Managerial Behavior, Takeovers and Employment Duration

David N. Margolis *

May 22, 2008
Preliminary Version, please do not cite

*Paris School of Economics, CNRS, CREST and IZA. Mailing address: Centre d'Economie de la Sorbonne, Université Paris 1 Panthéon - Sorbonne, 106-112 boulevard de l'Hôpital, 75647 Paris Cedex 13, France. E-mail: David.Margolis@univ-paris1.fr. Telephone: +33 (0)1 44 07 82 62. Fax: +33 (0)1 44 07 82 47. The author would like to thank Douglas Frank, Jean-Marc Robin, Antoine Terracol and participants at the 2007 Society of Labor Economists, Université Paris 1 Séminaire Interne de Microéconomie Appliquée and ISER-University of Essex seminar for helpful discussions and comments but retains responsibility for all remaining errors.

Abstract

This paper uses a unique data set of linked employer-employee data in which asset transfers between firms can be identified to study the manner in which employment policy changes in the aftermath of a merger or acquisition (M&A). Using parametric duration models with unobserved heterogeneity, it appears that the employment policy of the firm changes radically after an M&A with respect to the "steady state". Not only do various observed characteristics impact the probability of continued employment in a different manner after an M&A, but the distribution of unobserved characteristics that affect employment changes - reflecting differences in the stock of previous employees and the flow of new hires - as does the impact of this heterogeneity on employment durations.

Resumé

Ce papier exploite une base de données employeur-employée unique qui permet lídentification des transferts d'áctifs entre entreprises pour analyser la manière dans laquelle la politique de gestion de main d'oeuvre change autour dúne fusion ou acquisition. Avec des modèles de durée paramétriques incorporant de l'hétérogénéité inobservé, on trouve que la politique de gestion de main d'oeuvre change radicalement après une fusion ou acquisition par rapport à son état stationnaire. Les caractéristiques observés qui affectent la probabilité de rester en emploi changent avec le changement de propriété, et la manière que les inobservables affectent la durée d'emploi changent aussi.

Key Words: Managerial Behavior, Employment duration, Mergers and Acquisitions, Linked Employer-Employee Data.
Mots clés: Comportement de managers, durée d'emplois, fusions et acquisitions, données longitudinales appariées employeur-employées.

JEL Codes / Codes JEL: J31, C23

Please Note: The econometric model specified in the text is currently being estimated. The empirical results presented here come from a simplified version of the model. The only differences between the model currently being estimated and the model presented below are as follows.

- The text presents a model with time varying covariates; the estimated results are based on a model with covariates fixed at the first observation date for the employment spell.
- The text presents a piecewise constant baseline hazard; the estimated results use a Weibull baseline hazard.
- The probability of a "high" type in the unobserved heterogeneity distributions is estimated as a linear function of the probability of a future M&A; in the text it is constant due to identification issues.
- The unobserved heterogeneity parameters are correlated in the model presented in the text; they are supposed independent in the estimations.

1 Introduction

Although the corporate finance literature on the different reasons for mergers and acquisitions (M&As) is quite voluminous [1], much of the literature has focused on the Jensen and Meckling (1976)-style explanation under which takeovers serve as a market-based mechanism for controlling non-shareholder-maximizing managerial behavior [2]. This approach emphasizes the importance of managers behaving in a manner that maximizes private utility without necessarily maximizing shareholder value, and some recent empirical work (Bertrand and Mullainathan (2003), Margolis (2006)) finds that such principal-agent problems may be particularly manifest in the firm's compensation policy and human resource management practices. Without denying the potential importance of alternative explanations, this paper focuses on the role of private managerial incentives by examining the manner in which employment relations are sensitive to corrections resulting from an ownership change via M&As.

When considering a firm that will be subject to an M&A, one needs to distinguish the managerial team before the transaction occurs from the one that is present afterwards. Since the incumbent management team may be responding to private incentives whereas the new (post-transaction) management team is intent on increasing shareholder value in order to repay the debt incurred to finance the transaction, one might expect to see the determinants of continued employment change as the management team changes. In particular, the new management team has to decide what to do with the stock of employees it inherits from the pre-existing firms. If the previous management team chose to hire "too many" workers with particular characteristics (men, young people, highly skilled workers, people with anti-union attitudes, ...), or simply too many (or too few) workers period, then the new management team may find itself restructuring the workforce in the years following the transaction, in addition to structuring its hiring of new workers so that its workforce more closely corresponds to the optimal workforce structure.

One implication of this reasoning is that the determinants of continued employment with the pre-transaction management team may differ from those with the post-transaction management team for incumbent workers. The new management team may target particular workers based on their observable (to the econometrician) or unobservable characteristics, or they may simply opt to accelerate turnover of existing staff. Furthermore, the unobserved characteristics that determine continued employment under a self-serving management team may be completely different from those that determine continued employment with a shareholder returns-maximizing team.

These considerations have implications for the theoretical model that will be described in section 2 and for the way that the theoretical model is taken to the data, detailed in section 3. [3] After briefly describing the data in section 4, section 5 presents the results of estimating our model from section 3. Section 6 extends the interpretation of these results and situates them with respect to particular managerial objectives in the corporate finance literature, while section 7 concludes.

1. See Margolis (2006) for a brief discussion of various alternative models.
2. See Tirole (2006) for a recent summary of this literature.
3. This paper does not estimate the structural model of section 2 but rather reduced form models that can be used to examine its implications.

2 A Bargaining Model of Wages and Employment with Capital Investment and Management with Private Objectives

This section presents a theoretical framework from which one can draw reduced-form implications for employment levels and investigate how various managerial preferences should affect employment levels relative to the optimal employment probabilities. The point of this section is not to present a theoretical model from which one can estimate a structural model (we will avoid specifying the precise functional forms that would be necessary to estimate a fully structural model), but rather to present a theoretical framework whose principal attributes are well known in the literature and from which intuition about comparative statics can be derived.

After presenting the outline of a general framework in which employment and wages are bargained and capital investment is subject to holdup, we suggest how the model might be modified to accommodate a variety of private managerial preferences. By examining the changes in the empirical determinants of continued employment before and after a transaction (in section 6) in light of this model, one can draw implications for which private incentives seem to be the most important determinants of the sub-optimizing behavior that drives mergers and acquisitions in France.

2.1 A Bargaining Model with Holdup on Capital Investment

Consider a (slightly) extended version of the sequential bargaining model proposed by Manning (1987), chosen because of its flexibility and ability to nest various models of employment and wage determination. The timing of the model proceeds as follows. In a preliminary step, the firm's managers choose the level of capital investment that they wish to employ during production. They then negotiate with workers over wages and, once the wage negotiations are settled, negotiate over employment. [4] The model is solved by backward induction, although the point of this section is not to present an analytical solution to this model but rather to describe the mechanisms by which wages, employment and capital are determined.

When allowing for heterogeneous labor, the second stage (employment) negotiations solve the following maximization problem:

$$\max_{L_1,\cdots,L_M} \left[R\left(L_1,\cdots,L_M,k\right) - rk - \left(\sum_{m=1}^{M} w_m L_m \right) \right]^{(1-\gamma_L)} \tag{1}$$

$$\left[\prod_{m=1}^{M} \left(L_m \left(w_m - \overline{w_m} \right) \right)^{\beta_m} \right]^{\gamma_L},$$

4. Manning (1987) shows that when employment negotiations occur before wage negotiations, the parties necessarily agree on the employment level that maximizes quasi-rents and then only bargain over how to share the surplus. In other words, bargaining in the opposite order - employment then wages - generates results equivalent to those of a strongly efficient contracting model. By bargaining over wages before employment, the model changes from cooperative - maximizing the surplus and then dividing it by bargaining - to non-cooperative.

where wages $w_1, \cdots, w_M$ are fixed in a first stage bargain, capital k is considered chosen by the firm prior to the wage negotiations (and is thus subject to holdup, see Grout (1984), Manning (1994)), β_m is the weight put on expected utility gains for workers of type m in the union's objective function $\left(\sum_{m=1}^{M} \beta_m = 1\right)$ and γ_L is the union's bargaining power during the negotiations over employment. [5]

The negotiated employment levels $\hat{L}_1(w_1, \cdots, w_M, k), \cdots, \hat{L}_M(w_1, \cdots, w_M, k)$ are functions of the wage rates (negotiated in the first round) and capital. The round one wage negotiations solve the following maximization problem:

$$\max_{w_1,\cdots,w_M} \left[R\left(\hat{L}_1, \cdots, \hat{L}_M, k\right) - rk - \left(\sum_{m=1}^{M} w_m \hat{L}_m\right) \right]^{(1-\gamma_w)} \left[\prod_{m=1}^{M} \left(\hat{L}_m \left(w_m - \overline{w_m}\right)\right)^{\beta_m} \right]^{\gamma_w}, \tag{2}$$

where γ_w is the union's bargaining power during the negotiations over wages. Manning (1987) shows that this model nests the main models of employment determination in the literature:

- $\gamma_L = \gamma_w$ generates the strongly efficient contracting model
- $\gamma_L = 0$ and $\gamma_w = 1$ yields the monopoly union model
- $\gamma_L = 0$ and $0 < \gamma_w < 1$ gives the "right to manage" model
- $\gamma_L = 0$ and $\gamma_w = 0$ produces the standard (without bargaining) labor demand model

Any other combination of γ_L and γ_w yields what Manning (1987) calls an "inefficient contracting" model. [6]

Finally, in a preliminary stage the firm decides on the optimal level of capital investment by solving

$$\max_k R\left(\hat{L}_1, \cdots, \hat{L}_M, k\right) - rk - \left(\sum_{m=1}^{M} \hat{w}_m \hat{L}_m\right) \tag{3}$$

Note that at this stage there is no bargaining, but since management knows how its choice of investments predetermines the results of the subsequent wage and employment negotiations, one typically observes suboptimal levels of capital investment.

5. This specification of the union's objective function corresponds to assuming that the overall union objective is the result of an internal Nash bargain among the various types of workers in the unionized workforce. Alternatively, one could assume that the union's overall objective is the result of probabilistic voting among the different components of the workforce, each of whom's objective is $L_m\left(w_m - \overline{w_m}\right)$ (Lindbeck and Weibull (1987)). In this case, the winning "platform" would imply an overall union objective function of the form $\sum_{m=1}^{M} \beta_m L_m\left(w_m - \overline{w_m}\right)$.

6. There have been very few direct tests of this bargaining framework in the literature. One paper (Hosken and Margolis (1997)) finds that, in the case of public schools in New York State, one can always reject the right to manage, monopoly union and labor demand models, whereas the strongly efficient contracting model cannot be rejected at standard confidence levels for over 17% of New York State school districts.

2.2 Modifications to the Bargaining Model to Accommodate Private Managerial Motives

The theoretical framework presented in section 2.1 can accomodate the incorporation of private motives of managers in a straightforward manner. How the model needs to be modified depends on the source of the private incentives. Tirole (2006) suggests the following possible private motivations of managers that could lead to deviations from shareholder value maximizing behavior.

2.2.1 Empire Building

In this case, managers derive private utility from being in charge of a large firm. Depending on whether large is defined in terms of capital or employment, one could replace the expression

$$R\left(L_1, \cdots, L_M, k\right) - rk - \left(\sum_{m=1}^{M} w_m L_m\right), \tag{4}$$

which represents the manager's objective function in equations 1, 2 and 3 with

$$R\left(L_1, \cdots, L_M, k\right) - \left(1 - \alpha^{EmpBuild}\right)\left[rk - \left(\sum_{m=1}^{M} w_m L_m\right)\right].$$

When $\alpha^{EmpBuild} > 0$, the manager reduces the weight of the costs associated with more capital and more labor in his or her objective function. This will lead to a higher level of capital and employment than would have been the case were $\alpha^{EmpBuild} = 0$.

2.2.2 A Lot of Cash on Hand

It has been suggested that one reason why takeovers occur is that managers keep too much cash on hand as opposed to investing it or distributing it to shareholders through dividend payments. In the context of the manager's objective function presented in equation 4, the manager acts as if he or she wants to retain a share of revenues before paying creditors, workers and residual claimants. This can be modeled either as

$$\left(1 - \alpha^{CashHand}\right) R\left(L_1, \cdots, L_M, k\right) - rk - \left(\sum_{m=1}^{M} w_m L_m\right) \tag{5}$$

with $\alpha^{CashHand} > 0$, or as

$$R\left(L_1, \cdots, L_M, k\right) - C^{CashHand} - rk - \left(\sum_{m=1}^{M} w_m L_m\right). \tag{6}$$

Equation 5 models the cash on hand motivation as if the manager retains a share of revenues before paying costs and dividends, while equation 6 models the motive as if the manager puts aside a fixed amount ($C^{CashHand}$) of revenues for other uses before

paying costs and dividends. Both formulations can induce employment levels different from those of a shareholder value maximizing manager, even the model with a fixed amount withdrawn, due to the fact that the union's objective enters into the determination of the negotiated employment and wage levels.

2.2.3 Peaceful Labor Relations

It may be the case that a manager dislikes the bad publicity associated with strikes or with conflictual labor relations more generally. Such labor peace can be facilitated by adopting a more lenient stance during negotiations with workers. In the context of equations 1 and 2, this would correspond to replacing γ_L with $\left(1+\alpha^{LaborPeace,L}\right)\gamma_L$ and γ_w with $\left(1+\alpha^{LaborPeace,w}\right)\gamma_w$. When

$$\alpha^{LaborPeace,L} > 0 \text{ and/or } \alpha^{LaborPeace,w} > 0,$$

the union's bargaining power in negotiations is increased and the negotiated wages and employment are closer to its preferred outcome.

2.2.4 Esteem from the Workforce / Playing Favorites

Some managers like to be appreciated by their employees and gain utility from being seen as generous with their workers. This is a pure insider issue, in that managers prefer to give more wages to fewer workers, thereby making those who remain happier even if it means reducing overall employment due to higher labor costs per worker. In this case the manager integrates the fact that he will be giving extra bonuses, perhaps concentrated on certain types of workers, *after* the wage bargaining takes place. These unanticipated windfall gains for certain workers are thus not directly integrated into the union's objective function, but the manager negotiates over employment and wages as if the objective function in equation 4 were really

$$R\left(L_1,\cdots,L_M,k\right)-rk-\left(\sum_{m=1}^{M}\left(1+\alpha_m^{Esteem}\right)w_mL_m\right)$$

with $\alpha_m^{Esteem} \geq 0 \; \forall \; m$ and with $\sum_{m=1}^{M}\alpha_m^{Esteem} > 0$. Not all of the α_m^{Esteem} need be identical; the manager could prefer to be seen as generous to the blue collar workers and simply pay the white collar workers what was negotiated, or alternatively to give lots of bonuses to middle or upper management while neglecting blue collar workers. But in the cases where $\alpha_m^{Esteem} > 0$, the actually negotiated wage w_m is topped up by an extra α_m^{Esteem} percent.

2.2.5 Kickbacks from Suppliers

In some cases, managers may receive kickbacks, either as monetary compensation or through other "gifts", when they choose a particular supplier over the competition. As these kickbacks are costly for the supplier, the price charged to the firm for the goods provided (intermediate materials, capital, maintenance services...) is typically higher

than the least expensive alternative. Such behavior can be modeled in a straightforward manner by replacing the manager's objective from 4 with

$$R\left(L_1, \cdots, L_M, k\right) - \left(1 + \alpha^{Kickbacks}\right) rk - \left(\sum_{m=1}^{M} w_m L_m\right).$$

Although such a specification will induce a substitution of productive capacity away from capital and towards labor, the manager expects to recover a share of the extra $\alpha^{Kickbacks} r$ per unit of capital that is paid to the supplier.

2.2.6 Non-Risk Neutrality

When a firm's recent performance has been sub-par, it has been suggested that a manager may choose to take excessive risks in the hopes of provoking a turnaround that could save his or her job. Alternatively, the manager of a firm whose recent performance has been exceptionally good may choose to adopt an overly risk-averse strategy so as to "protect" the recently acquired gains. In both of these contexts, the manager is not behaving in a risk neutral manner, whereas (in principle) it is for the shareholders to adjust the riskiness of their portfolios based on the fundamentals of firm performance and not the idiosyncratic risk-loving or risk-averse behavior of its managers. Incorporating these concerns into the objective function 4 means that the manager's objective becomes

$$U\left(s\left[R\left(L_1, \cdots, L_M, k\right) - rk - \left(\sum_{m=1}^{M} w_m L_m\right)\right] + w_0\right),$$

where w_0 is the manager's base salary and s is his or her share of profits. If $U(\cdot)$ is a constant absolute risk aversion utility function, variations in the non-random component (the base salary) do not affect decision making about the stochastic component (the profit share).

3 An Econometric Model of Employment with Privately - Motivated Managers and Takeovers

The econometric model estimated in this paper is based on the reduced form implications of the theoretical model presented in section 2. The theoretical model implies separate processes for determining the probability of continued employment with the pre-transaction and post-transaction management teams since the presence of non-zero α terms and/or a non-linear $U(\cdot)$ function lead to different equilibria of the bargaining model.

The determinants of the equilibrium employment levels under each management regime can be empircally modeled using duration models of the probability of continued employment. In particular, we estimate piecewise constant proportional hazards duration models with time varying covariates, different parameter vectors across management regimes (pre- and post-M&A) and differing distributions of unobserved

heterogeneity (modeled as 2 mass point discrete distributions in the style of Heckman and Singer (1984)). The probability that a firm is about to participate in a merger or acquisition at any point in time is also modeled (as a simple probit) and this estimated probability is used to weight individual spells in the overall likelihood function. Appendix A provides the expression for the full likelihood function; below, while the rest of this section discusses particularities of its construction.

3.1 The Duration of Employment Models

Because our data cover a relatively short period of time, [7], stock sampling among the pre-transaction spells is a particularly important issue. Although some pre-transaction employment spells are flow-sampled in that they began after 1991 and before the year of the transaction, roughly 43 percent of the observed pre-transaction employment spells were already in progress in 1991. Stock sampling is not an issue for the post-transaction spells however since we are interested in the determinants of the length of the employment spell with the new management team for those individuals employed by the acquired firm at the time of the transaction. [8] These individuals all have (by construction) zero accrued seniority *with the new management team* at the time of the takeover, and thus are treated as if there was a massive inflow of workers at the transaction date with the new team. That said, since the seniority accumulated prior to the transaction likely provided these individuals with specific human capital, previous seniority should be a determinant of the duration of the spell with the new management team.

3.2 Mass Points

The object being studied in this paper is the duration of an employment spell. As such, the length of the spell depends on both employer and employee willingness to continue the employment relation. Since either side can terminate the spell of employment, [9] the unobserved heterogeneity is best considered to be match-specific, rather than specific to the individual or to the firm. One could explicitly model the matching of individuals to firms based on expectations of individual turnover, [10] but an advantage of the Heckman and Singer (1984) approach is precisely that it does not require a complicated structural

7. We only consider transactions taking place between 1993 and 1999, although we observe employment from 1991 to 1999.

8. The determinants of continued employment for individuals hired after the transaction is likely to reflect the new management team's strategy, and should thus be different from the determinants of the decision to keep a pre-existing worker who may not fit in the new management's plans - or the determinants of frustration that pre-existing workers might have with any changes in corporate culture introduced by new management that might induce them to quit.

9. This paper exploits data from France, where the legal framework governing employment relations is contractual (rather than "employment-at-will" as in the United States), sometimes specifying a particular contract length and imposing penalties on whichever side breaks the relation at a point in time other than that specifically foreseen in the contract. Nevertheless, quits and layoffs are possible (albeit likely more costly than in North America, see Bender et al. (2002) for details of French legislation governing layoffs) and thus one can transpose the "employment-at-will" reasoning to the French case for the sake of this issue.

10. See Margolis (1995) for such a model.

model to justify a distribution of heterogeneity, instead it simply fits, in reduced form, what amounts to the equilibrium allocation of worker types to firm types.

An additional point worth noting is that, if the M&A was indeed undertaken to improve the management team and the previous management team was adopting sub-optimal human resource management practices, there is no reason to assume that a low mass point individual (disproportionately long expected tenure) in the pre-transaction firm should also be a low mass point individual after the transaction, since the other partner to the match has changed and may value different unobserved characteristics than the previous management. Likewise, there is no reason to assume that the *values* of the mass points should be the same pre and post transaction. It may be the case that the previous management "played favorites" to a much higher extent than the incoming management team, and thus the same individual might be less protected against layoff, even if he or she is appreciated by is or her supervisors, in the world after the transaction than in the world prior to it.

3.3 The Model of Mergers and Acquisitions

This paper focuses on the employment durations of individuals employed by firms that are involved in M&A activity. As such, the sample is already selected (see section 4 below) to include only those spells related to firms with (at most) a single M&A event during the sample period. The spells are categorized as either pre-transaction or post-transaction spells, and the characteristics concerning each spell are allowed to time vary at annual intervals.

Since our data do not provide information on the identities of individual members of the management team, we are unable to tell precisely if the management team in place at the transaction date was already present at any point in time prior to the transaction date. As a result it is entirely possible that some of the pre-transaction spells where carried out under a management team that was actually maximizing shareholder value, but that the team changed (after the individual left) and the firm was eventually acquired. If this is the case, it would be inappropriate to consider the determinants of continued employment for this pre-transaction spell as if they were the same as the determinants of continued employment under the self-serving management team.

In order to reduce the importance of such a problem, we estimate in a preliminary step [11] a binary probit model of an M&A transaction in the following year. [12] The estimated coefficients of this binary model allow one to calculate the probability that any given pre-transaction observation, corresponds to a firm whose characteristics and employment structure resemble those of a firm that is about to be involved in a transaction. The higher the estimated probability resulting from this model, the more likely that the management team in place was behaving in the same manner as management teams that were soon to be involved in M&A activity. This estimated probability is used to

11. This probit is estimated in a preliminary step instead of simultaneously with the rest of the model for practical reasons; time to convergence was extremely long when attempting to estimate all of the equations simultaneously.

12. Since we do not know the creation date of the firm, this is equivalent to assuming that, for a given vector of observable covariates, the probabilty of takeover is constant over the life of the firm. The role of the management team is to change the values of the observed covariates.

calculate the expected hazard rate at each date, with the hazard function associated with a "target" management team receiving a weight equal to the probabilty of takeover in the following period and the probability of a "buyer" management team receiving a weight equal to one minus this probability.

4 The Data

For a detailed description of the data used in this paper, see appendix B. This section briefly describes the additional sample restrictions imposed on the data and details how the spell data was constructed and classified as pre- or post-transaction.

One point worth noting is that we examine all firms with at most one M&A transaction in our sample window (1993-1999). The data that we use to identify M&A events (the MDST file from SUSE 3, see appendix B) identify asset transfers between firms and classify each firm according to whether it received or transferred away assets. In the case of an acquisition, it may be optimal for tax or other reasons for the acquiring firm to transfer all of its assets to the target firm or to not consolidate the accounts of the two firms, [13] instead of recovering all of the target firms assets in its own accounts as might be expected. For this reason we do not exploit the information about the direction of the asset transfer that is available in the data and thus examine all firms involved in the M&A event. This implies that many pre-transaction spells are actually associated with acquiring firms, but the reweighting of these spells by the expected probability of an M&A event is intended to reduce the importance of this issue.

Since the goal of the paper is to examine employment durations before and after an M&A, the main selection criterion imposed was the elimination of all employment spells associated with firms that were involved in multiple M&A events over the time frame, for two reasons. First, since an acquired firm ceases to exist as an independent entity after the first acquisition (even though the firm code may remain alive if the acquiring firm does not consolidate both firms' accounts), at least part of the firm in question is likely to be an acquirer for at least one of the transactions. Since the analysis aims at studying employment durations in target firms before and after an merger or acquisition, we chose to eliminate these firms. The second reason is more prosaic; an employment spell that is considered post-transaction for the first M&A would be considered pre-transaction for the second event. This would result in double counting of the same episode and would lend too much weight to firms with multiple transactions.

Concerning transaction dates, the MDST data used to identify mergers and acquisitions (see appendix B) provide several dates for each transaction, including an "effective" date, a record year and a "decision" date. [14] Although these three variables are usually consistent with respect to the transaction year, missing data for at least one

13. (Acquiring firms whose ownership share exceed 50% are not required by law to consolidate the accounts of the acquired firm with their own provided the stake remains below a threshold. The obligation to make an offer for all remaining shares only applies once the stake in one firm held by another entity exceeds this threshold.

14. In constructing the transaction year, we used the year reported in the "effective date". If that was missing, we used the record year. Only if that variable was also missing did we use the year recorded in the "decision date" variable.

of the variables is not infrequent and when all variables are available and provide information on the year, month and day of the transaction (which is not sytematically the case), they often record different months and days within the year. To minimize the risk of measurement error we considered pre-transaction spells to be censored on January 1 of the transaction year and we required individuals to be present in the firm on December 31 of the year preceding the transaction year in order to be considered post-transaction spells.

As noted above, in order for a spell to be considered a post-transaction spell, the individual had to be present in the firm on December 31 of the year before the transaction year. The spell was considered to have begun at that that date and to have lasted for as long as it was visible in the data, possibly being right-censored by the end of the observation window. Since we do not exploit the information on the day within the year that the transaction occurred, it is thus possible that some spells that are classified as post-transaction actually finished before the transaction took place. [15] This possibility might be underlying the presence of the first mode in the non-parametric hazard graphs (see figure 2), as it might reflect either individuals quitting the firm in anticipation of (or just after) the transaction or the new management proceeding with large, immediate layoffs. If this is the case, one would expect a relatively large distance between the two mass points in the unobserved heterogeneity distribution, since a very high initial mass point would capture this advance departure and the second mass point would reflect the standard turnover behavior of individuals in the firm post-transaction.

5 The Estimation Results

Table 1 presents the results of estimating the model described by equation 15, while table 2 presents the results from the model identifying post-transaction spells that was estimated prior to estimating the duration models. [16] It is worth recalling that the coefficients in the top part of the table are elements of the proportionality factor that multiplies the baseline hazard, and thus a positive coefficient increases the baseline hazard and decreases the expected duration, ceteris paribus.

The main things to note from table 2 are that the exogenous instrument (the indicator variable for whether or not the number of firms in the 4-digit sector decreased for 3 consecutive years) is significant and that the model coefficients as a whole are jointly significant as well. The chances of an episode being classified as post-transaction increase with the share of equity in the balance sheet, decrease with the return on assets and value added per worker, and increase with firm size, be it measured by employment, assets per worker or sales per worker. On the other hand, relatively few individual characteristics are associated with the likelihood of a spell being classified post-transaction.

15. A worst-case scenario would have spells of up to 2 years minus 1 day being erroneously classified as post-transaction spells. This would occur if an individual was employed on January 1st of the year preceding the transaction and left the firm on December 30 of the transaction year, with the transaction occurring on December 31st.

16. It should be noted that the standard errors in table 1 have not been adjusted to account for the fact that the index used in the estimation of this model (see equations 11, 12 and 15) was itself the result of a prior estimation.

Table 1: Estimation results : Duration Models

Variable	**Pre-Transaction**		**Post-Transaction**	
	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
Pre-Transaction Job Seniority			0.002	(0.002)
=1 if Same 4-Digit Sector			0.317	(0.034)
$\frac{\text{Fixed Assets}_t - \text{Fixed Assets}_{t-1}}{\text{Fixed Assets}_{t-1}}$	0.000	(0.001)	0.000	(0.009)
$\text{Log}\left(\frac{\text{Total Debt}}{\text{Total Assets}}\right)$	-0.210	(0.024)	0.323	(0.045)
Log(Return on Assets)	-0.025	(0.019)	-0.011	(0.012)
$\text{Log}\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.010	(0.029)	-0.187	(0.055)
Log(Employment)	0.020	(0.008)		
Log($\sum$ Employment, All Firms)			0.113	(0.011)
$\text{Log}\left(\frac{\text{Total Fixed Assets}}{\text{Worker}}\right)$	-0.012	(0.008)	-0.134	(0.013)
$\text{Log}\left(\frac{\text{Sales}}{\text{Worker}}\right)$	-0.073	(0.018)	0.224	(0.024)
Paris Region	0.141	(0.035)	0.112	(0.039)
Skilled Blue Collar	-0.076	(0.036)	-0.099	(0.039)
Unskilled Blue Collar	0.228	(0.036)	0.062	(0.044)
Male	-0.155	(0.069)	0.059	(0.134)
Age	-0.114	(0.000)	-0.121	(0.062)
$\frac{\text{Age}^2}{100}$	0.099	(0.000)	0.088	(0.168)
$\frac{\text{Age}^3}{1000}$	0.003	(0.000)	0.012	(0.015)
$\frac{\text{Age}^4}{10000}$	0.000	(0.000)	0.000	(0.000)
Individual Fixed Effect	0.001	(0.000)	0.000	(0.000)
Firm-Specific Starting Earnings	-0.272	(0.015)	-0.309	(0.053)
Firm-Specific Return to Seniority	-0.217	(0.000)	0.510	(0.402)
Match-Specific Earnings Residual	0.032	(0.000)	-0.129	(0.034)
Intercept	-1.252	(0.000)	-0.483	(0.000)
Log(Weibull Shape Parameter)	-0.398	(0.009)	-0.039	(0.012)
P(High Type): Constant	-0.500	(0.043)	-1.500	(0.209)
P(High Type): Slope	-0.500	(0.060)	-0.500	(0.255)
Value of High Mass Point	1.750	(0.000)	1.000	(0.000)

Source: MDST, FUTE, DADS, EDP and author's calculations.

Notes: Table presents coefficients from scaling factor in proportional hazard models as well as the other coefficients from equation 15. The pre - transaction model also includes controls for 7 5-year hiring intervals and the post - transaction model includes controls for 8 merger years. Both models also control for 8 educational categories and 10 1-digit sectors.

Table 2: Estimation results : Probit of Merger Status

Variable	**Coefficient**	**(Std. Err.)**
=1 if No. of Firms in 4-Digit Sector Declined last 3 Years	-0.190	(0.029)
=1 if Return on Assets Declined last 3 Years	-0.020	(0.031)
$\frac{3 \cdot \text{Return on Assets}_t}{\sum_{\tau=t-3}^{t-1} \text{Fixed Assets}_\tau}$	0.000	(0.000)
$\frac{\text{Fixed Assets}_t - \text{Fixed Assets}_{t-1}}{\text{Fixed Assets}_{t-1}}$	0.004	(0.002)
$\text{Log}\left(\frac{\text{Total Debt}}{\text{Total Assets}}\right)$	-0.245	(0.013)
Log(Return on Assets)	-0.094	(0.005)
$\text{Log}\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.096	(0.022)
Log(Employment)	0.025	(0.004)
$\text{Log}\left(\frac{\text{Total Fixed Assets}}{\text{Worker}}\right)$	0.026	(0.006)
$\text{Log}\left(\frac{\text{Sales}}{\text{Worker}}\right)$	0.068	(0.011)
Paris Region	-0.075	(0.017)
Skilled Blue Collar	0.014	(0.018)
Unskilled Blue Collar	-0.169	(0.018)
Male	0.111	(0.071)
Age	0.160	(0.005)
$\frac{\text{Age}^2}{100}$	-0.184	(0.007)
$\frac{\text{Age}^3}{1000}$	0.002	(0.000)
$\frac{\text{Age}^4}{10000}$	0.000	(0.000)
Individual Fixed Effect	0.000	(0.000)
Firm-Specific Starting Earnings	0.277	(0.016)
Firm-Specific Return to Seniority	-0.164	(0.110)
Match-Specific Earnings Residual	-0.020	(0.011)
Intercept	-3.518	(0.263)

Source: MDST, FUTE, DADS, EDP and author's calculations.

Notes: Table presents coefficients from the latent index in the probit model for identifying the year before a merger. The model also controls for 5 transaction year groups for 8 educational categories and 10 1-digit sectors.

It appears that the this probability is concave in the age of workers in the firm, decreasing in the share of unskilled blue collar workers (relative to white collar workers) and increasing in the share of men. One can find reasonable economic justifications for all of these coefficients, but since they are not the focus of this paper we will not dwell on them here.

One can note a large number of coefficients that appear to affect employment duration differently in the pre- and post-transaction period, and the mass points of the unobserved heterogeneity distributions seem quite different as well, when one considers the duration models in table 1. Table 3 presents a series of hypothesis tests comparing the estimated coefficients in the two sub-models and tests, in particular, the hypothesis that all common coefficients are the same across sub-models. This hypothesis is clearly rejected by the data. Furthermore, table 3 shows that the baseline hazard distributions are significantly different from each other, and as expected the Weibull shape parameter is significantly lower for pre-transaction spells than for post-transaction spells, suggesting a more steeply declining hazard function and thus longer expected durations for those individuals with the firm pre-transaction.

It may however be the the case that this is counterbalanced by the much larger distance between the support points of the unobserved heterogeneity distribution in the pre-transaction population and the significantly higher constant in the function that determines the unobserved type of the individual, both of which suggest (for equivalent values of the index function derived from estimating the model described in section 3.3) that there are more fast-hazarding people in the pre-transaction population than in the post-transaction population. Given the way in which the spells are classified as pre- or post-transaction (see section 4,) it may be the case that a larger share of pre-transaction spells are comprised of short, fixed-term employment contracts. Since type of contract is unobservable in our data, this may explain the differences in the shares of fast and slow separating workers in the pre- and post-transaction periods.

6 Interpreting the Results Within the Context of Corporate Finance Theory

One can interpret the results of table 1, and in particular the hypothesis tests in table 3, in the context of the different corporate finance explanations in section 2.2. We treat each explanation in turn below.

6.1 Empire Building

Recall that, in this case, managers derive private utility from being in charge of a large firm. This implies that the coefficients on firm size in the probit model of merger status should be positive (first on the edge of being taken over should be larger on average) and that the probability of exit from a large firm (measured by employment) should be higher during post-M&A spells than in pre-M&A spells (since the extra people were not necessary for production). Looking at the results in tables 1 - 3, we see that all of these implications hold in the data.

Table 3: Hypothesis Tests

Coefficient / Hypothesis	**Test Statistic**	χ^2 **DF**	**P-Value**
Identity of Common Coefficients			
$\frac{\text{Fixed Assets}_t - \text{Fixed Assets}_{t-1}}{\text{Fixed Assets}_{t-1}}$	0.00	1	1.0000
$\text{Log}\left(\frac{\text{Total Debt}}{\text{Total Assets}}\right)$	106.99	1	0.0000
Log(Return on Assets)	0.41	1	1.0000
$\text{Log}\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	8.04	1	0.1510
$\text{Log}\left(\frac{\text{Total Fixed Assets}}{\text{Worker}}\right)$	63.90	1	0.0000
$\text{Log}\left(\frac{\text{Sales}}{\text{Worker}}\right)$	102.60	1	0.0000
Paris Region	0.29	1	1.0000
Skilled Blue Collar	0.19	1	1.0000
Unskilled Blue Collar	8.68	1	0.1061
Male	2.01	1	1.0000
Age	0.01	1	1.0000
$\frac{\text{Age}^2}{100}$	0.00	1	1.0000
$\frac{\text{Age}^3}{1000}$	0.42	1	1.0000
$\frac{\text{Age}^4}{10000}$	0.47	1	1.0000
Individual Fixed Effect	25.4	1	0.0000
Firm-Specific Starting Earnings	0.47	1	1.0000
Firm-Specific Return to Seniority	3.26	1	1.0000
Match-Specific Earnings Residual	22.67	1	0.0001
All Common Coefficients Jointly	1041.92	32	0.0000
Weibull Shape Parameters			
$\theta^B = \theta^A$	459.24	1	0.0000
$\theta^B = 0$	2792.78	1	0.0000
$\theta^A = 0$	10.38	1	0.0013
Unobserved Heterogeneity Parameters			
$\alpha_0^B = \alpha_0^A$	22.00	1	0.0000
$\alpha_1^B = \alpha_1^A$	0.00	1	1.0000

6.2 A Lot of Cash on Hand

The "Cash on Hand" motive suggests that managers keep too much cash on hand as opposed to investing it or distributing it to shareholders through dividend payments. As one can see in table 2, as the ratio of debt to assets falls, implying that as the net value of the firm increases, the probability of being taken over the following year increases. Although cash on hand is a component of the net value of the firm, it is far from the only determinant and thus although the results in table 2 are not inconsistent with this theory, they do not represent strong support either.

6.3 Peaceful Labor Relations

The means suggested in section 2.2 for accommodating the "Labor Peace" motive was simply to allow the managers to act as if the union's bargaining power was higher than it actually is. This results in higher wages and more employment under the pre-M&A management than under the post-M&A management. Accordingly, one should see higher hazard rates out of employment for individuals employed in firms that were bigger and paying more in the pre-transaction period. Although, as noted above, the results in table 1 show higher hazard rates from ex-ante bigger firms, according to table 3 there is no significant change in hazard rates based on the firm's specific compensation policy, be it through starting wages or through returns to seniority. Thus this explanation, at least as we suggest it be modeled, meets with only limited success.

6.4 Esteem from the Workforce / Playing Favorites

In this context, management pays some workers above what they would otherwise earn. This should be reflected in a positive coefficient on the match specific effect in the table 2 results. Furthermore, these higher wages can induce substitution away from the privileged factor and, as the workers laid off as a result of this substitution are no longer insiders, their layoffs (or slower replacement hiring) do not negatively affect the esteem of the in-place workforce for management. The empirical implication is that exit rates should change differently for different types of workers. Unfortunately, none of these explanation finds significant support in the data, as higher match-specific residual workers are less likely to be found in firms that are about to be part of an M&A event and there are no significant changes in exit rates for either skilled or unskilled blue collar workers, relative to white collar workers, when comparing pre- and post-transaction spells.

6.5 Kickbacks from Suppliers

When kickbacks are paid to suppliers, the costs of the goods they sell increases relative to other factors, inducing a substitution away from capital and towards labor. In terms of empirical implications, kickbacks would imply a faster exit rate post-transaction for workers employed by firms with a relatively low capital-labor ratio. This is precisely what is suggested by table 1, and the hypothesis test in table 3 shows that this difference is highly significant.

6.5.1 Non-Risk Neutrality

Without specifying further the sort of non risk-neutrality that the manager employs, it is difficult to draw empirically testable conclusions. However, our model of the probability of having an M&A event the following period is estimated as a function of the returns on assets in t relative to the average return on assets in the three previous years and an indicator variable for whether return on assets has been falling for three consecutive years. Both of these variables are measures that might inspire the manager to undertake excessively risky, or excessively risk-averse behavior. The results in table 2, however, suggest that neither measure has a statistically significant relation to the chance of being involved in a transaction the following period.

7 Conclusion

In this paper, we have investigated the hypothesis that managerial actions are at the origin of mergers and acquisitions. We began by by presenting a theoretical framework drawn from the labor economics literature in which many different private managerial motives can be incorporated. We then presented an econometric framework that allowed us to investigate the empirical implications of each of these modifications for the probability of continued employment of individuals with the firm after the transaction takes place relative to the determinants of continued employment with the previous management team.

Our empirical results suggest that there are significant differences in the determinants of continued employment under the new and old management teams. The distributions of unobserved heterogeneity under each management team is significantly different, suggesting different human resource management strategies for each management team.

When considering particular private managerial motives, the "empire building" and "kickbacks" motives seem to find the strongest support in the data. The motivations based on employment ("labor peace" and "worker esteem / playing favorites") find only lukewarm support, as does the "excessive cash on hand" motive, for which our data cannot provide a clean test. Non risk-neutrality, on the other hand, finds no significant support in our data.

In conclusion, the literature cited by Tirole (2006) concentrates on the the importance of managerial behavior in determining whether an M&A will take place, and where. If so, the change in the management team should bring about a change in managerial behavior. Such a change is indeed visible our data covering mergers and acquisitions in France, and allows us to identify which corporate finance explanations of private managerial incentives ("empire building" and "kickbacks") seem the most credible.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, March 1999, *67* (2), 251–333.

Baker, Michael and Angelo Melino, "Duration dependence and nonparametric heterogeneity: A Monte Carlo study," *Journal of Econometrics*, June 2000, *96* (2), 357–393.

Beaudry, Paul and John DiNardo, "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Micro Data," *Journal of Political Economy*, August 1991, *99* (4), 665–688.

Bender, Stefan, Christian Dustmann, David Margolis, and Costas Meghir, "Worker Displacement in France and Germany," in Peter Kuhn and Randall Eberts, eds., *Losing Work, Moving On: International Comparisons of Worker Displacement*, Kalamazoo, Michigan: The Upjohn Institute, 2002.

Bertrand, Marianne and Sendhil Mullainathan, "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences," *Journal of Political Economy*, October 2003, *111* (5), 1043–1075.

Gibbons, Robert and Lawrence F. Katz, "Layoffs and Lemons," *Journal of Labor Economics*, October 1991, *9* (4), 351–380.

Grout, Paul A., "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach," *Econometrica*, March 1984, *52* (2), 449–460.

Heckman, James J. and Burton Singer, "A method for minimising the impact of distributional assumptions in econometric models for duration data," *Econometrica*, March 1984, *52* (2), 271–320.

Hosken, Daniel S. and David N. Margolis, "The Efficiency of Collective Bargaiing in Public Schools," Working Paper 9755, CREST December 1997.

INSEE, *Guide d'utilisation des fichiers SUSE 3*, Paris: INSEE, 1995.

Jensen, Michael and William Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, October 1976, *3* (4), 305–360.

Lindbeck, Assar and Jürgen W. Weibull, "Balanced-Budget Redistribution as the Outcome of Political Competition," *Public Choice*, 1987, *52* (3), 273–297.

Manning, Alan, "An Integration of Trade Union Models in a Sequential Bargaining Framework," *Economic Journal*, 1987, pp. 121–139.

_ , "How Robust Is the Microeconomic Theory of the Trade Union?," *Journal of Labor Economics*, July 1994, *12* (3), 430–59.

Margolis, David N., "Firm Heterogeneity and Workers Self-Selection Bias Estimated Returns to Seniority," Working Paper 9502, Université de Montréal Département de Sciences Economiques January 1995.

_ , "Should Employment Authorities Worry About Mergers and Acquisitions?," *Portuguese Economic Journal*, August 2006, *5* (2), 167–194.

Tirole, Jean, *The Theory of Corporate Finance*, Princeton, NJ: Princeton University Press, 2006.

A The Full Likelihood Function

A.1 The Duration of Employment Models

Consider an individual i with job seniority t_i and observable characteristics $X_{i,t}$ who began his or her spell of employment at date T_0.[17] Adopting the notation of the duration econometrics literature, one can write the hazard function in the absence of unobserved heterogeneity as $h\left(t_i, X_{i,t}|\theta,\beta\right)$, the survivor function as $S\left(t_i, X_{i,t}|\theta,\beta\right)$ and the density function as $f\left(t_i, X_{i,t}|\theta,\beta\right)$ with the relations

$$f\left(t_i, X_{i,t}|\theta,\beta\right) = h\left(t_i, X_{i,t}|\theta,\beta\right) S\left(t_i, X_{i,t}|\theta,\beta\right)$$

and

$$S\left(t_i, X_{i,t}|\theta,\beta\right) = exp\left(-\int_0^{t_i} h\left(s, X_{i,t}|\theta,\beta\right) ds\right).$$

Assuming a proportional hazards specification, we introduce unobserved heterogeneity into the model in the same way as Heckman and Singer (1984). That is, we suppose each individual is of a particular (unobserved) type with a given probability, and that the type of the individual affects the hazard function in the same manner as the observables, namely multiplicatively with respect to the baseline hazard. One can rewrite the hazard function as

$$h\left(t_i, X_{i,t}|\theta,\beta\right) = exp\left(X_{i,t}\beta + \mu_k\right) h_0\left(t_i|\theta\right) \tag{7}$$

with the μ_k being the value of the mass point (at least one of which must be normalized to zero if there is a constant in $X\beta$). When deriving the likelihood function one takes expectations over the various possible contributions; one does not know a priori the value of (the non-normalized) μ_k nor which μ_k is relevant so one estimates the probability that individual i is of type k as p_k.[18]

Given the concerns expressed by Baker and Melino (2000) about identification of the mass points and coefficients when working within a semiparametric context, as well as the stability of the estimates of the mass points and their probabilities as their number increases, we restrict the general specification presented in equation 7. First, as is often done in the literature when using the Heckman and Singer (1984) approach, we limit the number of mass points to two. Second, we adopt a parametric (but relatively flexible) baseline hazard, choosing a piecewise constant distribution with pieces at $0 < t_i \leq 1$, $1 < t_i \leq 2$, $2 < t_i \leq 3$, $3 < t_i \leq 4$, $4 < t_i \leq 5$, $5 < t_i \leq 10$ and $t_i > 10$. This functional form assumption for the baseline hazard leads to a very simple expressions for the hazard function

$$h\left(t_i, X_{i,t_i}|\theta,\beta,\mu_k\right) = \exp\left(X_{i,t_i}\beta + \mu_k + \sum_{j=1}^{7} \theta_j \mathbf{1}_{\tilde{t}_{j-1}<t_i\leq\tilde{t}_j}\right), \tag{8}$$

17. The time varying covariates in $X_{i,t}$ are measured at annual intervals.

18. In principle one could make p_k a function of observables, $p_k = p\left(Z\zeta_k\right)$, but semiparametric identification of β and ζ_k require the set of observables Z not be linear combinations of X. As it is impossible to define theoretically justifiable exclusion restrictions (any observable that affects the probability of being a particular unobserved type should directly affect the probability of continued employment in the reduced form), one generally models these probabilities as constants.

with $\tilde{t}_0, \cdots, \tilde{t}_7 = \{0, 1, 2, 3, 4, 5, 10, \infty\}$, and for the survivor function: [19]

$$S\left(t_i, X_{i,1}, \cdots, X_{i,t_i} | \theta, \beta, \mu_k\right) = e^{\left(-\left(e^{\left(X_{i,t_i}\beta + \mu_k + \sum_{j=1}^{7} \theta_j (t_i - \underline{t}_i) \mathbf{1}_{\tilde{t}_{j-1} < t_i \leq \tilde{t}_j}\right)} + \sum_{\tau=1}^{\underline{t}_i} e^{\left(X_{i,\tau}\beta + \mu_k + \sum_{j=1}^{7} \theta_j \mathbf{1}_{\tilde{t}_{j-1} < \tau \leq \tilde{t}_j}\right)}\right)\right)} \tag{9}$$

where $\underline{t}_i$ corresponds to t_i rounded down to the nearest integer. Note that these expressions incorporate the fact that the time-varying $X_{i,t}$ variables change at annual intervals and are measured at the end of the year, with the exception of the final part-year in which the individual is employed. Thus the sum in the survivor function is from 0 to $\underline{t}_i$ for the full-year blocks and the integrated baseline hazard part of last part-year block is multiplied by $t_i - \underline{t}_i$

Exploiting equation 9, accommodating stock sampling corresponds to integrating the hazard from the stock sampled seniority to the final seniority. We thus rewrite equation 9 as

$$S\left(t_i, X_{i,t_i^S}, \cdots, X_{i,t_i} | \theta, \beta, \mu_k\right) = e^{\left(-\left(e^{\left(X_{i,t_i}\beta + \mu_k + \sum_{j=1}^{7} \theta_j (t_i - \underline{t}_i) \mathbf{1}_{\tilde{t}_{j-1} < t_i \leq \tilde{t}_j}\right)} + \sum_{\tau=t_i^S}^{\underline{t}_i} e^{\left(X_{i,\tau}\beta + \mu_k + \sum_{j=1}^{7} \theta_j \mathbf{1}_{\tilde{t}_{j-1} < \tau \leq \tilde{t}_j}\right)}\right)\right)} \tag{10}$$

where t_i^S is the level of seniority the individual had when first sampled.

A.2 Why Two Mass Points?

In this paper, we adopt a two mass point specification for two main reasons. First, some papers have attempted to estimate models with three mass points but they tend to converge poorly on real world data. To the best of our knowledge, no paper exploiting Heckman and Singer (1984) style heterogeneity with four mass points or more has ever been published.

Second, as can be seen in figures 1 and (especially) 2, a graph of the nonparametric hazard for post-transaction durations suggests two distinct modes. Such a situation can be accommodated with two mass points, whereas three mass points would appear to overfit the raw data. Given the convergence issues associated with 3 points in general, such problems seem likely to be aggravated when the data only "wants" two mass points whereas one tries to estimate a third (and its probability).

A.3 The Full Likelihood Function

The main problem in ourt data is the fact that we can not observe the management team in place at any point in time. We thus use the M&A model to estimate the probability

19. Since our observables change at annual intervals and the baseline hazard is defined with blocks that end at round years, the integrated hazard is the the sum of the annual blocks of the hazard for all years up to the last plus the fractional year for the final observation if the job spell does not year an exact year multiple after the start date.

that the management team is of a "target" type versus a "buyer" type, and we use the expected hazard rate when modeling the employment duration. Note that since the probability of having a "target" management team can change over time, the weight put on each type of management team in the hazard function can change with the "target" management team's contribution to the expected hazard function being the highest in the year preceding the takeover (by construction).

and will be higcomplete likelihood function integrates all the duration models described above. For each employment episode j out of a total J_i episodes for individual i in the data, there are three possible configurations observable in the data.

1. The individual was present before the transaction and left the firm before the transaction occurred, or alternatively the individual was employed by a firm for which no transaction was observed. In this case, the contribution to the likelihood function of the corresponding employment spell is weighted by the probability that the following year is a takeover year (from the M&A model) and we can use equations 8 and 10 to write

$$\begin{aligned} L_{i,j}^{Target}\left(\mu_k^{Target}\right) = & \; L\left(t_i, t_i^S, c_i, X_{i,t_i^S}^{Target}, \cdots, X_{i,t_i}^{Target}\right) \\ = & \; \frac{P(\text{M\&A})}{P(\text{M\&A})} h\left(t_i, X_{i,t_i}^{Target} | \theta^{Target}, \beta^{Target}, \mu_k^{Target}\right)^{1-c_i} \\ & S\left(t_i, X_{i,t_i^S}^{Target}, \cdots, X_{i,t_i}^{Target} | \theta^{Target}, \beta^{Target}, \mu_k^{Target}\right) \end{aligned} \tag{11}$$

where $c_i = 1$ if the spell is right-censored, $X_{i,t}^B$ is the set of covariates that determine the duration of employment with the pre-transaction management team and $t_i^S = 1$ if the spell is not stock sampled.

2. The individual was hired by the firm in the year of the transaction. Since we do not use transaction year data because of measurement problems, [20] the contribution of the employment spell to the likelihood function will only have a post-transaction component, i.e.

$$\begin{aligned} L_{i,j}^{Buyer}\left(\mu_k^{Buyer}\right) = & \; L\left(t_i, c_i, X_{i,1}^{Buyer}, \cdots, X_{i,t_i}^{Buyer}\right) \\ = & \; h\left(t_i, X_{i,t_i}^{Buyer} | \theta^{Buyer}, \beta^{Buyer}, \mu_k^{Buyer}\right)^{1-c_i} \\ & S\left(t_i, X_{i,1}^{Buyer}, \cdots, X_{i,t_i}^{Buyer} | \theta^{Buyer}, \beta^{Buyer}, \mu_k^{Buyer}\right) \end{aligned} \tag{12}$$

where X_i^A is the set of covariates for the post-transaction model.

3. The individual was hired before the transaction occured and did not leave before the transaction. In this case the employment spell contributes a pre-transaction

20. The transaction date is coded in the data in an inconsistent manner. For example, some transactions will have $date = 22031995$, others might have $date = 395$ and others might still have $date = 1995$. As it is always possible to extract the transaction year, this is the most detailed information used and accordingly employment during the transaction year cannot be credibly divided betwen pre- and post-transaction management teams. For consistency we consider all employment spells in progress during the transaction year to have employment with the post-transaction management team starting on January 1 of that year.

and a post-transaction component to the likelihood:

$$L_{i,j}^{Both}\left(\mu_k^{Target},\mu_k^{Buyer}\right) = \begin{array}{l} L\left(t_i^{Target},t_i^S,1,X_{i,t_i^S}^{Target},\cdots,X_{i,t_i^{Target}}^{Target}\right) \\ L\left(t_i^{Buyer},c_i,X_{i,1}^{Buyer},\cdots,X_{i,t_i^{Buyer}}^{Buyer}\right) \end{array} \tag{13}$$

It is possible that an individual in our data has several episodes of employment in the data, and thus the individual's contribution to the likelihood is the product of the likelihood contributions for each episode. Letting $d_{i,j}^1 = 1$ if the episode contains a pre-transaction spell and $d_{i,j}^2 = 1$ if the episode contains a post-transaction spell, the contibution to the overall likelihood function for an andividual with unobserved individual heterogeneity $\left(\mu_k^{Target},\mu_k^{Buyer}\right)$ can be written using equations 11, 12 and 13 as

$$L_i\left(\mu_k^{Target},\mu_k^{Buyer}\right) = \Pi_{j=1}^{J_i} \begin{array}{l} \left\{ \frac{w_{i,j}}{\overline{w}} L_{i,j}^{Target}\left(\mu_k^{Target}\right)^{d_{i,j}^1\left(1-d_{i,j}^2\right)} \right. \\ L_{i,j}^{Buyer}\left(\mu_k^{Buyer}\right)^{\left(1-d_{i,j}^1\right)d_{i,j}^2} \\ \left. L_{i,j}^{Both}\left(\mu_k^{Target},\mu_k^{Buyer}\right)^{d_{i,j}^1 d_{i,j}^2} \right\} \end{array} \tag{14}$$

Equation 14 is conditional on the unobserved type of the individual. The overall likelihood function takes the log of the expected value of the overall likelihood function with respect to the different types.

$$\mathcal{L} = \sum_i \log\{ \begin{array}{l} P\left(\mu_L^{Target},\mu_L^{Buyer}\right) L_i\left(\mu_L^{Target},\mu_L^{Buyer}\right) \\ +P\left(\mu_L^{Target},\mu_H^{Buyer}\right) L_i\left(\mu_L^{Target},\mu_H^{Buyer}\right) \\ +P\left(\mu_H^{Target},\mu_L^{Buyer}\right) L_i\left(\mu_H^{Target},\mu_L^{Buyer}\right) \\ +\ P\left(\mu_H^{Target},\mu_H^{Buyer}\right) L_i\left(\mu_H^{Target},\mu_H^{Buyer}\right) \} . \end{array} \tag{15}$$

A.4 The Explanatory Variables for Each Model

Since the three main sub-models estimated here refer to entirely different circumstances (pre-transaction employment, post-transaction employment and the probability of being involved in an M&A the following year), the sets of covariates used to explain each model is different. We describe below the different sets of covariates, each of which naturally leads itself to a set of exclusion restrictions with respect to the other sub-models. Each set of covariates includes both firm-level and individual-level covariates, reflecting the fact that an employment relation is a match and either side can decide to break it for reasons related to the match itself or for other reasons (poor economic conditions, transfer of a spouse, etc...).

A.4.1 The Common Covariates

All of the models share a set of common covariates, reflecting the fact that certain variables can influence the quality of the employer-employee match in any circumstances,

and these variables can also be related to the decisions taken by the management team.

Among the firm-specific covariates that are common to all three models, we include a series of balance sheet and income statement variables. In particular, we include measures of the rate of investment in fixed assets, the (log of the) debt to asset ratio, the log of return on assets, the log of value added per worker, the log of sales per worker, the log of the capital-labor ratio, the sector of activity at the 1-digit level, an indicator variable for the job being in the Paris region, and a measure of the firm's specific starting wage and returns to job seniority. [21]

We also included several individual specific characteristics in all three sub-models. These characteristics were the level of education (in 8 degree categories), the skill level (white collar, skilled blue collar or unskilled blue collar), sex, age through age^4 and a measure of the individual-specific fixed effect in earnings. [22]

Finally, we included the value of the residual from the earnings equation for the first observation in the spell as a common regressor. The residual is net of fixed individual and firm specific observed and unobserved heterogeneity, calculated by the persons-first, firms-first projection technique described in Abowd et al. (1999). The residual from the earnings equation can be interpreted in this context as being a noisy measure (due to the presence of idiosyncratic shocks) of the value of the match between the firm and the employee at that point in time.

A.4.2 Pre-Transaction Employment

The employment spells observed prior to the transaction correspond to spells that may have begun at any point prior to the transaction, including points in the distant past. As was shown by Beaudry and DiNardo (1991), labor market conditions at the start of an employment relation can have a significant impact on later career events. We thus include a set of indicator variables for being hired in a particular time period (prior to 1970, 1970-1974, 1975-1979, 1980-1984, 1985-1989, 1990-1994 and 1995-1999). We also include a variable to account for the size of the firm as this is related to the codetermination powers that employees have with respect to mass layoffs and other human resource management issues.

A.4.3 Post-Transaction Employment

The employment spells of individuals present at the time of the transaction are likely to be affected by general macroeconomic conditions at the time the transaction took place, and thus we include a set of indicator variables for each of the possible transaction years. As mentioned above, job seniority acquired prior to the transaction is also likely to be positively correlated with the specific human capital acquired by an individual, and therefore his or her value to the firm (and subsequent employment probability), so we include job seniority the year before the transaction as an additional variable. Finally, we also control for the total stock of workers present in *all* firms that were

21. These latter two measures are estimated in a preliminary step according to the persons-first, firms-first projection technique described in Abowd et al. (1999).

22. As before, the individual fixed effect (which is *not* a match specific effect) is estimated in a preliminary step according to the persons-first, firms-first projection technique described in Abowd et al. (1999).

involved in the M&A transaction, as this is a measure of the size of the workforce that needs to be integrated after the transaction occurs and larger transactions may be associated with more layoffs.

A.4.4 The M&A Equation

The goal of this equation being to measure variables that are likely to signal an impending M&A, we chose to include (in addition to firm employment and the observation year, as there can be merger waves that carry along firms that otherwise would have fallen on the other side of the margin) a set of variables relating to changes in firm performance and market structure. First, we included a measure of returns on assets in the observation year relative to average return on assets in the previous three years as an indicator for a firm's management performing particularly poorly in a given year. [23] Along similar lines, we also included an indicator variable for whether or not the firm had declining returns on assets for each of the three previous years. Finally, we included an (in principle exogenous) instrument, namely whether the number of distinctly identified firms (as defined by a time-invariant tax code) in the firm's 4-digit industry was declining each of the previous 3 years. This variable is intended to measure a sector-specific concentration wave that could also result in a firm being involved in an M&A. Although this paper focuses on managerial behavior motives for mergers and acquisitions, these last three variables can also be seen as informative about alternative merger motives found in the literature.

B Data Appendix

This data appendix provides details on the individual data sets that are used in the creation of the merged sample that is the basis of the empirical analysis. Each component data set is discussed separately.

B.1 The MDST Data

The first, and most original, data set is called the Modification of Structure, or MDST, file. This file is part of the SUSE 3 system (INSEE (1995)) and although it began in 1986, this paper only exploits data from 1993 to 1999, which includes 12,226 observations involving 17,078 distinct firms.

The objective of the data collectors is to cover all asset transfers of a minimum size, [24] providing the identifier of the firm (or firms) that transferred away the assets and the identifier of the firm (or firms) that received the assets. Asset transfers are categorized according to the point of view of the transferring (cédante, or CD below)

23. Our data do not contain the information necessary to construct standard ROE measures, and given that it is likely that an important share of the firms in our sample are privately held, restricting our attention to quoted firms could introduce a form of selection bias into our estimation results.

24. The threshold levels were 50 million French Francs in 1990, 10 million in 1991 and 8 million from 1992 through 1999, which (using the June 30 exchange rate or the closest date) is roughly equivalent to $8.9M, $1.6M, $1.6M, $1.4M, $1.5M, $1.6M, $1.6M, $1.4M and $1.3M for each year between 1990 and 1998, respectively .

and receiving (bénéficiaire, or BF below) firm and a small number of variables (including the effective date of the transaction) concerning the transaction are included with each record. Appendix Table 1 provides a breakdown of the different types of asset transfers covered by the MDST data. These data are used to date transactions, identify taken over and taking over firms and to know the identifier of the successor firm in any such transaction. In the interest of generating cleanly interpretable results, all firms that were involved in more than one takeover during the 1991-1999 period covered by the data were eliminated from the analyses. [25]

The analysis undertaken here focuses on transactions classified as either mergers or acquisitions, which corresponds to any of the following codes in appendix table 1: 23, 31 or 33 for the transferring enterprise; 24 or 43 for the receiving enterprise. The effective date of the transaction, as opposed to the announcement date, is used for dating purposes as the acquiring firm can only directly influence managerial decisions after the transaction becomes effective. [26] All firms that are listed as transferring firms in transactions defined by one of the above-listed codes are considered to be acquired and all firms for which the firm identifier is listed as the receiving firm are considered to be acquiring firms. One implication of this strategy is that, in a true "merger of equals", the acquiring firm will not have existed prior to the transaction. [27] In all other cases, one of the pre-existing firms' identifiers is maintained after the transaction and this firm is considered to be the acquiring firm. Individuals who were employed by an acquired firm prior to the transaction and by its acquiring firm after the transaction are considered to have stayed with the new entity throughout the takeover. Firms involved in MDST activity that do not involve any of the aforementioned codes are not considered as having been involved in a merger or acquisition, and thus appear in the control (non takeover firms) group. [28]

25. These firms are not treated as BF or CD firms in our analyses, and thus the observations associated with individuals employed by these firms do not enter into our calculations as being in taken over or taking over firms. This exclusion restriction eliminates 13 percent of the firm identifiers in the MDST sample. However, if an individual was employed by a firm whose only MDST activity was to have been taken over by another firm, and the acquiring firm had undertaken several MDST transactions, the post-takeover data is retained for the analyses.

26. Most stock-market based analyses of takeovers attempt to focus on announcement dates, as the market factors expected future decisions into the price at that point in time, and thus one avoids having a biased estimate of the pre-takeover price. Apart from the fact that this variable is more often missing than the effective date (which is available for all records in the data), it is not relevant for the means by which this paper evaluates the different theoretical models.

27. Note that, in this case, a single transaction will generate several MDST records, one for each pre-merger firm with that firm's identifier listed as the CD firm. The identifier listed as the BF firm in each of these records will be the same, and will correspond to the identifier of the (newly created) firm that springs from the merger. The fact that all of the records correspond to a single transaction can be established by their sharing a common transaction identification code.

28. The case of partial divestitures could potentially be added to the set of firms classed as takeover firms, but the decision was made explicitly not to do so for two reasons. First, some employees of the firm that divests itself of a portion of its activity will continue to be employed after the transaction in the CD firm, which contaminates the set of CD firms used in the analysis in section 5. Second, there is a clear selection issue (in the Gibbons and Katz (1991) sense) that intervenes in this context, as a firm that is about to divest itself of a division make seek to transfer its best workers to other divisions prior to the divestiture. For both of these reasons, attention is thus restricted to full mergers or acquisitions.

B.2 The FUTE Data

The second data source is called the FUTE data, and is also drawn from the SUSE 3 system (INSEE (1995)). The FUTE data set contains all of the information available on a firm's balance sheet, income statement and statement of flow of funds. In addition, it contains additional variables drawn from the Annual Enterprise Survey (EAE), in particular concerning employment, and information drawn from the firm's tax returns. This data set provides the largest number of firm variables in the SUSE 3 system, although it does not sample all firms. [29] The data cover the period 1989-2000, which generates 1,225,700 observations, of which 922,500 come from the 1991-1999 period common with the MDST data. This amounts to roughly 102,500 firms per year. Given the thresholds of the MDST data, all firms not in the "sample" (see appendix footnote 4) will appear in MDST when taken over, although it is possible that some "sample" firms will be excluded in the case of a takeover due to a lack of assets. Conversely, there may be MDST firms that are not covered by the FUTE as they are not part of the "sample" and are too small to be non-"sample" firms.

The FUTE data provide the control variables that we use to control for other reasons for takeovers besides compensation policies or human resource management practices. In particular, the analysis uses transformations of the following variables: total assets, total fixed assets, depreciation and amortization, total debt, value added and total employment.

B.3 The DADS-EDP Data

The final two data sources are the Annual Declarations of Social Data (DADS) and the Permanent Demographic Sample (EDP). A detailed description of both of these data sets and their basic construction is provided in the data appendix to Abowd et al. (1999). The data used here cover the period from 1991 through 1999, although exploitable panel data exist as far back as 1976.

The DADS data used here constitute a 1/25th random sample of the French population. The data consist of employer records filed by firms with the government on behalf of employees for the purposes of calculating retirement benefits (among other things), and thus contain identifiers for both the individual and the employing firm. The data provide information on total gross earnings and number of days worked during the course of each year for each employer for whom a sampled individual worked, as well as information on the age and sex of the individual, the département (geographic region) in which the person worked, the sector of the employer and the occupation and type of job (e.g. full or part time employment) held by the employee. They also provide the first and last days worked during each year, a criterion upon which we base our analysis sample. As each employment spell (other than non-salaried self-employment

29. The sampling scheme depends on whether the data was part of the "sample" of firms whose data is sent to INSEE before the full treatment of all firms (for the purpose of generating advanced indicators) or not. A firm that is part of the "sample" appears in the FUTE data if it has more than 10 employees, more than 3.5M French Francs in sales or more than 5M French Francs in assets. The non-sample firms are included in the FUTE if they have over 20 employees, over 100M francs in sales or over 200M francs in assets. Although the "sample" covered less than 5% of all French enterprises in 1992, it represented 76% of employment and 82% of sales by all French firms (INSEE (1995)).

and central government employment) generates an observation in the DADS, these data not only provide us with a large source of linked employer-employee data, but a potentially a 24 year panel on the individual side. [30] The linked panel aspect is also what allows us to calculate job seniority for each sample individual. [31] One weakness of the DADS data is the lack of individual specific information, notably the absence of data on education. The EDP data, which cover roughly 1/10 of the individuals in the DADS data, help to remedy this situation. The EDP data consist of information drawn from census reports, birth reports, marriage declarations and death reports in which a sample individual can be identified. The paper uses the information on highest degree obtained (available in the census reports) to measure education for individuals in the DADS-EDP overlap, and it imputes education to the remaining 9/10 of the sample on the basis of a multinomial logit model. [32]

30. The dimension of the panel on the firm side depends on the appearance of a sampled individual in a firm. Firms with at least one sample individual in each of the years of our data will also be available for 26 years. Given the sampling scheme, if there were a purely random redistribution of individuals across firms each year, there would be a better than 50% chance that a firm with at least 17 employees will have at least one sample individual in any given year. Given that individuals tend not to switch employers every year, the probability of having a sample individual in year $t + 1$ given the presence of such an individual in year t is significantly higher.

31. Left-censoring of job spells is dealt with by estimation of the pre-sample job seniority using data drawn from yet another source, the Salary Structure Survey. See Abowd et al. (1999) for details.

32. See Abowd et al. (1999) for details. Results of the multinomial logit estimation on the extended sample are available upon request.

Figure 1: Flow-Sampled Pre-Transaction Spells, Smoothed Hazard Function

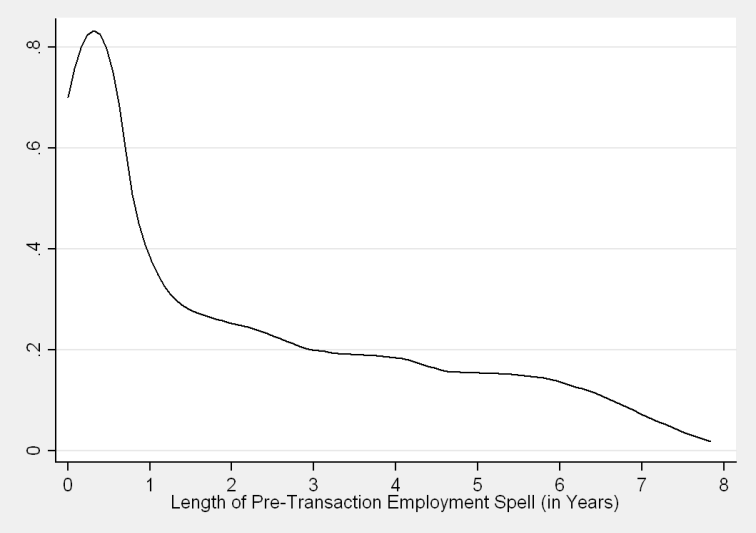

Figure 2: Flow-Sampled Post-Transaction Spells, Smoothed Hazard Function