ON THE TRADE-OFF BETWEEN EFFICIENCY IN JOB ASSIGNMENT AND TURNOVER: THE ROLE OF BREAK-UP FEES*

ARIJIT MUKHERJEE† AND LUÍS VASCONCELOS‡

Abstract. Firms often specify break-up fees in their employment contracts where a worker is obligated to compensate the firm if he leaves to take up employment with a competitor. We highlight the role of such break-up fees in the presence of asymmetric information about the worker's quality between the current employer and the outside labor market. Waldman (1984) argues that if the market attempts to learn the worker's quality from the firm's job assignment (or "promotion") decision, it raises the wage of a promoted worker leading to inefficiently few promotions. We argue that break-up fees can mitigate such inefficiencies by shielding the worker from the potential raiders. But in the presence of firm-specific matching, break-up fees thwart efficiency in turnover by muting the market's incentive to bid for the worker. We characterize the optimal contract and show that the optimality of a break-up fee depends on the relative size of the worker's expected productivity in the pre- and post-promotion jobs. It is *never* optimal to specify a break-up fee if the difference between the worker's expected productivity levels in the two jobs is sufficiently large. Otherwise, the optimal contract stipulates a break-up fee even though it may lead to market foreclosure for a better matched raider.

1. Introduction

Firms often specify a break-up fee in their employment contracts in an attempt to dissuade their workers from leaving for the competing employers. Such break-up fees, also known as "golden handcuffs," are a contractual obligation for the employee to pay back a compensation, or "damage fee", to the firm should the employee choose to leave and join a competing firm in the industry.

A typical example of such break-up fees is deferred compensations in terms of retirement contributions and stock options. Often, such compensation is paid out at a pre-specified future date *conditional* on the continuing employment relationship between the firm and the worker. If the worker voluntarily leaves the employment relationship, he may forfeit his claim on a part of his compensation. For example, the employee's retirement plan may not be vested or he may not be able to execute his stock options until he completes a certain length of tenure with the firm. Indeed, any back-loaded compensation plan where the employee forfeits her claim to a portion of her compensation should she decide to quit sooner than later can be conceived as a contract with break-up fees.[1]

Date: November 12, 2012.

*We especially thank Jay Pil Choi, Carl Davidson, Maria Goltsman, Larry Martin, José Mata, Gary Solon, Jan Zabojnik and the seminar participants at Freie Universität Berlin, London Business School, Michigan State University, RAND Corporation, University of Essex, and University of Western Ontario. Any errors that may remain are ours.
†Department of Economics, Michigan State University, 110 Marshall Adams Hall, East Lansing, MI 48823, USA. Email: arijit@msu.edu. URL: www.amukherjee.net.

‡Department of Economics, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK. E-mail: lvasco@essex.ac.uk.

[1]Several authors have highlighted the role of such back-loaded compensation in "locking-in" the key professional. See, e.g., Jackson and Lazear (1991) and Scholes (1991) for a discussion on the role of stock options as retention device. Mehran and Yermack (1999) show empirical evidence that stock options reduce CEO turnover. Also, Garmaise (2011) finds that non-compete clause helps to reduce turnover of top executives. (Such a clause can be interpreted as an employment contract with steep break-up fee since the worker can make a buyout offer in order to be able to absolve herself from any legal binding while switching employers.)

This paper seeks to highlight and analyze a novel trade-off associated with such break-up fees when there is asymmetric learning between the initial employer and the outside labor market about the workers' productivity. We argue that in such an environment the use of employment contracts with break-up fees improves the efficiency in job assignment (or promotion) but hinders the efficiency in turnover. This trade-off emanates from the interplay of the following two economic effects.

First, when the worker's quality (i.e., productivity) is gradually revealed, the initial employer is likely to be more informed (compared to the outside labor market) about its worker's quality. When the quality of the worker is not publicly observable, a typical channel through which the outside labor market attempts to infer the worker's quality is by observing the firm's job assignment, or "promotion" decisions (Waldman, 1984). Promotions are more visible publicly than the actual quality of the worker, and the workers with higher quality are more likely to be promoted. Hence, the outside labor market may take promotion as a signal of high quality of a worker and make him an appropriately high wage offer in an attempt to raid him. Waldman (1984) argues that this effect makes promotion more expensive for the firm since the firm must increase the wage of the promoted worker accordingly in order to retain him. Consequently, the firm may find it unprofitable to promote a worker unless his quality is sufficiently high to warrant the higher post-promotion wage. Thus, too few workers are promoted compared to what is socially efficient.[2] A contract with break-up fee can alleviate this inefficiency by specifying a payment that the worker must pay back to the firm if he decides to leave once he is promoted.[3] The break-up fee offsets the firm's need to pay a steep wage to retain the promoted worker—the worker may continue to stay with his initial employer since his outside wage offer *net of break-up fee* may be dominated by his current wage offer. As a result, for the firm, the "cost" of promotion decreases and the firm may promote more workers.

Second, if the productivity of a worker is governed by firm-specific matching, a break-up fee has its own cost. A high break-up fee may discourage an outside firm from bidding for the worker unless the matching gains from turnover are sufficiently high. Thus, contracts with break-up fees may lead to too few turnovers leading to a matching inefficiency. Such matching inefficiencies, in turn, hurt the firm's profit since the firm could extract the matching gains up-front from the worker.

Thus, a contract with break-up fee enhances the efficiency in job assignment at the cost of increased matching inefficiency, and the optimal contract must balance this trade-off.

To capture this trade-off, we consider a simple two-period principal-agent model where the firm (principal) has two types of job, 1 and 2. In period one, the firm hires an agent with unknown ability level (a) and assigns him to job 1. Let the productivity of the worker in job 1 be ψ_1. The initial contract specifies a wage and a break-up fee d payable to the firm should the worker decide to leave. In period two, the actual ability level of the worker is revealed to the firm, and the firm decides whether to promote the worker and assign him to job 2. In job 2, a worker with ability a produces $\psi_2 a$. The workers with higher level of ability are more productive in job 2 compared to job 1. Once the promotion decision is made, it is publicly observed and potential raiding firms—where the worker might be better matched—compete in wages to hire the worker. The initial employer can make a counteroffer upon observing the raiders' offers. The worker chooses the employer who offers the highest wage *net of break-up fee* (if any such fee is stipulated in the initial contract).

Consider the role of break-up fee in the light of the above framework. Such a fee would create a wedge between what the market offers a promoted worker and what the firm needs to pay to retain him. Consequently, promotion becomes less expensive (for the firm) and the firm would have a higher incentive to promote a worker. Thus, the worker-job matching inefficiency (as highlighted

[2]DeVaro and Waldman (2011) offers some empirical support to this argument. Also see Baker et. al (1994a, 1994b) and McCue (1996) for empirical evidence that promotion is often associated with large wage increases.

[3]Of course, in equilibrium, the employment contract with break-up fee must also ensure the worker's participation; i.e., the contract must offer the worker an expected wage that is at least as much as his outside option.

by Waldman (1984)) is reduced. But on the other hand, it affects the efficiency in turnover. This happens for two reasons: First, the raiding firms now correctly expect the ability of the promoted workers to be lower than before as the firm has lowered its threshold for promotion. As a result, the market reduces its bid and it becomes more likely that the firm would find it profitable to match such offer. Therefore, the worker may stay back with the firm even when he is more productive with the raiders. Second, if the break-up fee is sufficiently high, the raiders may be foreclosed from the labor market. The raiders need to compensate the worker for the steep break-up fee and may find it unprofitable to do so unless the matching gains are sufficiently large. As a result, they may refrain from bidding altogether even when they are a better match for the worker.[4] The trade-off between the efficiencies in job assignment and turnover shapes the optimal contract.

Our key finding is that the optimality of a break-up fee depends on the *relative size* of the worker's expected productivity in the two jobs. More specifically, it is optimal to stipulate a break-up fee if and only if ψ_1 is sufficiently close to $\psi_2 \mathbb{E}(a)$, i.e., the difference between the (expected) productivities of the worker in the two jobs is small. Moreover, in this case the inclusion of a break-up fee also increases the aggregate social surplus.

The intuition behind this finding is the following: when ψ_1 is small, the firm already has a strong incentive to promote most of the workers since they are much more productive in job 2 than in job 1. The workers who are inefficiently kept in job 1 are of low productivity and would have had little (though positive) gains in productivity had they been assigned to job 2. Thus, in such a setting, the marginal gains from more efficient promotion that is brought about by stipulating a break-up fee is relatively small. However, such a break-up fee would hinder the efficient turnover of all promoted workers. And as most of the workers are promoted (all of whom should leave for the raiders when there are matching gains), the marginal loss due to inefficient turnover is relatively large. Thus, it is optimal not to stipulated such a fee. But when ψ_1 is high, the firm would promote very few workers (those with sufficiently high ability). Also, the marginal worker who misses the promotion would have been considerably more productive if he were promoted. Thus, the marginal gain from improved worker-job matching is high whereas the marginal loss from reduced turnover is low. Therefore, it becomes optimal to stipulate a break-up fee as it eases the inefficiency in promotion but costs little in terms of matching inefficiency it creates.

We further show that in equilibrium, the optimal break-up fee forecloses the market if the productivity gains from promotion are relatively small (i.e., when ψ_1 is sufficiently *large*). In such a setting the firm has little incentive to promote a worker. The firm raises the ability threshold for promotion as the worker is already highly productive in job 1. Also, the raiders bid more aggressively for a promoted worker as they correctly infer that the firm is now more selective in offering promotion. Consequently, promotion becomes more expensive for the firm. Therefore, the inefficiencies with worker-job matching aggravate. To ensure countervailing incentives for more efficient promotion the worker must be shielded from the raiders through a steeper break-up fee. But when the break-up fee is too high, successful raids become more costly, and it is not profitable for the raiders to bid for the worker unless the matching gains are sufficiently large. In other words, some raiders may be foreclosed from the market even when they would have been a better match for the worker.

Related literature: As discussed above, any deferred or "back loaded" compensation plan can be conceived as a contract with break-up fee (as the employee typically loses part of the compensation should he decide to quit). And it is has been long established that back loaded compensations play a significant role in various key aspects of an employment relationship, such as, human capital accumulation (Becker, 1964), effort incentive throughout the employment tenure (Lazear, 1979), and worker retention (Salop and Salop, 1976).

[4]This effect is similar in spirit with the role of long-term contracts in bilateral trading as discussed in Aghion and Bolton (1987).

The key contribution of our paper is to highlight a novel trade-off between worker-job and worker-firm matching that may emanate from the use of such break-up fees. The environment where this trade-off appears has two salient features, both of which are well acknowledged in the existing literature: (i) Asymmetric information among employers leads to inefficient turnover (Greenwald, 1986; Lazear, 1986; Gibbons and Katz, 1991; also see Gibbons and Waldman, 1999, for a survey). (ii) The initial employer's (publicly observable) decisions—e.g., promotions, outcome of a rank-order tournament, etc.,—may signal the outside labor market about a worker's quality (Waldman, 1984, 1990; Bernhardt and Scoone, 1993; Zábojník and Bernhardt, 2001; Golan, 2005; Mukherjee, 2010, 2008a, 2008b; Ghosh and Waldman, 2010; Koch and Peyrache, 2011).

In the current literature on asymmetric information and learning in labor markets, our paper is most closely linked to Waldman (1984) (as discussed earlier). In a framework similar to Waldman (1984), Bernhardt and Scoones (1993) considers a more general model of promotion and turnover in the presence of firm-specific matching gains. The authors assume that the raiders can eliminate all information asymmetries if they invest in a costly information acquisition process. They argue that in order to dissuade the raiders from investing in information acquisition (as it increases the risk of losing the worker), the firm may promote the worker with a preemptive high wage. The wage signals a potentially good match between the worker and the firm and discourages the raiders to acquire information (as they anticipate a lower likelihood of successful raid). The assumption that the outside market can acquire the exact same information that the initial employer possesses is crucial for this finding. In our model such direct information acquisition is not feasible and the initial employer always enjoys some degree of information advantage. In many settings this is perhaps a more realistic assumption as the worker's productivity is often a "soft" information that can only be learned through close observation of the worker performance over a considerable duration.

Another article that is closely related with our work is that of Burguet et al. (2002). Burguet et al. examine the role of the break-up clause when the firms compete for talented workers. They find that in the presence of complete information, the firms set high break-up fees to restrain the workers' mobility in order to extract the maximum rent from a more efficient rival. Similar to the role of damage payments for breach of contract (see, Aghion and Bolton, 1987; Spier and Whinston, 1995), exclusive rights help the worker-firm coalition to capture a larger share of the surplus gained from efficient turnover. Burguet et al. study the link between the level of transparency about the worker's ability and the use of exclusive employment contracts as a rent extraction mechanism. In contrast, we consider an environment where the firm's decision on its job assignment reveals information to the market about the worker's ability, and we focus on the interplay of two contrasting roles of a break-up fee: shielding a productive worker from the raiders and rent extraction from the outside labor market when there is turnover.

It is interesting to note that our main finding on the optimality of the break-up fees in employment contracts is somewhat contrary to the role of such fees in the product market. In the product market context, an influential article by Aghion and Bolton (1987) and the literature that followed from it (see, for example, Bernheim and Whinston, 1998; Rasmusen et al., 1991) argue that break-up fees are generally inefficient as they may foreclose the market for a more efficient entrant. While this effect is also present in our model as break-up fees reduce efficiency in worker-firm matching (conditional on the promotion rule), our model also highlights a countervailing effect. In our case, the worker-job matching efficiency is also important and we argue that the use of break-up fees can increase such efficiency.[5]

[5]It is also worth mentioning that the exclusive employment contracts, which can be interpreted as contracts with prohibitively high break-up fees, have been studied extensively both by the legal scholars (Bishara, 2006; Gilson, 1999; Posner et. al, 2004; Rubin and Shedd, 1981) and by the labor economists (Burguet, et al., 2002; Franco and Mitchell, 2005; Kräkel and Sliwka, 2009). This literature is also closely related to the exclusive contracts literature in antitrust (see Posner, 1976; Aghion and Bolton, 1987; Bernheim and Whinston, 1998; Rasmusen et al., 1991) and

The rest of the paper is organized as follows. Section 2 elaborates on the baseline model that captures the key trade-off between the efficiencies in worker-job and worker-firm allocations. In light of this model, Section 3 explores the role of a break-up fee in a firm's equilibrium job assignment policy. Section 4 elaborates on the inefficiencies in worker-job and worker-firm allocations that emerge in equilibrium and how they relate with one another. The optimal break-up fee is discussed in Section 5. Section 6 discusses some robustness issues related to our key findings. A final section draws a conclusion. All proofs are given in the Appendix.

2. The Model

We consider a two-period principal-agent model that formalizes the environment discussed in the Introduction. The model is described below in terms of its five key components: *players, technology, contracts and job assignment, raids and counteroffer,* and *payoffs.*

Players. A firm (or "principal"), F, hires a worker (or "agent"), A, at the beginning of period one. The worker works for F in the first period of his life, but he may get raided in period two by the outside labor market where two identical raiding firms, R_1 and R_2, may bid competitively for the worker.

Technology. The technology specification of the firm is similar in spirit to that in the model used by Waldman (1984). The firm (F) has two types of jobs: job 1 and job 2. Job 1 is the entry level job where the worker (A) is assigned in period one.[6] The worker's productivity in job 1 is assumed to be fixed at ψ_1 (> 0). However, in job 2, the worker's productivity depends on his ability, or "type", $a \in [0, 1]$. The productivity of A in job 2 (with F) is solely driven by his ability a where he produces $\psi_2 a$.[7] For algebraic simplicity we will assume that $\psi_2 \geq 2\psi_1$. At the beginning of period one, the worker's ability is unknown to all players (including the firm, the raider and the worker himself) but is known to follow a uniform distribution on $[0, 1]$.

Job 1 is not available with the raiding firms, but they can employ the worker in job 2. However, the worker's productivity with the outside labor market depends not only on his ability but also on the firm-specific matching factor, m, where he produces $\psi_2 a (1 + m)$. The matching factor m is unknown to all players at the beginning of the game and it is assumed to be distributed uniformly on $[-1, 1]$. Note that $m > 0$ implies that the worker is a better match with the outside labor market—i.e., *a priori*, the worker as likely to be a better match with his initial employer as with the outside labor market. The exact value of m is revealed in period two and we will elaborate on this shortly.

Contracts and job assignment. At the beginning of period one, F makes a take-it-or-leave-it offer (w_1, d) to A where w_1 is the period-one wage and d is a break-up fee that A has to pay to F if A decides to leave in period two for the raiders. Note that one can re-interpret d as a deferred compensation. Assuming no time discounting, the above contract specification is equivalent to the scenario where A receives $w_1 - d$ upon accepting the employment and gets the remaining part of his period-one wage (i.e., d) only if he decides to stay with the firm in period two.

has mostly focused on the role of such contracts in fostering investments in human capital and its implications on labor mobility.

[6]For expositional clarity, we are ruling out the possibility that the firm assigns its new hire directly to job 2. One may justify such a specification by assuming that job 1 offers some on-the-job training that is essential to perform in job 2.

[7]We abstract away from the role of the worker's effort in the production process as the moral hazard issues are not the central focus of our article.

At the end of period one, the ability of the worker is observed by F (but not by the raiders) and F decides whether to assign or "promote" the worker to job 2.[8] Period-two wages are set by the spot market at the beginning of the period through an offer-counteroffer game as described below.[9]

Raids and counteroffer. At the beginning of period two, the raiding firms (R_1 and R_2) observe F's job assignment decision. For expositional clarity, we assume that it is never optimal for the raiders to bid for a worker who is not promoted.[10] After the promotion decision is made, the (identical) raiders observe the matching factor m for a promoted worker and make simultaneous wage bids b_i, $i = 1, 2$.[11] We will maintain the convention that $b_i = 0$ when the raiders refrain from bidding. Observing the bids, F makes a counteroffer; let w_2^i be the period-two wage that F offers to A in job i, $i = 1, 2$. The worker chooses the employer who offers the highest wage *net* of the break-up fee. In case of a tie, the worker stays with the initial employer.

Payoffs. We assume that all players are risk neutral. Upon successfully hiring the worker, the firm's payoff in period one is $\pi_1 = \psi_1 - w_1$. But in period two, the payoff depends on the ability of the worker, whether the worker is promoted, and if promoted, whether the worker is retained. So, the firm's payoff in period two from a worker with ability a is:

$$\pi_2(a) = \begin{cases} \psi_1 - w_2^1 & \textit{if } A \textit{ is not promoted} \\ \psi_2 a - w_2^2 & \textit{if } A \textit{ is promoted and retained by } F \\ d & \textit{if } A \textit{ is promoted but successfully raided} \end{cases}.$$

The firm's aggregate payoff from hiring a worker with ability a is $\Pi = \pi_1 + \pi_2(a)$. Similarly, the worker's payoff in period one is $u_1 = w_1$ but the period-two payoff, u_2, depends on the promotion decision of the firm and the offer/counteroffer received upon promotion. That is,

$$u_2 = \begin{cases} \max\left\{b_1 - d, b_2 - d, w_2^2\right\} & \textit{if } A \textit{ is assigned to job } 2 \\ w_2^1 & \textit{otherwise} \end{cases},$$

and the worker's aggregate payoff is $U = u_1 + u_2$. Finally, the raider's payoff from a worker with ability a is:

[8] Note that this specification implies that there is always a vacancy in job 2. One can also consider a more general setting where a vacancy in job 2 arises with probability $p \in (0, 1)$ and the job assignment is made only if there is an openning. The qualitative nature of our findings continue to hold in this general setting.

[9] Here, we are implicitly assuming that long-term contracts are not feasible in the sense that the firms cannot commit to period-two wages at the beginning of period one. The infeasibility of long-term contracts and the spot market wage setting in period two are common assumptions in this literature (see, for example, Zabojnik and Bernhardt, 2001; DeVaro and Waldman, 2009) and we will revisit the role of long-term contracts later in Section 6. Note that we are also abstracting away from the possibility that the firm can announce an "initial" period-two wage to a promoted worker *before* the raiders make their offers. Such a wage, even if it may get revised in the offer-counteroffer stage, may serve as an additional signal of the worker's underlying ability (see, Bernhardt and Scoones (1993) for a model on such signalling role of wage offers).

[10] One can motivate this assumption as the equilibrium behavior of the raiders under a slight variation of the aforementioned technology: suppose that there exists $\epsilon > 0$ sufficiently small such that a worker with ability $a \in [0, \epsilon]$ is only productive in job 1 and produces $-K$ if assigned to job 2. Now, for K sufficiently large, it is never optimal for the raiders to hire a worker who remains in job 1.

[11] The assumption that m is revealed to the raiders after the promotion decision is made only as a modeling convenience. The key issue is that m is not known to the firm when it makes the promotion decision. Our findings are robust to alterative modeling specifications as long as this key assumption is maintained. For example, one may assume that the raider knows m from the beginning of the game (the raiders may own certain complementary inputs that make the worker more productive) and it is revealed to the firm only through the raiders' bids for a promoted worker.

$$\pi_{R_i} = \begin{cases} \psi_2 a\,(1+m) - b_i & \text{if } R_i \text{ successfully raids the worker} \\ 0 & \text{otherwise} \end{cases}.$$

We assume that both the worker and the firm have a reservation payoff of 0.

Time Line. The timing of the game is as follows.

- **Period 1.0.** F offers a contract (w_1, d) to A. If accepted, the game proceeds but ends otherwise.
- **End of Period 1.** Period-one output realized. Firm observes ability and decides on job assignment. R_1 and R_2 observe job assignment and matching factor m becomes public.
- **Period 2.0.** R_1 and R_2 make simultaneous bids for A and offer b_1 and b_2 respectively.
- **Period 2.1.** After observing b_1 and b_2, F decides whether to make a counteroffer and the period-two wages w_2^1 and w_2^2 are set.
- **Period 2.2.** A chooses which employment contract to accept.
- **End of Period 2.** Period-two output is realized, wages are paid and the game ends.

Strategies and equilibrium concept: The strategy of F has three components: (i) at the beginning of period one, choose the initial contract offer (w_1, d); (ii) at the end of period one, decide on job assignment upon observing the worker's ability, and (iii) decide on the counteroffer $(w_2^i,\ i = 1, 2)$ upon observing the raiders' offers. The worker's strategy has two components: (i) accept or reject the firm's initial contract, and (ii) choose period-two employer given the raiders' offer and the firm's counteroffer. Finally, the raiders' strategy is to choose a wage bid b_i given the matching factor and the firm's job assignment decision. We use *perfect Bayesian Equilibrium (PBE)* as a solution concept.

3. Equilibrium job assignment policy of the firm

In order to derive the optimal contract for the firm, we first need to characterize the continuation game for a given break-up fee (d). In this vain, we discuss below the firm's equilibrium job assignment policy and analyze the offer-counteroffer subgame for an arbitrary value of d specified by the firm in period one. Our analysis also elucidates on the key trade-off between the efficiencies in the worker-job and the worker-firm matching.

But before we present the equilibrium analysis, it is instructive to consider the first best allocation of the worker as an efficiency benchmark.

3.1. **First best allocation of the worker.**

The first-best allocation of the worker requires efficiency in both worker-job and worker-firm matching. Ex-post, when there is no uncertainty about ability and matching gains, the first-best allocation is straightforward. When the worker is a better match for the firm (i.e., $m < 0$), the worker stays with the firm and is promoted if and only if he is more productive in job 2 than in job 1, i.e., if and only if $\psi_2 a \geq \psi_1$ or $a \geq \psi_1/\psi_2$. In contrast, when the worker is a better match with the raiders (i.e., $m > 0$), the worker is promoted *and* leaves for the raiders if $\psi_2\,(1+m)\,a \geq \psi_1$ or $a \geq \psi_1/\psi_2\,(1+m)$. Otherwise, the worker stays with the firm in job 1.

However, as the firm makes its job assignment decision *before* observing the matching gains, one may consider the ex-ante efficient job allocation as a benchmark for evaluating the extent of allocative inefficiency in equilibrium. The ex-ante efficient promotion rule is the one that maximizes total production (i.e., aggregate surplus), assuming that following promotion turnover is efficient.

Note that as the worker's productivity in job 2 (i.e., $\psi_2 a$) is increasing in a while it is constant (ψ_1) in job 1, the optimal promotion decision must follow a cut-off rule. Consider any arbitrary cut-off rule where a worker is assigned to job 2 if and only if his ability $a \geq x$. Assuming efficient turnover following promotion, the ex-ante aggregate expected surplus under such a policy is:

$$S(x) := \psi_1 x + \int_x^1 \psi_2 a \left[\frac{1}{2} \int_{-1}^{0} dm + \frac{1}{2} \int_0^1 (1+m) dm \right] da.$$

The ex-ante efficient (or "first best") promotion policy, a^{FB} (say), is the one that maximizes S. That is,

$$S'\left(a^{FB}\right) = 0, \text{ or } a^{FB} = \frac{4\psi_1}{5\psi_2}. \tag{1}$$

Note that under the ex-ante efficient policy more workers are promoted to job 2 than what the firm would promote in the absence of any raiders. All workers with ability $a \in [4\psi_1/5\psi_2, \psi_1/\psi_2]$ are more productive in job 1 than in job 2 when working for their initial employer, but should be assigned to job 2 under ex-ante efficient promotion rule due to the potential matching gains from turnover (recall that promotion to job 2 is necessary to realize the matching gains as a worker in job 1 is never raided).

We now consider the equilibrium job assignment and turnover and explore how the extent of inefficiency is affected by the break-up fee.

3.2. **Equilibrium job assignment and turnover.** Given that the worker's wage in period two is determined in the spot market, and the outside market does not observe the actual ability level of the worker, a worker's wage *conditional* on job assignment is independent of his ability. So, the firm's payoff from offering promotion is increasing in a while denying promotion yields a constant payoff (given the production technology). Thus, as in the case of first-best allocation rule, the firm's promotion decision also follows a cut-off rule in equilibrium where the firm promotes a worker if and only if his ability is greater than a cut-off value a^*. So, one can solve for the cut-off value a^* as a function of the break-up fee (d) by using backward induction.

First, note that in period-two, if there is no market offer (i.e., $b_i = 0$ for all i) the firm offers a wage of 0. In other words, a worker who stays in job 1 as well as a promoted worker who does not receive any market offer earns $w_2^1 = w_2^2 = 0$ as the firm simply matches the worker's outside option.

Now consider the case where a promoted worker receives a market offer. In this case the firm's optimal counteroffer decision needs a more careful study. Let b denote the highest bid that the worker receives (i.e., $b = \max\{b_1, b_2\}$). Throughout this article we refer to b as the market bid for the promoted worker. If $b \leq d$ then the worker's outside option of 0 is better than his payoff from paying the break-up fee and joining the raider. Thus, the firm retains the worker by matching his outside option and offers $w_2^2 = 0$. But if $b > d$, the firm has two options. The firm can either retain the worker by making a counteroffer $w_2^2 = b - d$ and earn a profit of $\psi_2 a - (b - d)$, or it can let the worker go and earn d. Thus, the firm will make a counteroffer if and only if $\psi_2 a - (b - d) \geq d$, or, equivalently, $a \geq b/\psi_2$.

Note that the break-up fee reduces the retention wage of a promoted worker. Furthermore, if the market bids $b \leq d$ or $b \leq \psi_2 a^*$, it fails to raid the worker irrespective of his ability. But if $b > \max\{\psi_2 a^*, d\}$, the market successfully raids some types of the worker.[12] More specifically, if $b < \max\{\psi_2, d\}$, the market raids all workers with ability $[a^*, b/\psi_2)$. That is, among the pool of promoted workers, only the relatively low ability workers leave the firm. If the market bids even higher, i.e., $b > \max\{\psi_2, d\}$, it raids all the workers who are promoted. So, to sum up, when the firm uses the promotion cut-off a^*, it retains every promoted worker with ability $a \geq \widehat{a}(b)$ where

[12]For brevity of exposition, in what follows, we will refer to different "types" (or ability levels) of a worker simply as different "workers."

$$\widehat{a}(b) = \begin{cases} a^* & \textit{if } b \leq \max\{\psi_2 a^*, d\} \\ b/\psi_2 & \textit{if } \max\{\psi_2 a^*, d\} < b \leq \psi_2 \\ 1 & \textit{if } b > \max\{d, \psi_2\} \end{cases} .$$

Note that the function $\widehat{a}(b)$ captures the firm's optimal policy of worker retention.

Given the firm's optimal counteroffer decision, raider i's expected gross profit (i.e., profit ignoring wage payment) from bidding b *conditional* on the worker's choice of period-two employer is:

$$\begin{aligned} \widehat{\pi}_{R_i}(b, m, d; a^*) &:= \mathbb{E}[\psi_2 a(1+m) \mid a \in [a^*, \widehat{a}(b))] \\ &= \begin{cases} 0 & \textit{if } b \leq \max\{\psi_2 a^*, d\} \\ \psi_2(1+m)\frac{1}{2}(a^* + \frac{b}{\psi_2}) & \textit{if } \max\{\psi_2 a^*, d\} < b \leq \psi_2 \\ \psi_2(1+m)\frac{1}{2}(a^*+1) & \textit{if } b > \max\{d, \psi_2\} \end{cases} . \end{aligned}$$

Since raiders compete in wages, they make zero expected profit in equilibrium and bid the entire expected value of the worker. That is, the raiders' equilibrium wage bids must be $b_1^* = b_2^* = b^*$ where b^* solves the equation $b = \widehat{\pi}_{R_i}(b, m, d; a^*)$ for all i. The equilibrium bid b^* critically depends on the value of the break-up fee:

(2)

if $d < \psi_2$, then

$$b^*(m, d; a^*) = \begin{cases} 0 & \textit{if } m \leq 0 \textit{ or } a^* \leq \frac{d}{\psi_2}\frac{1-m}{1+m} \\ \psi_2 a^* \frac{1+m}{1-m} & \textit{if } m > 0 \textit{ and } \frac{d}{\psi_2}\frac{1-m}{1+m} < a^* \leq \frac{1-m}{1+m} \\ \psi_2(1+m)(\frac{a^*+1}{2}) & \textit{otherwise} \end{cases} ,$$

and if $d \geq \psi_2$, then

$$b^*(m, d; a^*) = \begin{cases} 0 & \textit{if } m \leq 0 \textit{ or } a^* \leq \frac{2d}{\psi_2(1+m)} - 1 \\ \psi_2(1+m)(\frac{a^*+1}{2}) & \textit{otherwise} \end{cases} .$$

Note that when the break-up fee is too high or the promotion cut-off is too low the raiders refrain from bidding for the worker even when the worker is a better match for them. The argument is straightforward. When the break-up fee is too high then the market cannot profitably raid the worker. And if a^* is too small, then promotion is not quite informative about the worker's ability. So the market does not place any bid as it correctly anticipates attracting only a pool of sufficiently low ability workers. But as a^* increases, promotion becomes a stronger signal of quality and the market finds it worthwhile to bid for the promoted workers.[13]

Given the market's bidding strategy, one can plug b^* in the cut-off function $\widehat{a}(b)$ and derive the firm's retention threshold as follows:

[13] It is worth noting that the above characterization of the equilibrium bidding strategies implicitly assumes that the raiders do not play weakly dominated strategies. Otherwise, there may exist other equilibria where the raiders bid more than the expected value of the worker (to the raiders) if the firm is expected to retain the worker with certainty by making a counteroffer (this can happen if $m < 0$). One may rule out out such equilibria as they are not "trembling hand perfect"—if there is a small probability that the worker may mistakenly accept the raiders' bid, then the raider is strictly better off by not placing a bid that is higher than its valuation for the worker. Such equilibria in domoinated strategies also do not survive the "market-Nash" refinement of Waldman (1984).

(3)
if $d < \psi_2$, then
$$\widehat{a}(b^*(m,d;a^*)) = \begin{cases} a^* & \textit{if } m \leq 0 \textit{ or } a^* \leq \frac{d}{\psi_2}\frac{1-m}{1+m} \\ a^*\frac{1+m}{1-m} & \textit{if } m > 0 \textit{ and } \frac{d}{\psi_2}\frac{1-m}{1+m} < a^* \leq \frac{1-m}{1+m} \\ 1 & \textit{otherwise} \end{cases},$$
and if $d \geq \psi_2$, then
$$\widehat{a}(b^*(m,d;a^*)) = \begin{cases} a^* & \textit{if } m \leq 0 \textit{ or } a^* \leq \frac{2d}{\psi_2(1+m)} - 1 \\ 1 & \textit{otherwise} \end{cases}.$$

When $b^* > d$, the firm retains a (promoted) worker if his ability $a \geq \widehat{a}(b^*(m,d;a^*))$ by matching the raiders' bid net of the break-up fee (i.e., offers $w_2^2 = b^* - d$) but lets him leave otherwise (i.e., offers $w_2^2 = 0$). So, in any equilibrium, if a worker in job 2 receives a market offer $b^* > d$ the offer matching policy of the firm is given as follows:

$$w_2^2 = \begin{cases} b^* - d & \textit{if } a \geq \widehat{a}(b^*(m,d;a^*)) \\ 0 & \textit{otherwise} \end{cases}. \tag{4}$$

Now, we can also derive the firm's profit from promoting the "marginal" worker, i.e., the worker with ability a^*. This profit, π_p (say), depends on whether the firm will retain the worker or not, and, in case of retention, the wage it has to pay to the worker. From the analysis above, we obtain the following:

(5)
if $d < \psi_2$, then
$$\pi_p(a^*,m,d) = \begin{cases} \psi_2 a^* & \textit{if } m \leq 0 \textit{ or } a^* \leq \frac{d}{\psi_2}\frac{1-m}{1+m} \\ d & \textit{otherwise} \end{cases},$$
and if $d \geq \psi_2$, then
$$\pi_p(a^*,m,d) = \begin{cases} \psi_2 a^* & \textit{if } m \leq 0 \textit{ or } a^* \leq \frac{2d}{\psi_2(1+m)} - 1 \\ d & \textit{otherwise} \end{cases}.$$

As we have argued before, when $m \leq 0$ or if a^* is sufficiently small relative to d, the raiders do not bid for the promoted workers. So, the firm retains every worker it promotes, including the marginal worker, and pays zero wage. In all other cases, the firm either let all workers go or retains only the more able workers (among the promoted ones). Therefore, the marginal worker is never retained and the firm makes d on him.

Since the productivity of the worker in job 1 is ψ_1, the cut-off ability level for promotion, a^*, must solve $\mathbb{E}_m \pi_p(a^*,m,d) = \psi_1$. The following proposition characterizes the solution.

Proposition 1. *When the employment contract includes a break-up fee* (d)*, the firm promotes a worker if and only if his ability* $a > a^*(d)$*, where*

$$a^*(d) = \begin{cases} (2\psi_1 - d)/\psi_2 & \textit{if } 0 \leq d < \psi_1 \\ \psi_1 d/(2\psi_2 d - \psi_1\psi_2) & \textit{if } \psi_1 \leq d < \psi_2 \\ \left(\psi_1 - d + \frac{d^2}{\psi_2}\right)/(2d - \psi_1) & \textit{if } \psi_2 \leq d < \psi_2 + \psi_1 \\ \psi_1/\psi_2 & \textit{otherwise} \end{cases}. \tag{6}$$

Proposition 1 indicates how the equilibrium promotion rule changes with the break-up fee specified in the contract: unless the specified break-up fee is significantly large, an increase in the break-up fee always induces the firm to promote more workers—for $d < \psi_2$, the cut-off of ability, $a^*(d)$, (above which the firm promotes the worker) is decreasing in d. However, if the break-up fee is sufficiently large ($\psi_2 \leq d < \psi_2 + \psi_1$), an increase in the fee may restrict promotion, and, at the extreme ($d > \psi_2 + \psi_1$), break-up fee does not have any impact on the promotion rate (see Figure 1).

To see the intuition behind the equilibrium promotion policy, note that increasing d has two effects on the firm's profit: an increase in d increases the compensation that the firm gets in case a promoted worker is raided (and leaves the firm), but it also increases the probability of retaining the promoted worker. The first effect always increases the firm's expected profit from promotion. The second effect may increase or decrease expected profit depending on the attractiveness of retaining the worker relative to losing him to a competitor. When d is not too large, retaining the worker is more attractive than losing him (since the break-up fee earned due to turnover is moderate). Hence, as d increases, both effects increase the firm's profit from promoting a worker. This implies that the firm's incentive to promote workers also increases with d. So, $a^*(d)$ decreases in d.

In contrast, when d is large, the break-up fee is sufficiently lucrative and retaining the worker is less attractive than losing him to a raider. In this case, the second effect lowers the firm's profit (it restricts turnover, and hence, the firm fails to collect the break-up fee) and may dominate the first effect. When that happens, the firm's profit from promoting a worker decreases with d, meaning that its incentives to promote a worker also decreases in d (i.e., $a^*(d)$ *increases* in d). Finally, for d significantly large (i.e., $d > \psi_1 + \psi_2$) neither effect is relevant, as the market never attempts to raid the promoted workers. In this case the break-up fee has no effect on the firm's profit from promotion and, as a consequence, the incentives to promote the worker remain unchanged with d.

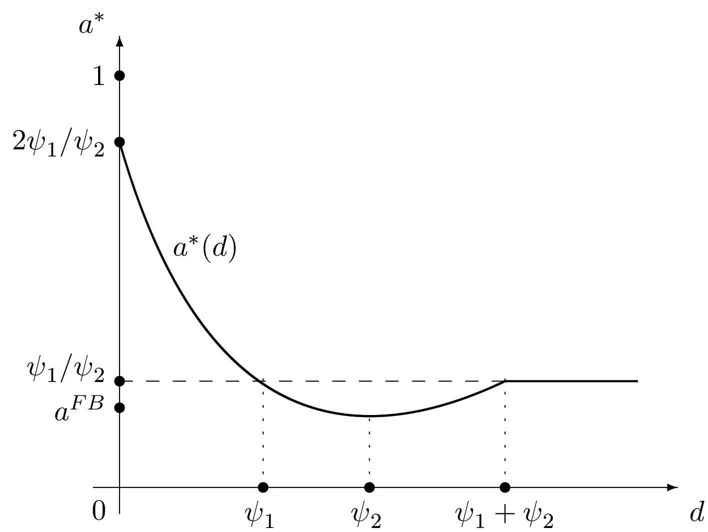

Figure 1. The optimal cut-off for promotion as a function of the break-up fee (d)

Having characterized the promotion rule for a *given* break-up fee, we can now address the question of the *optimal* break-up fee. But before we do so it is instructive to discuss the allocative inefficiencies arising from a given promotion policy. The optimal break-up fee is simply the one that induces a promotion policy that minimizes these inefficiencies.

4. The nature of allocative inefficiencies

In this section we elaborate on the nature of allocative inefficiencies that arise with an arbitrary promotion policy (given the offer-counteroffer game that follows the promotion decision). We do so with the help of Figure 2 below that plots the range of matching gains (m) and the worker's ability (a). The following discussion elucidates on the key economic effects that shape the firm's optimal contract and facilitates the characterization of the optimal break-up fee.

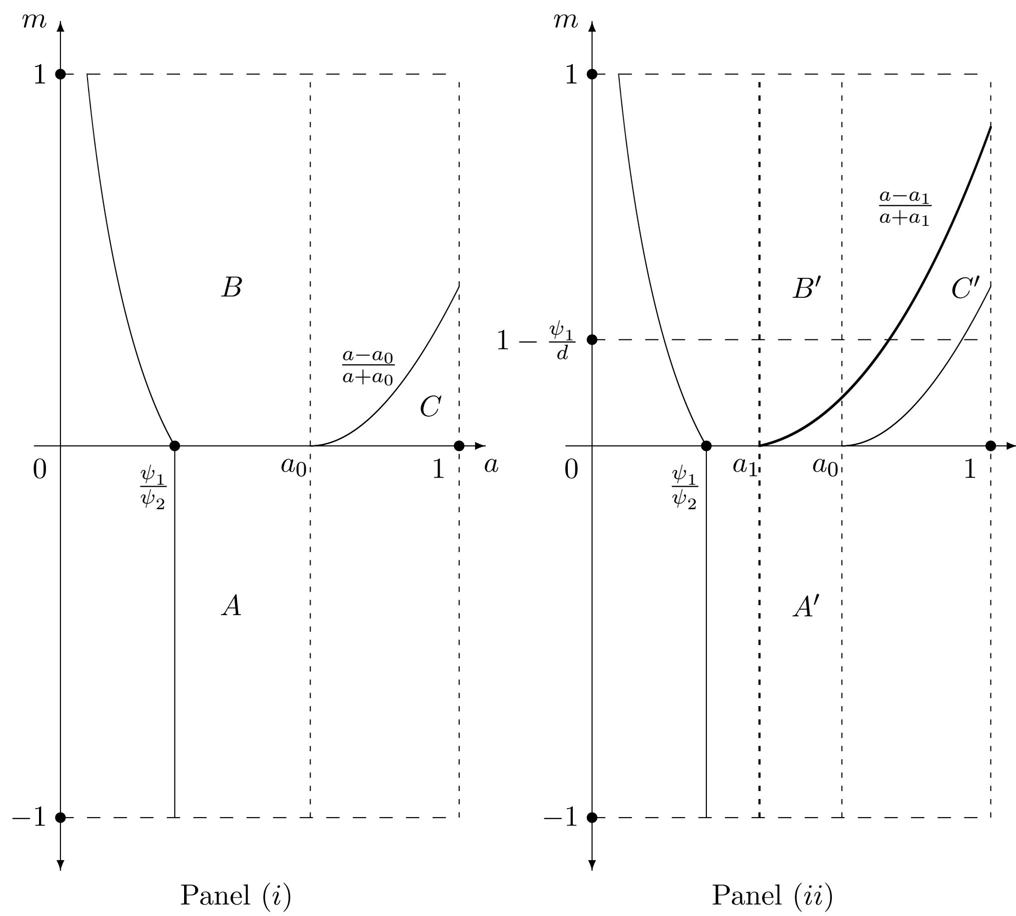

Figure 2. The allocative inefficiencies associated with a given promotion policy

Consider an arbitrary promotion policy where the firm assigns a worker to job 2 if and only if his ability $a \geq a_0 > \psi_1/\psi_2$ (see panel (i)). There are three potential sources of inefficiencies: first, for $m < 0$, there is a worker-job matching inefficiency—all workers with $a \in [\psi_1/\psi_2, a_0]$ should have been assigned to job 2 but were kept in job 1 instead (shown by area A). Second, when $m > 0$, for the the set of workers with $a \in [\psi_1/\psi_2 (1+m), a_0]$ there is both worker-firm and worker-job inefficiencies (shown by area B)—it is socially efficient for all of these workers to work in job 2 at the raiding firm but they remain in job 1 at the initial employer. Finally, even among the promoted workers there is a set of workers who are inefficiently matched with their initial employer. Given the raiders' bidding strategy, the raiders make a bid for all workers in job 2 when $m > 0$ but the firm matches the offer if the worker's ability is sufficiently high, i.e., if $a \geq \hat{a}(b^*(m, d; a_0))$, which is equivalent to $m \leq (a - a_0)/(a + a_0)$ (using equation (3)). Thus, among the promoted workers there is an inefficient worker-firm matching when $m \leq (a - a_0)/(a + a_0)$ (shown by area C). Note that this effect is similar to the "winner's curse" problem in common value auctions—the raiding firms lower their bids as a successful raid may carry a negative signal about the worker's ability, namely, the initial employer did not find the worker productive enough to warrant a matching wage offer. Thus, the initial employer finds it profitable to match the raiders' offer even though the worker would have been more productive at the raiding firm.

Accounting for these three sources of inefficiencies, the aggregate expected surplus (in period-two) under the above promotion policy can be written as:

$$
\begin{aligned}
\widehat{S}(a_0) \quad &:= \psi_1 \Pr[\textit{no promotion}] + \\
&\qquad \mathbb{E}[\psi_2 a \mid \textit{promotion, no turnover}] \Pr[\textit{promotion, no turnover}] + \\
&\qquad \mathbb{E}[\psi_2 (1+m) a \mid \textit{promotion, turnover}] \Pr[\textit{promotion, turnover}] \\
&= \psi_1 a_0 + \int_{a_0}^{1} \psi_2 a \left[\tfrac{1}{2}\int_{-1}^{0} dm + \tfrac{1}{2}\int_0^1 (1+m) dm\right] da - \int_{a_0}^{1} \psi_2 a \left[\int_0^{\frac{a-a_0}{a+a_0}} \tfrac{1}{2} m dm\right] da.
\end{aligned}
\tag{7}
$$

The optimal promotion policy *given* the information asymmetry in the offer-counteroffer game is the one that maximizes $\widehat{S}(a)$.

Note the marginal effects of promotion threshold (a_0) on the expected surplus: suppose that the promotion threshold is lowered from a_0 to a_1, say (panel (ii)). This change leads to a more efficient worker-job and worker-firm matching (areas A' and B'). But the improved worker-job matching comes at a cost of worse worker-firm matching that results from an aggravated winner's curse problem. As the ability threshold for promotion lowers, the expected productivity of the promoted worker decreases. And so does the equilibrium bid. Thus, the firm will retain a higher share of the workers: now a worker of ability a is successfully raided only if $m > (a - a_1)/(a + a_1) > (a - a_0)/(a + a_0)$ (the increased worker-firm matching inefficiency is shown by area C').

The promotion policy that maximizes $\hat{S}$ must balance the trade-off between improved worker-job matching and worsened worker-firm matching. Let us denote the policy that maximizes $\hat{S}$ as the "second best" promotion policy, or a^{SB} (in contrast with the "first best" policy discussed earlier in Section 3.1 where turnover following job assignment is always assumed to be efficient).

The allocative inefficiencies discussed above illustrate the costs and benefits of using break-up fees. Note that in absence of any break-up fee, even the second-best promotion policy may not be attained as $a^*(0) \neq a^{SB}$.[14] Also note that the raiders make zero profit due to competition and the firm can ensure zero rents for the worker by sufficiently lowering his first-period wage. Thus, the firm apropriates the entire surplus that is generated by the coalition of the firm, worker and the raiders. Consequently, the problem of choosing the optimal break-up fee can be conceived as the problem of choosing d such that equilibrium promotion rule $a^*(d)$ (as given in Proposition 1) coincides with the second-best optimal policy, a^{SB}.[15]

However, the equilibrium promotion policy indicates that if a sufficiently high break-up fee is needed to implement the second-best promotion, it may create an additional source of inefficiency. As the following lemma shows, when d is large, the firm may partially foreclose the labor market for the raiding firms even when there are positive matching gains.

[14] Such inefficiency in job assignment is similar in spirit to the one discussed in Waldman (1984).

[15] Recall that $a^*(d)$ is decreasing in d for $d < \psi_2$. We will also argue later that the optimal contract always specifies a $d < \psi_2$.

Lemma 1. *The raiders bid for the promoted worker if and only if $m > \widehat{m}(d)$, where*

$$\widehat{m}(d) = \begin{cases} 0 & \textit{if } 0 \le d < \psi_1 \\ 1 - \psi_1/d & \textit{if } \psi_1 \le d < \psi_2 \\ (3d - 2\psi_1 - \psi_2)/(d + \psi_2) & \textit{if } \psi_2 \le d < \psi_2 + \psi_1 \\ 1 & \textit{otherwise} \end{cases} .$$

The market foreclosure effect emanates from the fact that when the break-up fee is sufficiently high, the winner's curse effect becomes too severe and the raiders refrain from bidding unless the matching gains are also sufficiently large. Thus, if the firm needs to specify a sufficiently large break-up fee ($d > \psi_1$) in order to implement the second-best promotion policy, implementing such a policy is no longer optimal for the firm. In this case the firm must also account for the loss of surplus due to market foreclosure, and the promotion policy associated with the optimal contract falls short of even the second-best level.

In what follows, we elaborate on the optimal break-up fee in the light of the above discussion.

5. Characterization of the optimal break-up fee

The above discussion suggests that the optimal break-up fee is the one that implements a^{SB} if such a value of the fee is feasible and if it does not lead to market foreclosure. But when is such a value of d feasible? And when would one expect market foreclosure to take place in equilibrium? In this section we directly formulate the firm's optimal contracting problem to address these questions. It turns out that the answers critically hinge on the worker's relative productivity in jobs 1 and 2.

Consider the firm's optimal contracting problem at the beginning of period one (when the worker's ability is unknown to all parties). Recall that the firm's payoff in period one is simply $\pi_1 = \psi_1 - w_1$ (the worker is assigned in job 1 at a wage of w_1). However, the firm's expected payoff in period two, $\mathbb{E}\pi_2$, needs a more careful study. Similar to the aggregate social surplus (equation (7)), $\mathbb{E}\pi_2$ also depends on likelihood of promotion and turnover. But the expression for $\mathbb{E}\pi_2$ differs from that of the aggregate social surplus for two reasons: (i) if there is turnover (that is, there is a market offer and the firm decides not to make a counteroffer), the firm earns only the break-up fee (d) and (ii) if there is no turnover, the firm's payoff depends on whether there is market offer or not. If there is no market offer, the firm makes $\psi_2 a$ on the worker (since wage stays at 0). But if there is market offer then the firm pays $b - d$ and earns a profit of $\psi_2 a - (b - d)$. So, drawing parallel to equation (7), one obtains:

$$\begin{aligned} \mathbb{E}\pi_2(a) = {} & \psi_1 \times \Pr[\textit{no promotion}] + \mathbb{E}[\psi_2 a \mid \textit{promotion, no offer}] \Pr[\textit{promotion, no offer}] \\ & + d \times \Pr[\textit{promotion, offer, no counteroffer}] \\ & + \mathbb{E}[\psi_2 a - (b - d) \mid \textit{promotion, offer, counteroffer}] \Pr[\textit{promotion, offer, counteroffer}]. \end{aligned} \tag{8}$$

Now, in equilibrium, the raiders' bid is given by equation (2). Moreover, the raiders' bidding function in conjunction with the firm's equilibrium job assignment policy (equation (6)) and the offer-matching policy (equation (4)) determines the (joint) probability of a worker being promoted, receiving a wage offer from the raiders, and receiving a counteroffer from the firm. Thus, the firm's optimal contracting problem boils down to maximizing its aggregate expected profit $\pi_1 + \mathbb{E}\pi_2$ by choosing period-one wage (w_1) and the break-up fee (d) subject to the worker's participation constraint or, individual rationality constraint (IR) given the raiders' bidding function and the firm's job assignment and counteroffer policies. That is, the firm solves:

$$\max_{w_1,d}\ \Pi = \pi_1 + \mathbb{E}\pi_2$$

$$\textit{subject to equations } (2),\ (4),\ (6),\ \textit{and}$$

$$w_1 + \mathbb{E}\left[b - d \mid \textit{promotion, offer}\right] \times \Pr\left[\textit{promotion, offer}\right] \geq 0. \quad (IR)$$

Because the worker's (IR) constraint always binds in equilibrium (else the firm can lower w_1 and increase its profit), one can plug the (IR) constraint in the firm's objective function to eliminate w_1. Let the resulting profit function be $\tilde{\Pi}(d)$. Hence, the firm's optimal contracting problem boils down to an unconstrained maximization problem of solving $\max_d \tilde{\Pi}(d)$. The following lemma offers a useful characterization the function $\tilde{\Pi}$. (Recall that $\widehat{S}(a^*)$ is the aggregate expected surplus (in period-two) given the promotion cut-off a^*, as given in equation (7).)

Lemma 2. *The firm's expected profit function $\tilde{\Pi}$ is continuous in d and given by the following functional form:*

$$\tilde{\Pi}(d) = \begin{cases} \psi_1 + \widehat{S}\left(a^*\left(d\right)\right) & \textit{if } 0 \leq d < \psi_1 \\ \psi_1 + \widehat{S}\left(a^*\left(d\right)\right) - f(a^*\left(d\right), d) & \textit{if } \psi_1 \leq d < \psi_2 + \psi_1 \\ \left(\psi_1 + \psi_2\right)^2 / 2\psi_2 & \textit{otherwise} \end{cases},$$

where $f : [0,1] \times \mathbb{R}_+ \to \mathbb{R}$ and $f(a^*(d), d) > 0$ for all $d \in (\psi_1, \psi_2 + \psi_1)$.

Lemma 2 suggests that the firm's profit as a function of the break-up fee (d) has the following characteristics: for small values of d $(< \psi_1)$, the effect of the break-up fee on the firm's profit can be completely characterized by the break-up fee's impact on the equilibrium promotion rule, $a^*(d)$. In this case, the firm's profit reflects only the winners' curse effect of the break-up fee. In contrast, for d sufficiently large $(d > \psi_1 + \psi_2)$, the break-up fee has no impact on the profit since the market is foreclosed and there is no turnover. But for all intermediate values of d the firm's profit reflects both the winner's curse effect and the market foreclosure effect of the break-up fee. More specifically, the market foreclosure effect is captured by the function f. Also note that in absence of market foreclosure, the firm's profit is simply equal to the aggregate expected surplus generated across the two periods by the coalition of the firm, worker and the raiders, given the firm's promotion policy (ψ_1 in period one and $\widehat{S}(a^*)$ in period two).

Given the characterization of the firm's profit function, the first question we ask is under what circumstances is it optimal for the firm to specify a break-up fee? The following proposition addresses this question.

Proposition 2. *There exists a value of ψ_1 in $(0, \psi_2/2)$, say $\underline{\psi}_1$, such that the firm's optimal contract specifies a strictly positive break-up fee if and only if $\psi_1 > \underline{\psi}_1$. Moreover, under this condition the inclusion of a break-up fee is socially optimal as it increases the aggregate surplus.*

The above proposition suggests that the optimality of a break-up fee is driven by the relative productivity of the worker in the two jobs: it is never optimal to use the break-up fee in the employment contract if the worker's productivity in job 1 (i.e., ψ_1) is too low compared to his expected productivity in job 2 (i.e., $\psi_2/2$). Otherwise, it is always optimal to specify a break-up fee in the employment contract.

The intuition behind this finding is as follows. As discussed above, the use of a break-up fee improves worker-job matching but worsens worker-firm matching. When ψ_1 is small, the marginal gain from the former effect is lower than the marginal loss from the latter effect. To see this, note that the equilibrium promotion rule is such that when ψ_1 is low, a^* is also low even in the absence of any break-up fee (see Proposition 1). In other words, most workers (i.e., most "types" of the worker) are promoted when their productivity in job 1 is low. This means that the marginal worker who "misses" promotion has a relatively low ability and assigning him to job 2 (as efficiency in worker-job allocation dictates) has only a small impact on productivity. Therefore, while the introduction of a break-up fee does improve worker-job matching, its impact on aggregate surplus is small. In contrast, its impact resulting from a less efficient worker-firm matching is still large. As almost all workers are promoted, the introduction of a break-up fee reduces the likelihood of turnover for many workers. Hence, when ψ_1 is small, the marginal positive effect from a break-up fee (in terms of efficient promotion) is more than offset by the marginal negative effect (in terms of reduced turnover) and it is optimal not to use such a fee in the employment contract.

But when ψ_1 is high, the opposite happens—the marginal gain from worker-job matching dominates the marginal loss from inefficient worker-firm matching. When ψ_1 is large, very few workers are promoted in equilibrium if a break-up fee is not used. Note that $a^*(0) \to 1$ as $\psi_1 \to \psi_2/2$. Thus, the marginal worker who misses promotion has high ability and the gain in productivity from (efficiently) promoting him is relatively high. In contrast, the loss for reduced turnover from introducing a break-up fee is minimal. This is due to the fact that very few workers are promoted in the first place and those are the only workers whose turnover is affected by the existence of a break-up in the labor contract. Hence, when ψ_1 is large, the firm can increase its profit by stipulating a break-up fee that ensures more efficient promotion.

Now consider the optimality of using break-up fees from a social perspective. Since the firm extracts the entire surplus generated by the worker, if the inclusion of a break-up fee is profit-enhancing for the firm, it is also socially optimal—it increases the aggregate social surplus generated by the coalition of the firm, worker and the outside labor market.

As we have discussed in the previous section, even when it is feasible for the firm to implement a^{SB}, it may not always be optimal to do so. If the associated break-up fee leads to market foreclosure (i.e., if $d > \psi_1$) the optimal break-up fee must also account for this additional source of inefficiency. But when do we observe market foreclosure in equilibrium? The next proposition addresses this question.

Proposition 3. *There exists a value of ψ_1 in $(\underline{\psi}_1, \psi_2/2)$, say $\overline{\psi}_1$, such that for $\underline{\psi}_1 < \psi_1 \leq \overline{\psi}_1$, the optimal break-up fee $d^* \in (0, \psi_1]$, $a^*(d^*) = a^{SB}$, and the firm's profit $\Pi^* = \psi_1 + \widehat{S}(a^{SB})$. But for $\psi_1 > \overline{\psi}_1$, $d^* \in (\psi_1, \psi_2)$ and $\Pi^* < \psi_1 + \widehat{S}(a^{SB})$.*

Proposition 3 suggests that as long as ψ_1 is not too large (i.e., $\psi_1 < \overline{\psi}_1$), the optimal break-up fee never forecloses the market (as $d^* < \psi_1$). So, only the winner's curse effect remains, and as discussed earlier, whenever feasible (i.e., when $\psi_1 > \underline{\psi}_1$), the optimal d is the one that implements a^{SB}. But when ψ_1 is high $\left(> \overline{\psi}_1\right)$ there is direct foreclosure of the market since $d^* > \psi_1$. In this case, the associated profit of the firm falls short of the second-best due to the additional inefficiency (i.e., market foreclosure) that the break-up fee creates.

Propositions 2 and 3 allude to the fact that in equilibrium, higher values of break-up fee may be associated with higher values of ψ_1. This observation leads to a more general question of whether the optimal break-up fee is increasing in ψ_1. While such a comparative statics exercise appears to be analytically intractable, one can make an intuitive argument why d^* is expected to be increasing in ψ_1. A brief discussion is presented below; a more technical and detailed analysis is given in Appendix B.

When ψ_1 is large, a worker remains highly productive even when he is inefficiently assigned to job 1. Also, the market infers that the ability of a promoted worker must be sufficiently high as the firm has promoted him to job 2 even when he would have been highly productive in job 1. So, the market bids more aggressively for the promoted workers and promotion becomes more expensive. As a result, the firm becomes more inclined to (inefficiently) retain the worker in job 1 (rather than promoting him and risking turnover). Consequently, a higher d^* is needed to create a countervailing incentive for the firm to promote more efficiently by shielding the worker from potential raiders and felicitating retention. As a result, it improves the efficiency in job assignment in period two, which, in turn, enhances the firm's payoff.[16]

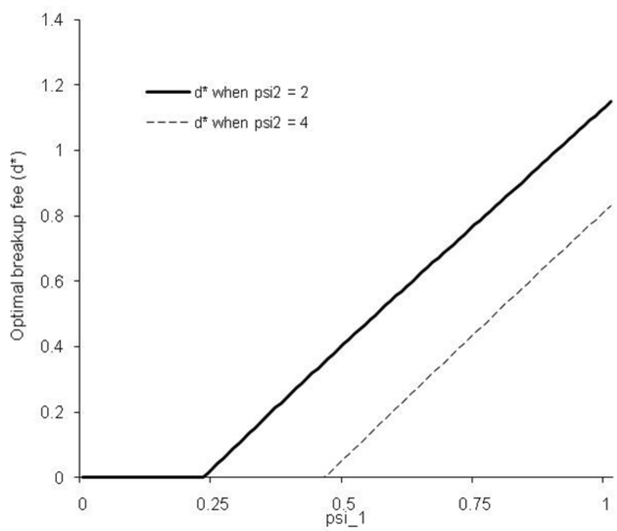

Figure 3. Optimal break-up fee as a function of ψ_1
(for $\psi_2 = 2$ and $\psi_2 = 4$).

Figure 3 above presents a numerical solution of the optimal break–up fee as a function of ψ_1. In conformity with the argument presented above, we find d^* to be increasing in ψ_1. Also, note that an increase in ψ_2 decreases the optimal break-up fee for any given ψ_1 as an increase in ψ_2 for a given ψ_1 has the same qualitative effect of lowering ψ_1 for a given ψ_2.

We conclude this section with a remark on the testable implications of our key finding. Our findings indicate that the optimality of break-up fees is driven by the difference in (expected) productivity of the worker in the pre- and post-promotion jobs. If this difference is too large, such fees are never optimal. Else, the use of such fee is indeed optimal, and moreover, the size of such fee is likely to grow as the difference in productivity narrows. This finding can also be interpreted as one linking the underlying production technologies of the two jobs with the nature of the employment contract: if the production technologies in the pre- and post-promotion jobs are not too different (e.g., they involve similar tasks), and hence, the expected productivity gains from promotion is not too large, it is always optimal to stipulate break-up fees in the employment contract.

6. Discussion and extensions

The analysis above highlights the trade-off with worker-job and worker-firm matching efficiencies that originates with the used of break-up fees in employment contracts. Recall that the inefficiencies in turnover arise due to two reasons: (i) the use of break-up fees makes the winner's curse problem with the raiding game more severe and (ii) when the break-up fee is sufficiently large, it may

[16]Note that a higher break-up fee may also reduce turnover and create worker-firm matching inefficiencies. But, as discussed in Appendix B, in the face of an increased ψ_1 a higher d^* improves the overall allocative efficiencies by trading off improved worker-job matching with turnover inefficiencies.

directly foreclose the raiders from the market. In this section, we consider a few extensions of our model that further elaborate on the foreclosure effect. We will first argue that when the matching gains are sufficiently important, in equilibrium, foreclosure may arise more frequently. We will then consider some different contracting environments such as the possibility for renegotiation and use of severance payments that do not have a market foreclosure effect even though the inefficiencies in both worker-job and worker-firm matching persist.

6.1. **Role of matching gains.** What happens to the optimal break-up fee as the potential matching gains become larger? In our model it is straightforward to parameterize the range of matching gains as $[-\mu, \mu]$, say, where $\mu > 0$. The model analyzed so far corresponds to the case where $\mu = 1$. As the following proposition shows, the firm would use higher break-up fees when matching gains are potentially large. And as a consequence, market foreclosure becomes more likely in equilibrium (recall that the market may be foreclosed whenever the equilibrium break-up fee $d^* > \psi_1$).

Proposition 4. *For any ψ_1, there exists a value of $\mu > 1$, say, $\hat{\mu}$, such that for all $\mu > \widehat{\mu}$, $d^* \geq \psi_1$.*

The intuition behind this finding is simple. For large μ, the gains from efficient worker-firm matching is also large. In other words, when d is raised, the expected loss of surplus due to the winner's curse effect is smaller.[17] Thus, the firm now has a greater incentive to increase d to ensure improved worker-job matching. As a consequence, the optimal d increases, and when μ is large enough (for a given ψ_1), the optimal d will always be above ψ_1 leading to market foreclosure.

Note that in our original model, market foreclosure arises in equilibrium when ψ_1 is sufficiently large ($\psi_1 > \underline{\psi}_1$). Proposition 4 suggests that if one allows for a broader range of matching gains, market foreclosure may occur even for low values of ψ_1.

6.2. **Renegotiation of break-up fee.** Observe that the market foreclosure effect stems from the fact that when the break-up fee is set at a sufficiently high level, the raiders need to raise their bid significantly in order to successfully hire the worker. And unless the matching gains are substantially large, it is not worthwhile for the raiders to do so. But note that if it is efficient for the worker to leave for the raiders, it would be profitable for the firm to let the worker go provided the firm can extract the matching gains generated through turnover. One way to do so is to renegotiate the initial contract if the (promoted) worker receives a better offer from the market. In what follows, we explore the role of renegotiation in our model and argue that with renegotiation break-up fees never foreclose the market.[18]

Suppose that the firm and worker can renegotiate the amount of the break-up fee if the worker receives an external offer. All other aspects of the model are kept unchanged and we assume that the firm continues to have the entire bargaining power even at the renegotiation stage. Note that the possibility of renegotiation makes a difference in our initial analysis only in the case where $b < d$. In our initial model, if $b < d$, the worker necessarily stays with the firm. But with renegotiation, the firm would lower d and let the worker leave if it is optimal for the firm-worker coalition to do so. This happens whenever $b > a\psi_1$, i.e., the bid exceeds the worker's value with the firm. At the renegotiation stage the firm sets $d = b$ to extract the matching gain and the worker leaves for the raiding firm. So, *irrespective of the value of d*, the worker stays with the firm if and only if $\psi_2 a > b$, or $a \geq \hat{a}(b) := b/\psi_2$. Note that with renegotiation, the market's bid need not exceed the break-up fee for the raid to be successful. Whenever the market offers $b > a\psi_2$, it will successfully raid workers with ability $a \in [a^*, \widehat{a}(b))$, a^* being the ability threshold for promotion.

[17] That is, for μ sufficiently large, the ex-ante probability that a and m are such that the firm will retain a promoted worker (in equilibrium) becomes relatively small.

[18] However, turnover continues to be inefficient due to the information asymmetries in the offer-counteroffer game.

Given the above observation, the subsequent derivation of the optimal contract parallels our analysis of the initial model. As before, competition ensures that the market's equilibrium bid, b^*, is equal to its expected payoff from bidding, i.e., $b^* = \mathbb{E}[\psi_2 a(1+m) \mid a \in [a^*, \widehat{a}(b^*))]$, or,

$$b^*(m, a^*) = \begin{cases} 0 & \text{if } m \leq 0 \\ \psi_2 a^* \frac{1+m}{1-m} & \text{if } 0 < m < \frac{1-a^*}{1+a^*} \\ \psi_2(1+m)\frac{1}{2}(a^*+1) & \text{if } m > \frac{1-a^*}{1+a^*} \end{cases} .$$

Two issues are important to note: first, b^* does not depend on d when renegotiation is allowed. This finding is intuitive as the market's bid no longer has to exceed d to raid the worker. Second, the market is never foreclosed in equilibrium—the raiders always bid for the worker as long as there are matching gains. The latter observation is an immediate implication of the former; as the bid need not have to exceed the break-up fee for a successful raid, it is always optimal for the raiders to place a bid if there are matching gains. This finding is reminiscent of the result discussed in Spier and Whinston (1995) who argue that in a model of bilateral trade with potential entrants, any break-up fee specified by the seller *does not* foreclose the market for a more efficient entrant if the buyer and the seller can renegotiate the break-up fee up on entry.

Now, as break-up fees never foreclose the market, our earlier discussion on the optimal break-up fee suggests that the firm chooses d such that $a^*(d) = a^{SB}$, whenever such a value of d is feasible. And whenever such a value of d is feasible in our initial model, it is also feasible even when renegotiation is allowed. To see this, note that given b^*, we can compute the firm's profit from promoting the marginal worker (with ability a^*) as:

$$\pi_p(a^*, m, d; a^*) = \begin{cases} \psi_2 a^* & \text{if } m \leq 0 \\ \min\{b^*(m, a^*), d\} & \text{if } m > 0 \end{cases} .$$

Observe that π_p is always larger than its "no-renegotiation" counterpart.[19] This is because with renegotiation, when $\psi_2 a^* < b^* < d$, the worker leaves the firm and the firm collects b^* whereas without renegotiation, the worker stays back and the firm earns only $\psi_2 a^*$. As the expected profit from promotion is higher with renegotiation, the firm has a stronger incentive to promote a worker, i.e., with renegotiation the equilibrium promotion threshold $a^*(d)$ is always lower than that without renegotiation. But we have already argued that in the absence of renegotiation, it is feasible to set $a^*(d) = a^{SB}$. Since $a^*(0)$ is the same with or without renegotiation (trivially, renegotiation does not play any role when no break-up fee is specified) and for any $d > 0$, a^* is lowered when renegotiation is allowed, it must be still feasible to set $a^*(d) = a^{SB}$.

So, one may conclude that in the presence of renegotiation, a better matched raider is never foreclosed from the market. However, both worker-job and worker-firm matching continue to remain inefficient (i.e., in equilibrium $a^*(d) \neq a^{FB}$) due to the winner's curse problem at the offer-counteroffer stage.[20]

[19] The derivation of π_m^p is straightforward. If $m < 0$ the firm makes $a^*\psi_2$ on him and when $m > 0$, the firm makes $\min\{b, d\}$ while the worker always leaves for the raider (if the market offers $b > d$, the firm collects d and if $b < d$, renegotiation implies that the firm sets $d = b$).

[20] Spier and Whinston also note that even with renegotiation, the market foreclosure effect reappears if the seller needs to make relationship specific investments and the entrant has some market power. In the context of our model, this finding suggests that if the initial employer invests in its worker for firm-specific human capital accumulation and if the raider can make take-it-or-leave-it offer, then contract renegotiation need not rule out the possibility of market foreclosure. A complete analysis of this issue is beyond the scope of this article and remains an interesting topic for future research.

6.3. **Contracts with severance payments.** The key role of the break-up fee that we highlight here is that it shields the promoted worker from the outside labor market, and, as a result, improves the worker-job matching efficiencies. But the break-up fees need not be the only contracting device that achieves this goal. The same can be achieved with, for example, severance payments. The firm may commit to make these lump-sum payments to the worker (depending on his job assignment) when the employment relation terminates in period two. However, the payments are made *irrespective* of whether the worker stays with the firm in period two (and leaves at the end of period) or leaves at the beginning of the period to join the raider's firm.

Let s_1 and s_2 be the severance payments in job 1 and 2 respectively. Relative to our initial model, we now rule out break-up fees but keep all other aspects of the model unchanged. Note that as the severance payments are made regardless of whether the worker stays or not, these payments do not affect the worker's decision on whether to switch employers. So the worker's choice of period two employer depends solely on the wage proposed by the firm in period two and the wage offer made by the raiders. The severance payments also do not affect the firm's counteroffer.[21] So, in order to derive the equilibrium promotion policy, we can continue to use our initial analysis and set $d = 0$.

This observation has two important implications: (i) Plugging $d = 0$ in the bidding function (equation (2)), we obtain:

$$b^*(m;a^*) \;=\; \begin{cases} 0 & \textit{if } m \leq 0 \\ \psi_2 a^* \frac{1+m}{1-m} & \textit{if } m > 0 \textit{ and } a^* < \frac{1-m}{1+m} \\ \psi_2(1+m)(\frac{a^*+1}{2}) & \textit{otherwise} \end{cases}$$

$$b^*(m;a^*) \;=\; \begin{cases} 0 & \textit{if } m \leq 0 \\ \psi_2 a^* \frac{1+m}{1-m} & \textit{if } m > 0 \textit{ and } a^* < \frac{1-m}{1+m} \\ \psi_2(1+m)(\frac{a^*+1}{2}) & \textit{otherwise} \end{cases} .$$

But the bidding behavior implies that there is no market foreclosure. The raiders always make a bid whenever there are matching gains. (ii) The equilibrium promotion rule depends on the difference of the severance payments across the two jobs, $\Delta s := (s_2 - s_1)$. To see this, note that the firm's profits associated with promoting and not promoting the marginal worker are given by $\pi_p = \frac{1}{2}a^*\psi_2 - s_2$ and $\pi_{np} = \psi_1 - s_1$, respectively. And the equilibrium promotion rule solves $\pi_{np} = \pi_p$, which implies that

$$a^* = \frac{2\psi_1}{\psi_2} + \frac{2\Delta s}{\psi_2}.$$

So, by choosing Δs the firm can implement any promotion rule (a^*) in equilibrium. As the market foreclosure effects are absent, it is always optimal to choose Δs such that $a^* = a^{SB}$.

In this context, it is important to note the following. First, the optimal contract with severance payment is a (weakly) more efficient than the optimal contract with break-up fees as it never forecloses the market and always guarantees the second-best. But note that similar to the case of renegotiation, both worker-job and worker-firm matching remain inefficient due to the winner's curse problem in the offer-counteroffer stage. Second, in equilibrium $\Delta s < 0$; that is, the firm commits to a larger severance pay in job 1 compared to job 2. As a result, at the beginning of period two, the firm creates a stronger incentive for itself to promote the worker. Finally, even though the use of the severance payments appear to be more efficient than the use of break-up fees, it has its own issues. The contract with severance payment is profitable provided that the firm can ex-ante recover such payments by lowering the period-one wage of the worker. As these

[21] Indeed, if we assume that the severance payments are paid immediately after the promotion decision, it is clear that they will not affect subsequent behavior of the firm and worker.

payments are made to all workers irrespective of their ability and job assignments, it would require the firm to significantly lower the worker's period-one wage to extract all rents. So, if the worker has liquidity constraints, such a low period-one wage may not be feasible and the optimal contract may still fall short of achieving even the second-best promotion policy.[22]

7. Conclusion

Break-up fee and more generally, deferred compensation, is a contracting tool that firms frequently use to retain their key employees. Such contracts dissuade potential raiders by making successful raids more expensive—the raiders need to compensate the employee for the break-up fee if they were to induce him to switch employers. This article highlights a novel trade-off associated with the use of such break-up fees and draws out the impact of such fees on efficiency and social welfare.

As argued by Waldman (1984), if there is an information asymmetry between the employer and the outside labor market regarding the quality of the workers, the firm's promotion decision may signal the market about a worker's quality. As promotion signals higher quality, the market bids up the wage of a promoted worker. Consequently, promotion becomes more expensive and too few workers are promoted in equilibrium. Contracts with break-up fees can resolve such worker-job matching inefficiencies by creating a wedge between the market offer and what the firm needs to pay in order to retain the worker. But in the presence of firm-specific matching gains, the improved worker-job matching comes at the cost of inefficient worker-firm matching. With a high break-up fee, the firm becomes more indiscriminate in its promotion decision and consequently, the expected quality of the promoted workers decreases. In response, the market lowers its bid for the promoted workers as it continue to suffer from an informational disadvantage. As a result, the firm retains too many workers by making counteroffers. And when the break-up fee is sufficiently large, it may directly foreclose the market for the raiders—unless the matching gains are significantly large, the raiders may find it unprofitable to bid for the worker. The optimal break-up fee balances this trade-off.

Our key finding is that the optimality of the break-up fee depends on the relative size of the worker's expected productivity across jobs. If there is substantial (expected) productivity gains from promotion, then it is never optimal to specify any break-up fee in the employment contract. Our analysis also suggests that the less disparate is the worker's (expected) productivity across jobs, the higher is the optimal break-up fee likey to be. Moreover, the use of break-up fee increases the aggregate social welfare by trading off gains in job assignment efficiencies with the loss from inefficient turnover.

There are several other economic effects that are interesting and relevant in our environment albeit beyond the scope of our model. One may assume that to be productive in the "post-promotion" job, it is necessary that the worker (and/or the firm) invests in human capital. How would the presence of break-up fees affect the incentives for investment? The answer to this question depends on whether the human capital is general or firm-specific and who undertakes the investments.[23] Also, as noted earlier, in the presence of investments in human capital, contracts with break-up fees may fail to achieve second-best even if renegotiation is allowed. Another interesting generalization of our model is to allow the firm to set the break-up fee *after* observing the ability of the worker. Such flexibility in setting the break-up fee may make the contracts more effective in shielding the workers from the raider (e.g., the firm may stipulate a higher fee for the more productive workers). The analysis of such a case, however, must also account for an additional

[22] Liquidity constraints can be less binding under contracts with break-up fee as the worker may have lower rents in period two (hence, period-one wage need not have to be lowered as much to ensure complete rent extraction).

[23] Golan (2005) addresses these issues in a related environment but does not consider break-up fees or matching gains with the outside labor market. Also see Bernhardt and Scoones (1998) for a related discussion on the incentives to invest on human capital.

effect: the signaling role of the contracts. If the break-up fee is set after observing the worker's ability, the market gets an additional (and perhaps more precise) signal on the worker's productivity from the contract (i.e., the stipulated break-up fee) he has with his initial employer.[24] It would also be interesting to consider the case where the market can screen the promoted workers (see Ricart i Costa (1988) for a related model on managerial job assignment). Here, the firm's promotion policy continues to play an important role as it can affect that information rent that the worker earns from the market (which, in turn, can be extracted by the initial employer). Finally, if there is a moral hazard problem in the production process, the use of break-up fee may create an additional cost: it mutes work incentives by dampening the raiders' bid, and therefore, lowering the prospect to future wage increments (see, Kräkel and Sliwka (2009) for a similar discussion).

The issues raised above offer useful directions for future research and may offer additional insights into the firm's job assignment policies. However, the key trade-off between the worker-job and the worker-firm matching that we highlight in this article continues to play a critical role in all these setting and we expect our findings to be informative in analyzing such complex environments.

Appendix A

This appendix contains the proofs omitted in the text.

Proof of Proposition 1. Using (5) we can obtain $\mathbb{E}_m\pi_p(a^*, m, d)$. When $d < \psi_2$, $\mathbb{E}_m\pi_p(a^*, m, d)$ $= a^*\psi_2 \times \Pr[m \leq \max\{0, (d - a^*\psi_2)/(d + a^*\psi_2)\}] + d \times \Pr[m > \max\{0, (d - a^*\psi_2)/(d + a^*\psi_2)\}]$. The exact values of the probabilities above depend on whether $d - a^*\psi_2$ is positive or not. By considering the two cases, we obtain that

$$\mathbb{E}_m\pi_p(a^*, m, d) = \begin{cases} 2d\psi_2 a^*/(d + a^*\psi_2) & \text{if} \quad a^* \leq d/\psi_2 \\ (a^*\psi_2 + d)/2 & \text{if} \quad a^* > d/\psi_2 \end{cases} .$$

Consider now the case where $d > \psi_2$. In this case, $\mathbb{E}_m\pi_p(a^*, m, d) = a^*\psi_2 \Pr[m \leq 2d/(\psi_2(1 + a^*))$ $- 1] + d \times \Pr[m > 2d/(\psi_2(1 + a^*)) - 1]$. When $d > \psi_2(1 + a^*)$, $m \leq 2d/(\psi_2(1 + a^*)) - 1$ for all possible realizations of m. Hence, when $d > \psi_2$,

$$\mathbb{E}_m\pi_p(a^*, m, d) = \begin{cases} a^*\psi_2 & \text{if} \quad a^* \leq d/\psi_2 - 1 \\ da^*/(1 + a^*) + d - d^2/(\psi_2(1 + a^*)) & \text{if} \quad a^* < d/\psi_2 - 1 \end{cases} .$$

Using the above characterization of $\mathbb{E}_m\pi_p(a^*, m, d)$, it is easy to obtain that $a^*(d)$ (as presented in the proposition) is the solution to $\mathbb{E}_m\pi_p(a^*, m, d) = \psi_1$ for each value of d. ■

Proof of Lemma 1. Suppose first that $d < \psi_2$. We know from (2) that $b^*(m, d; a^*) = 0$ if and only if $m \leq 0$ or $a^* \leq d(1-m)/(\psi_2(1+m))$. The second inequality is equivalent to $m \leq (d - a^*\psi_2)/(d + a^*\psi_2)$. Thus, $b^*(m, d; a^*) = 0$ if and only if $m \leq \max\{0, (d - a^*\psi_2)/(d + a^*\psi_2)\}$. From Proposition 1, it follows that in equilibrium $a^* = (2\psi_1 - d)/\psi_2$ if $d < \psi_1$ and $a^* = \psi_1 d/(2d\psi_2 - \psi_1\psi_2)$ if $\psi_1 \leq d < \psi_2$. Replacing a^* with its equilibrium value in this condition, it becomes $m \leq \max\{0, (d - \psi_1)/\psi_1\}$ if $d < \psi_1$ and $m \leq \max\{0, 1 - \psi_1/d\}$ if $\psi_1 \leq d < \psi_2$. When $d < \psi_1$, then $(d - \psi_1)/\psi_1 < 0$, which implies that $b^*(m, d; a^*(d)) = 0$ if and only if $m \leq 0$. If $\psi_1 \leq d < \psi_2$, then $1 - \psi_1/d \geq 0$, which implies that $b^*(m, d; a^*(d)) = 0$ if and only if $m \leq 1 - \psi_1/d$.

Suppose now that $d \geq \psi_2$. We follow the same steps as above. From (2) we know that $b^*(m, d; a^*) = 0$ if and only if $m \leq \max\{0, 2d/(\psi_2(1 + a^*)) - 1\}$. Clearly, when $d \geq \psi_2$, then $2d/(\psi_2(1 + a^*)) - 1 \geq 0$. Hence, $b^*(m, d; a^*) = 0$ if and only if $m \leq 2d/(\psi_2(1 + a^*)) - 1$. From Proposition 1, it follows that in equilibrium $a^* = \left(\psi_1 - d + d^2/\psi_2\right)/(2d - \psi_1)$ when $\psi_2 \leq d < \psi_1$. Replacing a^* with its equilibrium value, that inequality becomes $m \leq (3d - 2\psi_1 - \psi_2)/(d + \psi_2)$. Finally, When $d > \psi_1 + \psi_2$, in equilibrium $a^* = \psi_1/\psi_2$ and $2d/(\psi_2(1 + a^*)) - 1 > 1$. ■

[24]See Bernhardt and Scoone (1993) for a discussion on such signaling role of contracts in reducing turnover.

Proof of Lemma 2. In what follows, we will prove a more general version of this lemma that would be useful later in proving subsequent results (we do not present this version in text for expositional clarity). We will show that the firm's expected profit function $\tilde{\Pi}$ is continuous in d and given by the following functional form:

$$\tilde{\Pi}(d) = \begin{cases} \psi_1 + \widehat{S}\left(a^*\left(d\right)\right) & \text{if } 0 \leq d < \psi_1 \\ \psi_1 + \widehat{S}\left(a^*\left(d\right)\right) - H(a^*\left(d\right), d) & \text{if } \psi_1 \leq d < \psi_2 \\ \psi_1 + \widehat{S}\left(a^*\left(d\right)\right) - J(a^*\left(d\right), d) & \text{if } \psi_2 \leq d < \psi_2 + \psi_1 \\ \left(\psi_1 + \psi_2\right)^2 / 2\psi_2 & \text{otherwise} \end{cases},$$

where $H : [0,1] \times \mathbb{R}_+ \to \mathbb{R}$ and $J : [0,1] \times \mathbb{R}_+ \to \mathbb{R}$. Furthermore, (i) $H(a^*\left(d\right), d) > 0$ for all $d \in (\psi_1, \psi_2)$, (ii) $J(a^*\left(d\right), d) > 0$ for all $d \in [\psi_2, \psi_2 + \psi_1)$, and (iii) for any $d_1 \in [\psi_1, \psi_2)$ and $d_2 \in [\psi_2, \psi_2 + \psi_1)$ such that $a^*(d_1) = a^*(d_2)$, $J(x, d_2) > H(x, d_1)$.

Under the optimal contract, w_1 is such that the agent's individual rationality constraint is binding. Moreover, in any equilibrium, the raiders' bid the entire expected production of the workers they successfully raid. Hence, both the worker's expected utility and the raiders' expected profit are zero, which implies that $\tilde{\Pi}(d)$ is always equal to the aggregate expected surplus. The aggregate expected surplus depends on the firm's promotion policy, the raiders' equilibrium decision to bid for a promoted worker and the firms' decision to make a counter-offer and retain the worker. The remainder of the proof consists of the following for steps.

Step 1: $\tilde{\Pi}$ *when* $d < \psi_1$. It follows from Lemma 1 that when $d < \psi_1$ the raiders bid for a promoted worker if and only if $m > 0$. Hence, there is no market foreclosure. This implies that the expected aggregate period-two surplus is $\widehat{S}(a^*(d))$ and $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d))$.

Step 2: $\tilde{\Pi}$ *when* $\psi_1 \leq d < \psi_2$. From Lemma 1, it follows that when $\psi_1 \leq d < \psi_2$ the raiders bid for a promoted worker if and only $m > 1 - \psi_1/d$. There is partial foreclosure, since a promoted worker is retained by F whenever $m \leq 1 - \psi_1/d$. The allocative difference between this case and that underlying the second best aggregate surplus is that when the realization of a and m is such that $a^*(d) \leq a \leq ((2d - \psi_1)/\psi_1)a^*(d)$ and $(a - a^*(d))/(a + a^*(d)) < m \leq 1 - \psi_1/d$ the worker is retained by F instead of joining a raider firm where he is more efficient by a factor of m. Hence, in this case, $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d)) - H(a^*(d), d)$, where

$$H(x, y) := \int_{x}^{\min\{\frac{2y-\psi_1}{\psi_1}x, 1\}} \int_{\frac{a-x}{a+x}}^{1-\frac{\psi_1}{y}} a\psi_2 m \frac{1}{2} dm da.$$

Clearly, $H(x, y) > 0$ for all $x \in [0, 1)$ and $y > \psi_1$.

Step 3: $\tilde{\Pi}$ *when* $\psi_2 \leq d < \psi_1 + \psi_2$. From Lemma 1, it follows that when $\psi_2 \leq d < \psi_1 + \psi_2$ the raiders bid for a promoted worker if and only $m > (3d - 2\psi_1 - \psi_2)/(d + \psi_2)$. There is partial foreclosure, since a promoted worker is retained by F whenever $m \leq (3d - 2\psi_1 - \psi_2)/(d + \psi_2)$. The allocative difference between this case and that underlying the second best aggregate surplus is that when the realization of a and m is such that $a^*(d) \leq a \leq 1$ and $(a - a^*(d))/(a + a^*(d)) < m \leq (3d - 2\psi_1 - \psi_2)/(d + \psi_2)$ the worker is retained by F instead of joining a raider firm where he is more efficient by a factor of m. Hence, in this case, $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d)) - J(a^*(d), d)$, where

$$J(x, y) := \int_{x}^{1} \int_{\frac{a-x}{a+x}}^{\frac{3y-2\psi_1-\psi_2}{y+\psi_2}} a\psi_2 m \frac{1}{2} dm da.$$

Observe that $a^*(d) > a^*(\psi_2) = \psi_1/(2\psi_2 - \psi_1)$ when $d > \psi_2$. Observe also that $(3y - 2\psi_1 - \psi_2)/(y + \psi_2) > (a - x)/(a + x)$ for all $a \in [x, 1)$, when $x \in [\psi_1/(2\psi_2 - \psi_1), 1)$ and $y > \psi_2$. Hence, $J(x, d) > 0$ for all $x \in [\psi_1/(2\psi_2 - \psi_1), 1)$ and $y > \psi_2$.

Step 4: $\tilde{\Pi}$ *when* $d \geq \psi_1 + \psi_2$. It follows from Lemma 1 that when $d \geq \psi_1 + \psi_2$ the raiders bid for a promoted worker if and only $m > 1$. Hence,There is full market foreclosure since all promoted workers are retained by F. The allocative difference between this case and that underlying the second best aggregate surplus is that when the realization of a and m is such that $a^*(d) \leq a \leq 1$ and $(a - a^*(d))/(a + a^*(d)) < m \leq 1$ the worker is retained by F instead of joining a raider firm where he is more efficient by a factor of m. Hence, in this case, $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d)) - L(a^*(d), d)$, where

$$L(x, y) := \int_x^1 \int_{\frac{a-x}{a+x}}^1 a\psi_2 m \frac{1}{2} dm da.$$

Clearly, $\tilde{\Pi}(d) = \psi_1 + \psi_1 a^*(d) + \int_{a^*(d)}^1 \int_{-1}^1 \frac{1}{2}\psi_2 a dm da = \psi_1 + \psi_1 a^*(d) - \frac{1}{2}\psi_2 \left(a^*(d) - 1\right)\left(a^*(d) + 1\right)$. Since $a^*(d) = \psi_1/\psi_2$ when $d \geq \psi_1 + \psi_2$, we obtain that $\tilde{\Pi}(d) = \left(\psi_1 + \psi_2\right)^2 / 2\psi_2$.

Step 5: *Comparing* $H(x, y)$ *with* $J(x, y)$. To compare $H(x, y)$ with $J(x, y)$ one needs to compare the upper limits of integration. Clearly, $\min\{x(2y - \psi_1)/\psi_1, 1\} \leq 1$ and $1 - \psi_1/y < \left(\psi_2 - \psi_1\right)/\psi_2 = \min_{\widehat{y} \geq \psi_2}(3\widehat{y} - 2\psi_1 - \psi_2)/(\widehat{y} + \psi_2)$ for all $y < \psi_2$. Since both in $H(x, y)$ and in $J(x, y)$ the integrand is the same and is positive, we obtain that $H(x, y_1) > J(x, y_2)$ for all $y_1 \in [\psi_1, \psi_2)$, $y_2 \geq \psi_2$, and $x \in [0, 1]$. ■

Proof of Proposition 2. The proof is given in the following steps.

Step 1: *There exists* $\underline{\psi}_1$ *such that the optimal damage fee is strictly positive if* $\psi_1 > \underline{\psi}_1$. Using Lemma 2 we obtain that $\tilde{\Pi}'(0) = \widehat{S}'(a^*(0))a^{*\prime}(0)$. We know that $a^{*\prime}(0) = -1/\psi_2 < 0$. Moreover, computing the integrals in $\widehat{S}$, we obtain that

$$\widehat{S}(x) = \psi_1 x - \frac{1}{2}\frac{\left(2x + 1\right)^2 \left(x - 1\right)\psi_2}{(x + 1)} + 2x^2\psi_2 \ln\frac{2x}{x + 1}. \tag{9}$$

Using (9), we obtain that

$$\widehat{S}'(a^*(0)) = \widehat{S}'(2\psi_1/\psi_2) = \frac{\left(\psi_2^3 - 28\psi_1^3 + 5\psi_1\psi_2^2 - 12\psi_1^2\psi_2\right)}{\left(2\psi_1 + \psi_2\right)^2} + 8\psi_1 \ln\frac{4\psi_1}{2\psi_1 + \psi_2},$$

which, in turn, can be used to obtain that $\lim_{\psi_1 \to 0} \widehat{S}'(a^*(0)) = \psi_2 > 0$ and that $\lim_{\psi_1 \to \psi_2/2} \widehat{S}'(a^*(0))$ $= -3\psi_2/4 < 0$. Next, note that $\widehat{S}'(a^*(0))$ is continuous and decreasing in ψ_1 in $(0, \psi_2/2]$. Thus, there exists $\underline{\psi}_1 \in (0, \psi_2/2]$ such that $\widehat{S}'(a^*(0)) < 0$ if and only if $\psi_1 > \underline{\psi}_1$. Hence, for $\psi_1 > \underline{\psi}_1$, $\tilde{\Pi}'(0) > 0$, meaning that the optimal damage fee is strictly positive.

Step 2: *The optimal damage fee is strictly positive only if* $\psi_1 > \underline{\psi}_1$. To show this, we show the equivalent statement that the optimal damage fee is zero if $\psi_1 \leq \underline{\psi}_1$. Lemma 2, together with the fact that for all $d_1 \in (\psi_2, \psi_2 + \psi_1]$ there exists $d_2 \in [\psi_1, \psi_2)$ such that $a^*(d_1) = a^*(d_2)$ (see the version of Lemma 2 given in this Appendix), implies that $d > \psi_2$ is never optimal. Hence, it remains to show that $0 < d \leq \psi_2$ is not optimal either. If $\psi_1 \leq \underline{\psi}_1$, then $\widehat{S}'(a^*(0)) \geq 0$ (see the analysis in Step 1). Moreover, observe that $\widehat{S}''(x) = \psi_2\left(-2\left(x + 1\right)^{-3}\left(x + 4x^2 + 2x^3 - 2\right) + 4\ln 2 + 4\ln\frac{x}{x+1}\right) < 0$ for all $x \in [0, 1]$, meaning that $\widehat{S}(x)$ is concave. Hence, when $\psi_1 \leq \underline{\psi}_1$, $\widehat{S}'(x) > 0$ for all $x < a^*(0)$. Since $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d))$ and $a^*(d) < a^*(0)$ for all $d \in (0, \psi_1)$, then $\tilde{\Pi}'(d) = \widehat{S}'(a^*(d))a^{*\prime}(d) = \widehat{S}'(a^*(d)) \times (-1/\psi_2) < 0$ for all $d \in (0, \psi_1)$, which means that the optimal $d \notin (0, \psi_1)$. Consider now the case of $\psi_1 \leq d < \psi_2$. In this case, $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d)) - H(a^*(d), d)$ where $H(a^*(d), d) \leq 0$

(see the proof of Lemma 2). Hence, $\tilde{\Pi}(d) \leq \psi_1 + \widehat{S}(a^*(d))$. Moreover, concavity of $\widehat{S}(x)$, and the fact that $a^*(d) < a^*(0)$ and $a^{*\prime}(d) < 0$ when $\psi_1 \leq d < \psi_2$, implies that $\psi_1 + \widehat{S}(a^*(d))$ is decreasing in d in when $\psi_1 \leq d < \psi_2$. Thus, $\tilde{\Pi}(d) < \psi_1 + \widehat{S}(a^*(\psi_1)) < \psi_1 + \widehat{S}(a^*(0)) = \tilde{\Pi}(0)$ if $\psi_1 \leq d < \psi_2$. Finally, because $\tilde{\Pi}$ is continuous at $d = \psi_2$, then $\tilde{\Pi}(\psi_2) = \lim_{d \to \psi_2} \tilde{\Pi}(d) < \tilde{\Pi}(0)$. This, together with Step 1 above, establishes that the optimal damage fee is strictly positive if and only if $\psi_1 > \underline{\psi}_1$.

Step 3: *Aggregate surplus increases with inclusion of damage fee when* $\psi_1 > \underline{\psi}_1$. This follows from the fact that the firm's profit is identical to the expected aggregate surplus. ■

Proof of Proposition 3. In what follows, we use $f'(x^-)$ and $f'(x^+)$ to denote, respectively, the left and the right derivative of a function f at point x. The proof is given in the following steps.

Step 1: $\tilde{\Pi}$ *is differentiable at* $d = \psi_1$ *and* $\tilde{\Pi}'(\psi_1) = \widehat{S}'(a^*(\psi_1)) \times (-1/\psi_2)$. Since $\widehat{S}$ is differentiable at $a^*(\psi_1)$, $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d))a^*(d)$ when $d < \psi_1$, and $\tilde{\Pi}$ is continuous at ψ_1, then

$$\tilde{\Pi}'(\psi_1^-) = \widehat{S}'(a^*(\psi_1))a^{*\prime}(\psi_1^-) = \widehat{S}'(a^*(\psi_1)) \times (-1/\psi_2).$$

When $d \geq \psi_1$, $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d)) - H(a^*(d), d)$, where H is as defined in the proof of Lemma (2). Hence,

$$\tilde{\Pi}'(\psi_1^+) = \widehat{S}'(a^*(\psi_1))a^{*\prime}(\psi_1^+) - H_x'(a^*(\psi_1), \psi_1^+)a^{*\prime}(\psi_1^+) - H_y'(a^*(\psi_1), \psi_1^+).$$

Differentiating H with respect to x and to y,we obtain, respectively,

$$H_x' = -\int_0^{1-\frac{\psi_1}{y}} x\psi_2 m \frac{1}{2} dm + \int_x^{\frac{2y-\psi_1}{\psi_1}x} \left(\frac{(a-x)\,\psi_2 a^2}{(a+x)^3} \right) da \tag{10}$$

and

$$H_y' = \int_x^{\frac{2y-\psi_1}{\psi_1}x} \frac{\psi_1}{y^2} a\psi_2 (1 - \frac{\psi_1}{y}) \frac{1}{2} da. \tag{11}$$

(Note that we are interested in the behavior of H in a neighborhood of $(x, y) = (a^*(\psi_1), \psi_1)$ where $(2y - \psi_1)x/\psi_1 < 1$. So, in the above computation of the partial derivatives of H we simply assume that $\min\{(2y - \psi_1)x/\psi_1, 1\} = (2y - \psi_1)x/\psi_1$.) Using, (10) and (11), we obtain that $H_x'(a^*(\psi_1), \psi_1^+) = H_y'(a^*(\psi_1), \psi_1^+) = 0$. Since, $a^{*\prime}(\psi_1^+) = -1/\psi_2 = a^{*\prime}(\psi_1^-)$, we obtain that

$$\tilde{\Pi}'(\psi_1^+) = \widehat{S}'(a^*(\psi_1)) \times (-1/\psi_2) = \tilde{\Pi}'(\psi_1^-),$$

which implies that $\tilde{\Pi}$ is differentiable at $d = \psi_1$ and $\tilde{\Pi}(\psi_1) = \widehat{S}'(a^*(\psi_1)) \times (-1/\psi_2)$.

Step 2: *There exists* $\overline{\psi}_1 > \underline{\psi}_1$ *such that* $\widehat{S}'(a^*(\psi_1)) > 0$ *if* $\psi_1 < \overline{\psi}_1$, $\widehat{S}'(a^*(\psi_1)) < 0$ *if* $\psi_1 > \overline{\psi}_1$, and $\widehat{S}'(a^*(\psi_1)) = 0$ if $\psi_1 = \overline{\psi}_1$. Concavity of $\widehat{S}$ and the fact that $a^*(\psi_1) < a^*(0)$ implies that $\widehat{S}'(a^*(\psi_1)) > \widehat{S}'(a^*(0))$. Since $\widehat{S}'(a^*(0)) \geq 0$ when $\psi_1 = \underline{\psi}_1$ (see proof of Proposition 2), it follows that $\widehat{S}'(a^*(\psi_1)) > 0$ when $\psi_1 = \underline{\psi}_1$. We next analyze $\widehat{S}'(a^*(\psi_1))$ when $\psi_1 \to \psi_2/2$. Using (9), we obtain that

$$\widehat{S}'(a^*(\psi_1)) = \widehat{S}'(\psi_1/\psi_2) = \psi_1 + \frac{(\psi_2^3 - 4\psi_1^3 + 2\psi_1\psi_2^2 - 4\psi_1^2\psi_2)}{(\psi_1 + \psi_2)^2} + 4\psi_1 \left(\ln \frac{2\psi_1}{\psi_1 + \psi_2} \right).$$

Using this result, we obtain that $\lim_{\psi_1 \to \psi_2/2} (\widehat{S}'(a^*(\psi_1))) = \psi_2 \left(36 \ln(2/3) + 13\right)/18 < 0$. Next, note that $\widehat{S}'(a^*(\psi_1))$ is continuous and decreasing in ψ_1 in $(0, \psi_2/2]$. Hence, there exists $\overline{\psi}_1 \in (\underline{\psi}_1, \psi_2/2]$ such that $\widehat{S}'(a^*(\psi_1)) > 0$ if $\psi_1 < \overline{\psi}_1$, $\widehat{S}'(a^*(\psi_1)) < 0$ if $\psi_1 > \overline{\psi}_1$, and $\widehat{S}'(a^*(\psi_1)) = 0$ if $\psi_1 = \overline{\psi}_1$.

Step 3: *For* $\underline{\psi}_1 < \psi_1 \leq \overline{\psi}_1$, *the optimal break-up fee* $d^* \in (0, \psi_1]$ *and* $a^*(d^*) = a^{SB}$. We know from Proposition 2 that $d^* > 0$ when $\psi_1 > \underline{\psi}_1$. We next show that if $\psi_1 \leq \overline{\psi}_1$, then

$d^* \notin (\psi_1, \psi_2)$. Let $\psi_1 \leq \overline{\psi}_1$ and $\psi_1 < d < \psi_2$. Since $\psi_1 \leq \overline{\psi}_1$, then $\widehat{S}'(a^*(\psi_1)) \geq 0$ (see Step 2). Concavity of $\widehat{S}(x)$ together with the fact that $a^*(d) < a^*(\psi_1)$ implies that $\widehat{S}'(a^*(d)) > 0$. Since $a^{*\prime}(d) < 0$, then $\widehat{S}'(a^*(d)) \times a^{*\prime}(d) < 0$, meaning that $\psi_1 + \widehat{S}(a^*(d))$ decreases with d. This, together with the fact that $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d)) - H(a^*(d), d)$ where $H(a^*(d), d) < 0$, implies that $\tilde{\Pi}(d) < \psi_1 + \widehat{S}(a^*(d)) < \psi_1 + \widehat{S}(a^*(\psi_1)) = \tilde{\Pi}(\psi_1)$. Hence $d^* \notin (\psi_1, \psi_2)$. Since $d^* < \psi_2$ by Lemma 2, this means that $d^* \in (0, \psi_1]$. We next show that $a^*(d^*) = a^{SB}$. When $\underline{\psi}_1 < \psi_1 \leq \overline{\psi}_1$, $\widehat{S}'(a^*(0)) < 0$ and $\widehat{S}'(a^*(\psi_1)) \geq 0$. Because $\widehat{S}$ is concave, $a^{SB} \in [a^*(\psi_1), a^*(0))$. Finally, since $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d))$ when $0 \leq d < \psi_1$, then $d^* = a^{SB}$ necessarily.

Step 4: *For $\psi_1 > \overline{\psi}_1$, $d^* \in (\psi_1, \psi_2)$ and $\Pi^* < \psi_1 + \widehat{S}(a^{SB})$.* We first show that $d^* > \psi_1$. $\widehat{S}'(a^*(\psi_1)) < 0$ when $\psi_1 > \overline{\psi}_1$. Concavity of $\widehat{S}$ together with the fact that $a^*(d) \geq a^*(\psi_1)$ for all $d \leq \psi_1$ implies that $\widehat{S}'(a^*(d)) < \widehat{S}'(a^*(\psi_1))$. Hence $\widehat{S}'(a^*(d)) < 0$ for $d \leq \psi_1$. Furthermore, $a^{*\prime}(d) < 0$ for $d \leq \psi_1$. Hence, $\tilde{\Pi}'(d) = \widehat{S}'(a^*(d))a^{*\prime}(d) > 0$ for $d \leq \psi_1$. Thus, $d^* > \psi_1$. Since $d^* < \psi_2$ by Lemma 2**,** then $d^* \in (\psi_1, \psi_2)$. Finally, observe that $\Pi^* = \psi_1 + \widehat{S}(a^*(d^*)) - H(a^*(d^*), d^*) < \psi_1 + \widehat{S}(a^*(d^*)) \leq \psi_1 + \widehat{S}(a^{SB})$, where the second inequality follows from the fact that a^{SB} is a maximizer of $\psi_1 + \widehat{S}(x)$. ■

Proof of Proposition 4. The raiders' equilibrium bidding, F's decision to retain a promoted worker, and F's profit from promoting the marginal worker remain the same as when $\mu = 1$. That is, they are given by (2), (3), and (5), respectively. In what follows, we focus on the case where $d < \psi_2$ and $\mu > 1$, as this is the relevant case for the proposition. Using (2), (3), and (5), we can obtain F's expected profit of promoting the marginal worker. This profit depends on μ, as the decision to promote a worker is taken before the realization of m. We consider two cases separately depending on the value of d. If $d < a^*\psi_2$, then $\mathbb{E}_m\pi_p(a^*, m, d) = a^*\psi_2 \times \Pr[m \leq 0] + d \times \Pr[m > 0] = (a^*\psi_2 + d)/2$. If $a^*\psi_2 \leq d < \psi_2$, then $\mathbb{E}_m\pi_p(a^*, m, d) = a^*\psi_2 \times \Pr[m \leq (d - a^*\psi_2)/(d + a^*\psi_2)] + d \times \Pr[m > (d - a^*\psi_2)/(d + a^*\psi_2)]$, which means that

$$\mathbb{E}_m\pi_p(a^*, m, d) = \begin{cases} a^*\psi_2 & \text{if} \quad \mu < (d - a^*\psi_2)/(d + a^*\psi_2) \\ (a^*\psi_2 - d)(\frac{d - a^*\psi_2}{d + a^*\psi_2} + u)/(2u) & \text{if} \quad \mu \geq (d - a^*\psi_2)/(d + a^*\psi_2) \end{cases}.$$

Using this characterization of the expected profit of promoting the marginal worker, we obtain the equilibrium promotion cutoff. It satisfies $\mathbb{E}_m\pi_p(a^*, m, d) = \psi_1$. The solution is given by

$$a^*(d) = \begin{cases} (2\psi_1 - d)/\psi_2 & \text{if} \quad d < \psi_1 \\ \frac{d\mu + d - \mu\psi_1 - \sqrt{\mu\left(4d^2 - 4d\psi_1 + \mu\psi_1^2\right)}}{\psi_2(1-\mu)} & \text{if} \quad \psi_1 \leq d < \min\{2\psi_1\mu/(\mu - 1), \psi_2\} \\ 0 & \text{if} \quad \min\{2\mu\psi_1/(\mu - 1), \psi_2\} \leq d < \psi_2 \end{cases}.$$

The second-best aggregate surplus is given by

$$\begin{aligned} \hat{S}(x) &= \psi_1 x + \int_x^1 \psi_2 a \left[\frac{1}{2}\int_{-\mu}^0 dm + \frac{1}{2}\int_0^\mu (1+m)dm\right] da - \int_x^1 \psi_2 a \left[\int_0^{\frac{a-x}{a+x}} \frac{1}{2} m dm\right] da \\ &= x\psi_1 + \frac{1}{4}\psi_2\left(1 - x^2\right) - \frac{1}{4\mu}\psi_2\left(4x - 3x^2 - 1\right) + \frac{1}{8\mu}\psi_2(\mu - 1)(\mu + 3)\left(1 - x^2\right) \\ &\quad + \frac{1}{2\mu} x \frac{\psi_2}{x+1}\left(x - 5x^2 + 4\right) + \frac{2}{\mu} x^2 \psi_2 \ln 2\frac{x}{x+1}. \end{aligned}$$

As before, $\tilde{\Pi}(d) = \psi_1 + \hat{S}(a^*(d))$ when $d < \psi_1$ and $\tilde{\Pi}(d) = \psi_1 + \hat{S}(a^*(d)) - H(a^*(d), d)$ when $\psi_1 \leq d < \psi_2$ (see the proof of Lemma 2). The remainder of the proof is similar to the proof of Proposition 3 and therefore omitted. We simply note here that using the above characterization of $\hat{S}(x)$ we obtain that $\lim_{\mu\to\infty} \hat{S}'(a^*(\psi_1)) = \lim_{\mu\to\infty} \hat{S}'(\psi_1/\psi_2) = -\infty$. Hence, since $a^{*\prime}(\psi_1) = -1/\psi_2 < 0$, then $\lim_{\mu\to\infty} \tilde{\Pi}'(\psi_1) = +\infty$, which implies that exists $\widehat{\mu}$ such that $\tilde{\Pi}'(\psi_1) > 0$ for $\mu > \widehat{\mu}$. ■

Appendix B

As discussed in section 5, intuition suggests that the equilibrium break-up fee (d^*) is increasing in ψ_1. While such an argument is also corroborated by numerical solutions, the comparative statics properties of d^* remain analytically intractable for most parts. In this appendix, we further elaborate on the intuition developed in section 5. The discussion below succinctly draws out the salient economic effects that may make d^* increasing in ψ_1. We also show that for low values of ψ_1, this claim can indeed be analytically proved even though it remains algebraically intractable for high values of ψ_1.

In order to elaborate on the comparative statics property of d^*, we first consider the case were $\psi < \overline{\psi}_1$. Recall that in this case, the equilibrium $d^* < \psi_1$ and there is no market foreclosure. As discussed in section 4, in this case the optimal break-up fee solves $a^*(d) = a^{SB}$. Now, as ψ_1 increases, keeping the worker in job 1 becomes more lucrative; in other words, the ability threshold for promotion (a^*) increases. But for a given d^*, one can argue that a^* increases more than a^{SB}. To see this, note that a^* is the ability threshold for which the firm is indifferent between promoting the worker and keeping him in job 1, i.e., at a^*, we must have

$$\psi_2 a^* \Pr(m < 0) + d \Pr(m > 0) = \psi_1. \tag{$A1$}$$

The above equation suggests that the marginal impact of a^* on the firm's payoff from promoting the worker is muted by the fact that a^* affects the firm's payoff only when the worker is a worse match for the market (i.e., $m < 0$). When the worker is a better match for the market (i.e., $m > 0$), the firm earns d irrespective of the value of a^* (i.e., the marginal worker's ability does not matter).

In contrast, a^{SB} equalizes the (ex-ante) *social* return on the worker across the two jobs (subject to inefficiencies in offer-counteroffer game). So, similar to equation ($A1$) above, the promotion threshold a^{SB} sets the the expected productivity of the worker in job 2 (conditional on being promoted) is equal to ψ_1.

When $m < 0$, it is efficient for the worker to stay with the firm where he produces $\psi_2 a^{SB}$. In this case, the marginal impact of a^* on the firm's payoff is the same as the marginal impact of a^{SB} on the social return—both are equal to ψ_2. But when $m > 0$, an increase in a^* does not change the the firm's return on the worker which is (in this case) fixed at d while an increase in a^{SB} continues to affect the expected productivity of a promoted worker—the worker's expected productivity in job 2 increases and the extent of the inefficiency in turnover (conditional of being promoted) diminishes.

So, the marginal impact of a^{SB} on the *social return* from promotion is higher than the marginal impact of a^* on the *firm's return* (or payoff) from promotion. Consequently, as ψ_1 increases, a^* needs to increase more than a^{SB} in order to maintain the equality between the return on the worker across the two jobs. As the optimal contract requires $a^* = a^{SB}$, and a^* is decreasing in d over the relevant range of values (i.e., $d < \psi_2$), the firm must increase d in order to bring down a^* to a^{SB}.

Indeed, when $\psi < \overline{\psi}_1$, using the above line of argument we can analytically show that d^* is increasing in ψ_1.

Proposition 5. *If* $\psi_1 \in (\underline{\psi}_1, \overline{\psi}_1)$, d^* *is increasing in* ψ_1.

Proof. When $\psi_1 \in (\underline{\psi}_1, \overline{\psi}_1)$, $d^* \in (0, \psi_1)$ and satisfies $\tilde{\Pi}'(d^*) = 0$. Since ψ_1 affects $\tilde{\Pi}$ and, consequently, $\tilde{\Pi}'$, this condition implicitly defines d^* as a function of ψ_1. Using the Implicit Function Theorem, we obtain that $\partial d^*/\partial \psi_1 = -\tilde{\Pi}_{d\psi_1}/\tilde{\Pi}_{dd}$. Since $d^* \in (0, \psi_1)$, for the relevant values of d, $\tilde{\Pi}(d) = \psi_1 + \widehat{S}(a^*(d))$ and $a^*(d) = (2\psi_1 - d)/\psi_2$. Hence,

$$\tilde{\Pi}_{dd} = \widehat{S}_{a^*a^*} a_d^{*2} + \widehat{S}_{a^*} a_{dd}^*.$$

Using (9) we obtain that $\widehat{S}_{a^*a^*} < 0$. Also note that $a_{dd}^* = 0$. Hence, $\tilde{\Pi}_{dd} = \widehat{S}_{a^*a^*} a_d^{*2} < 0$. (This result is important not only because it helps obtaining the sign of $\partial d^*/\partial \psi_1$ but also because it confirms the second-order condition that validates the claim that d^* is determined by $\tilde{\Pi}'(d^*) = 0$.)

We now analyze $\tilde{\Pi}_{d\psi_1}$. Taking into account that ψ_1 affects $\tilde{\Pi}$ by its indirect effect on $\widehat{S}$ and a^*, we obtain

$$\tilde{\Pi}_{d\psi_1} = \left[\widehat{S}_{a^*\psi_1} + \widehat{S}_{a^*a^*}a^*_{\psi_1}\right]a^*_d + a^*_{d\psi_1}.$$

Since $a^*(d) = (2\psi_1 - d)/\psi_2$, then $a^*_{\psi_1} = 2/\psi_2$, $a^*_d = -1/\psi_2$, and $a^*_{d\psi_1} = 0$. Using (9) we obtain that $\widehat{S}_{a^*\psi_1} = 1$. We also obtain that (i) $\partial^3\widehat{S}/\partial a^{*3} = 2(2-a^*)\psi_2/((a^*+1)^4 a^*) > 0$ for all $a^* \in [0,1]$ and (ii) $\lim_{x\to 1}\widehat{S}_{a^*a^*} = -5\psi_2/4$, which implies that $\widehat{S}_{a^*a^*}(a^*) \leq -5\psi_2/4$ for all $a^* \in [0,1]$. Hence, $\tilde{\Pi}_{d\psi_1} = \left[1 + \widehat{S}_{a^*a^*} \times (2/\psi_2)\right](-1/\psi_2)$, and $1 + \widehat{S}_{a^*a^*} \times (2/\psi_2) \leq 1 + (-5\psi_2/4) \times (2/\psi_2) = -3/2$, which implies that $\tilde{\Pi}_{d\psi_1} > 0$. ■

The same effects are in play when $\psi_1 > \overline{\psi}_1$, or, equivalently, $d^* > \psi_1$, which would tend to increase d^* as ψ_1 increases. But in addition, there is now direct market foreclosure, and in equilibrium, $a^*(d^*) > a^{SB}$ as it is too costly for the firm to bring a^* all the way down to a^{SB} (due to loss of surplus stemming from foreclosure). This is why the proof of the above proposition cannot be extended to the case where $\psi_1 \geq \overline{\psi}_1$ and the problem loses algebraic tractability. But note that an increase in ψ_1 weakens the market foreclosure effect. The market is foreclosed when $m < 1 - \psi_1/d$, and for a given d this threshold is decreasing in ψ_1. In other words, an increase in ψ_1 not only continues to generate incentives for the firm to raise d but also lowers the "cost" of doing so (i.e., the loss from foreclosure). Thus, one would expect the optimal break-up fee to be increasing in ψ_1 even when there is market foreclosure in equilibrium.

References

[1] Aghion, P., and P. Bolton. (1987) "Contracts as a Barrier to Entry," *The American Economic Review*, Vol. 77, pp. 388-401.

[2] Becker, G. (1964) "Human Capital." University of Chicago Press: Chicago.

[3] Baker, G., M. Gibbs, and B. Holmstrom (1994a) "The Wage Policy of a Firm." *Quarterly Journal of Economics*, Vol. 109, pp. 881–919.

[4] ——, ——, and ——. (1994a) "The Internal Economics of the Firm: Evidence from Personnel Data." *Quarterly Journal of Economics*, Vol. 109, pp. 921–955.

[5] Bernhardt, D., and D. Scoones. (1993) "Promotion, Turnover, and Preemptive Wage Offers," *American Economic Review*, Vol. 83, pp. 771-91.

[6] ——., and ——. (1998) "Promotion, Turnover, and Discretionary Human Capital Acquisition," *Journal of Labor Economics,* Vol. 16, pp. 122-41.

[7] Bernheim, B. and M. Whinston. (1998) "Exclusive Dealing," *Journal of Political Economy,* Vol. 106, pp. 64-103.

[8] Bishara, N. (2006) "Covenants not-to-Compete in a Knowledge Economy: Balancing Innovation from Employee Mobility Against Legal Protection for Human Capital Investment," *Berkeley Journal of Employment & Labor Law,* Vol. 27 (2), pp. 289-325.

[9] Burguet, R., Caminal, R., and C. Matutes. (2002): "Golden Cages for Showy Birds: Optimal Switching Costs in Labor Contracts," *European Economic Review,* Vol. 46, pp. 1153-1185.

[10] DeVaro, J., and M. Waldman. (2011) "The Signaling Role of Promotions: Further Theory and Empirical Evidence." Forthcoming, *Journal of Labor Economics.*

[11] Franco, A. and M. Mitchell. (2005): "Covenants not-to-Compete, Labor Mobility, and Industry Dynamics." Mimeo, University of Iowa.

[12] Garmaise, M. (2011) "Ties that Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment," *Journal of Law Economics and Organization.* Forthcoming.

[13] Gibbons, R., and L. Katz. (1991) "Layoffs and Lemons," *Journal of Labor Economics,* 9(4), pp. 351–380.

[14] Gilson, R. (1999) "The Legal Infrastructure of High Technology Industrial Districts: Silicon valley, Route 128, and Covenants not to Compete," *New York University Law Review,* Vol. 74, pp. 575–629.

[15] Ghosh, S., and M. Waldman. (2010) "Standard Promotion Practices versus Up-or-out Contracts," *RAND Journal of Economics*, Vol 41, pp. 301–325.

[16] Golan, L. (2005) "Counteroffers and Efficiency in Labor Markets with Asymmetric Information," *Journal of Labor Economics*, Vol. 23, pp. 373–393.

[17] Greenwald, B. (1986) "Adverse Selection in Labor Market," *Review of Economic Studies,* Vol. 53, pp. 325-347.

[18] Jackson, M., and E. Lazear. (1991) "Stocks, Options, and Deferred Compensation." *Research in Labor Economics*, Vol. 12, pp. 41–62.

[19] Koch, A. K. and E. Peyrache. (2011) “Aligning Ambition and Incentives,” *Journal of Law, Economics & Organization*, Vol. 27 (1).

[20] Kräkel, M., and D. Sliwka. (2009) “Should you Allow your Agent to Become your Competitor?–On non-compete Agreement in Employment Contracts.” *International Economic Review,* Vol. 50, pp. 117-141.

[21] Lazear, E. (1979) “Why is There Mandatory Retirement?” *Journal of Political Economy*, Vol. 87 (6), pp. 1261-1284.

[22] ——. (1986) “Raids and Offer Matching.” In R. Ehrenberg, eds., *Research in Labor Economics*, Vol 8. Greenwich, Conn.: JAI.

[23] McCue, K. (1996) “Promotions and Wage Growth.” *Journal of Labor Economics*, Vol. 14, pp. 175–209.

[24] Mehran, H., and D. Yermack. (1997) “Compensation and top Management Turnover.” New York University, Department of Finance Working paper series No. FIN 98-051.

[25] Mukherjee, A. “The Optimal Disclosure Policy when Firms Offer Implicit Contracts,” *RAND Journal of Economics*, Vol. 41, pp. 549–573.

[26] ——. (2008a) “Career Concerns, Matching, and Optimal Disclosure Policy,” *International Economic Review*, Vol. 49, pp. 1211–1250.

[27] ——. (2008b) “Sustaining Implicit Contracts When Agents Have Career Concerns: The Role of Information Disclosure,” *RAND Journal of Economics*, Vol. 39, pp. 469–490.

[28] Posner, E., Triantis, A. and G. Triantis (2004) “Investing in Human Capital: The Efficiency of Covenants Not to Compete,” *University of Chicago Law & Economics, Olin Working Paper No. 137.*

[29] Posner, R. (1976) *Antitrust law: An economic perspective*. University of Chicago Press, Chicago.

[30] Rasmusen, E., M. Ramseyer and J. Wiley (1991) “Naked Exclusion,” *American Economic Review* Vol. 81, pp. 1137-1145.

[31] Rubin, P., and P. Shedd. (1981) “Human Capital and Covenants not to Compete,” *Journal of Legal Studies,* Vol. 10, pp. 93-110.

[32] Salop, J., and S. Salop. (1976) “Self-selection and Turnover in the Labor Market,” *Quarterly Journal of Economics*, Vol. 90, (4), pp. 619-627.

[33] Scholes, M. (1991) “Stock and Compensation,” *Journal of Finance,* Vol. 46, pp. 803–823.

[34] Spier, K., and M. Whinston. (1995) “On the Efficiency of Privately Stipulated Damages for Breach of Contract: Entry Barriers, Reliance, and Renegotiation,” *RAND Journal of Economics*, Vol. 26, pp. 180–202.

[35] Waldman, M. (1984) “Job Assignments, Signalling, and Efficiency Job Assignments, Signalling, and Efficiency,” *RAND Journal of Economics*, 15(2), pp. 255-267.

[36] ——. (1990) “Up-or-Out Contracts: A Signaling Perspective,” *Journal of Labor Economics*, Vol. 8, pp. 230–250.

[37] Zábojník, J., and D. Bernhardt. (2001) “Corporate Tournaments, Human Capital Acquisition, and the Firm Size-wage Relation,” *Review of Economic Studies,* Vol 68, pp. 693–716.