Employer Collusion and Employee Training[*]

Pedro S. Martins[†]
Queen Mary University of London
& NovaSBE & IZA & GLO

Jonathan P. Thomas[‡]
University of Edinburgh

January 3, 2023

Abstract

Firms may enter into employers' associations for a number of reasons. In this paper, we investigate one such reason: the incentives for firms to enter into a voluntary agreement to restrict employee poaching between each other. We examine theoretically the potential benefits and costs, including the impact on training, and also analyse the optimal size of such agreements. In an empirical application, we ask if employers' associations in Portugal may be operating tacit no-poach agreements. We find evidence from matched data consistent with such agreements: inter-firm worker mobility is significantly lower between EA-affiliated firms and training is considerably higher in such firms.

Keywords: Employers organisations, No-poach agreements, Worker mobility.

JEL Codes: J53, J62, L40.

[*]Earlier versions of this paper were circulated under the title "Collusion and Training". We thank comments from Thomas Breda, Claudio Lucifora, Matloob Piracha, Daphne Nicolitsas, Klaus Zimmerman and participants in EmpRep and GLO online seminars. We are also grateful for data access provided by the Ministry of Employment and the National Statistics Agency of Portugal, and research assistance by Joana Saraiva. thank funding from the European Union (EmpRep action, grant VS/2020/0122). This study reflects only the authors' views. The European Commission is not responsible for any use that may be made of the information that the study contains. All errors are our own.

[†]Corresponding author. Email: p.martins@qmul.ac.uk. Address: School of Business and Management, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom. Web: https://sites.google.com/site/pmrsmartins/.

[‡]Email: Jonathan.Thomas@ed.ac.uk. Address: Department of Economics, University of Edinburgh, Edinburgh, UK.

1 Introduction

Firms can engage in no-poach agreements to reduce their labour costs (from wages and turnover) and to increase their returns from training. These no-poach agreements (NPAs) have recently been documented in the US (Krueger & Ashenfelter (2018)),[1] leading to an executive order (White House (2021)) seeking to ban or limit non-compete agreements and to prevent employers from collaborating to suppress wages.[2]

One reason for firms to collude may be to mitigate "poaching (or quitting) externalities". This idea can be traced at least as far back as Pigou (1912) and arises when some of the returns from investment in training may accrue to an outside firm if a worker quits their original firm.[3]

In this paper, we investigate the incentives for firms to enter into a voluntary agreement to restrict poaching between each other. The theoretical contribution is that we investigate the effect of such restriction on training and profits in a simple but novel model that is general enough to capture a range of market and bargaining structures. In the model, firms invest in general human capital, and trained workers may receive outside offers. By participating in an NPA, a firm can invest in human capital formation in the knowledge that trained workers will receive fewer outside opportunities. This has a number of implications. First, training is generally higher, which is likely to improve profits. At the same time, although workers in a participating firm benefit from additional training, they also face reduced opportunities to further their career outside of the training firm, and this may make working for a participating firm less attractive. This has implications for the wage-tenure-training contract that needs to be offered to attract untrained workers, and implies that participation in an NPA may have an ambiguous effect on profits.

Moreover, as more firms join an NPA, the reduction in outside offers will be enhanced, and the costs and benefits of participation in an NPA change. We examine conditions under which some but not all firms would join an NPA: under what circumstances will none, or

[1] As reported in New York Times (2018), 'Seven major restaurant chains, including McDonald's, agreed to drop a hiring practice that critics say may be keeping tens of thousands of fast-food workers locked in low-wage jobs. The provisions prohibit workers at one franchise from going to another franchise of the same restaurant chain. No-poach clauses have drawn scrutiny over whether they hold down pay for restaurant employees'. See also U.S. Department of Justice (2021) for further evidence of wage collusion across employers.

[2] In the executive order, White House (2021), the U.S. President encourages the U.S. competition agency to ban or limit non-compete agreements. The executive order also seeks to strengthen antitrust guidance to prevent employers from collaborating to suppress wages by sharing wage information with one another.

[3] 'Franchise owners say the clauses help protect their investments of time and money in training employees' New York Times (2018).

some, or all firms choose to enter into a no-poach agreement? We also consider overall welfare implications of NPAs. In our model human capital investment is increased, which is efficiency enhancing; conversely, by restricting outside offers, some productivity enhancing moves do not occur. Hence there may be mismatch allocative inefficiency due to NPA participation to offset higher levels of training.

In an empirical application, we investigate the potential role of employers' associations (EAs) in promoting NPAs. EAs are the counterparts to trade unions in collective bargaining (OECD (2019)) in many countries. EAs typically provide many additional public goods, including representation, lobbying, dissemination of information across their members but may also promote *collusion* amongst affiliated firms, with detrimental effects (Krueger & Ashenfelter (2018), Hijzen & Martins (2020), Patault & Valtat (2020)). In this paper, we investigate the potential role of EAs in the specific case of NPAs, namely through tacit agreements that EA firms do not hire employees from each other.

Specifically, we analyse of role of EAs in worker mobility between firms (Buchinsky et al. (2010), Hijzen et al. (2013)) and training (and wages). We draw on matched employer-employee panel data from Portugal, 2009-2011, including information on EA affiliation and employee training of each firm and each worker. We find a number of results that are consistent with NPAs. In particular, we find that EA workers are less likely to move to another firm *of the same EA* and that EA workers tend to receive (much) more training than other workers. However, their wage levels are similar to those of workers in non-EA firms.

Our paper contributes to a number of literatures. First, as indicated above, on the theoretical modelling of training investments. Empirically, our paper contributes to the recently growing literature on monopsony (Azar, Marinescu & Steinbaum (2020), Azar, Marinescu, Steinbaum & Taska (2020)). We believe this is the first paper that examines empirically no-poaching agreements, even if doing so indirectly, through the role of EAs.

2 Theoretical Literature

There is of course a vast theoretical literature on human capital investment. We mainly restrict the discussion below to models of general human capital investment provided by firms and which consider the effects of varying the mobility of trained workers, and especially work that considers whether lower mobility might be attractive to individual firms.

Higher *exogenous* worker turnover will lead to lower training in many models. Acemoglu & Pischke (1999) show this in a basic two-period model, in which, after training is sunk in period 1, the period 2 wage will be the outcome of a Nash bargain where the worker's outside option is the wage (or utility) they would get from quitting (and the firm's is normalized to zero if the worker was to quit). In their "constrained case" where the wage in period 1 cannot be cut to make the worker effectively finance the training, it is shown that as turnover increases, training decreases. This follows straightforwardly as the firm does not get any benefit from investment if the worker quits, so the marginal return to training falls as turnover increases. They argue that this may explain why evidence that high turnover economies such as the U.S. have lower formal training than low turnover economies such as Germany.[4,5]

Moving away from exogenous separations, in Stevens (1994) investment may be in specific or general human capital, and the latter may be transferable to outside firms in differing degrees: specifically, some dimension of investment in human capital potentially benefits output in the training firm and a subset of outsider firms equally (but randomly). Varying the size of the subset of outsiders, which would be one way to model the possibility of increased outside opportunities, actually leads in her model to no effect on general training. The logic here is that although the leakage of surplus to outside firms is increasing in the subset size up to some point, and workers will be more likely to leave, leakage does not vary with the *level* of investment.[6]

[4]Firms do not benefit from the extra wages a trained worker may get after separation. However they also consider a "full competition" regime where firms compete in period 1 to hire workers and the worker pays upfront, through a lower period 1 wage for any additional period 2 wages. Acemoglu (1997) shows that in even in the full competition case, training is too low from a societal point of view because the firm-worker problem doesn't take account of any surplus that accrues to outside firms in frictional labour markets. This effect is however smaller than in the constrained case discussed above, as there is still *some* benefit to a separated worker of more human capital. In the directed search model of Moen & Åsa Rosén (2004) training is efficient so long as a firm and its employees can coordinate efficiently internally (maximizing expected joint income), for example if long-term contracts can be commited to. They argue a key difference with undirected search is congestion effects of low productivity workers searching in the same market as higher productivity workers. Stevens (2001) looks at possible institutions such as training levies that might correct for such externalities in a general model.

[5]Acemoglu & Pischke (1999) investigate whether active poaching by outside firms might be profitable with wage compression: poaching firms would potentially be paying less than a worker's marginal product, and so there appears to be an incentive to poach. If there is Bertrand competition between the poacher and the current firm, however, they argue this would push the wage up to productivity so there would be no gain to the poacher. (This assumes that productivity does not vary across firms.) In the asymmetric information model of Acemoglu & Pischke (1998) there would be a winner's curse at work, again making poaching unprofitable. They do identify one case where poaching would be profitable, where there is a union determined wage which not only incentivises training but which also means the incumbent firm cannot respond to an outside offer.

[6]A paper without mobility in equilibrium which explicitly addresses whether a noncompete agreement can increase training is Meccheri (2009). In his reduced-form model a noncompete reduces the worker's outside option. He applies the outside option bargaining principle to the second period bargaining (see also Balmaceda (2005) for a similar model). By reducing the frequency of a binding outside option, the noncompete increases the return to training.

Another paper that explicitly analyzes the impacts of restrictions on mobility ("covenants not to compete") is Posner et al. (2004). While it excludes the pure general human capital case, it considers in an incomplete contracting model the trade-offs between enhancing investment incentives by restricting mobility and achieving efficient ex post mobility. However it argues, that restrictions on worker movements post-training – modelled by not allowing worker mobility to certain other firms – can substantially enhance efficiency; if renegotiation is permitted (relaxing the restriction that workers cannot move to certain other firms when it is ex-post efficient to do so) then first-best outcomes can be achieved with these restrictions.

We share features with some of these models, in particular our market structure is similar to that considered by Acemoglu & Pischke (1999) in their full competition regime[7] but with a restriction on the period 1 wage as in their restricted regime. A key difference is that in addition to training, the firm can *commit* to the period 2 wage;[8] this means even when the period 1 wage is constrained, the firm may choose to satisfy the worker's period 1 participation constraint by varying the period 2 wage. Because turnover in our model is endogenous, and higher turnover does not automatically reduce the return to training as in, e.g., Acemoglu & Pischke (1999), the firm may prefer higher turnover as the extra outside opportunities relax the participation constraint. Our base model does not feature ex post renegotiation as in Posner et al. (2004) or offer matching as in, e.g., Lentz & Roys (2015), but we extend the model to consider this latter possibility.

In addition to the above, there is a substantial literature on non-competes which deals with similar issues, but in a context where an employee who leaves may be in a position to compete with the initial employer, bringing in an additional effect. See Wickelgren (2018) for a discussion of this literature.

3 Model

In order to focus on the main research questions, we will make the following simplifying assumptions. We will assume that there is an elastic supply of untrained workers in a com-

[7]See footnote 4.

[8]Lentz & Roys (2015) consider a multi-period model where firms commit to wages (subject to their participation constraint) with randomly arriving outside offers and heterogeneous firm productivities. With risk-neutral workers, general human capital investment is (socially) efficient due to the assumption of offer matching (as in Postel-Vinay & Robin (2004)), so the firm-worker pair can extract the full return from outside firms whenever a better match is made. In a calibrated version with risk-averse workers (which has similar implications to a minimum wage), a higher contact rate leads to lower training in partial equilibrium, but in general equilibrium this may be reversed. Our base model does not have offer matching but we discuss an extension below.

petitive hiring market, so that firms will have to offer a (fixed) minimum level of utility to hire such workers. In our modelling we allow firms to design a contract to best retain surplus from human capital investments – we assume there is firm commitment (but not worker commitment) We will assume that training is in general human capital and this has the same additive effect on productivity both at the training firm and at any outside firm. A further simplifying assumption is that all workers and firms are *ex ante* identical. This means that when we come to consider endogenous NPA adoption, all firms will have the same incentive to adopt.

The key parameter of interest is the arrival rate of contacts with outside firms that a trained worker makes. In our simple two-period model, firms invest in training in period 1 and in period 2 the trained worker faces a probability ϕ of making a contact with an outside firm. In the partial equilbrium framework of this section, we will interpret belonging to an NPA as equivalent to a reduction in ϕ. Our main result here is to show that when there is a binding lower bound on the period 1 wage, that training will be decreasing in ϕ, i.e., it will be higher for an NPA firm.

In detail: We consider a two-period model. There is a fixed number of identical firms. Each can employ one worker in period 1. We assume that workers and firms are risk neutral and there is no discounting. If a firm succeeds in hiring a worker in period 1, period 1 output is $y_1 > 0$, and the firm invests $\tau \geq 0$ in training. In period 2 the worker produces $y_2(\tau)$ if they remain with the firm, where $y_2(0) \geq y_1$, $y_2'(\tau) > 0$ and $y_2''(\tau) < 0$, and with $\lim_{\tau \to 0} y_2'(\tau) \to \infty$. We can define efficient training τ^* as maximizing $y_2(\tau) - \tau$, i.e., satisfying: $y_2'(\tau^*) = 1$. (We assume that $\tau^* < \infty$.) In period 1 there is a competitive labour market, and the firm must offer utility $\underline{U}$ in order to hire a worker. At the start of period 2, an employed worker makes a (single) contact with an outside firm with probability $\phi \leq 1$ and may quit. In terms of the main focus of the paper, ϕ is the critical variable.

We describe wages and the decision whether to accept an outside offer next. The firm offers a wage-tenure-training contract (w_1, w_2, τ) at the beginning of period 1 to which it is committed. If the worker receives an outside contact in period 2, we assume that the worker's outside productivity is $y_2(\tau) + \theta$, where θ is distributed according to the continuous distribution function $F(\theta)$ on $[\underline{\theta}, \overline{\theta}]$ where $\underline{\theta} \leq 0$ and $\overline{\theta} \geq 0$, with density $f(\theta)$, and such that $f > 0$ on $\left(\underline{\theta}, \overline{\theta}\right)$. Thus τ reflects investment in general human capital, while θ reflects (additive) random match quality in the potential new match. If $\theta > 0$, the new match is more productive (and it is ex post efficient for the worker to move). θ is known by worker and outside firm at the point of contact. We assume that the incumbent employer does not make counter offers. Burdett & Coles (2003) argue such an approach could be justified if the outside offer is non-verifiable, or if the firm employs multiple workers and internal equity considerations imply equal treatment with workers who do not receive an offer.[9] We also assume there are no costs, contractual or otherwise, for a worker who leaves. If $w_2 > y_2(\tau) + \theta$, there is no surplus in the new match and the worker will remain with the original firm. Otherwise the worker will leave, and we assume that she captures a fraction β, $0 \leq \beta \leq 1$, of the joint surplus $y_2(\tau) + \theta - w_2$ in the new match, receiving $w_2 + \beta(y_2(\tau) + \theta - w_2)$. The implicit assumption is that bargaining takes place before the worker quits, and failure to agree implies that the worker continues to earn w_2.

We make a couple of remarks on this structure. First, it allows for the fact that outside offers can be efficiency enhancing, and may increase the surplus available in the initial match, so that restricting mobility is not necessarily a dominant strategy. This also implies that the impact of increasing ϕ on τ does not follow automatically. With a higher arrival of efficiency enhancing outside offers, the returns to general human capital investment may be enhanced by these additional outside offers if the firm can exploit them. In fact it turns out that this is not the case in our model.[10] Secondly, it is general enough to capture the case where the worker receives the full value of general training in period 2: if $\phi = 1$, $\beta = 1$ and $F(\theta)$ is degenerate at 0, then she always receives an outside offer of $y_2(\tau)$. On the other hand, if $\beta = 1$ and $F(\theta)$ is not degenerate, the model is isomorphic to one where there is a known distribution of outside wage offers, $y_2(\tau) + \theta$, and the worker leaves for any superior offer.[11]

The firm solves the following problem:

$$\Pi(\phi) := \max_{w_1, w_2, \tau \geq 0} \{(y_1 - w_1 - \tau) + \mu(y_2(\tau) - w_2)\} \qquad \text{(Problem A)}$$

[9] Postel-Vinay & Robin (2004) show that it might be optimal for a firm to *commit* not to match outside offers to discourage moral hazard in the form of its workers exerting search effort to elicit wage increases from outside offers. We however allow for offer matching in an extension.

[10] However this is at least partly down to the additive or "modular" structure so that a worker's productivity will increase with general human capital to the same extent whether she moves or not. In a supermodular structure where more efficient matches also mean marginal human capital has higher value outside of the original match, there would be a force operating to increase τ with ϕ.

[11] Shimer (2006) has shown that when the wage can affect the quit rate, as here, the bargaining set at period 1 need not be convex, which is problematic for the axiomatic Nash bargaining approach. We model a competitive period 1 market so this is not an issue here.

subject to:

$$U(w_1, w_2, \tau) := \tag{1}$$

$$w_1 + \mu w_2 + (1-\mu) E_\theta [w_2 + \beta (y_2(\tau) + \theta - w_2) \mid y_2(\tau) + \theta \geq w_2] \geq \underline{U} \tag{2}$$

where

$$\mu(w_2, \tau) := (1-\phi) + \phi \Pr[y_2(\tau) + \theta < w_2] \equiv (1-\phi) + \phi F(w_2 - y_2(\tau)) \tag{3}$$

is the probability that the worker stays with the incumbent firm.[12]

Then, assuming that wages can be set at any level (including negative),[13]

Proposition 1 (Unrestricted wages) *For any values of ϕ, β, in any solution to Problem A, training is at the efficient level τ^*.*

In particular, τ does not vary with ϕ so restrictions to mobility have no impact on training. The intuition here is not that the initial firm-worker match can capture all the surplus even if the worker leaves (unless $\beta = 1$), but that the amount that goes to a poaching firm is constant as τ varies, and does not affect the return from investment in training. (For each level of τ there is an optimal w_2 that extracts maximum surplus from the potential outside opportunity. As τ is varied, the firm optimally keeps this amount constant by increasing w_2 in line with higher value of human capital.) Consequently to maximize match surplus $\tau = \tau^*$. Even though the firm pays the training cost, it can recoup this by reducing period 1 wages, so maximizing match surplus is optimal. This assumes that there is no restriction on wages, and may even require $w_1 < 0$.[14] Becker (1964) argued that when the worker receives the full value of general training in period 2 (see the discussion above Problem A) the worker should pay for that training, if necessary through reduced wages during training, and it should be provided at the efficient level. The proposition shows that a similar logic applies more generally in our model, even when some surplus goes to the outside firm – the point being that this loss of surplus does not increase as τ increases.

[12]Note that in (3) we are assuming, as in Shimer (2006), that the worker leaves when indifferent; this is for convenience. However if F is degenerate μ should be set equal to 0 at $w_2 = y_2(\tau)$.

[13]All proofs are in the Appendix.

[14]Likewise it requires risk-neutral workers; otherwise there would be a utility cost to having back-loaded wages.

3.1 Minimum wages

In what follows we will consider introducing a minimum wage. This will mean that the w_2 that extracts maximum surplus from outside offers when $\tau = \tau^*$, with w_1 set low enough to hit the worker's participation constraint, may not be feasible. The firm can nevertheless effectively make the worker pay for training by cutting instead w_2. However we shall see that this has implications on the returns to training the firm receives through increased worker loss in period 2.

We impose a minimum wage $\underline{w}$, so that

$$w_t \geq \underline{w} \quad t = 1, 2. \tag{4}$$

We define **Problem B** to be the same as Problem A but with the addition of the constraint (4).

Lemma 1 *If in a solution to problem B, profits are lower than in the unconstrained problem A (but nonnegative), there are two possibilities: (i)* $w_1 = \underline{w}$ *and* $w_2 = \underline{w}$*, or (ii)* $w_1 = \underline{w}$ *and* $w_2 > \underline{w}$*. In case (ii) the participation constraint (1) binds.*

Using this, we can show that when there is a binding minimum wage constraint (so that profits are below those in the solution to Problem A), training will always be inefficiently low, and hence lower than it would be in the absence of any mobility, when trivially $\tau = \tau^*$. To show that it is monotonically declining in mobility we impose an additional mild restriction on F.

Proposition 2 *[Restricted wages] If the minimum wage constraint binds, then for any values of* $\phi \in (0,1), \beta \in [0,1]$*, training is below the efficient level* τ^**. A sufficient condition for training to be strictly decreasing in* ϕ *is* $f(\theta) + \theta f'(\theta)/2 > 0$ *for* $\theta \in [\underline{\theta}, \overline{\theta}]$*.*[15]

In case (ii) of Lemma 1, given that the participation constraint binds, profits equal joint value net of the worker's outside value $\underline{U}$:

$$\Pi(\phi) = (y_1 - \tau) + y_2(\tau) + \phi S(y_2(\tau), w_2) - \underline{U}, \tag{5}$$

[15]There is one exception in which τ is constant rather than decreasing in ϕ: the minimum wage constraint only binds in period 1, $\beta = 1$ and $f(w_2 - y_2(\tau)) = 0$, i.e., w_2 is below the entire support of $y_2 + \theta$.

where $S(y_2(\tau), w_2)$ is the expected surplus that an outside contact generates for joint value, that is, the expected value of the wage, *net* of output $y_2(\tau)$ that is lost to the match, when the worker leaves, or zero when the worker stays. It is a function of $y_2(\tau)$ and w_2 because the worker's decision to stay or leave, given θ, depends only on y_2 and w_2:

$$S(y_2, w_2) = \int_{w_2 - y_2}^{\bar{\theta}} (-y_2 + w_2 + \beta(y_2 + \theta - w_2))\, dF. \tag{6}$$

Moreover S depends only on the difference $w_2 - y_2$. The proof shows that provided S is strictly concave in this difference, training is monotonically decreasing in ϕ.

As discussed in the introduction, the training decision of the firm has to take into account both whether the worker will leave, and the extent to which its investment can be recouped from outside firms. Paying a higher period 2 wage will discourage the worker from taking her human capital elsewhere, which may *per se* be desirable (*is,* in the sense that a contract to prevent the worker moving for certain outside offers would improve profits) but by doing so the firm is paying more than it needs to satisfy participation. Training at the efficient level would improve the worker's outside opportunities, and would require a high period 2 wage to prevent the worker from leaving from the point of view of joint surplus maximization. Since with a minimum wage, the firm cannot recoup the period 2 wage costs by paying a sufficiently low period 1 wage, it prefers to restrict training. Putting it differently, the firm could set $\tau = \tau^*$ and cut w_2 to satisfy the participation constraint – effectively getting the worker to finance the training in states where she stays. But this is sub-optimal because the worker will be more inclined to leave when she gets an offer – which is inefficient for low θ – and furthermore for $\beta < 1$ less surplus will be extracted from an outside firm in bargaining when she does leave.

As ϕ increases, these trade-offs worsen. Intuitively, the extra output from a marginal unit of investment either is retained by the firm-worker pair if the worker remains, or some of it is lost to the outside firm through bargaining. As w_2 is not increased (falls if $\beta > 0$), the outside firm gets, in all states where the worker leaves, $(1 - \beta)(y_2(\tau) + \theta - w_2)$, and so an *additional* $(1 - \beta) y_2'(\tau)$ surplus at least. Thus the firm-worker pair loses at least this surplus. At higher ϕ this occurs more often, so more of the extra output is "lost" to outsiders; hence τ is decreasing in ϕ.[16]

[16]If $\beta = 1$, τ is still decreasing in ϕ as there is a second effect described below.

There is a second effect when $\beta > 0$. The additional outside opportunities at a higher ϕ mean that to satisfy the participation constraint w_2 will be reduced, which leads to additional separations when the worker is more valuable inside the firm than outside. This means there is an *additional* loss of surplus so a similar logic leads to a further reduction in the return to τ.[17]

4 Extensions

4.1 Offer Matching

Thus far we have assumed that the incumbent firm does not respond to any outside contacts the worker has. Suppose instead that the incumbent firm can respond to an outside offer with a counter-offer (see Postel-Vinay & Robin (2002*b*), Postel-Vinay & Robin (2002*a*) and Dey & Flinn (2005)). In common with most of the literature we assume $\beta = 0$, so that the worker cannot bargain any of the additional surplus with an outside contact.[18] We assume that there is Bertrand competition and both firms will bid up to their valuation of the worker in a complete information setting, following Postel-Vinay & Robin (2002*b*). The worker would produce $y_2(\tau) + \theta$ in the outside firm, so this is the maximum it is prepared to offer, while the incumbent firm is prepared to pay up to $y_2(\tau)$. The worker is more valuable in the incumbent firm if (and only if) $\theta < 0$. Suppose first that $w_2 < y_2(\tau)$. Then if $w_2 < y_2(\tau) + \theta$, w_2 will be bid up to $y_2(\tau) + \theta$, and the worker stays with the firm, while if $w_2 \geq y_2(\tau) + \theta$ the outside firm does not bid and the wage is unchanged. If $\theta \geq 0$ the worker takes an outside offer of $y_2(\tau)$ and leaves. However if $w_2 \geq y_2(\tau)$ a new possibility arises not considered in earlier analyses.[19] If $w_2 > y_2(\tau) + \theta$ then the outside firm is not prepared to match w_2, even if $\theta > 0$, so the worker stays, at wage w_2, and leaves only if $w_2 \leq y_2(\tau) + \theta$, but at w_2 (as $\beta = 0$).

Proposition 3 *[Offer matching] If there is offer matching, and if the minimum wage constraint binds in period* 1*, then (a) for* $\underline{w} < w_2 \leq y_2(\tau)$*,* $\tau = \tau^*$*, while (b) for* $w_2 > y_2(\tau)$

[17] This requires certain mild restrictions on the density to be unambiguously signed, uniform being a sufficient condition.

[18] Dey & Flinn (2005) consider $\beta > 0$ as well.

[19] Previous analyses consider some form of commitment, either to a given wage unless renegotiated by mutual consent as in, e.g., Postel-Vinay & Robin (2002*b*), or to a tenure contract such as in Lentz & Roys (2015). Wages however are never greater than productivity since in the former analyses the firm would never renegotiate to a wage greater than productivity, and in the latter it is assumed that continuation profits cannot be negative. We allow the firm to commit to a wage at which it may make a loss in period 2, a previously unconsidered case to our knowledge.

training is below the efficient level τ^*, and a sufficient condition for training to be strictly decreasing in ϕ is if $F(x)$ satisfies $f(x) + xf'(x)/2 > 0$ on (a, b). A sufficient condition for case (b) to apply for all ϕ is that $\underline{U} - \underline{w} > y_2(\tau^*)$ so that the minimum wage constraint binds but is not too tight. Profits are constant in ϕ when case (a) applies, and increasing in ϕ in case (b).

Intuitively, if $w_2 > y_2(\tau)$, the argument follows the same logic as earlier. The wage is never renegotiated upwards, and the worker only leaves for an outside offer if $w_2 \leq y_2(\tau) + \theta$. It follows that with a binding minimum wage, increases in ϕ lead to additional turnover and loss of return to marginal investment. However, if $w_2 < y_2(\tau)$, the worker will leave when $\theta \geq 0$, at a wage of $y_2(\tau)$ as the incumbent will bid up to that amount. The full return on investment is retained within the match, and by setting w_2 to hold the worker to her participation constraint, the firm appropriates the full return so that $\tau = \tau^*$. What is crucial here is that in this range using w_2 to satisfy participation does not affect allocation, in contrast to the no offer-matching case. Consequently, no further surplus is lost from increased investment as ϕ increases, even though a wage above $y_2(\tau)$ would extract more surplus (but at a cost to the firm of paying the worker more than it needs to).

The sufficient condition for case (b) to apply, that $\underline{w}$ is not too tight so w_2 can be set above $y_2(\tau^*)$ (but not as high as the firm would like), is not necessary.[20]

4.2 Bargaining where unemployment is the outside option

We assume that when a contact occurs, the productivity of the worker with the outside firm is still observed by all parties, but if the worker quits, the wage she bargains with the new firm is determined with unemployment benefit b as the outside option. Thus she leaves only if $b + \beta(y_2(\tau) + \theta - b) \geq w_2$. This problem has the same maximand as in Problem A but with the following constraints:

$$U(w_1, w_2, \tau) := \tag{7}$$

$$w_1 + \mu w_2 + (1-\mu) E_\theta [b + \beta(y_2(\tau) + \theta - b) \mid \theta \geq b - y_2(\tau) + (w_2 - b)/\beta] \geq \underline{U}, \tag{8}$$

[20]When it fails to hold, case (b) can still apply above a threshold value of ϕ . For example, suppose θ is uniformly distributed on $[-0.5, 0.5]$ and training costs are quadratic, $y_2(\tau) = \tau^{1/2}$, $\overline{U} = 0.6$, and $\underline{w} = 0.15$. Then for $\phi \leq 0.26$, $w_2 \leq y_2(\tau)$, but for $\phi > 0.26$, $w_2 > y_2(\tau)$ so case (b) applies. There is a jump down in τ at $\phi = 0.26$ (from $\tau^* = 0.25$ to 0.22) and τ is strictly decreasing thereafter; note the sufficient condition does not hold as $w_2 = 0.47 < y_2(\tau^*) = 0.5$. At the switch point, the loss from moving to an inefficient level of training is just offset by additional surplus extracted from outside offers.

where the probability the worker stays is now

$$\mu(w_2,\tau) := (1-\phi) + \phi \Pr[y_2(\tau) + \theta < w_2] \equiv (1-\phi) + \phi F(b - y_2(\tau) + (w_2 - b)/\beta). \quad (9)$$

In this case, it does not follow that the worker's participation constraint always binds. For ϕ high enough the firm may want to pay an efficiency wage to reduce the probability of the worker taking an outside offer. In the previous case this possibility didn't arise as with a slack constraint the firm could cut y_2 and w_2 by the same amount without affecting the probability of a quit, and thus improve profits. Now, if $\beta < 1$, similar cuts would increase the probability of a quit as for each realization of θ the outside wage falls by less than does w_2. Moreover when the participation constraint doesn't bind, it is possible to construct examples where $d\tau/d\phi > 0$. Intuitively, as outside contact become more likely, it may be worthwhile to reduce the total probability of a quit, and this increases the return to investment.[21]

Proposition 4 *[Bargaining with outside option b] If the minimum wage and worker participation constraints bind, then a sufficient condition for training to be strictly decreasing in ϕ is $f'(\theta) \geq 0$ for $\theta \in [\underline{\theta}, \overline{\theta}]$.*

4.3 Including specific human capital

Suppose that in addition to general human capital investment, the firm also invests in specific human capital σ, and for simplicity we assume that the effect on the firm's period 2 output if the worker doesn't quit is additive and equals $y_2(\tau) + s_2(\sigma)$, where $s_2(\cdot)$ satisfies the same general properties as $y_2(\cdot)$, but the outside value is only dependent on $y_2(\tau)$.

The return to σ depends only on μ, the probability that the worker stays, as there is no benefit to the worker should she leave in terms of higher wages.[22] We can repeat the argument of the lemma to show that the participation constraint binds, and that of Proposition 1 that

[21] As an example, suppose that $y_2(\tau) = \sqrt{\tau}$, and that F is uniform on $[\underline{\theta}, \overline{\theta}]$. Suppose $\beta = 0.5$, $b = 0.006$, $\underline{\theta} = -0.003$, $\overline{\theta} = 0.115$, $\overline{U} = 0.1$, $\underline{w} = 0$. Then for $\phi > 0.74$ the PC constraint is slack, *and* $d\tau/d\phi > 0$ in this range. Such a case is more likely to arise if the support of F is concentrated and if β is small. The concentrated support implies that the retention probability is very responsive to wage increases, so a higher wage than needed to satisfy the PC may be profitable. If β is small, additional training doesn't increase the outside option very much and so w_2 doesn't need to be raised very much to retain the worker with high probability; the benefit from more efficient investment can come at a relatively low price in terms of delivering additional utility above $\underline{U}$. We can show, for example, that a case such as the above cannot arise if $\beta > 0.5$ (0.75) and $\overline{\theta} - \underline{\theta} > 0.25$ (0.09) (and the support of productivity $y_2(\tau) + \theta$ is nonnegative).

[22] This does not follow if there is offer-matching. In the model of Lentz & Roys (2015), as the incumbent firm is willing to pay more to attempt to retain the worker in the face of outside competition, then if the worker is poached the wage may be bid up higher if there is more specific capital..

if there is no binding minimum wage, $\tau = \tau^*$.[23] σ must satisfy

$$1 = s_2' \mu, \tag{10}$$

and so even where there is no binding minimum wage if the solution is interior ($0 < \mu < 1$), σ will generally vary with ϕ. In the absence of a binding minimum wage constraint, τ is at the efficient level as before; now however σ may be decreasing in ϕ.

Proposition 5 *If the minimum wage constraint does not bind, $\tau = \tau^*$. Moreover if match productivity shock θ is uniformly distributed and training costs are quadratic, $y_2(\tau) = \tau^{1/2}$, $s_2(\sigma) = \alpha\sigma^{1/2}$, then specific investment σ is strictly decreasing in ϕ or constant if the worker is always retained ($\mu = 1$).*[24]

When the minimum wage constraint binds we can no longer assert that τ or σ is always decreasing in ϕ.[25] It is possible to construct counterexamples where τ is increasing in ϕ. For example, take the uniform/quadratic-cost case as in Proposition 5, with $\beta = 0$ and $\alpha = 1$ (so specific and general investments are equally productive): $\tau^* = 0.25$ and, when $\phi = 0$, $\sigma = 0.25$. Suppose that even in the most productive outside match, there is no more than a doubling of productivity due to general human capital (i.e., $\overline{\theta} \leq 0.5$). Then if $\phi < 0.83$, $d\tau/d\phi < 0$. For higher values of ϕ however there is a small region where $d\tau/d\phi > 0$. If $\phi = 0.9$, then this region is limited to values of $\overline{\theta}$ between 0.25 and 0.4, and $\underline{\theta}$ between -0.03 and 0. Thus such cases seem to require a substantial upside to outside match productivity combined with a very high contact rate.

5 Endogenizing ϕ

We extend the model to allow for firms to choose whether to belong to an EA, in which case their workers will be restricted from job offers with other EA firms, and they will likewise be

[23]The key point is that the contract variations in both proofs hold $w_2 - y_2$ constant, which implies μ is unchanged. The return on σ depends on the contract only through μ, so both proofs go through *mutatis mutandis.*

[24]See the Appendix for technical details for this section.

[25]Surplus from outside contacts is now $S(w_2, y_2) - (1 - F(w_2 - y_2)) s_2$,where $S(w_2, y_2)$ is as defined in (6), and the new term reflects the loss of any return from σ when the worker quits (the quitting decision depends only on $w_2 - y_2$, as before). This implies that surplus falls in response to a marginal increase in y_2 by an additional $f \times s_2$ to reflect the extra quit probability. We cannot sign how this varies with ϕ and so the previous proof giving a sufficient condition for τ to be monotonically decreasing in ϕ fails. If s_2 is decreasing in ϕ then even if f is constant, our proof may fail.

restricted from hiring from such firms. Moreover we need to introduce a motive for firms to poach trained workers. We do this in the simplest way to illustrate some general points, and remain agnostic about the matching technology that generates contact probabilities.

Suppose that each firm now has an additional opportunity to employ a single additional worker in period 2 (and revealed in period 2) arising – if a contact with a trained worker is made – with probability ρ. This is independent of whether the incumbent worker leaves or not; if they do leave that position cannot be replaced. A newly hired worker with training τ produces $y_2(\tau) + \theta$ with θ distributed as above. Only workers trained within the industry can be hired in period 2.[26]

There is a fixed measure of one of firms in the industry. In period 0 firms simultaneously choose whether to join the EA or not. In period 1 they hire a worker on a competitive labor market, as in the previous section, where $\overline{U}$ is determined outside of the industry – there is a perfectly elastic industry labor supply.

Define γ_{EA} to be the proportion of firms joining the EA. Let $m_{EA}(\gamma_{EA})$ be the number of contacts for workers in EA firms, and $m_O(\gamma_{EA})$ for those in outside firms, and likewise $n_{EA}(\gamma_{EA})$ and $n_O(\gamma_{EA})$ be the corresponding number of contacts for firms in and out of the EA, respectively. Assume all are continuous functions. Each firm or worker has a maximum of one contact. Then we must have

$$m_{EA}(\gamma_{EA}) + m_O(\gamma_{EA}) = n_{EA}(\gamma_{EA}) + n_O(\gamma_{EA}) \tag{11}$$

and by assumption EA firms and workers can only have contacts with outside firms and workers: $m_{EA}(\gamma_{EA}) \leq n_O(\gamma_{EA})$, $n_{EA}(\gamma_{EA}) \leq m_O(\gamma_{EA})$.The contact rate for EA-firm workers is $\phi_{EA}(\gamma_{EA}) \equiv m_{EA}(\gamma_{EA})/\gamma_{EA}$, and that for non-EA-firm workers $\phi_O(\gamma_{EA}) \equiv m_O(\gamma_{EA})/(1-\gamma_{EA})$.[27] We impose the following further assumptions:

$$\phi_{EA}(\gamma_{EA}) < \phi_O(\gamma_{EA}) \quad \text{for } \gamma_{EA} > 0, \tag{12}$$

$$\phi'_{EA}(\gamma_{EA}) < 0,\ \phi'_O(\gamma_{EA}) \geq 0, \tag{13}$$

[26] An alternative approach would be to assume exogenous separations at the end of period 1, then period 2 labor demand (i.e., poaching) would be generated in a similar fashion by firms needing to replace separated workers. We avoided this to abstract from unemployment; however similar results should follow if this approach was taken.

[27] Search is non-directed so for "O"-workers and firms the probability of matching with a particular type of firm or worker will depend only the gross ratios of contacts.

and

$$\phi_{EA}(1) = 0 \quad \text{and} \quad \phi_O(1) = \overline{\phi} \leq 1, \quad \phi_{EA}(0) = \phi_O(0) = \underline{\phi} \leq 1. \tag{14}$$

Symmetric definitions apply for firm contact rates $\rho_{EA} \equiv n_{EA}(\gamma_{EA})/\gamma_{EA}$ and $\rho_O(\gamma_{EA}) \equiv n_O(\gamma_{EA})/(1-\gamma_{EA})$, and we make the corresponding assumptions (12) and (13) for the contact rates for firms. Note from (11) and the definitions of ϕ and ρ,

$$\rho_{EA}(1) = 0 \quad \text{and} \quad \rho_O(1) = \overline{\phi} \leq 1, \quad \rho_{EA}(0) = \rho_O(0) = \underline{\phi} \leq 1. \tag{15}$$

These assumptions reflect the idea that as more firms belong to the EA, as there are more EA-firm employees chasing contacts with the dwindling number of non-EA firms , their contact rate should fall; for non-EA firm employees, there are fewer facing a larger number of firms who can only hire from their number so contact rates would be expected to rise. In the limit if almost all firms belong to the EA, an EA firm employee will face almost no potential firm contacts so the contact rate will fall to zero. At the other extreme, if almost no firms belong to the EA then their employees will face approximately the same opportunities as the non-EA firm employees.[28]

We define an equilibrium to be a value for γ_{EA}, together with contracts for EA-firms and outside firms, $\left(\tau^{EA}, w_1^{EA}, w_2^{EA}\right)$ and $\left(\tau^O, w_1^O, w_2^O\right)$, respectively, such that the contracts maximise profits, denoted by $\Pi^{EA}(\gamma_{EA})$ and $\Pi^O(\gamma_{EA})$ for each category of firm, and such that a firm cannot increase profits by switching its membership – obviously this requires Π^{EA} $= \Pi^O$ whenever $0 < \gamma_{EA} < 1$. We call an equilibrium stable if a small perturbation of γ_{EA} to $\gamma_{EA} + \varepsilon$ implies that $\Pi^{EA}(\gamma_{EA}) < \Pi^O(\gamma_{EA})$ when $\varepsilon > 0$ and $\Pi^{EA}(\gamma_{EA}) > \Pi^O(\gamma_{EA})$ when $\varepsilon < 0$.

An EA-firm has contact rate $\phi_{EA}(\gamma_{EA})$, so $\Pi^{EA}(\gamma_{EA})$ is the solution for profits in Problem B, $\Pi(\phi)$, with $\phi = \phi_{EA}(\gamma_{EA})$, and in addition we suppose that there is an unmodelled benefit B from EA membership, where $B < 0$ would signify a cost.[29] So $\Pi^{EA}(\gamma_{EA}) \equiv \Pi(\phi_{EA}(\gamma_{EA})) + B$. There is also an additional profit from the period 2 poaching opportunity. However we consider initially the case where $\beta = 1$, where the poaching profit is zero as the worker receives all the surplus. Similarly $\Pi^O(\gamma_{EA}) \equiv \Pi(\phi_O(\gamma_{EA}))$. We can establish the

[28]For a simple example, suppose that there is a unique pairing between each firm and worker. However any pairings involving an EA firm and an EA worker is not allowed. So $\phi_{EA}(\gamma_{EA}) = \rho_{EA}(\gamma_{EA}) = (1-\gamma_{EA})$.

[29]We discuss the case where B varies with γ_{EA} below.

following based only on the underlying solution to Problem B and the general properties of the contact probabilities.

Proposition 6 *Assume* $\beta = 1$ *and that* $\Pi(\phi)$ *is strictly monotonic on* $[0, \overline{\phi}]$. *Consider first* $\Pi(\phi)$ *strictly increasing on* $[0, \overline{\phi}]$. *If* $B > 0$, *then there is a unique and stable equilibrium that is interior (i.e., with* $0 < \gamma_{EA} < 1$*) provided* B *is not too large,* $\Pi(0) + B < \Pi(\overline{\phi})$, *whereas if* $\Pi(0) + B \geq \Pi(\overline{\phi})$ *the unique stable equilibrium is at* $\gamma_{EA} = 1$. *If* $B \leq 0$, *then the unique (and stable) equilibrium is at* $\gamma_{EA} = 0$. *Next, consider* $\Pi(\phi)$ *strictly decreasing. If* $B < 0$ *and not too large in absolute size,* $\Pi(0) + B > \Pi(\overline{\phi})$, *then there is an unstable interior equilibrium as well as two stable equilibria at* $\gamma_{EA} = 0$ *and* $\gamma_{EA} = 1$, *whereas if* B *is sufficiently negative that* $\Pi(0) + B < \Pi(\overline{\phi})$ *the unique and stable equilibrium is* $\gamma_{EA} = 0$.[30] *If* $B \geq 0$, *then the unique (and stable) equilibrium is* $\gamma_{EA} = 1$ *except for an additional but unstable equilibrium at* $\gamma_{EA} = 0$ *if* $B = 0$.

The proposition implies that for monotonic profits, stable interior equilibria only arise in a fairly limited set of circumstances, namely when profits are increasing in ϕ and there is some unmodelled benefit, which is not too large, from being in an EA. Under what circumstances is $\Pi(\phi)$ increasing? This occurs when the minimum wage constraint is not too tight, so that w_2 is not too far from its optimum level. Intuitively, this means that the firm and its employee are jointly extracting a good deal of surplus from any outside offers. As the probability of outside offers increases, this addditional surplus feeds through to extra profits. If the minimum wage increases, this not only flattens the wage-tenure profile in case (ii) of Lemma 1 because w_1 is higher, but reduces w_2 at each value of ϕ (assuming $\overline{U}$ is unchanged). For w_2 sufficiently low, the worker is poached more frequently when his value inside the firm is greater than outside, thus reducing the ex ante surplus that is shared. Higher ϕ implies this loss of surplus is more frequent, and profits may fall.[31] If the minimum wage constraint binds in both periods, profits are strictly decreasing in ϕ.[32] Intuitively, there is no benefit

[30] In the knife-edge case $\Pi(0) + B = \Pi(\overline{\phi})$ there is also an unstable equilibrium at $\gamma_{EA} = 1$.

[31] Take a numerical example: the uniform- quadratic cost case where the distribution of θ is uniform on $[-1/2, 1/8]$, so on average the outside productivity is lower than the inside one, and suppose that $\overline{U} = 1/2$.Then for $\underline{w} < 1.25$, $\Pi(\phi)$ is increasing on $[0, 1]$, and for $\underline{w} > 1.46$, $\Pi(\phi)$ is decreasing on $[0, 1]$. (For high $\underline{w}$ such that the minimum wage constraint is binding in both periods, profits are decreasing as extra contacts lead to more quits, and any extra surplus is lost to the firm.) Note that the mean of θ is negative in this example; with $\beta = 1$ a positive mean leads to some extra expected surplus from a contact, so profits will generally be increasing in ϕ even for low w_2.

[32] Let $\phi' > \phi$, where optimal training at ϕ' is τ'. Consider profits at ϕ using τ' . From (3), $\mu(\underline{w}, \tau')$ at ϕ, $\mu(\underline{w}, \tau'; \phi)$ say, is greater than $\mu(\underline{w}, \tau', \phi')$ (recall that $F(\underline{w} - y_2(\tau)) < 1$ in any optimal contract). It follows

to the firm of more frequent contacts as it cannot recoup any surplus the worker gets from outside offers by cutting wages, and higher turnover increases the probability that it gets no return on investment.

This may offer an alternative explanation to the monopsony theory for why no poaching agreements seem to be common in franchise arrangements involving low-paid workers. If wages are at low levels, and *a fortiori* if they are at minimum wage levels, profits are likely to be decreasing in ϕ, and so even if, as the proposition suggests, full sector-level concentration in a no-poaching agreement would be stable, this may violate anti-trust law. However, it would be still be optimal to impose no poaching within the franchise.

For the case $\beta < 1$ we add profits from potential period-2 new hires. For EA firms this is:

$$\hat{\Pi}_{EA} \equiv \rho_{EA} \Pr[y_2\left(\tau^O\right) + \theta < w_2^O] E_\theta[(1-\beta)\left(y_2\left(\tau^O\right) + \theta - w_2^O\right) \mid y_2\left(\tau^O\right) + \theta \geq w_2^O].$$

For a non-EA firm it is slightly more complicated as they can poach workers from both types of firm, and the corresponding additional profit is

$$\hat{\Pi}_O \equiv \rho_O^O \Pr[y_2\left(\tau^O\right) + \theta < w_2^O] E_\theta[(1-\beta)\left(y_2\left(\tau^O\right) + \theta - w_2^O\right) \mid y_2\left(\tau^O\right) + \theta \geq w_2^O] +$$
$$\rho_O^{EA} \Pr[y_2\left(\tau^{EA}\right) + \theta < w_2^{EA}] E_\theta[(1-\beta)\left(y_2\left(\tau^{EA}\right) + \theta - w_2^{EA}\right) \mid y_2\left(\tau^{EA}\right) + \theta \geq w_2^{EA}],$$

where $\rho_O^{EA} = m_{EA}/\left(1-\gamma_{EA}\right)$ is the probability of a contact with an EA-firm employee (all employees of EA-firms have contacts with non-EA firms) and $\rho_O^O = \left(n_O - m_{EA}\right)/\left(1-\gamma_{EA}\right)$ is the probability of a contact with a non-EA firm employee. So now profits depend not only directly on γ_{EA}, but also on other firms' wage-training policies. In addition, we allow for B to vary with γ_{EA} and we write this as $B\left(\gamma_{EA}\right)$ (assumed continuous). We will not attempt as complete an anaysis of the general case, but it is straightforward to give sufficient conditions for a stable interior equilibrium to exist. Moreover, there is a sense in which they are more likely to be satisfied than when $\beta = 1$.

Proposition 7 *If $B\left(0\right) > 0$, then there is a stable equilibrium that is interior (i.e., with $0 < \gamma_{EA} < 1$) provided $\Pi\left(0\right) + B\left(1\right) < \Pi\left(\overline{\phi}\right) + \hat{\Pi}_O$, where $\hat{\Pi}_O$ is computed using $\rho_O^{EA} = \underline{\phi}$*

from $\Pi\left(\phi\right) \geq \left(y_1 - \underline{w} - \tau'\right) + \mu\left(\underline{w}, \tau'; \phi\right)\left(y_2\left(\tau'\right) - \underline{w}\right)$ and $y_2\left(\tau\right) > \underline{w}$ (otherwise profits would be negative), that $\Pi\left(\phi\right) > \Pi\left(\phi'\right)$.

and $\rho_O^O = 0$, *with* (w_2^{EA}, τ^{EA}) *the solution to Problem B when* $\phi = 0$.

Given that $\hat{\Pi}_O > 0$ whenever $\beta < 1$, this is weaker than the condition in Proposition 6 for a stable interior equilibrium. Note that we cannot rule out other interior equilibria.

6 Empirical results

6.1 Portugal: institutional background

The labour market of Portugal shares many similarities to those of other continental European countries, in particular in Southern Europe. One important dimension concerns the relevance of sectoral collective bargaining, which covers 86% of private-sector employees. Such bargaining is conducted by over 300 EAs together with different trade unions (see Martins20what for a detailed description of EAs).

Note that the EA affiliation is estimated at 43%, a figure in line with the OECD mean, but much below the coverage rate of sectoral agreements. This gap is explained by the pervasive nature of administration extension schemes, which widen the coverage of collective agreements to all firms and employees in each sector **?**.

Regarding NPAs, the labour code of Portugal states that 'agreements between employers that forbid the hiring of a current or former employee or that require the payment of compensation for such hires are null'. This indicates that NPAs in the country are illegal in the sense that they are not enforceable in a court of law. However, if two or more employers agree tacitly to pursue such arrangements and benefit from them, such NPAs will be sustainable from a practical point of view.

Firms mandated to provide 35 hours of training to each employee per year (exceptions apply)

6.2 Data

We study the case of Portugal using detailed data, covering the population of all private-sector firms in the country and information on employers' association affiliation for each firm. These data are made available in Personnel Records ('Quadros de Pessoal', QP), a compulsory survey of all firms in Portugal with at least one employee, conducted by the Ministry of Employment. This census also includes a number of additional variables about firms and their workers, such as identifiers, geographical location, industry (five-digit code), sales, employee headcount, and individual wages of each employee.

We focus on employers' association data for 2009, the latest year available with that variable, and wages and training data for 2010 and 2011 (training data is only available for those two years). We also assume that each firm's EA affiliation is not changed between 2009 and 2011.

6.3 Mobility

We exploit the comprehensive nature of the QP data set to construct a data set of inter-firm worker mobility. As QP covers the full population of employees in the country and in each year and also includes a time-invariant identifier for each employee, we cn establish all pairs of firms that were linked through mobility of their workers. As the latest year for which we have EA affiliation information at the firm level is 2009 and as the training data that we exploit later is only available for 2010 and 2011, we focus on inter-firm mobility between the last two years. Moreover, we assume that the 2009 affiliation status remained unchanged in 2010 and 2011.

We find a total of nearly 100,000 employees that move between different firms in the period above, namely that are employed in one firm in (October of) 2010 and are then employed in a different firm in (October of) 2011.[33] These 100,000 employees are employed by about 37,000 firms in 2010 and by about 15,000 firms in 2011. The difference in the last two figures indicates greater dispersion or less concentration across separating firms compared to hiring firms.[34]

Inter-firm mobility equations based on actual and *potential* mobility:

Actual: all workers that change firms between 2010 and 2011

Potential: (0.1%-5%) samples of not realised combinations between firms with actual mobility

Identified from population nature of matched data

[33] To ensure that these are not spurious moves driven by changes in the firm identifier because of mergers or acquisitions, for instance, we also require that the tenure counter of the worker is reset at the new firm. Moreover, we ignore inter-firm mobility spells that involve more than 25 employees moving between a specific pair of firms, as that may denote a displacement from the first firm.

[34] Although we have information in QP regarding the month when the contract with each firm started, we do not know when the contracts come to an end. This implies that our data set includes both separations and quits and both workers that move directly from one firm to the next and those that experience a spell of unemployment in between.

Actual and potential mobility: an illustration

Worker 1 moves from firm A in Oct 2010 to firm B in Oct 2011

Worker 2 moves from firm C in Oct 2010 to firm D in Oct 2011

Actual mobility spells: A$\longmapsto$B and C$\longmapsto$D

Tenure in new firm must be zero

Large flows (25+ employees) dropped (displacements)

Potential (but not realised) mobility spells: A$\longmapsto$D and B$\longmapsto$C

Only 'sending' and 'receiving' firms (not eg B$\longmapsto$A)

Table 1 presents the resulting data set, in which the left-hand-side panel considers only firm pairs in which worker mobility was observed (79,082 observations). In contrast, the right-hand-side panel considers all firm pairs, including a sample of those in which worker mobility is not observed (3.1 million observations). We find that the number of worker movers per firm pair in which mobility is observed is low, with an average of 1.25. In other words, most of the 79,000 mobility spells found involve only one worker.

7.6% of such spells take place between firms in the same EA, a figure that increases to 20.8% in our full sample of firm pairs (including potential but not realised mobility spells). The percentage of realised mobility spells that involve both firms in the same collective bargaining agreement is 29.9%, 55.6% are located in the same region, and 24.3% work in the same industry. In the full sample, including both realised and non-realised mobility, the three percentages are lower, at 8.1%, 10.7% and 4.7%, respectively.

Moreover, 51% of the mobility pairs correspond to EA-affiliated firms, while 28.7% correspond to case in which both firms are EA-affiliated (although not necessarily in the same EA). In the full sample, the equivalent percentages are 78% and 68%. Finally, realised mobility firms are large, with mean number of employees of about 830 workers both in the first and second year (2010 and 2011), while their full sample counterparts are much smaller, at about 65 workers.

These descriptive statistics may already point to restrictions in worker mobility between same-EA firms. On the one hand, we observe that EA firms are active in both separating and recruiting workers that move between firms. Moreover, operating in the same region, industry or collective agreement (which will all be the case of many same-EA firms) appears to be a strong predictor of inter-firm mobility, as expected given the importance of local labour markets and industry-specific skills. However, on the other hand, we find that same-EA mobility occurs only in a small percentage of realised mobility spells, despite the presumably large share of same-EA firms that operate in the same region, industry or collective agreement. Moreover, these statistics also indicate that EA firms correspond to a large share of firms with realised mobility.

Table 2 describes our data set at the level of the employee, pooling data for 2010 and 2011 and corresponding to a total of 5.1 million observations. On average, employees have 9.3 years of schooling, they are 39.2 years old, and have been with their firms for eight years. 45.5% are women. 55.4% are employed by EA-affiliated firms, with average employment of 1,054 workers and annual sales of 185 million euros. 47.2% of the observations correspond to 2011. 32% of the employees receive training in the year of observation and the average amount of training weeks (across all employees, including those who do not receive training) is 0.33. Average log earnings is 6.6

6.4 Inter-firm mobility results

Our main analysis, presented in this subsection, concerns the question of whether EA-affiliation has a negative effect on worker inter-firm mobility. As discussed above, we hypothesise that EAs can serve as coordination devices to reduce worker mobility between affiliated firms, thus allowing the latter to benefit more from their investments in worker training.

Our empirical analysis is based on all instances of inter-firm worker mobility between (October of) 2010 and (October of) 2011 and a sample of potential but not realised spells of inter-firm mobility. The full sample used is described in Table 1 (right-hand-side panel). Each observation corresponds to a pair of firms, in which the 'separation firm' is a firm from which at least one employee left (to another firm) in 2010 and in which the 'hiring firm' is a firm from which at least one employee left (to another firm) in 2010.

We estimate two types of models: the first one is focused on the extensive margin (whether there is or not worker mobility from a given firm to another given firm), while the second also considers the intensive margin (how many workers move between the two firms, including zero - no mobility - but also one, two, or any other number of workers). We estimate the first case using a simple linear probability model and the second using a Poisson model (and the algorithm of Correia20).

We also pay particular attention to a number of potential determinants of inter-firm worker mobility which could confound the role of the EA-related variables. From the limited literature on this particular type of worker mobility (including BUCHINSKY10), we seek to control for the role of local labour markets, which will greatly facilitate worker mobility while also potentially be correlated with same-EA status. Similarly, we also control for the industry where both firms operate can also facilitate mobility, given the role of industry-specific skills. The collective bargaining agreement of each firm can also be another form of similarity between the firms that can promote mobility while strongly correlated with EA affiliation and is controlled for in our equation.[35] We also control for the general EA status of each firm (affiliated or not in any EA), both individually and jointly (i.e. both separating and hiring firms being EA affiliated, although not necessarily in the same EA). These variables will control for systematic differences between EA-affiliated firms in terms of their separation and recruitment outcomes. Note that all previous variables above are also constructed in terms of whether they are matched between the (realised or not) separation and hiring firm.

More specifically, we estimate the following inter-firm mobility equation:

$$\begin{aligned} y_{i,j} =& \beta_1 SameEA_{i,j} + \beta_2 BothEA_{i,j} + \beta_3 SameRegion_{i,j} + \\ &+ \beta_4 SameCBA_{i,j} + \beta_5 SameIndustry_{i,j} + \alpha_i + \delta_j + u_{i,j}. \end{aligned} \tag{16}$$

The dependent variable, $y_{i,j}$, is a dichotomous variable equal to one if at least one worker from firm i in (October of) 2010 is employed by firm j in (October of) 2011 (linear probability model). Alternatively, $y_{i,j}$ is the count of workers that move from firm i in 2010 to firm j in 2011 (Poisson model). Each i, j observation is an actual or a potential (but not realised) match between two different firms: in all instances in each the match is not realised, $y_{i,j}$ is equal to zero.

The key explanatory variable is $SameEA_{i,j}$, a dummy variable equal to one if firms i and j are affiliated in the same EA and zero otherwise. Control variables include, depending on the specification: $BothEA_{i,j}$, a dummy variable equal to one if firms i and j are both EA affiliated (in the same or in a different EA); $SameIndustry(Region, CBA)_{i,j}$, a dummy variable equal to one if firms i and j are in same industry (region, CBA); and firm controls (total employment of each firm, in each year).

[35] Note that, as discussed in Section 6.1, because of the extensions mechanism, firms can apply a given collective agreement although they are not affiliated with the EA that bargained such agreement.

Finally, the specification may also include α_i and δ_j, which are separating and hiring firm fixed effects, respectively. These will control for systematic differences across firms in their separation and hiring outcomes. Note that controls for firm characteristics (as opposed to match characteristics) will be subsumed by the firm fixed effects, as we observe each firm only once in each year, as either separating or hiring. Standard errors are clustered at the separating and hiring firm levels.

Table 3 presents our results from the perspective of the extensive margin (linear probability model). The first two columns control for EA affiliation (of each firm individually and jointly) and for firm size (column 1) or firms fixed effects (column 2) but do not control for match characteristics, except for the key variable of same-EA status. These results indicate that same-EA combinations are more likely to lead to worker mobility. However, as discussed above, firm pairs that are affiliated to the same EAs may also operate in the same region, industry and or collective agreement. All such common characteristics may also influence positively mobility of workers between firms, leading to an estimate of the same-EA effect that is biased upwards.

Indeed, when we control for such common characteristics, we find (columns 3 and 4) that the same-EA coefficient switches sign and become larger in absolute terms. When controlling for firm characteristics (EA affiliation and size), the same-EA coefficient is -2.3% while, when controlling for firms fixed effects, it increases to -4.2%, in both cases statistically significant at the 0.1%. These results indicate that, consistently with our earlier discussion, firms that are in the same EA are less likely to have workers moving between them. In terms of their magnitude, the same-EA effects are approximately around half the size of the same-region or same-industry coefficients and two-thirds of the same-collective-agreement coefficient. Note that these same-EA effects are already stripped out of the direct EA effects, both in individual terms (through direct controls and firms fixed effects) and in match terms (through a both-EA-affiliated dummy variable).

We now turn a complementary analysis of the counts of workers moving between each pair of firms (zero, one, or more). Table 4 presents the results from our estimation of a Poisson model that captures both the extensive margin above but also the intensive margin in which several employees may be moving between a specific pair of firms. We find very similar results to those of the previous table in that the same-EA coefficients are positive when not controlling for the common region, industry and collective agreement characteristics of both firms, but these coefficients become negative when considering such variables. In the latter two cases, we find that same-EA effects are of around -70% and statistically significant at the 0.1%. The signs of the control variables are also the same as in Table 3.

Overall, these two sets of results support the view that firms that are affiliated are less likely to exhibit inter-firm worker mobility. This result emerges once we control for firms' possible common characteristics along other dimensions that may also influence worker mobility, which would otherwise have their effects picked up by the same-EA variable. Our findings are consistent with the view that EAs can facilitate coordination across affiliated firms towards diminished worker mobility, thus increase such firms' ability to fully benefit from their investments in the training of their workforce. In the next subsection, we examine the extent to which workers are effectively receiving more training in EA firms.

6.5 Training results

Our analysis of training differentials between EA and non-EA firms is similar to the approach of the previous subsection in that we consider both the extensive and intensive margins (train or no train vs different hours of training), using either linear probability or Poisson models. In this case, we consider the following equation:

$$tr_{e,i,t} = \beta_1 EAaffiliated_i + \beta_2 X_{e,i,t} + \beta_t + a_i + v_{e,i,t} \tag{17}$$

The dependent variable, $tr_{e,i,t}$, is either a dummy variable equal to one if worker e receives firm-provided training in firm i in year t, or the actual count of hours received by the work. As before, $EAaffiliated_i$ is a dummy variable equal to one if firm i is EA affiliated. $X_{e,i,t}$ is a set of worker and firm control variables (namely age, schooling, tenure, and female; and a 2011 dummy, number of workers and sales volume). These variables can explain differences in training across workers and also be correlated with the EA status of their firm. a_i denotes a worker fixed effect, exploiting the fact that our data includes instances of worker mobility between affiliated and non-affiliated firms. The key parameter is $beta_1$, which indicates the average difference between workers in affiliated and non-affiliated firms regarding the training they received.

In Table 5, we present our results concerning the extensive margin. We find positive and statistically significant coefficients across all specifications except in column 4, which includes worker fixed effects and firm controls (firm size measured in both number of workers and total sales). The last result may follow from the limited within-worker variation in EA-firm status, given the short, two-year period covered in our data. Another important aspect concerns the legal requirement (subject to several caveats that most employees should receive at least 35 hours of training per year, which could lead to limited variation across workers in the dependent variable in this model.

In this context, we now turn to Table 6, which presents the results of the Poisson model based on the number of hours of training per worker as the dependent variable. Here we find statistically significant, positive effects of EA firms across all specifications, including in specifications with worker fixed effects.[36] The coefficients vary between 0.152 and 0.318 and are always significant at least at the 1% level. These results indicate that the amount of training provided at EA firms is substantially, at least 15% higher at EA firms.[37]

6.6 Separations and wages results

In our main results above, we have established that inter-firm worker mobility is lower between EA firms and that the latter provide much higher levels of training to their employees. In this subsection, we investigate further two complementary dimensions of the main results above. First, we study the separation behaviour of employees in EA firms. Second, we examine the wage differentials of such firms.

In the first group of analysis, we estimate the following equation:

$$s_{e,i} = \lambda_1 EAaffiliated_i + \lambda_2 t_{e,i} + \lambda_3 X_{e,i} + \lambda_t + x_{e,i} \tag{18}$$

The dependent variable, $s_{e,i}$, is a separation dummy variable of worker e in firm i in (October of) 2010. We consider separations from the firm, regardless of the destination of the worker (workers that leave their 2010 firm in the sense of not appearing in the QP data set in 2011) and separations specifically to other firms (workers that appear in a different firm in 2011). 26.8% of the 2.54 million workers employed in 2010 As before, $EAaffiliated_i$ is the

[36]Note that the number of observations used in the latter case is substantially smaller than in models without worker fixed effects. This is because the estimation dropped 3.1 million observations that are either singletons or separated by a fixed effect Correia et al. (2020).

[37]The coefficients of the remaining control variables are also of general interest. They indicate that training tends to be lower for older and female workers, and higher for more educated and higher-tenure workers.

key explanatory variable, a dummy equal to one if firm i is EA affiliated, and $X_{e,i}$ are a set of worker and firm control variables. Some specifications also control for $t_{e,i}$, the amount of training received by the worker. Note that the t subscript does not appear in the variables above as the data considers only 2010 observations (2011 data is used only for the purpose of establishing the type of mobility, if any, of the worker).

Table 7 presents the results considering any type of mobility (to another firm or to outside the data set). We find that workers in EA firms in 2010 are less likely to leave their firms (as of 2011). The differentials are of around 3% in all specifications and always statistically significant at the 1% level. We also find that workers that received more training in 2010 are less likely to leave their firm (columns 3 and 4), although this gap does not seem to vary with EA affiliation.[38]

Table A1 presents the results considering a different dependent variable and sample. In this case, a separation occurs only when a worker leaves to a different firm. Moreover, the sample includes only individuals that stay in the same firm or that move to a different firm (workers that leave the sample to non-employment or employment outside the QP data base after 2010 are excluded from the analysis). The latter corresponds to only 5.8% of 1.97 million workers employed in 2010. The results indicate a smaller probability to separate for workers in EA firms but which is not statistically significant in most cases.

Finally, we investigate the relationship between wages, EA affiliation and training by estimating different wage equations as follows:

$$w_{e,i,t} = \lambda_1 EAaffiliated_{i,t} + \lambda_2 tr_{e,i,t} + \lambda_3 X_{e,i,t} + \lambda_t + x_{e,i,t} \qquad (19)$$

In all cases, the dependent variable, $w_{e,i,t}$, corresponds to the log monthly wages of worker e in firm i in (October of) year t. As before, $tr_{e,i,t}$ denotes the training hours of worker e in firm i in year t, and $EAaffiliated_{i,t}$ is a dummy variable to indicate if worker i in year t is in an EA-affiliated firm. $X_{e,i,t}$ represents a set of worker and firm control variables and λ_t is a 2011 dummy variable.

The results are presented in Tables 8 and 9. They indicate that workers that receive training are paid higher wages, with premiums of around 3% per week of training. Moreover, EA firms appear to either not pay higher wages than non-EA firms or to pay slightly higher

[38] The coefficients of the control variables indicate that older and more schooled workers are more likely to leave, while women and more tenured workers are less likely to do so.

wages (with premiums of less than 2% in all cases). All other coefficients (schooling, agem tenure, etc) present standard results.

7 Conclusions

Training can have important positive effects both for firms and workers. As workers enter the labour market, in many cases with little professional experience, or when industries are subject to technology shocks, training can greatly increase productivity and potentially wages too.

A potentially key driver of the incentives for firms to engage in training is the likelihood that their employees will not leave soon after receiving the training. This likelihood will be greatly influenced by the degree of competition in the labour market, namely in terms of potential poaching by other firms.

In this paper, we present a simple but novel model that is general enough to capture a range of market and bargaining structures, which we use to investigate the effect of labour market competition on training and profits.

We then tested some predictions of the model using matched panel data, including worker-level information on wages, training, and mobility across firms for all employees in Portugal. We examine the role of employers' associations (EAs), an important but little studied labour market institution. As their membership includes a large number of firms operating in the same industry, these EAs can promote tacit no-poach agreements, leading to lower levels of mobility of workers between EA firms. Such reduced worker turnover can promote increased investments in training and higher productivity, although not necessarily higher wages.

After constructing a large data set including all realised (and a sample of unrealised) spells of worker mobility between firms, we find that workers are much less likely to move between firms affiliated in the same EA. Moreover, again consistently with our theory, we find much higher levels of training in EA firms, even after controlling for a number of additional potential training drivers. Despite the above - and the significant training wage premiums observed in the data -, EA firms do not pay significantly higher wages to their employees.

We believe that this paper may have a number of policy implications. First, competition agencies and perhaps labour inspectorates may need to pay more attention to labour market collusion, including by EAs. While the literature was recently considered the case of covenants not to compete signed by employees, coordination amongst employers may also be a relevant source of restrictions in competition in labour markets.

Second, policy makers may want to incentivise training provided by employers but at the same minimise promote a more balanced sharing of its returns to employees. Collective bargaining may be a tool to promote such rent sharing. However, if such bargaining is conducted primarily at a sectoral level, involving EAs, the enhanced scope for employer collusion that can follow may also need to be addressed.

References

Acemoglu, D. (1997), 'Training and Innovation in an Imperfect Labour Market', *The Review of Economic Studies* **64**(3), 445–464.

Acemoglu, D. & Pischke, J.-S. (1998), 'Why do firms train? theory and evidence', *The Quarterly Journal of Economics* **113**(1), 79–119.

Acemoglu, D. & Pischke, J.-S. (1999), 'The structure of wages and investment in general training', *Journal of political economy* **107**(3), 539–572.

Azar, J. A., Marinescu, I., Steinbaum, M. I. & Taska, B. (2020), 'Concentration in US labor markets: Evidence from online vacancy data', *Labour Economics* **66**.

Azar, J., Marinescu, I. & Steinbaum, M. I. (2020), 'Labor Market Concentration', *Journal of Human Resources* .

Balmaceda, F. (2005), 'Firm-sponsored general training', *Journal of Labor Economics* **23**(1), 115–133.

Buchinsky, M., Fougere, D., Kramarz, F. & Tchernis, R. (2010), 'Interfirm mobility, wages and the returns to seniority and experience in the United States', *Review of Economic Studies* **77**(3), 972–1001.

Burdett, K. & Coles, M. (2003), 'Equilibrium wage-tenure contracts', *Econometrica* **71**(5), 1377–1404.

Correia, S., Guimarães, P. & Zylkin, T. (2020), 'Fast Poisson estimation with high-dimensional fixed effects', *Stata Journal* **20**(1), 95–115.

Dey, M. S. & Flinn, C. J. (2005), 'An equilibrium model of health insurance provision and wage determination', *Econometrica* **73**(2), 571–627.

Hijzen, A. & Martins, P. S. (2020), 'No Extension without Representation? Evidence from a Natural Experiment in Collective Bargaining', *IZA Journal of Labor Economics* **9**(5), 1–31.

Hijzen, A., Martins, P. S., Schank, T. & Upward, R. (2013), 'Foreign-owned firms around the world: A comparative analysis of wages and employment at the micro-level', *European Economic Review* **60**(C), 170–188.

Krueger, A. B. & Ashenfelter, O. (2018), Theory and evidence on employer collusion in the franchise sector, Working Paper 24831, NBER.

Lentz, R. & Roys, N. (2015), Training and Search On the Job, NBER Working Papers 21702, National Bureau of Economic Research, Inc.

Meccheri, N. (2009), 'A note on noncompetes, bargaining and training by firms', *Economics Letters* **102**(3), 198–200.

Moen, E. R. & Åsa Rosén (2004), 'Does poaching distort training?', *The Review of Economic Studies* **71**(4), 1143–1162.

New York Times (2018), 7 Fast-Food Chains to End No Poach Deals That Lock Down Low-Wage Workers, July 12.

OECD (2019), *Negotiating Our Way Up*, OECD, Paris.

Patault, B. & Valtat, A. (2020), Large firms' collusion in the labor market: evidence from collective bargaining, ENSAE, Paris, mimeo.

Pigou, A. C. (1912), *Wealth and Welfare*, London.

Posner, E. A., Triantis, G. G. & Triantis, A. J. (2004), 'Investing in human capital: The efficiency of covenants not to compete', *U Chicago Law & Economics, Olin Working Paper* (137), 01–08.

Postel-Vinay, F. & Robin, J.-M. (2002*a*), 'The distribution of earnings in an equilibrium search model with state-dependent offers and counteroffers', *International Economic Review* **43**(4), 989–1016.

Postel-Vinay, F. & Robin, J.-M. (2002*b*), 'Equilibrium wage dispersion with worker and employer heterogeneity', *Econometrica* **70**(6), 2295–2350.

Postel-Vinay, F. & Robin, J.-M. (2004), 'To match or not to match?: Optimal wage policy with endogenous worker search intensity', *Review of Economic Dynamics* **7**(2), 297–330.

Shimer, R. (2006), 'On-the-job search and strategic bargaining', *European Economic Review* **50**(4), 811–830.

Stevens, M. (1994), 'A theoretical model of on-the-job training with imperfect competition', *Oxford Economic Papers* pp. 537–562.

Stevens, M. (2001), 'Should Firms Be Required To Pay For Vocational Training?', *The Economic Journal* **111**(473), 485–505.

U.S. Department of Justice (2021), Former Owner of Health Care Staffing Company Indicted for Wage Fixing - Antitrust Division Remains Committed to Prosecuting Collusion in Labor Markets, Technical report.

White House (2021), Promoting competition in the American economy, Technical report.

Wickelgren, A. L. (2018), 'A novel justification for legal restrictions on non-compete clauses', *International Review of Law and Economics* **54**, 49–57.

Table 1: **Descriptive statistics: inter-firm mobility**

	Mobility type:			
	Realised		Realised+Potential	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Positive N. of movers (DV)	1.000		0.025	
N. of movers	1.254	1.313	0.032	0.288
Same EA	0.076		0.208	
Same CBA	0.299		0.081	
Same region	0.556		0.107	
Same industry	0.243		0.047	
EA affiliated (2010)	0.514		0.78	
EA affiliated (2011)	0.512		0.782	
EA affiliated (2010 and 2011)	0.287		0.68	
Employees (2010)	838.5	2777.13	64.84	539.747
Employees (2011)	826.1	2675.25	68.17	531.926
N (firm pairs)	79,082		3,106,783	

Notes:

Table 2: **Descriptive statistics: workers (2010 and 2011)**

Variable	**Mean**	**Std. Dev.**	**N**
Schooling	9.348	4.02	5113319
Age	39.293	11.092	5120851
Tenure	8.045	8.412	5126812
Female	0.455	0.498	5127627
EA firm	0.554	0.497	5127627
Firm employees	1054.631	3134.778	5127627
Firm sales	185.016	784.794	5127627
Year 2011	0.472	0.499	5127627
Training (0/1)	0.32	0.466	5127627
Training weeks	0.332	1.149	5127627
Log earnings	6.646	0.685	4840909

Notes:

Table 3: **Inter-firm mobility: extensive margin**

	(1)	(2)	(3)	(4)
Same EA	0.004***	0.011***	-0.023***	-0.042***
	(7.59)	(23.27)	(-33.66)	(-55.50)
EA affiliated	-0.038***	-0.036***	-0.035***	-0.032***
(2010 and 2011)	(-32.52)	(-36.10)	(-32.68)	(-34.74)
Employees (2010)	0.000***		0.000***	
	(7.36)		(7.66)	
Employees (2011)	0.000***		0.000***	
	(7.90)		(7.95)	
EA affiliated (2010)	-0.003**		-0.000	
	(-2.95)		(-0.45)	
EA affiliated (2011)	-0.005***		-0.002	
	(-3.80)		(-1.43)	
Same CBA			0.054***	0.065***
			(47.10)	(61.27)
Same region			0.098***	0.105***
			(80.14)	(101.35)
Same industry			0.089***	0.088***
			(51.67)	(58.81)
Constant	0.049***	0.047***	0.030***	0.035***
	(50.49)	(72.32)	(34.05)	(56.94)
Firm controls x2	X		X	
Firm FE x2		X		X
Observations	3106783	3106783	3106783	3106783

Notes: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4: **Inter-firm mobility: intensive margin**

	(1)	(2)	(3)	(4)
Same EA	0.201***	0.750***	-0.707***	-0.657***
	(3.67)	(15.02)	(-17.10)	(-13.49)
EA affiliated (2010 and 2011)	-1.206***	-1.111***	-1.037***	-0.958***
	(-14.54)	(-17.50)	(-12.70)	(-13.70)
Employees (2010)	0.000***		0.000***	
	(16.56)		(12.64)	
Employees (2011)	0.000***		0.000***	
	(15.87)		(14.20)	
EA affiliated (2010)	-0.204***		-0.106*	
	(-4.98)		(-2.57)	
EA affiliated (2011)	-0.220***		-0.115**	
	(-5.63)		(-2.86)	
Same CBA			1.175***	1.165***
			(36.73)	(26.26)
Same region			1.800***	2.116***
			(32.54)	(31.78)
Same industry			1.061***	1.154***
			(18.74)	(20.07)
Constant	-2.237***	-1.423***	-3.318***	-2.698***
	(-107.51)	(-88.52)	(-64.13)	(-57.65)
Firm controls x2	X		X	
Firm FE x2		X		X
Observations	3106783	3106783	3106783	3106783

Notes: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5: **Training: extensive margin**

	(1)	(2)	(3)	(4)
EA firm	0.074***	0.044***	0.041*	0.019
	(5.21)	(6.68)	(2.45)	(1.20)
Age	-0.003***	-0.003***	-0.001*	-0.001*
	(-8.31)	(-21.91)	(-2.21)	(-2.18)
Schooling	0.021***	0.012***	0.005***	0.006***
	(18.89)	(32.26)	(4.86)	(5.75)
Tenure	0.006***	0.003***	0.003***	0.002***
	(13.20)	(11.96)	(4.02)	(3.49)
Female	-0.020**	-0.010***	0.027**	0.007
	(-2.72)	(-4.75)	(2.87)	(0.85)
Year 2011	0.032***	0.027***	0.010*	0.011*
	(7.21)	(6.31)	(2.17)	(2.30)
Workers		0.000*		0.000
		(2.03)		(1.26)
Sales		0.000		0.000
		(1.06)		(1.06)
Constant	0.143***	0.246***	0.280***	0.293***
	(7.08)	(28.54)	(9.77)	(10.60)
Firm controls		X		X
Worker FE			X	X
Observations	5105988	5105987	4149389	4149387

Notes: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6: **Training: intensive margin**

	(1)	(2)	(3)	(4)
EA firm	0.216***	0.152**	0.318***	0.295***
	(3.86)	(2.59)	(6.87)	(6.50)
Age	-0.0185***	-0.0158***	-0.00976	-0.00963
	(-15.50)	(-15.04)	(-1.76)	(-1.65)
Schooling	0.0949***	0.0713***	0.0290***	0.0279***
	(23.48)	(26.70)	(3.77)	(3.52)
Tenure	0.0224***	0.0104***	0.00953*	0.00588
	(10.49)	(5.70)	(2.56)	(1.46)
Female	-0.141***	-0.0708***	0.0177	-0.0232
	(-5.77)	(-4.92)	(0.36)	(-0.52)
Year 2011	0.0761*	0.0606	-0.0108	-0.0130
	(2.21)	(1.76)	(-0.33)	(-0.40)
Workers		0.0000314		0.0000155
		(1.39)		(1.48)
Sales		-6.13e-11		-1.66e-10***
		(-0.66)		(-3.84)
Constant	-1.642***	-1.272***	0.170	0.275
	(-19.40)	(-17.20)	(0.74)	(1.13)
Firm controls		X		X
Worker FE			X	X
Observations	5105988	5105567	1914511	1914509

Notes: Dependent variable: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 7: **Separation (leaving firm)**

	(1)	(2)	(3)	(4)
EA firm	-0.036***	-0.028***	-0.027***	
	(-7.93)	(-7.27)	(-7.26)	
Workers	-0.000	-0.000	-0.000	
	(-1.72)	(-1.72)	(-1.68)	
Sales	-0.000	-0.000	-0.000	
	(-1.79)	(-1.53)	(-1.49)	
Age		0.003***	0.003***	0.002***
		(6.79)	(6.76)	(5.06)
Schooling		0.003***	0.003***	0.005***
		(3.44)	(3.78)	(5.60)
Female		-0.033***	-0.033***	-0.039***
		(-9.10)	(-9.22)	(-10.21)
Tenure		-0.006***	-0.006***	-0.003***
		(-10.55)	(-10.55)	(-5.99)
Training weeks			-0.014***	-0.013***
			(-6.31)	(-7.53)
EA firm*Training weeks				0.002
				(0.81)
Constant	0.295***	0.215***	0.218***	0.174***
	(101.52)	(9.21)	(9.39)	(6.17)
Firm controls	X	X	X	
Worker controls		X	X	X
Firm FE				X
Observations	2542887	2530584	2530584	2472335

Notes: Dependent variable: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 8: **Wage equation**

	(1)	(2)	(3)	(4)
Schooling	0.082***	0.059***	0.083***	0.060***
(years)	(56.41)	(86.85)	(54.17)	(89.01)
Age	0.045***	0.039***	0.045***	0.039***
	(26.14)	(28.21)	(26.56)	(27.88)
Age^2	-0.000***	-0.000***	-0.000***	-0.000***
	(-24.38)	(-24.68)	(-24.84)	(-24.43)
Tenure	0.039***	0.028***	0.039***	0.028***
	(27.78)	(49.12)	(27.64)	(48.68)
$Tenure^2$	-0.001***	-0.000***	-0.001***	-0.000***
	(-14.66)	(-27.67)	(-14.78)	(-27.82)
Female	-0.283***	-0.214***	-0.285***	-0.215***
	(-44.02)	(-61.41)	(-44.25)	(-61.04)
Training	0.036***	0.025***		
(weeks)	(8.37)	(8.00)		
Year 2011	-0.019***	-0.013***	-0.019***	-0.013***
	(-6.38)	(-5.65)	(-6.37)	(-5.61)
Workers		-0.000		-0.000
		(-1.33)		(-1.45)
Sales		0.000		0.000
		(1.95)		(1.90)
EA firm			-0.006	0.017**
			(-0.49)	(2.96)
Constant	4.731***	5.117***	4.739***	5.113***
	(110.51)	(171.74)	(110.09)	(168.70)
Worker controls	X	X	X	X
Firm controls		X		X
Firm FE				
Observations	4821831	4821830	4821831	4821830

Notes: Dependent variable: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 9: **Wage equation**

	(1)	(2)	(3)	(4)
Schooling	0.0816***	0.0593***	0.0513***	0.0513***
(years)	(54.89)	(86.84)	(76.96)	(76.80)
Age	0.0452***	0.0390***	0.0339***	0.0339***
	(26.34)	(28.18)	(20.83)	(20.86)
Age^2	-0.000437***	-0.000376***	-0.000315***	-0.000315***
	(-24.58)	(-24.65)	(-17.70)	(-17.73)
Tenure	0.0386***	0.0283***	0.0291***	0.0291***
	(27.65)	(48.08)	(43.05)	(43.00)
$Tenure^2$	-0.000553***	-0.000447***	-0.000485***	-0.000485***
	(-14.74)	(-27.41)	(-26.92)	(-26.92)
Female	-0.284***	-0.214***	-0.189***	-0.189***
	(-44.08)	(-61.44)	(-59.11)	(-59.19)
Training	0.0360***	0.0252***	0.0156***	0.0115***
(weeks)	(8.42)	(7.99)	(7.03)	(4.48)
EA firm	-0.00890	0.0160**		
	(-0.68)	(2.80)		
Year 2011	-0.0193***	-0.0135***	-0.0131***	-0.0130***
	(-6.39)	(-5.64)	(-6.74)	(-6.70)
Training *				0.00841*
EA firm				(2.17)
Constant	4.737***	5.109***	5.283***	5.283***
	(109.52)	(169.57)	(144.45)	(144.44)
Worker controls	X	X	X	X
Firm controls		X	X	X
Firm FE			X	X
Observations	4821831	4821830	4799637	4799637

Notes: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

A Appendix: Supplementary Figures and Tables

Table A1: **Separation (different firm)**

	(1)	(2)	(3)	(4)
EA firm	-0.009***	-0.004	-0.004	
	(-3.53)	(-1.68)	(-1.58)	
Workers	-0.000	-0.000	-0.000	
	(-0.07)	(-0.24)	(-0.23)	
Sales	-0.000	-0.000	-0.000	
	(-1.67)	(-1.43)	(-1.43)	
Age		-0.001***	-0.001***	-0.001***
		(-12.00)	(-12.08)	(-11.12)
Schooling		-0.000*	-0.000	-0.000
		(-2.05)	(-1.86)	(-1.63)
Female		-0.008***	-0.009***	-0.009***
		(-5.70)	(-5.74)	(-7.62)
Tenure		-0.003***	-0.003***	-0.002***
		(-8.73)	(-8.74)	(-8.44)
Training weeks			-0.002*	-0.006***
			(-2.21)	(-5.48)
EA firm*Training weeks				0.004*
				(2.46)
Constant	0.067***	0.148***	0.149***	0.131***
	(34.46)	(18.74)	(18.90)	(18.87)
Firm controls	X	X	X	
Worker controls		X	X	X
Firm FE				X
Observations	1974412	1967087	1967087	1901196

Notes: Dependent variable: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B Appendix: Proofs

Proof of Proposition 1:

Proof. Suppose otherwise, so that $\hat{\tau} \neq \tau^*$, where $(\hat{w}_1, \hat{w}_2, \hat{\tau})$ is a solution to problem A. Then consider changing the contract so that $\tau = \tau^*$, $w_2 = \hat{w}_2 + y_2(\tau^*) - y_2(\hat{\tau})$, and $w_1 = \hat{w}_1 - (y_2(\tau^*) - y_2(\hat{\tau}))$. From (3) μ is unchanged and the LHS of (1) is unchanged, so it continues to hold. Moreover the second term in the maximand remains constant, while

$$
\begin{aligned}
y_1 - w_1 - \tau^* &= y_1 - \hat{w}_1 + (y_2(\tau^*) - y_2(\hat{\tau})) - \tau^* \\
&> y_1 - \hat{w}_1 - \hat{\tau},
\end{aligned}
$$

by $y_2(\tau^*) - \tau^* > y_2(\hat{\tau}) - \hat{\tau}$. Hence profits are increased, a contradiction. ■

Proof of Lemma 1:

Proof. Suppose we are not in case (i), so that one of $w_1 \geq \underline{w}$ or $w_2 \geq \underline{w}$ is slack (if both are then profits would be at the unconstrained level). Then (1) cannot be slack. Suppose to the contrary it is. Let $(\hat{w}_1, \hat{w}_2, \hat{\tau})$.be a solution to problem B. If $\hat{w}_1 = \underline{w}$ and $\hat{w}_2 > \underline{w}$, consider the following contract $(\hat{w}_1, w_2, \tau)$, where $w_2 = \hat{w}_2 - \Delta$ and τ is such that $y_2(\tau) = y_2(\hat{\tau}) - \Delta$ for some $\Delta > 0$ small enough that $w_2 \geq \underline{w}$. Then from (3) μ is unchanged, and so profits increase as τ is cut and $\mu(y_2(\tau) - w_2)$ is unchanged, contradicting the supposition. If on the other hand $\hat{w}_1 > \underline{w}$, $\hat{w}_2 = \underline{w}$ and (1) is slack, a cut in w_1 will increase profits without violating any constraints, a contradiction again. Thus we can conclude that (1) is binding. Suppose that $\underline{w} \geq y_2(\hat{\tau})$. Then, by assumption on y_2, $\underline{w} \geq y_2(\hat{\tau}) \geq y_1$, implying profits are negative contrary to hypothesis, as either $w_1 > \underline{w}$ or $w_2 > \underline{w}$ in this case. Thus $\underline{w} < y_2(\hat{\tau})$. Now suppose $w_1 > \underline{w}$ and $w_2 = \underline{w}$. Then by the immediately preceding, $w_2 < y_2(\hat{\tau})$. In this case a small increase in w_2 must increase joint value, and a cut in w_1 will allow the firm to appropriate the increase, thus increasing profits contrary to initial optimality. In detail: joint value is

$$
J(w_1, w_2, \tau) := (y_1 - \tau) + \mu y_2(\tau) + (1 - \mu) E_\theta[w_2 + \beta(y_2(\tau) + \theta - w_2) \mid y_2(\tau) + \theta \geq w_2],
$$

so

$$\begin{aligned}
\frac{\partial J(w_1,w_2,\tau)}{\partial w_2} &= \frac{\partial \mu}{\partial w_2} y_2(\tau) - \frac{\partial \mu}{\partial w_2}(w_2+Z) + (1-\mu) + \qquad (20)\\
&(1-\mu)\frac{\beta(-1+F(w_2-y_2(\tau)) + Zf(w_2-y_2(\tau)))}{1-F(w_2-y_2(\tau))}\\
&= \frac{\partial \mu}{\partial w_2}(y_2(\tau)-w_2) + (1-\mu) - \phi\beta f(w_2-y_2(\tau))Z + \phi\beta(-(1-F(w_2-y_2(\tau))) + Zf(w_2 -\\
&= \frac{\partial \mu}{\partial w_2}(y_2(\tau)-w_2) + (1-\mu) - \phi\beta(1-F(w_2-y_2(\tau)))\\
&= \phi f(w_2-y_2(\tau))(y_2(\tau)-w_2) + \phi(1-\beta)(1-F(w_2-y_2(\tau)))\\
&> 0, \qquad (21)
\end{aligned}$$

unless $f(w_2-y_2(\tau))=0$ *and* $\beta=1$ (recall $f(0)>0$ so $F(w_2-y_2(\tau))<1$), where we write Z for $E_\theta[\beta(y_2(\tau)+\theta-w_2) \mid y_2(\tau)+\theta \geq w_2]$, and are using $\partial\mu/\partial w_2 = \phi f(w_2-y_2(\tau))$ and $(1-\mu)=\phi(1-F(w_2-y_2(\tau)))$.[39] This establishes part (ii). ■

Proof of Proposition 2:

Proof. If the constraint binds in the solution to Problem B, profits are lower than in the unconstrained problem A and so Lemma 1 applies. We consider first case (ii) of the lemma. Assuming that $(w_2-y_2(\tau))$ is in the interior of the support of F,[40]

$$\begin{aligned}
\frac{\partial U(w_1,w_2,\tau)}{\partial w_2} &= 1 - \frac{\partial \mu}{\partial w_2}Z + \\
&(1-\mu)\frac{\beta(-1+F(w_2-y_2(\tau)) + Zf(w_2-y_2(\tau)))}{1-F(w_2-y_2(\tau))}\\
&= 1-\phi\beta(1-F(w_2-y_2(\tau))),
\end{aligned}$$

where Z is as defined in the proof of Lemma 1. Likewise

$$\frac{\partial U(w_1,w_2,\tau)}{\partial \tau} = \phi\beta y_2'(\tau)(1-F(w_2-y_2(\tau))).$$

[39] We cannot rule out $w_1 > \underline{w}$ and $w_2 = \underline{w}$ in the case where $f(w_2-y_2(\tau))=0$ and $\beta=1$. However in this case profits are unchanged if $w_1=\underline{w}$ and $w_2>\underline{w}$ (an equivalent way of satisfying the participation constraint, with μ and τ unaffected). Hence we can assume $w_1=\underline{w}$ and $w_2>\underline{w}$ without loss of generality.

[40] Otherwise, w_2 is below the entire support of $y_2+\theta$ (it would be sub-optimal if it was above) and the result follows straightforwardly provided $\beta<1$. Intuitively, the worker always leaves if there is a contact, so the marginal return to training retained by the match is smaller than y_2' in contact states because – with fixed w_2 – surplus going to the outside firm increases as the $y_2+\theta$ distribution shifts to the right. Hence as ϕ increases, this happens more frequently, reducing the return. (In addition if $\beta>0$, so w_2 falls, this is magnified.) But when $\beta=1$, there is no additional loss of surplus, so the return is independent of ϕ and τ is constant rather than decreasing in ϕ.

Consequently FOCs are:

$$-1+\partial\mu/\partial\tau\left(y_{2}\left(\tau\right)-w_{2}\right)+\mu y_{2}'\left(\tau\right)+\lambda\phi\beta y_{2}'\left(\tau\right)\left(1-F\left(w_{2}-y_{2}\left(\tau\right)\right)\right)=0, \tag{22}$$

$$\phi f\left(w_{2}-y_{2}\left(\tau\right)\right)-\mu+\lambda\left(1-\phi\beta\left(1-F\left(w_{2}-y_{2}\left(\tau\right)\right)\right)\right)=0, \tag{23}$$

and

$$-1+\lambda+\xi=0, \tag{24}$$

where λ is the Lagrange multiplier on (3) and ξ that on (4) at $t=1$. From (22) and (23) $\lambda=1/y_{2}'\left(\tau\right)$. Hence from (24), $\xi=0$ implies $y_{2}'\left(\tau\right)=1$, so $\tau=\tau^{*}$, and we established that (1) is always binding, so we would have $\Pi=\Pi^{*}$, contrary to hypothesis. So $\xi>0$ and we get $y_{2}'\left(\tau\right)>1$, hence $\tau<\tau^{*}$ by $y_{2}''<0$.

To establish monotonicity, first-order conditions (differentiating (5) w.r.t. τ) are

$$1=y_{2}'\left[1+\phi\left(\frac{\partial S}{\partial y_{2}}+\frac{\partial S}{\partial w_{2}}\frac{dw_{2}}{dy_{2}}\right)\right]. \tag{25}$$

It is convenient to write

$$S\left(y_{2},w_{2}\right)=\hat{S}\left(w_{2}-y_{2}\right),$$

where

$$\hat{S}\left(x\right):=\int_{x}^{\overline{\theta}}\left(x+\beta\left(\theta-x\right)\right)dF. \tag{26}$$

Thus $\frac{\partial S}{\partial w_{2}}=-\frac{\partial S}{\partial y_{2}}=\hat{S}'$, so (25) becomes

$$1=y_{2}'\left[1-\phi\hat{S}'\left(1-\frac{dw_{2}}{dy_{2}}\right)\right]. \tag{27}$$

By the lemma $w_{1}=\underline{w}$ and the participation constraint binds, so

$$\underline{w}+w_{2}+\phi\int_{w_{2}-y_{2}}^{\overline{\theta}}\beta\left(y_{2}+\theta-w_{2}\right)dF=\underline{U} \tag{28}$$

and differentiating and rearranging yields

$$1-\frac{dw_{2}}{dy_{2}}=\frac{1}{1-\phi\beta\left(1-F\right)}.$$

Thus (38) becomes

$$1 = y_2' \left[1 - \phi \hat{S}' \left(\frac{1}{1 - \phi\beta (1 - F)} \right) \right]. \tag{29}$$

Consider now an increase in ϕ, and suppose to the contrary of the claim that y_2 does not fall. $F(w_2 - y_2)$ weakly decreases unless w_2 increases. However we can rule this out, as from (28),

$$\frac{dw_2}{d\phi} (1 - \phi (1 - F) \beta) = -\phi (1 - F) \beta \frac{dy_2}{d\phi} - \int_{w_2 - y_2}^{\overline{\theta}} \beta (y_2 + \theta - w_2) \, dF,$$

so $\frac{dy_2}{d\phi} \geq 0$ implies $\frac{dw_2}{d\phi} \leq 0$. Thus the term in round brackets in (29) does not decrease. $\hat{S}'' (x) = (-2 + \beta) f (x) - x f' (x) < 0$ by assumption, and $\hat{S}'$ does not fall, as its argument $w_2 - y_2$ does not increase ($\hat{S}'$ must be positive given $\tau < \tau^*$, i.e., $y_2' > 1$). Then as ϕ is higher, the term in square brackets has decreased, so that y_2' must increase. This contradicts the hypothesis that y_2 does not fall. We conclude then that τ falls under the assumption that guarantees $\hat{S}'' < 0$.

In the case of part (i) of the lemma, the return from investment when $w_1 = w_2 = \underline{w}$ is $\mu y_2' (\tau) + (\partial \mu / \partial \tau) (y_2 (\tau) - \underline{w})$, where $\partial \mu / \partial \tau = -\phi y_2' (\tau) f (\underline{w} - y_2 (\tau))$ assuming $\underline{w} - y_2 (\tau)$ lies in the interior of the support of F. Holding τ constant, and differentiating with respect to ϕ yields:

$$(F (\underline{w} - y_2 (\tau)) - 1) y_2' (\tau) - y_2' (\tau) f (\underline{w} - y_2 (\tau)) < 0.$$

So as ϕ increases, the return to investment falls and so τ is reduced. If $\underline{w} - y_2 (\tau)$ lies below the support of F then the expression equals $-y_2' (\tau) < 0$. (Note this does not require the restriction on F.) ■

Proof of Proposition 3.

Proof. (a) For $\underline{w} < w_2 \leq y_2 (\tau)$, a small increase in τ of say $\Delta\tau$ leads to period 2 output increasing by $(y_2 (\tau + \Delta\tau) - y_2 (\tau))$ when $\theta < 0$, while the worker leaves when $\theta \geq 0$ but receives an extra $(y_2 (\tau + \Delta\tau) - y_2 (\tau))$. Thus total match surplus rises by $(y_2 (\tau + \Delta\tau) - y_2 (\tau))$, and a cut in w_2 to satisfy the worker's participation constraint implies the firm receives the full return. So $\tau = \tau^*$. (b) As discussed in the text, for $w_2 > y_2 (\tau)$ the proof of Proposition 2 applies so $\tau < \tau^*$, and the same condition on F guarantees that τ is strictly decreasing in ϕ. (Note that in this case S simplifies to $(1 - F(w_2 - y_2)) (w_2 - y_2)$.)

Next, if $\underline{w}$ is such that $\underline{U} - \underline{w} > y_2 (\tau^*)$, then the participation constraint is satisfied with $w_1 = \underline{w}$ and $w_2 = \underline{U} - \underline{w}$ as the worker receives w_2 in all states given $w_2 > y_2 (\tau^*) > y_2(\tau)$.

In case (a) profits are constant in ϕ as $\tau = \tau^*$ and joint value is constant (the worker adds $y_2(\tau^*)$ to value whether she stays or leaves). In case (b) profits are increasing in ϕ as $S(y_2, w_2) > 0$ from (5) (profits rise if τ is held constant, although optimally τ may be cut).

Proof of Proposition 4

We follow a similar logic to the proof of Proposition 2. Given the PC is assumed to be binding, the firm maximizes the expression in (5) but where the surplus is given by

$$S(w_2, y_2) = \int_{b-y_2+(w_2-b)/\beta}^{\overline{\theta}} (-y_2 + b + \beta(y_2 + \theta - b))\, dF,$$

so (25) becomes

$$1 = y_2' \left[1 + \phi\left((1-F)(-1+\beta) + (-y_2 + w_2) f - (1/\beta)(-y_2 + w_2) f \frac{dw_2}{dy_2}\right)\right]. \tag{30}$$

The participation constraint binds, so[41]

$$\underline{w} + w_2 + \phi \int_{b-y_2+(w_2-b)/\beta}^{\overline{\theta}} (-w_2 + b + \beta(y_2 + \theta - b))\, dF = \underline{U}, \tag{31}$$

and differentiating and substituting back into (30), we get

$$1 = y_2' \left[1 + \phi\left\{-(1-F)(1-\beta) - (y_2 - w_2) f \cdot (1/(1 - \phi(1-F)))\right\}\right]. \tag{32}$$

Consider now an increase in ϕ, and suppose that y_2 does not fall. w_2 weakly decreases as from differentiating (30), after rearrangement:

$$\frac{dw_2}{d\phi}(1 - \phi(1-F)) = -\int_{b-y_2+(w_2-b)/\beta}^{\overline{\theta}} (-w_2 + b + \beta(y_2 + \theta - b))\, dF - \phi(1-F)\beta\frac{dy_2}{d\phi},$$

and $\frac{dy_2}{d\phi} \geq 0$ implies $\frac{dw_2}{d\phi} \leq 0$. $F(b - y_2 + (w_2 - b)/\beta)$ weakly decreases, $(y_2 - w_2)$ weakly increases and so the term in square brackets in (32) decreases if $f(b - y_2 + (w_2 - b)/\beta)$ does not fall, which follows given $f' \geq 0$. This implies that y_2 decreases, contradicting the assumption.

[41] Assuming below that $w_1 = \underline{w}$ and $w_2 > \underline{w}$. We can rule out $w_1 > \underline{w}$ and $w_2 = \underline{w}$ following the argument of the lemma, but it is possible that $w_1 = \underline{w}$ and $w_2 = \underline{w}$ *and* the PC binds (as τ can be varied to maintain the PC). However $d\tau/d\phi < 0$ in this case as worker utility is increasing in ϕ whenever the support of the outside wage doesn't lie entirely below w_2, so τ is reduced to maintain the PC.

B.1 Details for specific investment case

The argument in Lemma 1 can be repeated to show the participation constraint binds, so again profits equal joint value net of the worker's outside value $\underline{U}$:

$$\Pi(\phi) = (y_1 - \tau - \sigma) + y_2(\tau) + s_2(\sigma) + \phi S(y_2, s_2, w_2) - \underline{U}, \tag{33}$$

where $S(y_2, s_2, w_2)$ is now the expected value of the wage net of output $y_2(\tau) + s_2(\sigma)$ that is lost to the match when the worker leaves, or zero when the worker stays:

$$S(y_2, s_2, w_2) = \int_{w_2 - y_2}^{\overline{\theta}} \left(-(y_2 + s_2) + w_2 + \beta(y_2 + \theta - w_2)\right) dF. \tag{34}$$

First-order conditions (differentiating (33) w.r.t. τ, σ,respectively) are

$$1 = y_2' \left[1 + \phi\left(\frac{\partial S}{\partial y_2} + \frac{\partial S}{\partial w_2}\frac{dw_2}{dy_2}\right)\right], \tag{35}$$

and

$$1 = s_2' \left[1 + \phi\left(\frac{\partial S}{\partial s_2} + \frac{\partial S}{\partial w_2}\frac{dw_2}{ds_2}\right)\right], \tag{36}$$

where $\frac{dw_2}{dy_2}$ and $\frac{dw_2}{ds_2}$ arise from differentiating the PC,

$$w_1 + w_2 + \phi \int_{w_2 - y_2}^{\overline{\theta}} \beta(y_2 + \theta - w_2)\, dF = \underline{U}, \tag{37}$$

so $\frac{dw_2}{ds_2} = 0$. It is convenient to write

$$S(y_2, s_2, w_2) = \hat{S}(w_2 - y_2) - (1 - F(w_2 - y_2))\, s_2,$$

where $\hat{S}$ is as defined in (26):

$$\hat{S}(x) := \int_x^{\overline{\theta}} (x + \beta(\theta - x))\, dF.$$

Thus $\frac{\partial S}{\partial w_2} = -\frac{\partial S}{\partial y_2} = \hat{S}' + f \cdot s_2$ so (35) becomes

$$1 = y_2' \left[1 - \phi\left(\hat{S}' + f \cdot s_2\right)\left(1 - \frac{dw_2}{dy_2}\right)\right], \tag{38}$$

and (36) becomes

$$1 = s_2' \left[1 - \phi \left(1 - F\right)\right], \tag{39}$$

as $\frac{dw_2}{ds_2} = 0$.

Then when the minimum wage constraint binds we have $w_1 = \underline{w}$ and the participation constraint binds, so differentiating (37) and rearranging

$$1 - \frac{dw_2}{dy_2} = \frac{1}{1 - \phi\beta \left(1 - F\right)}.$$

Thus (38) becomes

$$1 = y_2' \left[1 - \phi \left(\hat{S}' + f \cdot s_2\right) \left(\frac{1}{1 - \phi\beta \left(1 - F\right)}\right)\right]. \tag{40}$$

In the case discussed in the text where where the match productivity shock θ is uniformly distributed and training costs are quadratic, $y_2(\tau) = \tau^{1/2}$, $s_2(\sigma) = \alpha\sigma^{1/2}$, and $\beta = 0$, then

$$s_2(\sigma) = \frac{\alpha^2(\overline{\theta} - \underline{\theta} - \overline{\theta}\phi + \phi(w_2 - y_2(\tau)))}{2(\overline{\theta} - \underline{\theta})}.$$

Using this, (40) can be solved for

$$y_2(\tau) = \frac{-\phi(\overline{\theta} - \underline{\theta})\left(\alpha^2 + 2\overline{\theta} - 4w_2\right) + 2(\overline{\theta} - \underline{\theta})^2 + \alpha^2\phi^2(\overline{\theta} - w_2)}{-\alpha^2\phi^2 + 4\phi(\overline{\theta} - \underline{\theta}) + 4(\overline{\theta} - \underline{\theta})^2},$$

and $w_2 = \underline{U} - \underline{w}$ from (37) when $\beta = 0$. These expressions are used for the computed examples.

Next, consider the case as in Proposition 5, where the minimum wage constraint is not binding, and in the uniform/quadratic costs case but with general β. Then maximizing the expression in (33) with respect to σ and w_2, we get for any interior solution ($0 < \mu < 1$)

$$s_2 = \frac{\alpha^2((\beta - 2)(\overline{\theta} - \underline{\theta}) + \phi\overline{\theta})}{\alpha^2\phi + 2(\beta - 2)(\overline{\theta} - \underline{\theta})}.$$

Then it is straightforward but tedious to show $ds_2/d\phi < 0$. For $\mu = \phi$, $ds_2/d\phi < 0$ from (39) as $F = 0$, while for $\mu = 1$, $F = 1$ so $ds_2/d\phi = 0$.

Proof of Proposition 6:

Proof. We prove the first claim. The remainder of the proposition can be establised straightforwardly using similar arguments. Assume that $\Pi(\phi)$ is strictly increasing on $[0, \overline{\phi}]$, and $\Pi(0)+B < \Pi(\overline{\phi})$ with $B>0$. First, at $\gamma_{EA}=1$, $\Pi(\phi_{EA}(1))+B < \Pi(\phi_O(1))$ using (14) and $\Pi(0)+B < \Pi(\overline{\phi})$. Hence $\Pi^{EA}(1) < \Pi^O(1)$ Likewise, at $\gamma_{EA}=0$, $\Pi(\phi_{EA}(0))+B > \Pi(\phi_O(0))$ by $\phi_{EA}(0)=\phi_O(0)$ (from (14)) and $B>0$, so $\Pi^{EA}(1) > \Pi^O(1)$ Using (13) and $\Pi(\phi)$ increasing , Π^{EA} is decreasing and Π^O increasing in γ_{EA}; so by continuity there is a unique γ^*_{EA}, $0<\gamma^*_{EA}<1$, with $\Pi^{EA}(\gamma^*_{EA})=\Pi^O(\gamma^*_{EA})$. For a small small perturbation of γ_{EA} to $\gamma^*_{EA}+\varepsilon$, $\varepsilon>0$, $\Pi^{EA}(\gamma_{EA}) < \Pi^O(\gamma_{EA})$ by Π^{EA} decreasing, Π^O increasing; if $\varepsilon<0$ then $\Pi^{EA}(\gamma_{EA}) > \Pi^O(\gamma_{EA})$. Hence the equilibrium is stable. ■

Proof of Proposition 7:

Proof. At $\gamma_{EA}=1$ we have $\hat{\Pi}_{EA}=0$ as $\rho_{EA}(1)=0$ by (15). Hence $\Pi(0)+B(1) < \Pi(\overline{\phi})+\hat{\Pi}_O$ implies that $\Pi^{EA}(1) < \Pi^O(1)$. At $\gamma_{EA}=0$, $\rho_{EA}(0)=\underline{\phi}$ by (15).and as all contacts for non-EA firms come from workers at other non-EA firms ($m_{EA}=0$),[42] $\rho^O_O=\rho_O(0)=\underline{\phi}$ and $\rho^{EA}_O=0$. It follows that at $\gamma_{EA}=0$, $\hat{\Pi}_{EA}=\hat{\Pi}_O$. So by $B(0)>0$ and $\phi_{EA}(0)=\phi_O(0)$ (from (14), which implies $\Pi(\phi_{EA}(0))=\Pi(\phi_O(0))$, we have $\Pi^{EA}(0) > \Pi^O(0)$. By continuity there must be at least one interior value for γ_{EA}, γ^*_{EA}, such that $\Pi^{EA}(\gamma^*_{EA})=\Pi^O(\gamma^*_{EA})$, and such that for γ_{EA} in a neighborhood of γ^*_{EA}, $\Pi^{EA}(\gamma_{EA}) > \Pi^O(\gamma_{EA})$ for $\gamma_{EA} < \gamma^*_{EA}$ and $\Pi^{EA}(\gamma_{EA}) < \Pi^O(\gamma_{EA})$ for $\gamma_{EA} > \gamma^*_{EA}$, so the equilibrium is stable.[43] ■

[42] Of course in the limit when $\gamma_{EA}=0$ there are no firms in the EA, but in a discrete model with a single firm in the EA all its period 2 hires would come from outside. Formally, given that $\phi_{EA}(\gamma_{EA}) \equiv m_{EA}(\gamma_{EA})/\gamma_{EA}$ has a limit of $\underline{\phi}$ as $\gamma_{EA}\to 0$, $m_{EA}(\gamma_{EA})\to 0$, and by continuity therefore $m_{EA}(0)=0$.

[43] Strictly, this is a generic statement since we cannot rule out $\Pi^{EA}(\gamma_{EA})=\Pi^O(\gamma_{EA})$ in some neighborhood of γ^*_{EA}. In such a case we could generalize the definition of stability to include a stable interval of values of γ_{EA}.