06	−0.43	0.52
Firm characteristics					
Log target assets	1520	11.07	1.35	6.92	14.61
Board size	1520	5.77	2.73	1.00	15.00
Log stock share of top 10	1520	−6.89	1.35	−8.39	−0.14
Target firm's relative size in assets	1520	0.27	0.23	0.00	0.99
Related	1520	0.48	0.50	0.00	1.00
Acquisition	1520	0.61	0.49	0.00	1.00
Acquirer or opponent ROA	1520	0.00	0.06	−0.68	0.60
Log top 10 share of acquirer or opponents	1520	−7.09	1.56	−8.89	0.00

Table 3: Summary statistics

the vector $\mathbf{D}_{fj}$ of target-firm manager characteristics other than tenure at the time of the M&A, the vector $\mathbf{X}_f$ that influences the marginal productivity of managers, which includes the target-firm characteristics and the transaction characteristics suggested by Wulf and Singh (2011)[16], and the vector $\mathbf{X}_{Z,f}$, which contains any variables that influence the baseline hazard.

The vector $\mathbf{D}_{fj}$ includes the variables that affect learning ability, the target-specific human capital, the acquirer-specific knowledge of the manager, and the manager's age at the time of the M&A (*Age*). For the variables in the vector $\mathbf{D}_{fj}$, other than age, we use the number of firms in which the target manager served as a board member (*No. of firms as a board member*), whether the target manager was a board member in the acquirer before the M&A (*Board member in the acquirer before the M&A*), and whether the target manager was the president or chairman of the board of directors (*Upper*). These variables are calculated from the *Directors data*, as explained in the previous section. We include the target firm's performance as a manager characteristic to measure the manager's ability. To represent the target firm's performance, we use the return on assets (ROA) in the year prior to the M&A, where the ROA is defined as operating income divided by assets, measured using the deviation from the industry median. Note that this measure is at the firm level, even though we consider it as a manager characteristic, as suggested by Wulf and Singh (2011).

The target firm's characteristics in $\mathbf{X}_f$ include firm size and the board size of target firms. To control for the target firm's size and its board size, we include the logarithm of the target firm's assets in the year prior to the merger (*Log target assets*) and the number of board members (*Board size*). The previous literature has considered the target firm's size to be important because larger and more complex firms may require managers with unique skills, which may increase the retention rate of managers (Finkelstein and Hambrick, 1989)[17]. We include the board size of the target firm because an increase in the number of managers may reduce the marginal productivity of managers in a target firm.

Following the previous literature, including Wulf (2004), Hartzell et al. (2004), and Barg-

[16] Wulf and Singh (2011) investigate the determinants of target CEO retention rates. Our analysis differs in that we include other directors with a title. As for the determinants of CEO retention rates, Wulf and Singh (2011) consider target-firm characteristics, CEO characteristics, transaction characteristics, and acquirer governance characteristics. They also include variables such as target-firm CEO compensation and the number of days as CEO until completion of the merger. Unfortunately, these variables are not available for our chosen context. However, we are able to include more detailed director characteristics because of the richness of our personnel data.

[17] The models regarding span of control (e.g., Lucas, 1978, Rosen, 1982) predict that a large firm should be operated by a talented person. Recent assignment theories of managerial compensation (e.g., Terviö, 2007, and Gabaix and Landier, 2008) make this assumption and find that this model can quantitatively explain the relation between CEO compensation and firm size. If this prediction is correct, the size of a firm can also be considered as a manager characteristic.

eron, Schlingemann, Stulz, and Zutter (2009), we consider a moral hazard problem relating to the process of the merger. This problem arises because a manager may control the sale of the firm and refuse to take a deal in the M&A process unless he/she is promised something personally in return, such as a position on the board of the new firm. Because our theory is based on the presumption that a manager cannot write an explicit contract on this type of promise, any promise made before the M&A cannot influence the negotiation that occurs after the M&A. However, if a target-firm manager and the opponent firms can write a binding contract about a post on the board of the new firm, this may influence the separation rate.

In order to take into account such a possibility, Bargeron, Schlingemann, Stulz, and Zutter (2009) control for a measure of the ownership structure, specifically, the level of insider ownership in a target firm. Following this idea, we include the logarithm of the stock share of the top 10 stock holders (*log stock share of top 10*) in $\mathbf{X}_f$. The impact of strategic retention would be smaller if the value of this variable were larger because large shareholders have more incentive to monitor managers under a concentrated ownership structure.

The transaction characteristics in $\mathbf{X}_f$ include the relative size of the target firm. If the firm is large, the postmerger integration may be more difficult and, therefore, we would expect the likelihood of manager retention to be higher (Zollo and Singh, 2004). To measure relative firm size, we calculate the ratio of the target firm assets to the acquirer assets if there is a surviving firm with the same stock code. For cases of M&As without any surviving firm, we use the total assets of all opponent firms in place of the acquirer assets. Following the previous literature, including Walsh (1989) and Buchholtz et al. (2003), we include a dummy variable that represents whether the target and the new firm operate in the same industry as a transaction characteristic. More specifically, we create an indicator variable that is equal to one if the new and target firms operate in the same three-digit Nikkei industry code (i.e., medium industry classification) and is zero otherwise (*Related*). We also include a full-ownership acquisition dummy to observe whether mergers and full-ownership acquisitions have different effects on the retention rate (*Acquisition*).

We control for the governance characteristics of the opponent firms. Specifically, we control for the ROA and the stock share of the top 10 stock holders of the acquirer if there is a surviving firm with the same stock code. For cases of M&As without any surviving firm, we use the average value of these variables for all opponent firms in place of the acquirer values ("Acquirer or opponent ROA" and "Log top 10 share of acquirer or opponents").

The summary statistics of these variables, which appear exclusively in the stratified Cox proportional hazards estimation, are shown in the last two groups in Table 3. In addition to these variables, we need tenure as a board member at $\hat{t}_f^m$, tenure as an employee, and age, for which summary statistics are shown in the first group of the table. Tenure and age variables are calculated as explained above.

In addition, for the vector $\mathbf{X}_{Z,f}$ in equation (16), we include two-digit industry codes for a new firm and a target firm and the year of the M&A. This means that we allow the baseline hazard functions to differ for these groups. Finally, we construct an initial dummy that assigns a value of one for the initial year ($t = 0$) and zero for any other separation time[18].

5 Results

Table 4 shows the results of our GMM estimation. In addition to the variables explained in the previous section, we include two-digit industry dummies and decade dummies in $\mathbf{N}_h\left(f, \hat{t}_j^*\right)$ and $\mathbf{N}_h\left(f, \hat{t}_f^m\right)$. We select a group of exogenous variables so that the estimation passes the over-identification test and use them as instruments, along with the instruments for the endogenous variables. The instrument variables we used are listed under Table 4.

We can see in Table 4 that the signs of the coefficients of the variables in the R function in equation (13) are all as expected. That is, a higher number of employees, higher wages, higher median tenure as a titled director, and higher median age seem to make it more difficult for a director to be appointed as a titled director, whereas poor performance of the current management makes it easier. The coefficient of *Direct* is negative, indicating that the human capital level required to be appointed to a titled director position is lower for managers from outside the firm. The coefficients for *Direct*, *Wage*, and *Median age* are statistically significant.

The results show that the coefficients for both tenure as an employee and outside experience are positive, implying that these variables raise individuals' human capital. However, because the values of these coefficients are smaller than one, their effect on human capital is smaller than that of tenure as a board member, i.e., experience as a board member seems to be more important for promotion to a management position than is experience as an employee. We estimate $1/(\zeta + 1)$ and obtain a value such that $0 < 1/(\zeta + 1) < 1$ without imposing any restrictions on estimation. This is consistent with the assumption of $\zeta > 0$ in Lemma 5.

[18]We expand our data in order to include this time-varying covariate. More specifically, each observation of titled directors who are retained in the new firms is expanded to two observations, one for the initial year and the other for the following years. Therefore, we have 2018 observations in total, as seen in our later estimation of the Cox proportional hazards model, comprising 588 observations with an initial dummy value of zero and 1520 observations with an initial dummy value of one.

The coefficients of the variables that affect the manager's human capital are all statistically significant. The coefficient of sales that appears in $\mathbf{N}_{\chi}\left(f, \hat{t}_f^m\right)$ is positive and significant, implying that, as the company size grows, the potential highest level of ability increases.

Using the estimated parameters, we obtain the unobserved ability, $\tilde{\chi}_{fj}$, for each manager in our sample from equation (15). Table 5 shows the summary statistics for individual unobserved ability, whereas Figure 2 represents the kernel density estimate of the estimated unobserved ability.

Variable	Coef.	Std. Err.	
Negative operating income	-0.636	0.404	
Direct	-1.964	1.030	*
No. of employees	0.000	0.000	
Wage	0.022	0.037	***
Median tenure as a titled director	0.025	0.028	
Median age	0.801	0.253	***
Constant ($\tilde{\phi}_0$)	73.183	13.474	***
Tenure as an employee (ϕ_e)	0.400	0.230	*
Outside experience ($\tilde{\phi}_o$)	0.382	0.224	*
$1/(\zeta+1)$	0.663	0.051	***
Sales	0.000	0.000	*
Industry code dummy	Yes		
Decade dummy	Yes		
Number of observations	1519		

Instruments for the equation: negative operating income, direct, no. of employees, wage, median tenure as a titled director, median age as a titled director, median tenure as a board member, median tenure as an employee, median outside experience, sales, and decade and industry dummy for both $\hat{t}_j^*$ and $\hat{t}_f^m$

Table 4: GMM estimation

	Obs.	Mean	Std. Dev.	Min.	Max.
$\tilde{\chi}_{fj}$	1520	-6.606	5.104	-44.191	5.654

Table 5: Estimated unobserved skill

Tables 6 and 7 show the results of our stratified Cox proportional hazards model estimation. All results pass the test of the proportional hazards assumption, based on the test of nonzero slope in a generalized linear regression of the scaled Schoenfeld residuals on a natural log of the analysis time at the 5% level. The bootstrap standard errors are shown in the tables for each coefficient.

Specification (1) in Table 6 is our benchmark specification. The result for specification (1) shows that the coefficient of tenure as a board member is not significant, whereas the coefficient of tenure as an employee is positive and significant. The coefficient of tenure as a board member

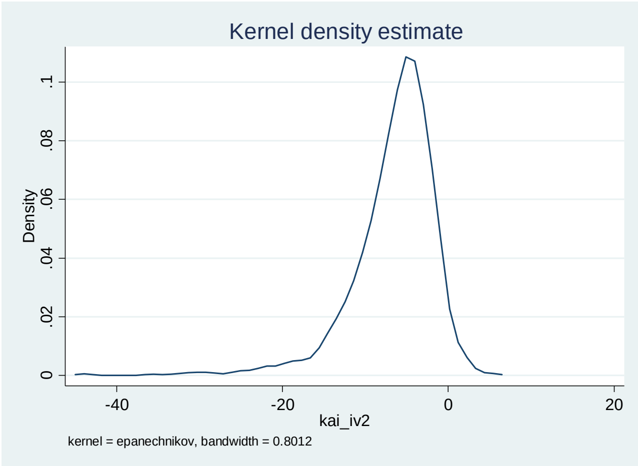

Figure 2: Kernel density of estimated unobserved skill

with an initial dummy is not significant, whereas the coefficient of tenure as an employee with an initial dummy is negative and significant. The sum of tenure as a board member and that with an initial dummy is negative, as is the sum of tenure as an employee and that with an initial dummy. A Wald test shows that the latter summed variable is statistically significant, whereas the former is not. The results indicate that, whereas the tenure as a board member has no impacts on the separation probability after M&As, irrespective of timing, a longer tenure as an employee increases the separation rate and the probability of appointment as a board member in a new firm, with statistical significance.

The coefficient for age is positive and significant, whereas that for age with an initial dummy is negative and significant. The sum of these variables is positive. Therefore, older managers are less likely to be retained either in the initial year after an M&A or in the years after.

Because many other control variables are insignificant in Specification (1), we suspect that there may be some multicollinearity among variables. Hence, we drop the variables with initial dummies for which both the level and interaction terms are insignificant in Specification (1). The result is shown as Specification (3) in Table 7. It is shown that the results on tenures are the same and, therefore, are robust.

The coefficients of many other control variables become significant in Specification (3). Specifically, the coefficient of board size is positive and significant. This may imply that the marginal productivity of managers decreases with the number of managers. The results also indicate that managers who serve as board members in different companies have a lower separation rate. We can see that managers who were in the acquirer before the M&A and managers who belonged to a target firm in the same industry as the new firm are likely to be retained on the board of the new firm. This may imply that acquirer-specific knowledge and industry-specific knowledge are appreciated by the new firm. We can also see that the target managers are less likely to be retained in cases of acquisition and that higher-level managers (*Upper*) are more likely to be retained.

Specifications (2) and (4) are the same as Specifications (1) and (3), respectively, except that the former do not include the estimated unobserved skill, $\tilde{\chi}_{fj}$. We can see that the results for Specifications (2) and (4) are almost the same as those for (1) and (3), respectively, except that the Wald test for the sum of the coefficients of tenure as a board member and its interaction term become negative and significant in both (2) and (4). This may imply that it is important to control for unobserved skill in our estimation. Whether unobserved skill is controlled for seems to most affect the coefficient of tenure as a board member.

Now, we wish to interpret the results using our theory. We first apply Proposition 2 to interpret the results of our estimation. Because the conditions in Proposition 2 must be satisfied by the coefficients of both tenure as a board member and tenure as an employee, if the coefficients of either type of tenure can reject the conditions, then the corresponding hypothesis in Proposition 2 is rejected. First, let us test hypotheses 1 and 2 in Proposition 2 using the result for tenure as an employee. Our estimated results on the coefficient of tenure as an employee, the coefficient of tenure as an employee with the initial dummy, and the sum of these two coefficients correspond to $\frac{\partial[-\lambda\pi]}{\partial\tau_e}$, $\frac{\partial[-\lambda\Delta y]}{\partial\tau_e}$, and $\frac{\partial[-\lambda(\Delta y+\pi)]}{\partial\tau_e}$, respectively. Our estimated coefficients are all statistically significant and the signs we obtained are $\frac{\partial[-\lambda\pi]}{\partial\tau_e} > 0$, $\frac{\partial[-\lambda\Delta y]}{\partial\tau_e} < 0$, and $\frac{\partial[-\lambda(\Delta y+\pi)]}{\partial\tau_e} < 0$. Hence, they reject both hypotheses $\tilde{\alpha}_T = \bar{\alpha}_T = 0$ and $\tilde{\alpha}_N = 0$ in Proposition 2. Therefore, the evidence suggests that new firms appreciate both specific human capital and the new skills that managers in target firms must learn after M&As.

Next, we separately investigate the role of tenure as a board member and that of tenure as an employee after M&As. Given that $\tilde{\alpha}_T \neq 0$ and $\bar{\alpha}_T \neq 0$, let us first apply Proposition 3 to the estimated results for the tenure as an employee. It shows that the evidence for tenure as an employee, $\frac{\partial[-\lambda\pi]}{\partial\tau_e} > 0$, $\frac{\partial[-\lambda\Delta y]}{\partial\tau_e} < 0$, and $\frac{\partial[-\lambda(\Delta y+\pi)]}{\partial\tau_e} < 0$, rejects the hypothesis that experience as an employee does not increase firm-specific skills ($\eta_e = 0$). Similarly, applying Proposition 4, given that $\tilde{\alpha}_N \neq 0$, the evidence for tenure as an employee, $\frac{\partial[-\lambda\pi]}{\partial\tau_e} > 0$, $\frac{\partial[-\lambda\Delta y]}{\partial\tau_b} < 0$, and $\frac{\partial[-\lambda(\Delta y+\pi)]}{\partial\tau_e} < 0$, rejects the hypothesis that $\omega_e > \mu$, although we cannot reject the hypothesis that $\omega_e < \mu$. The results consistently show that longer tenure as an employee of a target firm

	(1)			(2)		
Variable	Coef.	Std. Err.		Coef.	Std. Err.	
1: Tenure as a board member	0.006	0.021		−0.010	0.014	
2: Tenure as an employee	0.011	0.007	*	0.011	0.006	*
3: Tenure as a board member× initial dummy	−0.012	0.021		0.002	0.014	
4: Tenure as an employee × initial dummy	−0.014	0.007	**	−0.014	0.006	**
Age	0.112	0.027	***	0.100	0.027	***
Board size	0.097	0.082		0.105	0.075	
No. of firms as a board member	−0.166	0.213		−0.179	0.208	
Board member in the acquirer before the M&A	−0.106	0.504		−0.108	0.500	
Acquisition	−0.025	0.789		0.003	0.740	
Target ROA	−11.912	7.622		−11.335	7.193	
Related	0.117	0.558		0.034	0.521	
Target firm's relative size in assets	−0.990	0.542	*	−0.942	0.537	*
Log target assets	−0.026	0.216		−0.066	0.203	
Upper	−0.277	0.203		−0.264	0.192	
Log stock share of top 10	−0.460	0.315		−0.455	0.360	
Acquirer or opponents ROA	−3.164	5.814		−2.214	6.362	
Log top 10 share of acquirer or opponents	0.134	0.276		0.091	0.309	
$\tilde{\chi}_{fj}$	0.044	0.037				
Age $\times I(t=0)$	−0.072	0.027	***	−0.061	0.028	***
Board size $\times I(t=0)$	−0.073	0.082		−0.080	0.076	
No. of firms as a board member $\times I(t=0)$	−0.293	0.242		−0.281	0.230	
Board member in the acquirer before the M&A $\times I(t=0)$	−0.114	0.506		−0.112	0.499	
Acquisition $\times I(t=0)$	0.729	0.805		0.702	0.757	
Target ROA $\times I(t=0)$	11.813	7.631		11.198	7.258	
Related $\times I(t=0)$	−0.383	0.576		−0.300	0.533	
Target firm's relative size in assets $\times I(t=0)$	−0.210	0.595		−0.257	0.569	
Log target assets $\times I(t=0)$	0.012	0.222		0.054	0.208	
Upper $\times I(t=0)$	−0.007	0.212		−0.017	0.199	
Log stock share of top 10 $\times I(t=0)$	0.504	0.320		0.503	0.367	
Acquirer or opponents ROA $\times I(t=0)$	5.533	5.920		4.636	6.467	
Log top 10 share of acquirer or opponents $\times I(t=0)$	−0.161	0.276		−0.120	0.310	
$\tilde{\chi}_{fj} \times I(t=0)$	−0.037	0.039				
Wald Test	Coef.	Chi2(1)		Coef.	Chi2(1)	
1+3	−0.006	1.37		−0.008	4.04	**
2+4	−0.003	3.01	*	−0.003	2.91	*
Number of observations	2108			2108		

Stratified by year, acquirer industry code, and target industry code.

Table 6: Stratified Cox proportional hazards model estimation

	(3)			(4)		
Variable	Coef.	Std. Err.		Coef.	Std. Err.	
1: Tenure as a board member	−0.002	0.020		−0.016	0.013	
2: Tenure as an employee	0.010	0.006	*	0.010	0.006	*
3: Tenure as a board member × initial dummy	−0.005	0.020		0.007	0.013	
4: Tenure as an employee × initial dummy	−0.013	0.006	**	−0.012	0.006	**
Age	0.116	0.023	***	0.107	0.024	***
Board size	0.036	0.020	*	0.040	0.020	**
No. of firms as a board member	−0.358	0.097	***	−0.361	0.096	***
Board member in the acquirer before the M&A	−0.229	0.078	***	−0.232	0.076	***
Acquisition	0.579	0.153	***	0.578	0.157	***
Target ROA	−0.466	1.045		−0.463	1.001	
Related	−0.244	0.108	**	−0.252	0.107	**
Target firm's relative size in assets	−1.016	0.471	**	−1.006	0.443	**
Log target assets	−0.036	0.052		−0.040	0.052	
Upper	−0.279	0.061	***	−0.275	0.060	***
Log stock share of top 10	−0.009	0.060		−0.008	0.057	
Acquirer or opponents ROA	1.223	1.422		1.370	1.418	
Log top 10 share of acquirer or opponents	−0.028	0.051		−0.035	0.051	
$\tilde{\chi}_{fj}$	0.036	0.032				
Age × $I(t=0)$	−0.076	0.023	***	−0.069	0.024	***
Board size × $I(t=0)$						
No. of firms as a board member × $I(t=0)$						
Board member in the acquirer before the M&A × $I(t=0)$						
Acquisition × $I(t=0)$						
Target ROA × $I(t=0)$						
Related × $I(t=0)$						
Target firm's relative size in assets × $I(t=0)$	−0.294	0.522		−0.289	0.491	
Log target assets × $I(t=0)$						
Upper × $I(t=0)$						
Log stock share of top 10 × $I(t=0)$						
Acquirer or opponents ROA × $I(t=0)$						
Log top 10 share of acquirer or opponents × $I(t=0)$						
$\tilde{\chi}_{fj}$ × $I(t=0)$	−0.027	0.034				
Wald Test	Coef.	Chi2(1)		Coef.	Chi2(1)	
1+3	−0.006	1.27		−0.009	4.18	**
2+4	−0.003	3.16	*	−0.003	3.32	*
Number of observations	2108			2108		
Stratified by year, acquirer industry code, and target industry code.						

Table 7: Stratified Cox proportional hazards model estimation

hampers the ability to learn new skills in the new firm.

On the other hand, the relevant coefficients for tenure as a board member are not statistically significant, and we are not able to reject the hypothesis that experience as a board member does not increase firm-specific skills and does not influence learning capability. Moreover, applying Proposition 1, we cannot reject the hypothesis that experience as a board member improves general skills for managerial operations. Without rejections of any null hypotheses, we cannot make a strong argument in relation to the effects of experience as a board member. However, our result seems to be consistent with the argument that managerial experience involves general skills that are applicable across firms.

In summary, our results show that experience as a board member and experience as an employee may result in the accumulation of different types of skills. In particular, long experience as an employee enables the accumulation of valuable firm-specific skills but at the expense of learning capability, which is lost, whereas experience as a board member seems to involve general skills and does not result in acquisition or loss of this learning capability.

6 Conclusion

This paper examines how the tenure of managers influences the retention rate of management groups after M&As in Japanese companies. It explicitly models the following insights: (1) as managers gradually accumulate knowledge about a new firm after M&As, the benefits from a manager being a quick learner in the new firm accumulate over time; and (2) acquiring firms mostly need managers' firm-specific knowledge from the target firm immediately after a new firm is established through an M&A. We derive sets of empirically testable conditions that can be utilized to interpret our empirical results. In addition, we propose a novel empirical method to correct selection biases, utilizing information on the timing of appointments to manager positions. Our empirical results show that acquiring firms obtain benefits from skills that are specific to the target firm and from skills that are specific to the new firm, which must be acquired by managers from the target firm following an M&A. We also show that experience as an employee in a target firm, prior to becoming a manager, increases valuable specific skills but at the expense of managers' learning capability.

At this point we discuss whether the issue of entrenchment could influence our interpretation. In line with the discussion in the previous literature, including Wulf (2004), Hartzell et al. (2004), and Bargeron, Schlingemann, Stulz, and Zutter (2009), we note that managers with long tenures in target firms may become very powerful and attempt to protect "their position" as much as possible in the event of M&As. If so, these managers would remain very powerful in the initial period and the acquiring firms would find it difficult to fire them. However, eventually, the power of these managers is eroded and, finally, it becomes possible for the new firms to fire them. Thus, entrenchment may be able to explain the coefficient of tenure in our survival analysis.

Unfortunately, our model is not designed to explicitly distinguish this possibility. However, there are at least four reasons to believe that entrenchment is not important for our results. First, the entrenchment argument is consistent with our coefficient of tenure as an employee, but it is inconsistent with the coefficient of tenure as a board member. These results are peculiar because we would expect that tenure as a board member would be a more appropriate measure of entrenchment than tenure as an employee. Second, because it is impossible, in reality, to write an explicit contract for all possible contingencies, powerful managers should anticipate that they will eventually lose their power. If so, it is not theoretically clear why they would initially agree to a merger. Most mergers in Japan are friendly and, therefore, if the managers are very powerful, they could potentially oppose them. This suggests that managers do not possess the degree of power required to support the entrenchment argument. Third, one might argue that, even if a manager does not have enough power to oppose an M&A, he/she may have the ability to 'buy' a short-term position. Our results may reflect this possibility. In order to take into account this possibility, following Bargeron, Schlingemann, Stulz, and Zutter (2009), we control for a measure of the ownership structure using the logarithm of the stock share of the top 10 stock holders. However, it does not have any significant impacts on the separation rate. This suggests that entrenchment may not be a serious problem for M&As in the Japanese context. Finally, senior people typically have more authority in Japanese society than do younger people. Hence, it is possible that entrenchment could be captured by age rather than tenure. In fact, the coefficient of age is significantly positive, the interaction with the initial dummy is significantly negative, and the sum of the two coefficients is significantly positive, which indicates that senior managers are less likely to be appointed as board members of a new firm. This is to be expected when these managers are less productive, despite their power in the initial years. Therefore, we expect that, after controlling for age, entrenchment will have little effect on the coefficient of tenure.

There may be other factors that potentially influence the retention of managers after M&As that we have not considered in this paper. However, the importance of human capital in understanding the retention of managers cannot be ignored. Therefore, our model is a significant benchmark in understanding the the retention rate of managers after M&As in Japan. Moreover, because our model is flexible and amendable, it provides a sound basis for the development of a more complete theory of retention in the future.

7 Appendix

The Proof of Lemma 5: Define $\bar{F}(\tilde{\chi}) = 1 - F(\tilde{\chi})$ and $Q(x) = \bar{F}^{-1}(x)$. Then:

$$\begin{aligned} E[\chi|\varepsilon \leq x \leq x_p] &= \frac{\int_{\chi(x_p)}^{\chi(\varepsilon)} \chi F'(\chi)\, d\chi}{F(Q(\varepsilon)) - F(Q(x_p))} \\ &= \frac{-\int_{\chi(x_p)}^{\chi(\varepsilon)} \chi \bar{F}'(\chi)\, d\chi}{\bar{F}(Q(x_p)) - \bar{F}(Q(\varepsilon))} \\ &= \frac{\int_{\varepsilon}^{x_p} Q(x)\, \bar{F}'(Q(x))\, Q'(x)\, dx}{x_p - \varepsilon} \\ &= \frac{\int_{\varepsilon}^{x_p} Q(x)\, dx}{x_p - \varepsilon}. \end{aligned}$$

Because the top x percent position of $Q(x)$ can be approximated by $Q'(x) = -\chi_0 x^{\zeta-1}$, we can show that $Q(x) = Q(\varepsilon) - \frac{\chi_0[x^\zeta - \varepsilon^\zeta]}{\zeta}$. Hence:

$$\begin{aligned} E[\chi|\varepsilon \leq x \leq x_p] &= \frac{\int_{\varepsilon}^{x_p} \left\{ Q(\varepsilon) - \frac{\chi_0[x^\zeta - \varepsilon^\zeta]}{\zeta} \right\} dx}{x_p - \varepsilon} \\ &= Q(\varepsilon) + \frac{\chi_0 \varepsilon^\zeta}{\zeta} - \frac{\chi_0 \left[x_p^{\zeta+1} - \varepsilon^{\zeta+1} \right]}{(x_p - \varepsilon)\, \zeta\, (\zeta+1)}. \end{aligned}$$

Because $\zeta \geq 0$ and there exists a finite $Q(0) = \lim_{\varepsilon \to 0} \left[Q(\varepsilon) + \frac{\chi_0 \varepsilon^\zeta}{\zeta} \right]$, we can derive:

$$\begin{aligned} E[\chi|x \leq x_p] &= Q(0) - \frac{\chi_0 x_p^\zeta}{\zeta(\zeta+1)} \; and \\ Q(x_p) &= \lim_{\varepsilon \to 0} \left\{ Q(\varepsilon) - \frac{\chi_0 \left[x_p^\zeta - \varepsilon^\zeta \right]}{\zeta} \right\} = Q(0) - \frac{\chi_0 x_p^\zeta}{\zeta}. \end{aligned}$$

Combining the two equations, we can derive:

$$E[\chi|x \leq x_p] = \frac{\zeta Q(0) + Q(x_p)}{\zeta + 1}.$$

References

[1] Bargeron, Leonce L., Frederik P. Schlingemann, René M. Stulz and Chad J. Zutter (2009), "Do Target CEOs Sell Out their Shareholders to Keep Their Job in a Merger?" NBER Working Paper Series.

[2] Bertrand, Marianne (2009), "CEOs," Annual Review of Economics, 1:1.1-1.29.

[3] Bertrand, Marianne and Antoinette Schoar (2003), "Managing with Style: The Effect of Managers on Firm Policies," Quarterly Journal of Economics, 1169–1208.

[4] Buchholtz, A.K., B.A. Ribbens and I.T. Houle (2003), "The Role of Human Capital in Postacquisition CEO Departure," Academy of Management Journal, Vol. 46, 506–514.

[5] Caballero, Ricardo J. and Mohamad L. Hammour (1998), "The Macroeconomics of Specificity," Journal of Political Economy, Vol. 106, No. 4, 724–767.

[6] Cannella Jr., Albert A. and Donald C. Hambrick (1993), "Effects of Executive Departures on the Performance of Acquired Firms," Strategic Management Journal, Vol. 14, 137–152.

[7] Finkelstein, Sydney and Donald C. Hambrick (1989), "Chief Executive Compensation: A Study of the Intersection of Markets and Political Processes," Strategic Management Journal, Vol. 10, 121–134.

[8] Fiet, James O. (1996), "The Informational Basis of Entrepreneurial Discovery," Small Business Economics, Vol. 8, 419–430.

[9] Frydman, Carola (2005), "Rising Through the Ranks: The Evolution of the Market for Corporate Executives, 1936–2003," Mimeo.

[10] Frydman, Carola and Dirk Jenter (2010), "CEO Compensation," Annual Review of Financial Economics, Vol. 2, 75–102.

[11] Fukao Kyoji, Keiko Ito and Hyeog Ug Kwon (2005), "Do Out–In M&As Bring Higher TFP to Japan? An Empirical Analysis Based on Micro-Data on Japanese Manufacturing Firms," Journal of the Japanese International Economies, Vol. 19, 272–301.

[12] Gabaix, Xavier and Augustin Landier (2008), "Why has CEO Pay Increased So much?" Quarterly Journal of Economics, Vol. 123, 49–100.

[13] Hartzell, Jay C., Eli Ofek and David Yermack (2004), "What's In It For Me? CEOs Whose Firms Are Acquired," Review of Financial Studies, Vol. 17, No. 1, 37–61.

[14] Heckman, James (1979), "Sample Selection Bias as a Specification Error," Econometrica, Vol. 47, 153–161.

[15] Jacoby, Sanford M. (2005), *The Embedded Corporation: Corporate Governance and Employment Relations in Japan and the United States*, Princeton University Press: Princeton and Oxford.

[16] Jovanovic, Boyan and Yaw Nyarko (1995), "A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves," Brookings Papers on Economic Activity. Microeconomics, Vol. 1, 247–305.

[17] Jovanovic, Boyan and Peter L. Rousseau (2002), "The Q-Theory of Mergers," American Economic Review, Vol. 92, No. 2, 198–204.

[18] Kaplan, S. N. (1994), "Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States," Journal of Political Economy, Vol. 102, No. 3, 510–546.

[19] Kaplan, Steven N., Mark M. Klebanov and Morten Sorensen (2012) "Which CEO Characteristics and Abilities Matter?" Journal of Finance, Vol. LXVII, No. 3, 971–1005.

[20] Klepper, Steven (2001) "Employee Startups in High-Tech Industries," Industrial and Corporate Change, Vol. 10, No. 3, 639–674.

[21] Kubo, Katsuyuki (2004), "Gappei ni tomonau Jinji-Seido no Togo to Koyo · Shogu no Henka – Kojin data ni yoru Bunseki (Unification of Human Resource Management after Merger)," Japanese Journal of Labor Studies, No. 529, August, 24–36 (in Japanese).

[22] Kubo, Katsuyuki and Takuji Saito (2012), "The Effect of Mergers on Employment and Wages," Journal of the Japanese and International Economies, forthcoming.

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

[24] Martin, Kenneth J. and John J. McConnell (1991), "Corporate Performance, Corporate Takeovers and Management Turnover," Journal of Finance, Vol. 46, No. 2, 671–687.

[25] Mateos de Cabo, Ruth, Jens Hagendorff and Ricardo Gimeno (2014), "Jobs for the Boys? The Glass Ceiling and the Market for Corporate Control," Mimeo.

[26] Matsusaka, John G. (1993), "Takeover Motives during the Conglomerate Merger Wave " The RAND Journal of Economics, Vol. 24, No. 3 (Autumn), pp. 357–379.

[27] McGuckin, Robert H. and Sang V. Nguyen (1995), "On Productivity and Plant Ownership Change: New Evidence from the Longitudinal Research Database," Rand Journal of Economics, Vol. 26, 257–276.

[28] Mincer, Jacob and Yoshio Higuchi (1988), "Wage Structures and Labor Turnover in the United States and Japan," Journal of the Japanese and International Economies, Vol. 2, 97–133.

[29] Moriguchi, Chiaki and Emmanuel Saez (2008), "The Evolution of Income Concentration in Japan, 1986-2005: Evidence from Income Tax Statistics," Review of Economics and Statistics, Vol. 90, No. 4, 713–734.

[30] Murphy, Kevin J. and Ján Zábojník (2004), "CEO Pay and Appointments: A Market-Based Explanation for Recent Trends," American Economic Review, Vol. 94, No. 2, 192–196.

[31] Nelson, Richard R. and Edmund S. Phelps (1966), "Investment in Humans, Technological Diffusion and Economic Growth," American Economic Review, Vol. 56, No. 1/2, 69–75.

[32] Pissarides, Christopher A. (2000), *Equilibrium Unemployment Theory*, MIT Press: Cambridge.

[33] Rosen Sherwin (1982), "Authority, Control and the Distribution of Earnings," Bell Journal of Economics, 311–323.

[34] Saito, Takuji and Hiroyuki Odagiri (2008), "Intraboard Heterogeneity and the Role of Bank-dispatched Managers in Japanese Firms: An Empirical Study," Pacific-Basin Finance Journal, Vol. 16, 572–590.

[35] Shane, Scott (2000), "Prior Knowledge and the Discovery of Entrepreneurial Opportunities," Organization Science, Vol. 11, 448–469.

[36] Shleifer, Andrei and Lawrence H. Summers (1988), "Breach of Trust in Hostile Takeovers," in *Corporate Takeovers: Causes and Consequences*, edited by Alan J. Auerbach, University of Chicago Press: Chicago, 33–68.

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

[38] Takii, Katsuya (2003), "Prediction Ability," Review of Economic Dynamics, Vol. 6, 80–98.

[39] Terviö, Marko (2008), "The Difference That CEOs Make: An Assignment Model Approach," American Economic Review, Vol. 98, No. 3, 642–668.

[40] Veldkamp, Laura L. (2011), *Information Choice in Macroeconomics and Finance*, Princeton University Press: Princeton.

[41] Walsh, James P. (1988), "Top Management Following Mergers and Acquisitions," Strategic Management Journal, Vol. 9, 173–183.

[42] Walsh, James P. (1989), "Doing a Deal: Merger and Acquisition Negotiations and their Impact upon Target Company Top Management Turnover," Strategic Management Journal, Vol. 10, 307–322.

[43] Walsh, James P. and John W. Ellwood (1991), "Mergers, Acquisitions and the Pruning of Managerial Deadwood," Strategic Management Journal, Vol. 12, 201–217.

[44] Wulf, Julie (2004), "Do CEOs in Mergers Trade Power for Premium? Evidence from "Mergers of Equals"," Journal of Law, Economics, & Organization, Vol. 20, No. 1, 60–101.

[45] Wulf, Julie and Harbir Singh (2011), "How Do Acquirers Retain Successful Target CEOs? The Role of Governance," Management Science, Vol. 57, No. 12, 2101–2114.

[46] Zollo, Maurizio and Harbir Singh (2004), "Deliberate Learning in Corporate Acquisitions: Post-Acquisition Strategies and Integration Capability in U.S. Bank Mergers," Strategic Management Journal, Vol. 25, 1233–1256.