Kenjiro Hirata[1], *Ayako Suzuki*[2] *and Katsuya Takii*[3,*]

How general is managerial human capital?: Evidence from the Retention of Managers after M&As*

Abstract

This paper investigates the transferability of managerial human capital by examining how managers' tenures in target firms influence their probability of retention as board members after mergers or acquisitions in Japanese firms. It develops a general equilibrium model that distinguishes several hypotheses on managerial human capital based on the coefficients of tenure on separation, given several data limitations. In particular, the paper provides a novel method to correct for selection biases by utilizing the timing of selection in a selected sample, which does not require a random sample from the population. Our results suggest that Japanese firms value both target firm-specific and general human capital after M&As and that experience as an employee increases firm-specific skills, but at the expense of the accumulation of general skills. However, managerial experience does not have this effect.

Current version: April 11, 2023
Keywords: Tenure, Managerial Skill, Managerial Turnover after M&As, Selection Bias, Cox Proportional Hazards Model
JEL-codes: G34, J41, J63
Corresponding author: Katsuya Takii
takii@osipp.osaka-u.ac.jp

* We would like to thank Boyan Jovanovic, Kei Kawai, Ryosuke Okazawa, David Yermack, and numerous seminar participants for their helpful comments. In addition, we thank Yoko Kitagawa 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 18330051, 21243024, 25245040, 18H03636) is gratefully acknowledged. We are not able to share data in this research because the data are proprietary.

1 Faculty of Economics, Kobe International University, 9-1-6 Koyocho-naka, Higashinada-ku, Kobe, Hyogo 658-0032, Japan.

2 School of International Liberal Studies, Waseda University, 1-6-1, Nishi-Waseda, Shinjuku, Tokyo, 169-8050, Japan.

3 Osaka School of International Public Policy, Osaka University, 1-31, Machikameyama, Toyonaka, Osaka, 560-0043, Japan.

© The Author(s). 2023 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
Cite as: Hirata et al. *IZA Journal of Labor Economics* (2023) 12:08
https://doi.org/10.2478/izajole-2023-0008

1 Introduction

How general is managerial human capital? Market-based views on managerial compensations explain high CEO compensation in the U.S. as resulting from the competition to employ scarce managerial talent (e.g., Gabaix and Landier, 2008; Terviö, 2008). However, the validity of these theories critically depends on the answer to this question. This question is also related to the sensitivity of managerial turnover to poor performance by a firm (e.g., Eisfeldt and Kuhnen, 2013). Thus, it is one of the most important empirical questions in the literature on corporate governance.

The empirical studies on the retention of managers after M&As may provide some evidence to assist in answering this question. 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 group is important to the new firm[2]. 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; 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 adapting to the new environment - on the basis that the smaller number of firms in which their experience is gained reduces their opportunity to improve their general human capital - 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.

This paper examines how the tenure of managers influences the retention rate of a management group after M&As in Japanese companies during the period 1990-2006. It attempts to separate the two different hypotheses regarding the effect of tenure on separation probability. We show that negative and positive effects of tenure on the managers' separation rate coexist in the case of Japanese M&As.

The main challenge to achieving our goal is developing a way to cope with the limitations of the data. Ideally, we need a fairly rich dataset, which includes the random assignment of target managers and their history afterward, including not only their career paths, but also wage payments and obtained skills. Unfortunately, it is not possible to obtain such data for Japanese firms. These data restrictions force us to develop a new model and a new method that can extract meaningful information from the restricted dataset. We develop a general equilibrium model that is designed to analyze the separation of target-firm managers after M&As. It is shown that our structural model can be approximately estimated by a stratified

1 Some 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.

2 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.

Cox proportional hazards model. In addition, we propose a novel method to correct for selection biases arising from unobserved heterogeneity of managers.

Equipped with the constructed theoretical model and the new estimation method, we conduct survival analyses of Japanese titled directors, who are considered to be the top executives in Japanese companies[3]. Because the Japanese promotion 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.

We consider two sources of managerial human capital in this paper: experience as employees and managerial experience, which is measured by experience as a board member. 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 either appointment probability or separation probability. Through the lens of our theory, we can interpret these findings as follows: 1) Japanese firms after M&As value both the target firm-specific human capital and the general human capital of managers; 2) experience as an employee increases firm-specific skills, but at the expense of the accumulation of general human capital; and 3) Managerial experiences are likely to be more general than those as an employee.

There have been many attempts to understand the retention of managers after M&As[4]. This paper makes two main contributions to the literature. First, to the best of our knowledge, ours is the first paper that constructs a structural model of managers' retention after M&As. Second, we propose a novel method to correct for selection biases arising from the unobserved heterogeneity of managers, which does not require a random sample from the population. These two contributions are discussed in detail below.

We develop a structural model to disentangle the several mechanisms underlying the causal effects of tenure on the separation of target managers after M&As given several data limitations. However, in contrast to the standard structural estimation approach of, for instance, Rust (1987) and Hotz and Miller (1993), we do not attempt to estimate all structural parameters. Instead, we focus on particular parameters of interest, namely the coefficients of the effects of tenure on the separation rates. Our model is designed to clarify the conditions under which we can obtain economically meaningful interpretations from these parameters. Given the data limitations that we face, this strategy allows us to utilize weaker parametric assumptions for our identification, resulting in greater transparency of the source of variations for the identification[5].

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.

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 et al. (2004), and Bargeron et al. (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 et al. (2014) investigate how individual characteristics, including gender or membership of a minority group, influence appointments to directorships after M&As.

5 Similar to our empirical strategy, Chetty (2009) and Heckman(2010) discuss the benefits of estimating wellfocused parameters rather than full structural parameters.

Let us describe the intuition on how to identify some mechanisms from the coefficients of the estimated effect of tenure on turnover in conditions where the data are constrained. In particular, we are concerned about two data problems: the lack of data about skills and the lack of data about compensation. First, because we cannot observe skills, we must differentiate two different hypotheses regarding the length of tenure and the separation probability of managers. Our identification strategy is built on the recognition that, although target-specific human capital is important as an input for current production, employment of managers with the ability to learn in a new environment is an investment in future production. More concretely, on the one hand, because managers in an acquiring firm do not have much knowledge of the target firm's organization, the acquirer initially regards the specific skills held by the target firm managers as highly valuable 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[6]. It is shown that the timing of separation reflects the relative importance of each type of human capital.

This seemingly intuitive argument requires nontrivial theoretical consideration of learning capability. Although the argument in the previous paragraph implicitly assumes that a newly merged firm is willing to employ managers with high learning capability, this may not be true. Because the ability to learn in a new environment is a general ability, the market competition can raise the compensation of able managers, which may make the firm reluctant to employ them. Therefore, the validity of the above argument depends on the structure of the external labor market for managers. Because we developed a general equilibrium model, we can discuss how the frictions in the managerial external labor market influence the coefficients of the effects of tenure on the separation probability.

In addition, the coefficients of the effects of tenure might be influenced by the second data problem: the lack of compensation data for Japanese managers. It is not obligatory for publicly traded companies to provide information on the compensation of managers in Japan. As suggested in the literature (e.g., Wulf, 2004, Hartzell et al., 2004, and Bargeron et al., 2009), this might cause several biases if managers in target firms can negotiate the compensation and severance pay in the newly merged firm during the process of an M&A. Our theoretical model shows that if it is impossible to make a complete contract on treatment of the new managers in the new firm during the M&A and there are some explicit or implicit transfers after the new firm starts as a result of ex post bargaining, then the separation probability does not depend either on the compensation or on the severance pay of managers. That is, the model provides

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

a theoretical condition under which we can extract meaningful information from the coefficients of the impact of tenure on the separation rate even when compensation data are not available.

Next, we explain the second main contribution to the literature. In fact, this second contribution can be considered the more important contribution because it has a broader relevance than the post-M&A managerial retention literature. Note that because managers in a target company are not randomly assigned, 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.

Compared with the standard sample selection models that have appeared in econometrics textbooks (e.g., Wooldridge, 2002; Cameron and Trivedi, 2005), our approach has three advantages. First, because we utilize the timing of selection in a selected sample to correct for selection bias, we do not need to access a random sample from the population. We are not aware of any other papers that provide a tractable method to correct a selection bias inherent to personnel data using only the selected sample. Second, because the timing of observations is generally different from the timing of promotions, we can find several exclusive variables that are required to identify parameters and obtain reliable estimates. Finally, although a standard two-step estimator makes distributional assumptions to correct selection biases, we rely on an extreme value theory to obtain robust results against misspecifications of distributional assumptions. 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. We apply their argument and derive an explicit solution to the conditional expectation of unobserved ability under a selected sample, which is used to determine unbiased estimates of parameters under a selected sample.

As discussed before, this methodology is applicable to a broader literature than that of managerial retention after M&As. Although it is recognized that the characteristics of CEOs 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. Moreover, our methodology can be applied to data on internal promotions, when researchers are not able to access data on employees' whole careers. We hope that this novel approach can help researchers struggling with data limitations to evaluate the role of leaders.

In addition to these two main contributions to the literature, our results contribute to several streams of literature, including those on managerial compensation and Japanese M&As. We clarify the contribution of our results to these fields in Section 8.

The paper is organized as follows. Section 2 presents a general equilibrium model of the separation of target managers after M&As. It shows how target firm-specific human capital

and general human capital can influence the separation decision after M&As. Section 3 links the length of tenure and each type of human capital and provides a number of testable conditions that distinguish several hypotheses on managerial human capital based on the evidence of the effect of tenure on separation probability. Section 4 introduces our empirical model and discusses how to correct a selection bias using our selected dataset. Section 5 describes our Japanese dataset and Section 6 presents our estimation model and our results, interpreted through the lens of our theory. Section 7 discusses whether the moral hazard of managers could influence the interpretation of our results. Section 8 relates our results to several streams of literature that cover managerial compensation and Japanese M&As. Section 9 concludes the paper.

2 A General Equilibrium Model of Managerial Separation after M&A

In this section, we construct a general equilibrium model that is designed to examine the separation of target managers after M&As using the following procedure. First, we discuss the assumptions on the production functions, which describe how target firm-specific human capital and general human capital influence the production process. Second, using the production functions, we model the separation of a manager from a target firm after an M&A by taking the outside values of managers as given. Third, we model an external managerial labor market, which allows us to determine the outside values of the managers. Finally, we show how our structural model can be approximately represented by the proportional hazards model and discuss how target firm-specific human capital and general human capital influence the parameters of the proportional hazards model.

Target Firm-Specific Human Capital vs. General Human Capital

Suppose that when a firm is merged with another firm, managers from the target firm must learn new skills and/or new routines to manage the merged firm in its new environment. We assume that all learning takes place at time $t = 0$ and that normal firm operations start 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. Assume that a manager in a target firm has general human capital h^G and human capital specific to the target firm, h^T. Suppose that the productivity of the manager with human capital $\mathbf{h} = \{h^G, h^T\}$ is described by $p_t(\mathbf{h})$:

$$\begin{aligned} p_t(\mathbf{h}) &= f_{Gs}\left(h^G\right) + f_{Ts}\left(h^T\right) + \zeta\left(t : \mathbf{X}\right), \\ s &= l \text{ if } t = 0 \\ &= n \text{ if } t \geq 1, \end{aligned}$$

where the vector $\mathbf{X}$ includes any variables that can be controlled for in our empirical study and $\zeta(t: \mathbf{X})$ captures the transitional dynamics of productivity after the M&A that is unrelated to human capital accumulation[7].

7 Although we construct a model under a stationary environment in this paper, macroeconomic shocks can be easily incorporated by multiplying all variables by $A(x)$, which denotes macroeconomic shocks in year x. Given this specification, we can show that the introduction of $A(x)$ does not change the results at all.

The production technologies f_{Gs} and f_{Ts} depend on the state s. We use $s = l$ to denote a learning state, whereas $s = n$ denotes the normal operation state. We assume that $f_{Gs}\left(h^G\right) \geq 0$ and $f_{Ts}\left(h^T\right) \geq 0$ for any $s \in \{n, l\}$ and h^G and h^T. We also assume that $f'_{Tl}\left(h^T\right) \geq f'_{Tn}\left(h^T\right) \geq 0$ and $f'_{Gn}\left(h^G\right) \geq f'_{Gl}\left(h^G\right) \geq 0$.

The assumption of $f'_{Tn}\left(h^T\right) \leq f'_{Tl}\left(h^T\right)$ means that an increase in the target firm-specific human capital can initially increase benefits. When a merger occurs, managers in the acquiring firm do not have enough knowledge about the target firm. Hence, initially, the firm-specific knowledge relating to the target firm must be more valuable. The assumption of $f'_{Tn}\left(h^T\right) \leq f'_{Tl}\left(h^T\right)$ captures this effect.

The assumption of $f'_{Gn}\left(h^G\right) \geq f'_{Gl}\left(h^G\right) \geq 0$ may require more explanation. There are several plausible micro foundations for this assumption. Our preferred explanation is that the operation of a new firm requires new firm specific knowledge, which a manager with higher general human capital can more effectively learn. To clarify this point, we can construct an illustrative model, such as $f_{Gn}\left(h^G\right) = f_{Gl}\left(h^G\right) + f_N\left(h^N\right)$ and $h^N = g\left(h^G\right)$, where h^N is new firm-specific human capital and $g\left(h^G\right)$ is a learning function. We also assume that $f_N\left(h^N\right) \geq 0$, $g\left(h^G\right) \geq 0$ and $g'\left(h^G\right) \geq 0$. Under these assumptions, it can be shown that $f'_{Gn}\left(h^G\right) = f'_{Gl}\left(h^G\right) + f'_N\left(g\left(h^G\right)\right)g'\left(h^G\right) \geq f'_{Gl}\left(h^G\right)$.

The assumption of $g'\left(h^G\right) \geq 0$ attempts to capture the idea that general human capital can improve the speed of learning. Nelson and Phelps (1966) consider that education increases the speed of technology adoption because educated people are better 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. Gibbons and Waldman (2006) construct a model of promotion in which education can increase learning capability, arguing that this concept has significant empirical support. We attempt to incorporate this idea in our model.

The Separation of a Manager from a Target Firm after an M&A

Now, we construct a model of managerial separation. The present value of the expected profit sequences from employing a particular manager, $J_t(\mathbf{h})$ at $t \geq 1$, can be described as follows:

$$J_t\left(\mathbf{h}\right) = p_t\left(\mathbf{h}\right) - w_t + \beta E_\varepsilon \left[\begin{array}{c} I\left(W_{t+1}\left(\mathbf{h}\right) \geq \hat{W}^l\left(h^G\right) + P_{t+1} + \varepsilon_{t+1}\right)\max\left\{J_{t+1}\left(\mathbf{h}\right), \hat{J}^0 - P_{t+1}\right\} \\ + \left[1 - I\left(W_{t+1}\left(\mathbf{h}\right) \geq \hat{W}^l\left(h^G\right) + P_{t+1} + \varepsilon_{t+1}\right)\right]\left(\hat{J}^0 - P_{t+1}\right) \end{array} \right],$$

where w_t is the manager's compensation at time t, P_{t+1} is the severance pay at time $t + 1$, $\beta \in (0,1)$ 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, $E_\varepsilon[\cdot]$ is the expectation operator with respect to ε_{t+1}, $\hat{J}^0$ is the present value of the expected net profit sequences from employing a new manager, and $W_t(\mathbf{h})$ and $\hat{W}^l\left(h^G\right)$ 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. Note that when managers stay in the same firm, their value functions depend on both target firm-specific human capital and general human capital, $\mathbf{h}$, whereas when they leave, their value functions depend only on their general human capital, h^G.

We assume that firms review their managers every period. Assuming that the manager prefers to stay, which means that $I\left(W_{t+1}(\mathbf{h}) \geq \hat{W}^{l}\left(h^{G}\right)+P_{t+1}(\mathbf{h})+\varepsilon_{t+1}\right)=1$, the firm has to decide whether to retain the manager as a board member. If it retains the manager, it expects to obtain $J_{t+1}(\mathbf{h})$ in future from the manager. However, if it fires the manager, the firm must find a new manager and receives $\hat{J}^{0}$. Because the firm must pay severance pay P_{t+1} in the case of separation, it must incur P_{t+1}. Because a firm is assumed to maximize its profit, the expected profit would be $\max\left\{J_{t+1}(\mathbf{h}),\ \hat{J}^{0}-P_{t+1}\right\}$. If $W_{t+1}(\mathbf{h})<\hat{W}^{l}\left(h^{G}\right)+P_{t+1}(\mathbf{h})+\varepsilon_{t+1}$, the manager leaves, irrespective of the firm's decision and, therefore, the firm receives $\hat{J}^{0}-P_{t+1}$.

Similarly, the present values of the sum of the expected income flows of managers in the firm at the analysis time t, $W_{t}(\mathbf{h})$, can be described as follows:

$$W_{t}(\mathbf{h})=w_{t}+\beta E_{\varepsilon}\left[\begin{array}{c} I\left(J_{t+1}(\mathbf{h}) \geq \hat{J}^{0}-P_{t+1}\right)\max\left\{W_{t+1}(\mathbf{h}),\ \hat{W}^{l}\left(h^{G}\right)+P_{t+1}+\varepsilon_{t+1}\right\} \\ +\left(1-I\left(J_{t+1}(\mathbf{h}) \geq \hat{J}^{0}-P_{t+1}\right)\right)\left(\hat{W}^{l}\left(h^{G}\right)+P_{t+1}+\varepsilon_{t+1}\right) \end{array}\right],$$

when a manager stays in a firm, in every period that follows, he/she reassesses and decides whether to stay with the firm or leave. If the manager stays, he/she expects to obtain $W_{t+1}(\mathbf{h})$. If the manager leaves, he/she expects to obtain $\hat{W}^{l}\left(h^{G}\right)+P_{t+1}+\varepsilon_{t+1}$. Hence, if a firm decides to continue employing the manager, $I\left(J_{t+1}(\mathbf{h}) \geq \hat{J}^{0}-P_{t+1}\right)=1$, the expected value must be $\max\left\{W_{t+1}(\mathbf{h}), \hat{W}^{l}\left(h^{G}\right)+P_{t+1}+\varepsilon_{t+1}\right\}$. If the firm decides to fire the manager, $I\left(J_{t+1}(\mathbf{h}) \geq \hat{J}^{0}-P_{t+1}\right)=0$, the manager must leave and receive the outside benefits of $\hat{W}^{l}\left(h^{G}\right)+P_{t+1}+\varepsilon_{t+1}$, irrespective of his/her own decision.

We assume that:

$$w_{t}=\hat{w}_{t}+\tilde{w}_{t},$$

where $\hat{w}_{t}$ is the contracted compensation, whereas $\tilde{w}_{t} \in \mathbf{R}$ is the negotiable compensation. As suggested in the literature (e.g., Wulf, 2004, Hartzell et al., 2004, and Bargeron et al., 2009), a manager may control the sale of the firm and refuse to take a deal in the M&A process unless he/she is personally promised something in return, such as a position on the board of the new firm. As a result of this negotiation, a target-firm manager and the acquiring firm may write an explicit contract on $\left(\hat{w}_{t}, P_{t}\right)$ before the new firm starts its operation. However, we consider that the contract in the real world is far from complete and that it leaves room for negotiable transfers $\tilde{w}_{t}$ after the new firm starts. We consider that $\tilde{w}_{t}$ can be a negative implicit transfer because it may involve harassment and mental pressure. We show that the existence of $\tilde{w}_{t}$ makes the separation decision independent of $\left(\hat{w}_{t}, \tilde{w}_{t}, P_{t}\right)$.

We define the joint surplus of the manager and the new firm by $S_{t}(\mathbf{h}) \equiv J_{t}(\mathbf{h})+W_{t}(\mathbf{h})-\hat{J}^{0}-\hat{W}^{l}\left(h^{G}\right)$. The proof of the following proposition is provided in the Appendix.

Proposition 1 Suppose that the manager and the firm can negotiate $\tilde{w}_{t} \in \mathbf{R}$ at the beginning of time t. The separation occurs at time t if and only if $S_{t}(\mathbf{h}) < \varepsilon_{t}$ and the dynamics of $S_{t}(\mathbf{h})$ can be described by:

$$S_{t}(\mathbf{h})=p_{t}(\mathbf{h})-b\left(h^{G}\right)+\beta\left[S_{t+1}(\mathbf{h})+E_{\varepsilon}\left[I\left(\varepsilon_{t+1}>S_{t+1}(\mathbf{h})\right)\left(\varepsilon_{t+1}-S_{t+1}(\mathbf{h})\right)\right]\right], \tag{1}$$

where $b\left(h^{G}\right)=(1-\beta)\left[\hat{W}^{l}\left(h^{G}\right)+\hat{J}^{0}\right]$ is the sum of the expected flow benefits received by both a manager and a firm from the outside market.

Proposition 1 shows that the separation decision does not depend on $\left(\hat{w}_t, \tilde{w}_t, P_t\right)$, and that it depends only on the joint surplus of the target manager and the merged firm, $S_t(\mathbf{h})$. This means that, although the negotiation during an M&A may result in a complicated compensation scheme and severance pay, it does not influence the separation decision as long as a firm and a manager can flexibly negotiate the ex post transfer.

In addition, Proposition 1 shows what determines the joint surplus. For any t, maintaining the relationship brings a joint instantaneous surplus of $p_t\left(\mathbf{h}\right)-\left(1-\beta\right)\left(\hat{J}^0+\hat{W}^I\left(h^G\right)\right)$. If they maintain their relationship, the manager and firm can enjoy $S_{t+1}(\mathbf{h})$ in the next period. However, if the instantaneous random benefits from the outside market happen to be greater than the joint surplus, $\varepsilon_{t+1} > S_{t+1}(\mathbf{h})$, then they will decide to separate and obtain the additional value of $\varepsilon_{t+1} - S_{t+1}(\mathbf{h})$. 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 $E_\varepsilon\left[I\left(\varepsilon_{t+1} > S_{t+1}\left(\mathbf{h}\right)\right)\left(\varepsilon_{t+1}-S_{t+1}\left(\mathbf{h}\right)\right)\right]$.

Assume that ε_t is exponentially distributed with a parameter λ. The separation probability, q_t, is shown to be a function of $S_t(\mathbf{h})$, as follows:

$$q_t\left(\mathbf{h}\right)=E_\varepsilon\left[I\left(\varepsilon_t > S_t\left(\mathbf{h}\right)\right)\right]=e^{-\lambda S_t(\mathbf{h})}$$

This suggests that increases in the joint surplus, $S_t(\mathbf{h})$, reduce the probability of separation.

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

$$E_\varepsilon\left[I\left(\varepsilon_{t+1} > S_{t+1}\left(\mathbf{h}\right)\right)\left(\varepsilon_{t+1}-S_{t+1}\left(\mathbf{h}\right)\right)\right]=\frac{q_{t+1}\left(\mathbf{h}\right)}{\lambda}.$$

Inserting the expected gain from separation into the dynamics of the joint surplus, we can rewrite equation (1) and summarize how the two types of human capital influence the separation probability using the following equations:

$$q_t\left(\mathbf{h}\right)=e^{-\lambda S_t(\mathbf{h})}$$
$$S_t\left(\mathbf{h}\right)=p_t\left(\mathbf{h}\right)-b\left(h^G\right)+\beta\left[S_{t+1}\left(\mathbf{h}\right)+\frac{e^{-\lambda S_{t+1}(\mathbf{h})}}{\lambda}\right].$$

Note that both target firm-specific and general human capital influence the separation through an increase in productivity, $p_t(\mathbf{h})$. The general human capital has additional effects on the separation through a change in outside value, $b(h^G)$. Therefore, without knowing the structure of function $b(h^G)$, we cannot make any clear theoretical prediction on the impacts of general human capital on the separation probability.

The Model of the Managerial Labor Market

We model an external labor market and endogenize $b\left(h^G\right)=\left(1-\beta\right)\left(\hat{W}^I\left(h^G\right)+\hat{J}^0\right)$. Suppose that an external managerial labor market is segregated by the type of worker h^G. A firm can find a worker with h^G by paying a hiring cost $F(h^G)$ such that:

$$\hat{J}^0=\max_{h^G}\left\{\hat{J}^I\left(h^G\right)-F\left(h^G\right)\right\},$$

where:

$$\hat{J}^l\left(h^G\right)=f_{Gl}\left(h^G\right)-\hat{w}^l\left(h^G\right)+\beta\left[\hat{q}\hat{J}^l\left(h^G\right)+\left(1-\hat{q}\right)\hat{J}^n\left(h^G\right)\right],$$
$$\hat{J}^n\left(h^G\right)=f_{Gn}\left(h^G\right)-\hat{w}^n\left(h^G\right)+\beta\left[\hat{q}\hat{J}^l\left(h^G\right)+\left(1-\hat{q}\right)\hat{J}^n\left(h^G\right)\right].$$

$\hat{q}$ is the separation probability, and $\hat{w}^s\left(h^G\right)$ and $\hat{J}^s\left(h^G\right)$ are the compensation for a manager and the present value of the expected profit sequences from employing a manager who has a general skill h^G at the state s, respectively. We assume that $F(h^G) \geq 0$ and $F'(h^G) \geq 0$. These assumptions imply that it is difficult to find a manager and even more difficult to find an able manager. Similar to the situation for a new firm after an M&A, there is a learning state and a normal operation state for the production function. In contrast to the situation for the new firm after an M&A, productivity does not depend on $f_{Ts}(h^T)$ and $\zeta(t: \mathbf{X})$. This means that target-firm specific human capital is useless and there is no transition dynamics after an M&A for the firms that do not experience M&As. As a result, the environment is stationary and the separation probability $\hat{q}$ does not depend on t.

Suppose that managerial talent is a scarce resource and all firms competitively post an initial wage to attract a manager as long as profits are positive. Because of the price competition, $\hat{J}^0=0$ at the equilibrium. Hence, for all hirable $h^G, \hat{J}^l\left(h^G\right)=F\left(h^G\right)$ and, therefore:

$$\hat{w}^l\left(h^G\right)=f_{Gl}\left(h^G\right)+\beta\left(1-\hat{q}\right)\hat{J}^n\left(h^G\right)-F\left(h^G\right). \tag{2}$$

Similarly, the present values of the sum of the expected income flows of a manager who has a general skill h^G at the state $s: \hat{W}^s\left(h^G\right)$ are described as follows:

$$\hat{W}^l\left(h^G\right)=\hat{w}^l\left(h^G\right)+\beta\left[\hat{q}\hat{W}^l\left(h^G\right)+\left(1-\hat{q}\right)\hat{W}^n\left(h^G\right)\right] \tag{3}$$

$$\hat{W}^n\left(h^G\right)=\hat{w}^n\left(h^G\right)+\beta\left[\hat{q}\hat{W}^l\left(h^G\right)+\left(1-\hat{q}\right)\hat{W}^n\left(h^G\right)\right]. \tag{4}$$

We define the surplus under a normal operation as $\hat{S}^n\left(h^G\right)=\hat{J}^n\left(h^G\right)+\hat{W}^n\left(h^G\right)-\hat{W}^l\left(h^G\right)$. Then, it is shown that:

$$\hat{S}^n\left(h^G\right)=f_{Gn}\left(h^G\right)-f_{Gl}\left(h^G\right)+F\left(h^G\right)\geq 0.$$

Note that $\hat{S}^n\left(h^G\right)\geq 0$ means that there is a surplus that managers and firms can share. Hence, they can adjust $\hat{w}^n\left(h^G\right)$ to maintain their relationship. Note also that $\hat{S}^n\left(h^G\right)\geq 0$ as long as learning increases productivity $f'_{Gn}\left(h^G\right)=f'_{Gl}\left(h^G\right)+f'_N\left(g\left(h^G\right)\right)g'\left(h^G\right)\geq f'_{Gl}\left(h^G\right)$ and there is a hiring cost $F\left(h^G\right)\geq 0$.

Because $\hat{J}^0=0$ at the equilibrium, $b\left(h^G\right)=\left(1-\beta\right)\hat{W}^l\left(h^G\right)$. Substituting equation (2) into equation (3), we can derive:

$$b\left(h^G\right)=f_{Gl}\left(h^G\right)-F\left(h^G\right)+\beta\left(1-\hat{q}\right)\hat{S}^n\left(h^G\right).$$

Using this result, we can summarize the separation probability derived from our general equilibrium model of the managerial separation after an M&A in the following proposition.

Proposition 2 The separation probability of a target manager after an *M*&*A* is characterized by the following equations:

$$q_t(\mathbf{h}) = e^{-\lambda S_t(\mathbf{h})}$$
$$S_t(\mathbf{h}) = \pi^s(\mathbf{h}) + \zeta(t:\mathbf{X}) + \beta\left[S_{t+1}(\mathbf{h}) + \frac{e^{-\lambda S_{t+1}(\mathbf{h})}}{\lambda}\right]$$

where:

$$\pi^n(\mathbf{h}) \equiv f_{Tn}(h^T) + \left[1-\beta(1-\hat{q})\right]\hat{S}^n(h^G)$$
$$\pi^l(\mathbf{h}) \equiv f_{Tl}(h^T) + F(h^G) - \beta(1-\hat{q})\hat{S}^n(h^G)$$
$$\hat{S}^n(h^G) = f_{Gn}(h^G) - f_{Gl}(h^G) + F(h^G).$$

The proposition shows that not only the production function but also the hiring cost function influences the separation productivity of the target manager after an M&A.

Proportional Hazard Model of Separation

We show how our structural model can be approximately represented by the proportional hazards model. The advantage of the proportional hazards model is that we do not need any parametric assumption on the basic hazard and, therefore, the model does not make any assumptions on the shape of the hazard over time. It is shown that without making any assumption on $\zeta(t: \mathbf{X})$, we can conduct our empirical analysis.

We assume that after the merger, a new firm faces a transition period and eventually converges to an operation state that is similar to those of other firms that did not experience an M&A. We derive an approximate solution of $S_t(\mathbf{h})$. We decompose $\mathbf{X}$ and $\zeta(t: \mathbf{X})$ into two parts: $\mathbf{X} = (\mathbf{X}_Z, \mathbf{X}_V)$, $v^s(\mathbf{X}_V)$ and $z(t: \mathbf{X}_z)$ so that $\zeta(t:\mathbf{X}) = v^s(\mathbf{X}_V) + z(t:\mathbf{X}_Z)$, where $s = \{l, n\}$. Taking a first-order approximation of $\frac{e^{-\lambda S_{t+1}(\mathbf{h})}}{\lambda}$ around $e^{-\lambda S_{t+1}(\mathbf{h})} = \hat{q}$, we can derive an approximate solution of $S_t(\mathbf{h})$ for $t \geq 1$:

$$S_t(\mathbf{h}) \approx \Pi^n(\mathbf{h}) + Z(t:\mathbf{X}_Z)$$
$$\Pi^n(\mathbf{h}) = \frac{\pi^n(\mathbf{h}) + v^n(\mathbf{X}_V)}{1-\beta(1-\hat{q})}$$
$$Z(t:\mathbf{X}_Z) = \sum_{\tau=0}^{\infty}[\beta(1-\hat{q})]^\tau z(t+\tau:\mathbf{X}_Z) + \frac{\hat{\beta}(\hat{q})}{1-\beta(1-\hat{q})},$$

where $\hat{\beta}(\hat{q}) = \beta\left(\frac{1-\ln\hat{q}}{\lambda}\right)\hat{q}$.

Similarly, the present value of the stream of instantaneous joint surpluses at time 0, $S_0(\mathbf{h})$ can be expressed as:

$$S_0(\mathbf{h}) = \pi^l(\mathbf{h}) + v^l(\mathbf{X}_V) + z(0:\mathbf{X}_Z) + \beta\left[S_1(\mathbf{h}) + \frac{e^{-\lambda S_1(\mathbf{h})}}{\lambda}\right].$$

Therefore, using the same approximation, we can show that:

$$S_0(\mathbf{h}) \approx \Delta\pi(\mathbf{h}) + \Pi^n(\mathbf{h}) + Z(0:\mathbf{X}_Z),$$

where $\Delta\pi(\mathbf{h}) = \pi^l(\mathbf{h}) - \pi^n(\mathbf{h}) + v^l(\mathbf{X}_V) - v^n(\mathbf{X}_V)$.

The following proposition summarizes the results.

Proposition 3 Separation probability can be estimated using the following proportional hazards model:

$$q_t = B\left(t:\mathbf{X}_Z\right)e^{-\lambda\left[\Delta\pi\left(\mathbf{h}\right)I\left(t=0\right)+\Pi^n\left(\mathbf{h}\right)\right]}$$
$$\Pi^n\left(\mathbf{h}\right)=\frac{\pi^n\left(\mathbf{h}\right)+v^n\left(\mathbf{X}_V\right)}{1-\beta\left(1-\hat{q}\right)},$$

where:

$$\Delta\pi\left(\mathbf{h}\right)=f_{Gl}\left(h^G\right)-f_{Gn}\left(h^G\right)+f_{Tl}\left(h^T\right)-f_{Tn}\left(h^T\right)+v^l\left(\mathbf{X}_V\right)-v^n\left(\mathbf{X}_V\right),$$
$$\pi^n\left(\mathbf{h}\right)=f_{Tn}\left(h^T\right)+\left[1-\beta\left(1-\hat{q}\right)\right]\left[f_{Gn}\left(h^G\right)-f_{G0}\left(h^G\right)+F\left(h^G\right)\right],$$

$B\left(t:\mathbf{X}_Z\right)=e^{-\lambda Z\left(t:\mathbf{X}_Z\right)}$ 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. As we discussed, we derive the basic hazard $B(t: \mathbf{X}_z)$ as a function of $z(t + \tau: \mathbf{X}_z)$. Therefore, without making any parametric assumption on $z(t + \tau: \mathbf{X}_z)$, we can conduct the estimation of this proportional hazards model.

Let us examine how target firm-specific and general human capital influence the parameters on the proportional hazards model. Because, for $x = G$ or T, $\frac{\partial q_t}{\partial h^x}=q_t\left[-\lambda\frac{\partial\Pi^n\left(\mathbf{h}\right)}{\partial h^x}\right]$ for any $t \geq 1$ and $\frac{\partial q_0}{\partial h^x}=q_0\left[-\lambda\left(\frac{\partial\Delta\pi\left(\mathbf{h}\right)}{\partial h^x}+\frac{\partial\Pi^n\left(\mathbf{h}\right)}{\partial h^x}\right)\right]$, we can analyze the impact of human capital on the separation rate by $\frac{\partial-\lambda\Pi^n\left(\mathbf{h}\right)}{\partial h^x}$, $\frac{\partial-\lambda\Delta\pi\left(\mathbf{h}\right)}{\partial h^x}$ and $\frac{\partial-\lambda\left[\Delta\pi\left(\mathbf{h}\right)+\Pi^n\left(\mathbf{h}\right)\right]}{\partial h^x}$, where $x = G$ or T. We can obtain the following proposition.

Proposition 4 The impacts of target firm-specific human capital and general human capital on the proportional hazards model can be analyzed by the following equations:

1. The impacts of target firm-specific human capital:

$$\frac{\partial-\lambda\Pi^n\left(\mathbf{h}\right)}{\partial h^T}=\frac{-\lambda f'_{Tn}\left(h^T\right)}{1-\beta\left(1-\hat{q}\right)}\leq 0$$
$$\frac{\partial-\lambda\Delta\pi\left(\mathbf{h}\right)}{\partial h^T}=-\lambda\left[f'_{Tl}\left(h^T\right)-f'_{Tn}\left(h^T\right)\right]\leq 0$$
$$\frac{\partial-\lambda\left[\Delta\pi\left(\mathbf{h}\right)+\Pi^n\left(\mathbf{h}\right)\right]}{\partial h^T}=-\lambda\left[f'_{Tl}\left(h^T\right)+\frac{\beta\left(1-\hat{q}\right)f'_{Tn}\left(h^T\right)}{1-\beta\left(1-\hat{q}\right)}\right]\leq 0$$

2. The impacts of general human capital:

$$\frac{\partial-\lambda\Pi^n\left(\mathbf{h}\right)}{\partial h^G}=-\lambda\left[f'_{Gn}\left(h^G\right)-f'_{Gl}\left(h^G\right)+F'\left(h^G\right)\right]\leq 0$$
$$\frac{\partial-\lambda\Delta\pi\left(\mathbf{h}\right)}{\partial h^G}=-\lambda\left[f'_{Gl}\left(h^G\right)-f'_{Gn}\left(h^G\right)\right]\geq 0$$
$$\frac{\partial-\lambda\left[\Delta\pi\left(\mathbf{h}\right)+\Pi^n\left(\mathbf{h}\right)\right]}{\partial h^G}=-\lambda F'\left(h^G\right)\leq 0$$

The first result in Proposition 4 shows that a rise in h^T increases $\Pi^n(\mathbf{h})$ and $\Delta\pi(\mathbf{h}) + \Pi^n(\mathbf{h})$, and, therefore, always decreases the separation probability of managers. That is, because target firm-specific human capital always improves productivity, $f'_{Tl}\left(h^T\right) \geq 0$ and $f'_{Tn}\left(h^T\right) \geq 0$. A rise in h^T also increases $\Delta\pi(\mathbf{h})$ and, therefore, lowers $-\lambda\Delta\pi(\mathbf{h})$. This is because we assume that the target firm-specific human capital is initially more important, $f'_{Tl}\left(h^T\right) \geq f'_{Tn}\left(h^T\right)$.

The impacts of general human capital are more complicated. Because of the competition for general human capital, a rise in this type of human capital increases not only productivity, but also the cost of employing managers. Therefore, how general human capital influences the separation probability appears to be unclear. However, the second results in Proposition 4 provide clear theoretical predictions. That is, a rise in h^G increases $\Pi^n(\mathbf{h})$ and $\Delta\pi(\mathbf{h}) + \Pi^n(\mathbf{h})$, and, therefore, always decreases the separation probability of managers. A rise in h^G increases $\Delta\pi(\mathbf{h}) + \Pi^n(\mathbf{h})$ and, therefore, reduces q_0 if it is difficult to find more able managers $F'\left(h^G\right) \geq 0$. If the costs of finding talented managers are the same as those of finding ordinary managers, there are no benefits from employing managers with high general human capital because the competition for talent forces firms to pay high compensation rates for talented managers. The benefits of having managers with high general human capital increases in the later period if employing someone with high learning capability can be considered an investment $f'_{Gl}\left(h^G\right) \leq f'_{Gn}\left(h^G\right)$. If managers have a greater ability to learn in the new environment, this skill increases the firm's future outputs. However, it does not increase the initial output during a learning period. Because $\Delta\pi(\mathbf{h})$ simply reflects the initial instantaneous output, net of the future instantaneous output yielded by the manager, the impact of learning capability on $\Delta\pi(\mathbf{h})$ is negative. This unique prediction helps us to separate general learning capability from target firm-specific skills.

3 Tenure and Separation Probability

In this section, we explain how we utilize the timing of separation to obtain information to distinguish two hypotheses on managerial human capital in a firm (i.e., an increase in target firm specific human capital and a reduction in the general ability to learn in a new environment) from the coefficients of the effect of tenure on separation probability. For this purpose, we propose testable hypotheses to interpret the relationship between tenure and separation probability.

Taking a first-order approximation, it is shown that:

$$\Pi^n\left(\mathbf{h}\right) = \tilde{\alpha}_T h^T + \tilde{\alpha}_G h^G + \tilde{\alpha}'_X \mathbf{X}_V, \tag{5}$$

$$\Delta\pi\left(\mathbf{h}\right) = \bar{\alpha}_T h^T - \bar{\alpha}_G h^G + \bar{\alpha}'_X \mathbf{X}_V, \tag{6}$$

where $\tilde{\alpha}_T = \frac{f'_{Tn}\left(h^T\right)}{1-\beta\left(1-\hat{q}\right)} \geq 0$, $\tilde{\alpha}_G = f'_{Gn}\left(h^G\right) - f'_{Gl}\left(h^G\right) + F'\left(h^G\right) \geq 0$, $\tilde{\alpha}_X = \frac{\frac{\partial v^n(\mathbf{x}_V)}{\partial \mathbf{X}_V}}{1-\beta\left(1-\hat{q}\right)}$, $\bar{\alpha}_T = f'_{Tl}\left(h^T\right) - f'_{Tn}\left(h^T\right) \geq 0$, $\bar{\alpha}_G = f'_{Gn}\left(h^G\right) - f'_{Gl}\left(h^G\right) \geq 0$, and $\bar{\alpha}_X = \frac{\partial v^l(\mathbf{X}_V)}{\partial \mathbf{X}_V} - \frac{\partial v^n(\mathbf{X}_V)}{\partial \mathbf{X}_V}$.

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.

$$h^T = \eta_0 + \eta_b \tau_b + \eta_e \tau_e$$

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 firm-specific human capital[8].

On the other hand, as h^G is a general skill, labor economists typically assume that h^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_0. Therefore, we have:

$$h^G = \omega_0 + \omega_b \tau_b + \omega_e \tau_e + \mu \tau_o + \chi$$

where $\omega_x \geq 0 (x = b,e)$ are the parameters that capture the importance of tenure for learning and $\mu \geq 0$ is the parameter that captures the importance of other experiences[9]. This model generalizes the standard assumption, in that it allows for different experiences to vary in their effectiveness in accumulating general human capital. 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^G as a function of tenures and age:

$$h^G = \omega_0 + (\omega_b - \mu)\tau_b + (\omega_e - \mu)\tau_e + \mu a + \chi.$$

The parameters $\omega_x - \mu$ $(x = b,e)$ capture the relative productivity of tenure for the accumulation of general human capital compared with other experiences. Buchholtz et al. (2003) argue that longer tenure may hamper the accumulation of general skills that are needed for adaptation to a new environment. This implies that $\omega_x - \mu < 0$. Although this might be a reasonable assumption, we also allow for the possibility that $\omega_x - \mu > 0$: i.e., that experience in a target firm can assist in improving general human capital. 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 $\omega_x - \mu > 0$.

The following proportional hazards model can be derived:

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

8 Our human capital accumulation functions are linear if they do not change positions. Since a more general production function would greatly complicate our empirical predictions, we made the linear assumption in order to derive a meaningful interpretation from the empirical results. This assumption can be thought of as a linear approximation to the general human capital accumulation function.

9 Our human capital accumulation function is consistent with the following human capital accumulation function. For any year t

$h_t^T = h_{t-1}^T + \eta_b$ if a worker is board member,
$= h_{t-1}^T + \eta_e$, if a worker is an employee,

and

$h_t^G = h_{t-1}^G + \omega_b$, if a worker is board member,
$= h_{t-1}^G + \omega_e$, if a worker is an employee,
$= h_{t-1}^G + \omega_o$, if a worker does not belong to the firm.

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

$$\frac{d\left(-\lambda\Pi^{n}\left(\mathbf{h}\right)\right)}{d\tau_{x}}=-\lambda\left[\tilde{\alpha}_{T}\eta_{x}+\tilde{\alpha}_{G}\left(\omega_{x}-\mu\right)\right], \tag{7}$$

$$\frac{d\left(-\lambda\Delta\pi\left(\mathbf{h}\right)\right)}{d\tau_{x}}=-\lambda\left[\bar{\alpha}_{T}\eta_{x}-\bar{\alpha}_{G}\left(\omega_{x}-\mu\right)\right], \tag{8}$$

$$\frac{\partial\left[-\lambda\left(\Delta\pi\left(\mathbf{h}\right)+\Pi^{n}\left(\mathbf{h}\right)\right)\right]}{\partial\tau_{x}}=-\lambda\left[\left(\tilde{\alpha}_{T}+\bar{\alpha}_{T}\right)\eta_{x}+\left(\tilde{\alpha}_{G}-\bar{\alpha}_{G}\right)\left(\omega_{x}-\mu\right)\right], \tag{9}$$

where $x = b$ or e.

Investigating equations (7), (8), and (9) 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 value any specific skills of employees in the target firm and/or any ability to learn in a new environment, and under which a market friction does not depend on the general human capital.

Proposition 5 We assume that $\eta_{x} \geq 0$ for both $x = b$ and e.

1. Suppose that $f'_{Gl}\left(h^{G}\right) \leq f'_{Gn}\left(h^{G}\right)$ and $F'\left(h^{G}\right) \geq 0$ and, therefore, $\tilde{\alpha}_{G} \geq \bar{\alpha}_{G} \geq 0$. If $f'_{Tn}\left(h^{T}\right) = f'_{Tl}\left(h^{T}\right) = 0$, and, therefore, $\tilde{\alpha}_{T} = \bar{\alpha}_{T} = 0$, then, for both $x = b$ and e, we have:

$$\operatorname{sign}\left[\frac{d\left(-\lambda\Pi^{n}\left(\mathbf{h}\right)\right)}{d\tau_{x}}\right]=-\operatorname{sign}\left[\frac{d\left(-\lambda\Delta\pi\left(\mathbf{h}\right)\right)}{d\tau_{x}}\right]=\operatorname{sign}\left[\frac{\partial\left[-\lambda\left(\Delta\pi\left(\mathbf{h}\right)+\Pi^{n}\left(\mathbf{h}\right)\right)\right]}{\partial\tau_{x}}\right]. \tag{10}$$

2. Suppose that $f'_{Tl}\left(h^{T}\right) \geq f'_{Tn}\left(h^{T}\right) \geq 0$ and, therefore, $\bar{\alpha}_{T} \geq \tilde{\alpha}_{T} \geq 0$.

   (a) If $f'_{Gl}\left(h^{G}\right) = f'_{Gn}\left(h^{G}\right)$ and, therefore, $\bar{\alpha}_{G} = 0$, then, for both $x = b$ and e, we have:

$$\frac{d\left(-\lambda\Delta\pi\left(\mathbf{h}\right)\right)}{d\tau_{x}} \leq 0 \tag{11}$$

   (b) If $F'\left(h^{G}\right) = 0$ and, therefore, $\tilde{\alpha}_{G} - \bar{\alpha}_{G} = 0$, then, for both $x = b$ and e, we have:

$$\frac{\partial\left[-\lambda\left(\Delta\pi\left(\mathbf{h}\right)+\Pi^{n}\left(\mathbf{h}\right)\right)\right]}{\partial\tau_{x}} \leq 0 \tag{12}$$

   (c) If $f'_{Gl}\left(h^{G}\right) = f'_{Gn}\left(h^{G}\right)$ and $F'\left(h^{G}\right) = 0$ and, therefore, $\tilde{\alpha}_{G} = \bar{\alpha}_{G} = 0$, then, for both $x = b$ and e, we have:

$$\frac{d\left(-\lambda\Pi^{n}\left(\mathbf{h}\right)\right)}{d\tau_{x}} \leq 0, \frac{d\left(-\lambda\Delta\pi\left(\mathbf{h}\right)\right)}{d\tau_{x}} \leq 0, \frac{\partial\left[-\lambda\left(\Delta\pi\left(\mathbf{h}\right)+\Pi^{n}\left(\mathbf{h}\right)\right)\right]}{\partial\tau_{x}} \leq 0 \tag{13}$$

The intuition behind equation (10) in Proposition 5 is explained as follows. If target firm-specific human capital does not increase productivity, $f'_{Tn}\left(h^{T}\right) = f'_{Tl}\left(h^{T}\right) = 0$, 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, because an increase in tenure assists in improving

general human capital relative to other work experience, it increases the 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 a general skill that improves the learning capability are not initially realized. Hence, the initial products of tenure, net of future 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 the experiences that improve general skills in this case, it simply lowers the productivity of the manager in a merged company and, therefore, increases the separation probability at all times. However, because the initial productivity of the manager after the M&A is not subject to this negative effect, the initial products of tenure, net of future products, are positive. Therefore, the coefficient of the interaction with the $t = 0$ dummy must be negative. Hence, for both cases, equation (10) must be satisfied.

Similarly, we can explain the intuition behind equations (11) (12) and (13) in Proposition 5 as follows. If there are no benefits from learning capability $f'_{Gl}\left(h^G\right) = f'_{Gn}\left(h^G\right)$, an improvement in general human capital does not cause any productivity difference between the initial period and the later period. Because the human capital specific to the target firm possessed by the target-firm manager 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 (11) shows this condition.

If the cost of hiring a new manager does not depend on the general human capital of the manager, $F'\left(h^G\right) = 0$, as a result of market competition, there are no benefits from employing a manager with high general human capital. Therefore, as an increase in tenure raises target firm-specific human capital, it certainly lowers the initial separation probability. Equation (12) shows this condition.

If there are no benefits from learning capability $f'_{Gl}\left(h^G\right) = f'_{Gn}\left(h^G\right)$, and the cost of hiring a new manager does not depend on the general human capital of the manager, $F'\left(h^G\right) = 0$, an increase in tenure changes the separation probability only because it can raise the human capital specific to the target firm. In addition to this, for the reason explained above, the coefficient of the interaction of tenure with the $t = 0$ dummy must be negative. Equation (13) summarizes these conditions.

Note that Proposition 5 requires that equations (10) and (11) are 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 profitable within the merged firm, the coefficients of tenure 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 5.

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 6 Suppose that $f'_{Tl}\left(h^T\right) \geq f'_{Tn}\left(h^T\right) \geq 0$, and, therefore, $\bar{\alpha}_T \geq \tilde{\alpha}_T \geq 0$. Suppose that $f'_{Gl}\left(h^G\right) \leq f'_{Gn}\left(h^G\right)$ and $F'\left(h^G\right) \geq 0$ and, therefore, $\tilde{\alpha}_G \geq \bar{\alpha}_G \geq 0$. 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:

$$\operatorname{sign}\left[\frac{d\left(-\lambda\Pi^{n}\left(\mathbf{h}\right)\right)}{d\tau_{x}}\right]=-\operatorname{sign}\left[\frac{d\left(-\lambda\Delta\pi\left(\mathbf{h}\right)\right)}{d\tau_{x}}\right]=\operatorname{sign}\left[\frac{\partial\left[-\lambda\left(\Delta\pi\left(\mathbf{h}\right)+\Pi^{n}\left(\mathbf{h}\right)\right)\right]}{\partial\tau_{x}}\right]. \tag{14}$$

Equation (14) looks very similar to equation (10). The only difference is that Proposition 5 requires that equation (10) must be jointly satisfied for both $x = e$ and $x = b$, whereas Proposition 6 states that we can separately use equation (14) for $x = e$ and $x = b$. The intuition behind equation (14) is the same as that behind equation (10). 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 general human capital and experience that we wish to reject.

Proposition 7. Suppose that $f'_{Tl}\left(h^{T}\right)\geq f'_{Tn}\left(h^{T}\right)\geq 0$, and, therefore, $\overline{\alpha}_{T}\geq\tilde{\alpha}_{T}\geq 0$. Suppose that $f'_{Gl}\left(h^{G}\right)\leq f'_{Gn}\left(h^{G}\right)$ and $F'\left(h^{G}\right)\geq 0$ and, therefore, $\tilde{\alpha}_{G}\geq\overline{\alpha}_{G}\geq 0$.

1. If $\omega_{x}\geq\mu$, which means that a long tenure during x, where $x = b$ or e, does not hamper the accumulation of general human capital relative to other experiences, then:

$$\frac{d\left(-\lambda\Pi^{n}\left(\mathbf{h}\right)\right)}{d\tau_{x}}\leq 0,\frac{\partial\left[-\lambda\left(\Delta\pi\left(\mathbf{h}\right)+\Pi^{n}\left(\mathbf{h}\right)\right)\right]}{\partial\tau_{x}}\leq 0.$$

2. If $\omega_{x}\leq\mu$, which means that a long tenure during x, where $x = b$ or e, does not assist in the accumulation of general human capital relative to other experiences, then:

$$\frac{d\left(-\lambda\Delta\pi\left(\mathbf{h}\right)\right)}{d\tau_{x}}\leq 0.$$

If $\omega_{x}\geq\mu$, an increase in tenure assists in the accumulation of general human capital. 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 initial value of tenure is larger than its subsequent value. 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.

4 Dealing with Unobserved Heterogeneity

Suppose that $\mathbf{X}_{V}$ can be decomposed into the characteristics of the target-firm manager and those of the firm to which the manager belongs. 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:

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

where $\tau_{b,fj}$ is the j th person's tenure as a board member at firm f, $\tau_{e,fj}$ is the j th person's tenure as an employee at firm f, the vector $\mathbf{X}_{D,f,j}$ contains any other target-firm manager characteristics of the j th person in firm f, including the j th person's age when firm f was taken over, and the vector $\mathbf{X}_{F,f}$ contains any characteristics of the fth firm that can influence the productivity of managers. Finally, the vector $\mathbf{X}_{z,f}$ contains any firm-level variables that can influence baseline hazard and $\hat{B}\left(t : \mathbf{X}_{Z,f}\right)$ contains $B\left(t : \mathbf{X}_{Z,f}\right)$ and some constant terms.

This equation depends on the individual unobserved heterogeneity of a manager j at firm f, χ_{fj}. This unobserved heterogeneity may be the result of unobserved capabilities, or it may be the relationship with the founding family, main bank, or parent company. We do not know the cause of this heterogeneity, but we assume that the parameter χ_{fj} can summarize the effect of these heterogeneities. Without controlling for χ_{fj}, our estimates might be biased. As explained in footnote 3, we focus on titled directors in Japanese companies. However, talented workers may be appointed to titled director positions with less experience than less talented workers. Hence, χ_{fj} and $\tau_{b,fj}$ can be correlated.

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

Suppose that a person is appointed to a management position (titled director) in Japan if and only if:

$$
\begin{aligned}
&\hat{h}_{fj}\left(\hat{t}\right) \geq \mathbf{H}'\mathbf{N}_h\left(f, \hat{t}\right), \\
&\hat{h}_{fj}\left(\hat{t}\right) \equiv \left(1-\omega\right) h^T_{fj}\left(\hat{t}\right) + \omega h^G_{fj}\left(\hat{t}\right)
\end{aligned}
\tag{15}
$$

where $h^T_{fj}\left(\hat{t}\right)$ and $h^G_{fj}\left(\hat{t}\right)$ are the target firm-specific human capital and the general human capital of a person j at the fth firm in calendar year $\hat{t}, \omega$ is the weight on the general human capital 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 $\hat{t}$.

From the definition of $\hat{h}_{fj}\left(\hat{t}\right)$, we can find the parameters ϕ_o, ϕ_b, ϕ_e, and ϕ_o to describe the following human capital accumulation function:

$$
\hat{h}_{fj}\left(\hat{t}\right) = \phi_0 + \phi_b \tau_{b,fj}\left(\hat{t}\right) + \phi_e \tau_{e,fj} + \phi_o \tau_{o,fj} + \omega \chi_{fj}, \tag{16}
$$

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 j th person's length of experience before joining firm f. Combining equations (15) and (16), when an employee j is appointed to an executive position, the following equality must be satisfied:

$$
\tau_{b,fj}\left(\hat{t}^*_j\right) = \tilde{\mathbf{H}}'\mathbf{N}_h\left(f, \hat{t}^*_j\right) - \tilde{\phi}_0 - \tilde{\phi}_e \tau_{e,fj} - \tilde{\phi}_o \tau_{o,fj} - \tilde{\chi}_{fj}, \tag{17}
$$

where $\tilde{\mathbf{H}}=\frac{\mathbf{H}}{\phi_b},\tilde{\phi}_0=\frac{\phi_0}{\phi_b},\tilde{\phi}_e=\frac{\phi_e}{\phi_b},\tilde{\phi}_o=\frac{\phi_o}{\phi_b},\tilde{\chi}_{fj}=\frac{\omega\chi_{fj}}{\phi_b}$, 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:

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

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, $E\left[\tilde{\chi}_{fj}\mid I_b=1,f,\hat{t}\right]$ 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 (17) can be rewritten as follows:

$$\tau_{b,fj}\left(\hat{t}_j^*\right)=\tilde{\mathbf{H}}'\mathbf{N}_h\left(f,\hat{t}_j^*\right)-\tilde{\phi}_0-\tilde{\phi}_e\tau_{e,fj}-\tilde{\phi}_o\tau_{o,fj}-E\left[\tilde{\chi}_{fj}\mid I_b=1,f,\hat{t}\right]-\varepsilon_{fj}\left(\hat{t}\right),\forall\hat{t}. \tag{18}$$

Note that $E\left[\varepsilon_{fj}\left(\hat{t}\right)\mid I_b=1\right]=0$ 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 $E\left[\varepsilon_{fj}\left(\hat{t}\right)\mathbf{N}_f\mid I_b=1\right]=0$ for the vector of firm-level variables $\mathbf{N}_f$ for any $\hat{t}$. Therefore, if we know a functional form of $E\left[\tilde{\chi}_{fj}\mid I_b=1,f,\hat{t}\right]$, 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 $E\left[\tilde{\chi}_{fj}\mid I_b=1,f,\hat{t}\right]$. Because we use the sample of titled directors at target firms in the year when an M&A occurs, equation (15) 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 (16) 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:

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

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:

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

We need to estimate $E\left[\tilde{\chi}_{fj} \mid \tilde{\chi}_{fj} \geq R\left(f,j,\hat{t}_f^m\right), f, \hat{t}_f^m\right]$ 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 $E\left[\tilde{\chi}_{fj} \mid \tilde{\chi}_{fj} \geq R\left(f,j,\hat{t}_f^m\right), f, \hat{t}_f^m\right]$, 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, $E\left[\tilde{\chi}_{fj} \mid \tilde{\chi}_{fj} \geq R\left(f,j,\hat{t}_f^m\right), f, \hat{t}_f^m\right]$, is a function of $R\left(f,j,\hat{t}_f^m\right)$, which is determined by the parameters of our structural equation (18), $\tilde{\mathbf{H}}, \tilde{\phi}_0, \tilde{\phi}_e$, and $\tilde{\phi}_o$. Therefore, it might be possible to jointly estimate the parameters on the conditional expectation of unobserved ability with equation (18) 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 $F\left(\tilde{\chi}\right)$, where $F'\left(\tilde{\chi}\right)$ is differentiable in a neighborhood of the upper bound of its support $\overline{\chi} \in R \cup \{+\infty\}$, and there exists $\xi = \lim_{\tilde{\chi} \to \overline{\chi}} \frac{d}{d\tilde{\chi}} \frac{1-F(\tilde{\chi})}{F'(\tilde{\chi})}$ and $\xi < \infty$. We define $\overline{F}\left(\tilde{\chi}\right) = 1 - F\left(\tilde{\chi}\right)$ and $Q\left(x\right) = \overline{F}^{-1}\left(x\right)$. Gabaix and Landier (2008) apply an extreme value theorem and show that there exists a χ_0 and ζ such that:

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

Our appendix proves the following lemma.

Lemma 8 Suppose that $Q'\left(x\right) = -\chi_0 x^{\zeta-1}$, where $\zeta \geq 0$, and that there exists a $Q\left(0\right) = \lim_{\varepsilon \to 0}\left[Q\left(\varepsilon\right) + \frac{\chi_0 \varepsilon^{\zeta}}{\zeta}\right]$ and $Q(0) < \infty$. Then, for any small x_p, we have:

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

The assumption in Lemma 8 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 $x_p = \overline{F}\left(R\left(f,j,\hat{t}_f^m\right) : f, \hat{t}_f^m\right)$. Then, $\left\{x \leq x_p\right\} = \left\{\tilde{\chi}_{fj} \geq R\left(f,j,\hat{t}_f^m\right)\right\}$ 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 $E\left[\tilde{\chi}_{fj} \mid \tilde{\chi}_{fj} \geq R\left(f,j,\hat{t}_f^m\right) : f, \hat{t}_f^m\right]$ by:

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

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:

$$0 = E\left[\varepsilon_{fj}\left(\hat{t}_f^m\right)\mathbf{N}_f \mid I_b = 1\right], \tag{19}$$

$$\varepsilon_{fj}\left(\hat{t}_f^m\right) = R\left(f, j, \hat{t}_j^*\right) - \frac{\zeta \boldsymbol{\Psi}' \mathbf{N}_\chi\left(f, \hat{t}_f^m\right) + R\left(f, j, \hat{t}_f^m\right)}{\zeta + 1}, \tag{20}$$

$$R(f, j, \hat{t}) \equiv \tilde{\mathbf{H}}' \mathbf{N}_h\left(f, \hat{t}\right) - \tilde{\phi}_0 - \tilde{\phi}_e \tau_{e,fj} - \tilde{\phi}_o \tau_{o,fj} - \tau_{b,fj}\left(\hat{t}\right), \tag{21}$$

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, $\tilde{\mathbf{H}}, \tilde{\phi}_0, \tilde{\phi}_e$, and $\tilde{\phi}_o$, we can estimate $\tilde{\chi}_{fj}$ by:

$$\tilde{\chi}_{fj} = R\left(f, j, \hat{t}_j^*\right). \tag{22}$$

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

$$q_{fjt} = \hat{B}\left(t: \mathbf{X}_{Z,f}\right) e^{-\lambda\left[\Delta\pi_{fj} I(t=0) + \Pi_{fj}\right]} \tag{23}$$

$$\Pi_{fj} = \tilde{\alpha}_b \tau_{b,fj}\left(\hat{t}_f^m\right) + \tilde{\alpha}_e \tau_{e,fj} + \tilde{\alpha}_D' \mathbf{X}_{D,f,j} + \tilde{\alpha}_F' \mathbf{X}_{F,f} + \alpha_\chi^+ \tilde{\chi}_{fj}, \tag{24}$$

$$\Delta\pi_{fj} = \overline{\alpha}_b \tau_{b,fj}\left(\hat{t}_f^m\right) + \overline{\alpha}_e \tau_{e,fj} + \overline{\alpha}_D' \mathbf{X}_{D,f,j} + \overline{\alpha}_F' \mathbf{X}_{F,f} + \overline{\alpha}_\chi^+ \tilde{\chi}_{fj}, \tag{25}$$

where $\tilde{\alpha}_\chi^+ = \frac{\phi_b}{\omega}\tilde{\alpha}_\chi$ and $\overline{\alpha}_\chi^+ = \frac{\phi_b}{\omega}\overline{\alpha}_\chi$.

5 Data

5.1 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

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.

we refer to the post-merger 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 "other merged firms".

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

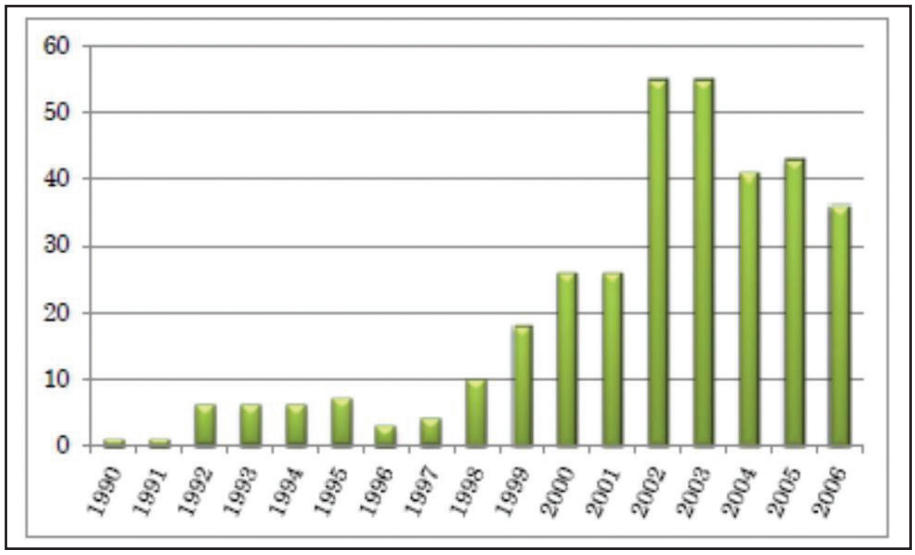
The exchanges on which the target firms are listed include the Tokyo, Osaka, Nagoya, and other local stock exchanges, as well as emerging and over-the-counter markets. However, we exclude cases in the JASDAQ market because director-level data for these firms are available only after 2000.

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 prior to the M&A if there is a surviving firm with the same stock code. For cases of M&As

Figure 1 Number of M&As each year during the period 1990-2006.



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

Table 1 Numbers of mergers and acquisitions by industry

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

without any surviving firm, we use the average value of these variables for all other merged firms instead to represent those of the acquirer.

5.2 **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

12 Auditors are excluded.

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*) 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.

5.3 Empirical specification

5.3.1 GMM estimation

Using the dataset explained above, we first conduct a GMM estimation using equations (19), (20), and (21) to obtain the unobserved individual ability defined in equation (22).

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 (16). *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

Table 2 Characteristics of target directors

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

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 other merged firm of an M&A survives as the new firm and the manager of the target firm was also a director in this other merged firm and continues to be a director in the new firm, we consider this manager to have been retained.

Table 3 Summary statistics

Variable	No.	Mean	S.D.	Min.	Max.
Tenure as a board member at $\hat{t}_i^*$ (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 $\boldsymbol{N}_h\left(f,\hat{t}\right)$, 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 $\hat{t}_j^*$ $\left(1000\right)$	1520	1.70	2.65	0.006	23.87
No. of employees at $\hat{t}_f^m$ $\left(1000\right)$	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 $\boldsymbol{N}_\chi\left(f,\hat{t}_f^m\right)$					
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 others ROA	1520	0.00	0.06	−0.68	0.60
Log top 10 share of acquirer or others	1520	−7.09	1.56	−8.89	0.00

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*, $\tau_{o,ff}$, 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. 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. If a firm has the high median age, it is likely that the firm set a high requirement to become a titled director. 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.

5.3.2 **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 (23), (24), and (25), which requires us to specify the vector $\mathbf{X}_{D,f,j}$ of target-firm manager characteristics other than tenure at the time of the M&A, the vector $\mathbf{X}_{F,f}$ that influences the productivity of managers, which includes the

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 $\hat{t}_j^*$.

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.

For the variables in the vector $\mathbf{X}_{D,f,f}$, we use the manager's age at the time of the M&A (*Age*), 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 $\mathbf{X}_{F,f}$ 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 productivity of managers in a target firm. The board size of the target company is also important for another reason: between 1990 and 2006, many companies reduced the size of their boards. If we do not control for board size, the coefficient on tenure may capture this effect as well. Our theory is based on the presumption that there are some explicit or implicit transfers that have to be negotiated after the M&A. Although we believe that this is a plausible assumption, some may argue that this may not be true. If so, the initial contracts resulting from the ex ante strategic bargaining during the M&A can influence the separation probability. To estimate the importance of ex ante strategic bargaining, Bargeron et al. (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 $\mathbf{X}_{F,f}$, so that we can minimize the bias even if our assumptions are not valid. We expect that 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 $\mathbf{X}_{F,f}$ 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

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; 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.

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 other merged 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 other merged 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 other merged firms in place of the acquirer values ("*Acquirer or others ROA*" and "*Log top 10 share of acquirer or others*").

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 (23), 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. Note that because hazard rate depends on year after M&A, controlling the year of the M&A helps avoiding the impact of economic conditions. 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].

6 Results

Table 4 shows the results of our GMM estimation using equations (19), (20), and (21). 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 (20) 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

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.

Table 4 GMM estimation

Variable	Coef.	Std. Err.	
Negative operating income	−0.636	0.404	
Direct	−1.964	1.030	*
No. of employees	0.000	0.000	
Wage	0.242	0.049	***
Median tenure as a titled director	0.025	0.028	
Median age	0.801	0.253	***
Constant $\left(\tilde{\phi}_0\right)$	73.183	13.474	***
Tenure as an employee $\left(\phi_e\right)$	0.400	0.230	*
Outside experience $\left(\tilde{\phi}_o\right)$	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		

1. 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 $\hat{t}_f^m$.
2. ***, ** and * denote statistical significance at the 1%, 5%, and 10% significance levels. Standard errors are calculated using bootstrap methods.

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 8. 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 (22). Table 5 shows the summary statistics for individual unobserved ability, whereas Figure 2 represents the kernel density estimate of the estimated unobserved ability.

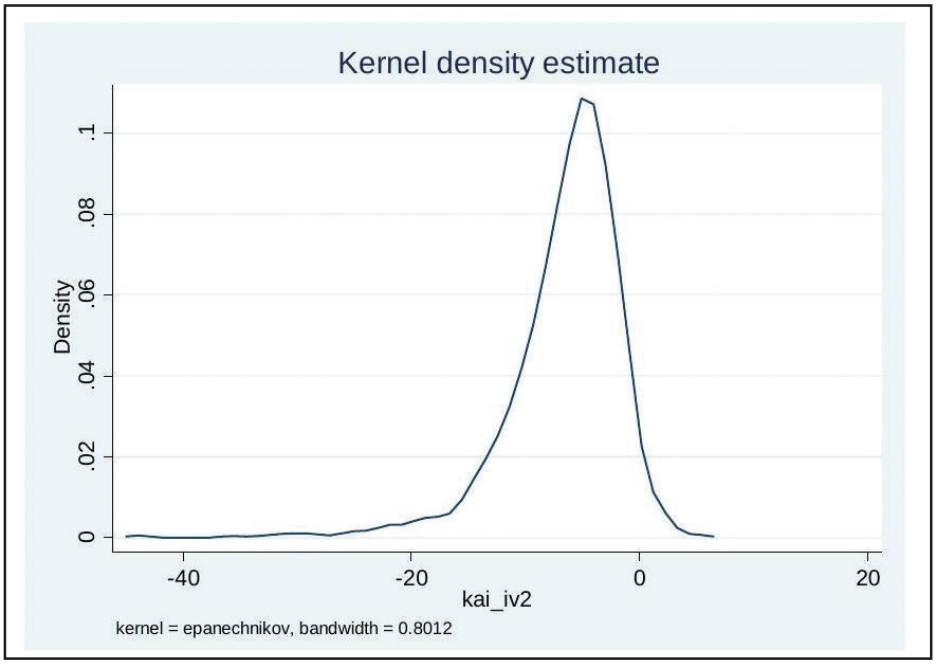
Tables 6 and 7 show the results of our stratified Cox proportional hazards model estimation using equations (23), (24), and (25). All results pass the test of the proportional hazards assumption, based on the test of nonzero slope in a generalized linear regression of the scaled Schoenfeld residuals on a natural log of the analysis time at the 5% level. For equations (1) and (3), the standard errors are obtained using 1000 bootstrap samples.

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

Table 5 Estimated unobserved skill

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

Figure 2 Kernel density of estimated unobserved skill.



of tenure as an employee is positive and significant. The coefficient of tenure as a board member 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 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

Table 6 Stratified Cox proportional hazards model estimation

	(1)			(2)		
Variable	**Coef.**	**Std. Err.**		**Coef.**	**Std. Err.**	
1: Tenure as a board member	0.006	0.024		−0.010	0.009	
2: Tenure as an employee	0.011	0.007	*	0.011	0.005	**
3: Tenure as a board member × initial dummy	−0.012	0.024		0.002	0.010	
4: Tenure as an employee × initial dummy	−0.014	0.007	**	−0.014	0.005	***
Age	0.112	0.031	***	0.100	0.016	***
Board size	0.097	0.081		0.105	0.049	**
No. of firms as a board member	−0.166	0.215		−0.179	0.137	
Board member in the acquirer before the M&A	−0.106	0.479		−0.108	0.308	
Acquisition	−0.025	0.797		0.003	0.495	
Target ROA	−11.913	7.626		−11.335	3.860	***
Related	0.117	0.562		0.034	0.374	
Target firm's relative size in assets	−0.990	0.530	*	−0.942	0.433	**
Log target assets	−0.026	0.216		−0.066	0.133	
Upper	−0.277	0.198		−0.264	0.132	**
Log stock share of top 10	−0.460	0.355		−0.455	0.194	**
Acquirer or others ROA	−3.164	6.143		−2.214	3.368	
Log top 10 share of acquirer or others	0.134	0.265		0.092	0.162	
$\tilde{\chi}_{fj}$	0.044	0.039				
Age × $I(t = 0)$	−0.072	0.031	***	−0.061	0.017	***
Board size × $I(t = 0)$	−0.073	0.082		−0.080	0.051	
No. of firms as a board member × $I(t = 0)$	−0.293	0.248		−0.281	0.184	
Board member in the acquirer before the M&A × $I(t = 0)$	−0.114	0.482		−0.116	0.312	
Acquisition × $I(t = 0)$	0.729	0.819		0.702	0.525	
Target ROA × $I(t = 0)$	11.813	7.691		11.198	3.910	***
Related × $I(t = 0)$	−0.383	0.577		−0.300	0.388	
Target firm's relative size in assets × $I(t = 0)$	−0.210	0.567		−0.257	0.464	
Log target assets × $I(t = 0)$	0.012	0.220		0.054	0.138	
Upper × $I(t = 0)$	−0.007	0.207		−0.017	0.143	
Log stock share of top 10 × $I(t = 0)$	0.504	0.357		0.503	0.200	**
Acquirer or others ROA × $I(t = 0)$	5.533	6.241		4.636	3.577	
Log top 10 share of acquirer or others × $I(t = 0)$	−0.161	0.269		−0.120	0.165	
$\tilde{\chi}_{fj} \times I(t = 0)$	−0.037	0.039				
Wald Test	Coef.	Chi2(1)		Coef.	Chi2(1)	
1 + 3	−0.006	1.28		−0.008	5.31	**
2 + 4	−0.003	2.71	*	−0.003	3.74	*
Number of observations	2108			2108		

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

***, **, * denote statistical significance at the 1%,5%, and 10% significance levels.

Standard errors are calculated by bootstrap methods.

Table 7 Stratified Cox proportional hazards model estimation

	(1)			(2)		
Variable	**Coef.**	**Std. Err.**		**Coef.**	**Err.ef.**	
1: Tenure as a board member	−0.002	0.022		−0.016	0.009	
2: Tenure as an employee	0.010	0.006	*	0.010	0.005	*
3: Tenure as a board member × initial dummy	−0.005	0.022		0.007	0.010	
4: Tenure as an employee × initial dummy	−0.013	0.006	**	−0.012	0.005	***
Age	0.116	0.026	***	0.107	0.016	***
Board size	0.036	0.020	*	0.039	0.046	**
No. of firms as a board member	−0.358	0.094	***	−0.361	0.082	***
Board member in the acquirer before the M&A	−0.229	0.074	***	−0.232	0.064	***
Acquisition	0.579	0.156	***	0.578	0.139	***
Target ROA	−0.466	1.000		−0.463	4.068	
Related	−0.244	0.106	**	−0.252	0.097	***
Target firm's relative size in assets	−1.016	0.437	**	−1.006	0.374	***
Log target assets	−0.036	0.053		−0.040	0.049	
Upper	−0.279	0.059	***	−0.275	0.133	***
Log stock share of top 10	−0.009	0.057		−0.008	0.195	
Acquirer or others ROA	1.223	1.439		1.370	1.219	
Log top 10 share of acquirer or others	−0.028	0.047		−0.035	0.043	
$\tilde{\chi}_{fj}$	0.036	0.033				
Age × $I(t = 0)$	−0.076	0.026	***	−0.069	0.017	***
Board size × $I(t = 0)$						
No. of firms as a board member × $I(t = 0)$						
Board member in the acquirer before the M&A						
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	0.484		−0.289	0.414	
Log target assets × $I(t = 0)$						
Upper × $I(t = 0)$						
Log stock share of top 10 × $I(t = 0)$						
Acquirer or others ROA × $I(t = 0)$						
Log top 10 share of acquirer or others × $I(t = 0)$						
$\tilde{\chi}_{fj} \times I(t = 0)$	−0.027	0.034				
Wald Test	Coef.	Chi2(1)		Coef.	Chi2(1)	
1 + 3	−0.006	1.18		−0.009	5.58	**
2 + 4	−0.003	2.95	*	−0.003	4.08	**
Number of observations	2108			2108		

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

***, **, * denote statistical significance at the 1%,5%, and 10% significance levels.

Standard errors are calculated using bootstrap methods.

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 estimated coefficients for Specifications (2) and (4) are almost the same as those for (1) and (3), respectively, except for the coefficients on tenure as a board member and its interaction term. 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), implying that longer board experiences increase the appointment probability of a board member in a new firm. However, this seems to reflect a selection bias resulting from an unobserved skill because these effects of board experiences disappear in (1) and (3), where we control for unobserved skills.

Now, we wish to interpret the results using our theory. We first apply Proposition 5 to interpret the results of our estimation. Because the conditions in Proposition 5 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 5 is rejected. First, let us test hypotheses 1 and 2 in Proposition 5 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 $\frac{\partial\left[-\lambda\Pi^{n}(\mathbf{h})\right]}{\partial\tau_{e}}$, $\frac{\partial\left[-\lambda\Delta\pi(\mathbf{h})\right]}{\partial\tau_{e}}$, and $\frac{\partial\left[-\lambda\left(\Delta\pi(\mathbf{h})+\Pi^{n}(\mathbf{h})\right)\right]}{\partial\tau_{e}}$, respectively. Our estimated coefficients are all statistically significant and the signs we obtained are $\frac{\partial\left[-\lambda\Pi^{n}(\mathbf{h})\right]}{\partial\tau_{e}}>0$, $\frac{\partial\left[-\lambda\Delta\pi(\mathbf{h})\right]}{\partial\tau_{e}}<0$, and $\frac{\partial\left[-\lambda\left(\Delta\pi(\mathbf{h})+\Pi^{n}(\mathbf{h})\right)\right]}{\partial\tau_{e}}<0$. Hence, they definitely reject hypothesis $f'_{Tn}\left(h^{T}\right)=f'_{Tl}\left(h^{T}\right)=0$ in Proposition 5. Therefore, the evidence suggests that new firms value firm-specific human capital. Although the directions of the coefficients cannot reject hypotheses $f'_{Gl}\left(h^{G}\right)=f'_{Gn}\left(h^{G}\right)$ and $F'\left(h^{G}\right)=0$ separately, they can reject the joint hypotheses of $f'_{Gl}\left(h^{G}\right)=f'_{Gn}\left(h^{G}\right)$ and $F'\left(h^{G}\right)=0$ in Proposition 5. This suggests that a new firm values general human capital either because new firms appreciate learning capability and/or because it is more difficult to hire managers with higher general human capital.

Next, we separately investigate the role of tenure as a board member and that of tenure as an employee after M&As. Let us first apply Proposition 6 to the estimated results for the tenure as an employee. It shows that the evidence for tenure as an employee, $\frac{\partial\left[-\lambda\Pi^{n}(\mathbf{h})\right]}{\partial\tau_{e}}>0$, $\frac{\partial\left[-\lambda\Delta\pi(\mathbf{h})\right]}{\partial\tau_{e}}<0$, and $\frac{\partial\left[-\lambda\left(\Delta\pi(\mathbf{h})+\Pi^{n}(\mathbf{h})\right)\right]}{\partial\tau_{e}}<0$, rejects the hypothesis that experience as an employee does not increase firm-specific skills ($\eta_e = 0$). Similarly, applying Proposition 7, the evidence for tenure as an employee, $\frac{\partial\left[-\lambda\Pi^{n}(\mathbf{h})\right]}{\partial\tau_{e}}>0$, $\frac{\partial\left[-\lambda\Delta\pi(\mathbf{h})\right]}{\partial\tau_{b}}<0$, and $\frac{\partial\left[-\lambda\left(\Delta\pi(\mathbf{h})+\Pi^{n}(\mathbf{h})\right)\right]}{\partial\tau_{e}}<0$, 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 hampers the accumulation of general human capital relative to other experiences.

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 have a larger influence on the accumulation of general human capital relative to other experiences. 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 experience as a board member does not involve firm-specific human capital and does not hamper the accumulation of general human capital. Hence, it is likely that managerial experiences are more general than those as an employee.

In summary, we can interpret our results as follows. 1) After an M&A, Japanese firms value both the target firm-specific human capital and the general human capital of managers. 2) Experience as an employee increases firm-specific skills, but at the expense of the accumulation of general human capital. 3) Managerial experiences are likely to be more general than those as an employee.

7 Discussion

In this section, we discuss whether the moral hazard of managers could influence our interpretation. In line with the literature, including Wulf (2004), Hartzell et al. (2004), and Bargeron et al. (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.

As this paper shows, as long as ex post bargaining is possible without any costs, we do not need to be concerned about this possibility. In fact, although we control for a measure of the ownership structure using the logarithm of the stock share of the top 10 stock holders, it does not have any significant impacts on the separation rate. This strongly suggests that the entrenchment story may not be a serious problem for M&As in the Japanese context.

Of course, this may not be taken as conclusive evidence to justify our presumption. If the ex post negotiation is costly, 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 would be eroded such that, finally, it would become possible for the new firms to fire them. Thus, this kind of entrenchment story may be able to explain the coefficient of tenure in our survival analysis.

To the extent that the power of these managers is the source of their productivity, there is no reason to distinguish their power from their target firm-specific skills. In this case, the above story can be consistent with the story that we have presented. However, it is possible that a powerful manager could be unproductive. Unfortunately, we do not have conclusive evidence that can reject this kind of entrenchment story. Nevertheless, there are at least three reasons to consider that the entrenchment of powerful but unproductive managers 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 power in a firm 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. Finally, senior people typically have more authority in Japanese society than do younger people. Hence, it is possible that power in a firm 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, the entrenchment story will have little effect on the coefficient of tenure.

8 Contribution of our Results to the Literature

In addition to our methodological contributions to the literature, our results have a number of important implications for several streams of literature, including those on managerial compensation and Japanese M&As. This section discusses how our results contribute to these fields.

First, 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, whereas others consider it to be the result of a competitive market for managerial talent. One of the key presumptions behind the market-based view (e.g., Gabaix and Landier, 2008; Terviö, 2008) is that important managerial skills are transferable across firms. Indeed, 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 et al. (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.-based literature; the transferability of managerial human capital depends on the source of managerial human capital. We find that whereas managerial human capital accumulated through experiences as an employee is firm specific, managerial experiences are likely to be more general than those gained as employees, 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[19] and that Japanese CEOs have very long tenures at their firms relative to CEOs in U.S. firms[20]. This suggests that, even if managerial experience 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 not only tasks that require managerial

19 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 wage income share of the top 1% of wage earners 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.

20 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.

experiences, but also those that require the type of experience gained as an employee[21], which make it more difficult to hire managers from outside the firm in Japan.

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 et al. 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 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 differ from the reasons for the separation of employees. Hence, the evidence in our paper provides valuable new information to understand how Japanese firms adapt to merger waves.

9 Conclusion

This paper examines how the tenure of managers influences the retention rate of management groups after M&As in Japanese companies. It develops a general equilibrium model of managerial separation after M&As that distinguishes several hypotheses about the effect of tenure on separation, given several data limitations. 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 the accumulation of managers' general human capital.

To confirm the external validity of the empirical results found in data-rich countries such as the U.S., it is important to analyze managerial human capital in multiple countries, including data-poor countries such as Japan. However, it is likely that the data availability would impose several limitations on such a study, as they have in our case. This paper provides useful methods to overcome these possible data limitations. In particular, we provide a novel method to correct for selection biases by utilizing the timing of selection in a selected sample, which does not require a random sample from the population. We hope that our approach can assist in accumulating empirical evidence in countries with limited data.

Of course, the development of the model is based on several assumptions. Hence, we do not claim to provide conclusive evidence on the retention rate of managers after M&As. However, because the importance of human capital in understanding the retention of managers cannot be denied, we believe that our results provide a reasonable explanation of the otherwise puzzling evidence on the retention rate of managers after M&As. Moreover, because our model is flexible and amendable, it can provide a sound basis for the development of a more complete theory of retention in the future.

21 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.

References

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.

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

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

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.

Cameron, A. Colin and Pravin K. Trivedi (2005), *Microeconometrics: Method and Applications*, Cambridge University Press: Cambridge.

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.

Chetty, Raj (2009), "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," Annual Review of Economics, Vol. 1, 451-487.

Eisfeldt, Andrea L. and Camelia M. Kuhnen (2013), "CEO Turnover in a Competitive Assignment Framework," Journal of Financial Economics, Vol. 109, 351-372.

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.

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

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

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

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.

Gabaix, Xavier and Augustin Landier (2008), "Why has CEO Pay Increased So Much?" Quarterly Journal of Economics, Vol. 123, 49-100.

Gibbons, Robert and Michael Waldman (2006), "Enriching a Theory of Wage and Promotion Dynamics Inside Firms," Journal of Labor Economics, Vol. 24, No. 1, 59-107.

Grossman, Sanford J. and Oliver D. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," Journal of Political Economy, Vol. 94, No. 4, 691-719.

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.

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

Heckman, James J. (2010), "Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy," Journal of Economic Literature Vol. 48, No.2, 356-398.

Hotz, Joseph V. and Robert A. Miller, (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models, " The Review of Economic Studies, Vol. 60, No. 3, pp. 497-531.

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

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

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.

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.

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.

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

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).

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

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

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

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

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

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.

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.

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.

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

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.

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

Rust, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," Econometrica, 55, 999-1033.

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.

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

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

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.

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

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

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

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.

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

Wooldridge, Jeffrey M. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press: Cambridge.

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.

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.

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.

Appendix

The Proof of Proposition 1 Note that, if both the firm and the manager prefer to maintain the relationship at time t, $\left(J_t(\mathbf{h}) \geq \hat{J}^0 - P_t \text{ and } W_t(\mathbf{h}) \geq \hat{W}^l\left(h^G\right) + P_t + \varepsilon_t\right)$, they can do so. However, if one party prefers to walk away, $J_t(\mathbf{h}) < \hat{J}^0 - P_t$ or $W_t(\mathbf{h}) < \hat{W}^l\left(h^G\right) + P_t + \varepsilon_t$, then they must separate from each other. This implies that they can maintain their relationship at t if and only if:

$$S_t(\mathbf{h}) \geq W_t(\mathbf{h}) - \hat{W}^l\left(h^G\right) - P_t \geq \varepsilon_t.$$

Applying equation (26) to the definition of S_t, the dynamics of S_t can be expressed as:

$$\begin{aligned} S_t(\mathbf{h}) = {} & p_t(\mathbf{h}) - (1-\beta)\left(\hat{J}^0 + \hat{W}^l\left(h^G\right)\right) \\ & + \beta E_t\left[\begin{array}{c} S_{t+1}(\mathbf{h}) + \\ \left[1 - I\left(S_{t+1}(\mathbf{h}) \geq W_{t+1}(\mathbf{h}) - \hat{W}^l\left(h^G\right) - P_{t+1} \geq \varepsilon_{t+1}\right)\right]\left(\varepsilon_{t+1} - S_{t+1}(\mathbf{h})\right) \end{array}\right] \end{aligned}$$

Hence, w_t^n cannot influence $S_t(\mathbf{h})$. We claim that if $S_t(\mathbf{h}) \geq \varepsilon_t$, it is mutually beneficial to find w_t^n and, therefore, W_t to satisfy equation (26). Suppose that $S_t(\mathbf{h}) \geq \varepsilon_t$. If $W_t(\mathbf{h}) - \hat{W}^l\left(h^G\right) - P_t < \varepsilon_t$, then separation occurs, the firm obtains $\hat{J}^0 - P_t$ and the worker receives $\hat{W}^l\left(h^G\right) + P_t + \varepsilon_t$. We define $W_t^*\left(h^G\right)$ such that $W_t^*\left(h^G\right) \equiv \hat{W}^l\left(h^G\right) + P_t + \varepsilon_t > W_t(\mathbf{h})$. Then, $J_t^*(\mathbf{h}) \equiv S_t(\mathbf{h}) - W_t^*\left(h^G\right) + \hat{W}^l\left(h^G\right) + \hat{J}^0 \geq \hat{J}^0 - P_t$. Hence, both parties have incentives to renegotiate and choose $W_t^*\left(h^G\right)$ to avoid separation. On the other hand, if $W_t(\mathbf{h}) - \hat{W}^l\left(h^G\right) - P_t > S_t(\mathbf{h})$, then the separation occurs, the firm obtain $\hat{J}^0 - P_t$ and worker receives $\hat{W}^l\left(h^G\right) + P_t + \varepsilon_t$. We define $W_t^{**}(\mathbf{h}) \equiv S_t(\mathbf{h}) + \hat{W}^l\left(h^G\right) + P_t$. Then, $W_t^{**}(\mathbf{h}) \geq \hat{W}^l\left(h^G\right) + P_t + \varepsilon_t$ and $J_t^{**}(\mathbf{h}) \equiv S_t(\mathbf{h}) - W_t^{**}(\mathbf{h}) + \hat{W}^l\left(h^G\right) + \hat{J}^0 = \hat{J}^0 - P_t$. Hence, both parties have incentives to renegotiate and choose $W_t^{**}(\mathbf{h})$ and avoid separation. Therefore, as long as $S_t(\mathbf{h}) \geq \varepsilon_t$, the parties will avoid separation. However, when $S_t(\mathbf{h}) < \varepsilon_t$, it is impossible to satisfy equation (26) and they must separate. The desired result is immediate.

The Proof of Lemma 8 Define $\bar{F}(\tilde{\chi}) = 1 - F(\tilde{\chi})$ and $Q(x) = \bar{F}^{-1}(x)$. Then:

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

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

$$E\left[\chi \mid \varepsilon \le x \le x_p\right] = \frac{\int_{\varepsilon}^{x_p}\left\{Q(\varepsilon) - \frac{\chi_0\left[x^{\zeta} - \varepsilon^{\zeta}\right]}{\zeta}\right\}dx}{x_p - \varepsilon}$$

$$= Q(\varepsilon) + \frac{\chi_0\varepsilon^{\zeta}}{\zeta} - \frac{\chi_0\left[x_p^{\zeta+1} - \varepsilon^{\zeta+1}\right]}{(x_p - \varepsilon)\zeta(\zeta+1)}$$

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

$$E\left[\chi \mid x \le x_p\right] = Q(0) - \frac{\chi_0 x_p^{\zeta}}{\zeta(\zeta+1)} \text{ and}$$

$$Q(x_p) = \lim_{\varepsilon\to 0}\left\{Q(\varepsilon) - \frac{\chi_0\left[x_p^{\zeta} - \varepsilon^{\zeta}\right]}{\zeta}\right\} = Q(0) - \frac{\chi_0 x_p^{\zeta}}{\zeta}$$

Combining the two equations, we can derive:

$$E\left[\chi \mid x \le x_p\right] = \frac{\zeta Q(0) + Q(x_p)}{\zeta+1}.$$