Matching, Wage Rigidities and Efficient Severance Pay

Giulio Fella

November 2012

Abstract

This paper studies the effect of mandated severance pay in a matching model featuring wage rigidity for ongoing, but not new, matches. Mandated severance pay matters only if binding real wage rigidities imply inefficient separation under employment at will. In such a case, large enough severance payments reduce job destruction, and increase job creation and social efficiency, under very mild conditions. Furthermore, mandated severance pay never results in inefficient labor hoarding. Whenever separation is jointly optimal, the parties agree to end the match with a spot severance payment below the statutory one. The marginal effect of mandated severance pay is zero when its size exceeds that which induces the same allocation that would prevail in the absence of wage rigidity. The results hold under alternative micro-foundations for wage rigidity.

JEL classification: E24, J64, J65.

Keywords: Severance pay, renegotiation, wage rigidity

1 Introduction

Mandated employment protection measures associate monetary and shadow costs to workforce adjustment. A common concern is that these “layoff” or “firing” costs result in inefficient labor hoarding and are responsible for the poor employment performance of a number of Continental European countries relative to the United States. This concern has sparked an extensive literature discussed below.

*I am indebted to Paola Manzini for many discussions and to Renato Faccini and two anonymous referees for useful comments.

†School of Economics and Finance; Queen Mary, University of London; Mile End Road, London E1 4NS, UK. Email: g.fella@qmul.ac.uk.
Within this literature, Lazear’s (1990) seminal contribution established that for mandated employment protection to alter the allocation of labor: (1) either employment protection measures must entail a tax component that is lost to the firm-worker pair; or (2) payments between firms and workers must be constrained.

For this reason, the literature on employment protection falls into two main categories. Papers in the first (and larger) category study the effect of firing costs when the latter take the form of a tax on separation. Papers in the second category consider the effect of severance payments in the presence of wage rigidities.

This paper belongs to the second category. It constructs a version of Mortensen and Pissarides’s (1994) matching model which encompasses alternative wage setting mechanisms for ongoing matches, but in which entry wages are fully flexible and (spot) side payments upon separation are unconstrained. It analyzes the two main micro-foundations for wage rigidity: right-to-manage bargaining and efficiency wages.

The paper derives two main results. First, if the real wage that applies to ongoing matches is downward rigid, severance payments increase job creation and social efficiency as long as they reduce job destruction. For large enough severance payments this is indeed the case under very mild conditions, up to the minimum efficient size that induces the same labor allocation as under flexible wages. Second, severance payments never result in privately inefficient labor hoarding. Any increase in severance payments beyond their minimum efficient size is neutral, as it is optimal for the parties to undo them by means of (spot) side payments upon separation. In a nutshell, job destruction is inefficiently high in the absence of severance pay, but severance payments cannot induce inefficient labor hoarding.

The first result can be understood in light of the following two assumptions. First, the wage that applies to ongoing matches is higher than the marginal worker’s reservation wage and is downward rigid, in the sense that the worker can neither offer to work for a lower wage nor, equivalently, rebate part of the wage to the employer by means of side payments. It follows that the job destruction rate is higher than under flexible wages and separation is ex post, hence ex ante, privately inefficient. To the extent that severance payments reduce separation, they increase the ex ante joint payoff from a job match, up to the point where termination is jointly optimal. Secondly, wages for new hires are flexible. As a consequence, firms capture a positive share of the increased ex ante surplus and job creation increases. More surprisingly, social efficiency also improves unambiguously despite that, as known from Hosios (1990), a priori higher job creation and lower job destruction do not necessarily improve efficiency in the presence of matching externalities.

If the equilibrium under employment at will is unique, severance payments below the minimum efficient size unambiguously reduce the separation rate at the margin under mild, model-specific conditions. In general though, multiple, Pareto-ranked, steady-state equilibria cannot be ruled out in the presence of binding real wage rigidity. Some of these equilibria feature a perverse comparative statics in which higher severance payments

---

1 Recently, Garibaldi and Violante (2005) and Fella (2007) have argued that the tax component of mandated employment protection measures is unlikely to be quantitatively important.

2 This two-tier wage structure is not only a common benchmark in the literature (see, e.g., Mortensen and Pissarides, 1999, Garibaldi and Violante, 2005) but it also usefully helps insulating issues of quantity flexibility from issues of entry wage flexibility.
increase job destruction and reduce job creation at the margin. Yet, only the well-behaved, (constrained-) efficient equilibrium with the lowest job destruction rate survives if severance payments are large enough. Therefore, large enough severance payments unambiguously reduce job destruction.

The paper’s second result—that severance payments never result in privately inefficient labor hoarding—stems from the assumption that the parties can negotiate one-off, spot side payments upon separation. To see this, consider the extreme case in which the mandated severance payment is infinite. The marginal worker’s reservation severance payment equals the difference between the present value of income if employed and if unemployed. The firm reservation severance payment is the one that gives it a separation payoff equal to the expected value of profits from the marginal worker. Whenever the joint payoff from separation exceeds the joint return from continuation, the two reservation payments define a non-empty set of transfers that leave one party at least indifferent, while making the other one strictly prefer separation to continuation. Therefore, even if it is not optimal for the firm to pay the mandated severance payment and lay the worker off, it is Pareto-optimal for the parties to agree to exchange a lower transfer and separate consensually. Therefore, severance payments do not affect the marginal separation payment and job destruction when they exceed the minimum efficient value that ensures efficient separation. Since the ex post spot payments are foreseen at the time a match is formed, job creation is also unaffected to the extent that entry wages are flexible.

There is an obvious tension among the three assumptions of wage rigidity for ongoing matches, flexible wages for new hires and the ability to negotiate ex post side payments upon separation. The role of each assumption and the rather supportive empirical evidence are discussed in depth in Section 5. Here it is worth highlighting the following two points.

First, one may argue that there is an unnatural asymmetry between the assumption that the parties can negotiate around inefficiently high severance pay, but not around wage rigidities. It would indeed be in the marginal job loser’s interest to rebate a fraction of the wage to the firm as long as separation is involuntary and inefficient. A crucial difference is that, while renegotiation of separation payments involves a one-off transfer, payments aimed at undoing wage rigidity would have to be made in each period in which productivity is low. In fact, if wage rigidity is due to a binding efficiency-wage constraint of the kind considered in Section 4.2, any expectation of a partial rebate of the future efficiency wage would reduce the worker’s expected surplus below its incentive-compatible level. A second difference is that voluntary redundancy agreements are legal and often explicitly contracted ex post, as attested by the evidence discussed in Section 5. Conversely, to the extent that wage rigidity is stemming from statutory wage minima, violations of such minima are illegal and expose employers to sanctions.

Secondly, though the assumption of flexible wages for new hires provides a useful benchmark, it is possible and even likely that entry wages are not flexible enough. Even in such a case, the paper implies that to the extent that wages in ongoing matches are downward rigid the constrained-efficient policy response is to increase the flexibility of

---

3Being sunk, a one-off transfer at unchanged wages would not alter firms’ incentives to retain. I am grateful to an anonymous referee for suggesting this point.

4Legally binding wage minima can stem not only from a national minimum wage but also from laws extending the coverage of union wages to non-unionized workers.
entry wages, or subsidize job creation, rather than to remove legislated severance payments. A reform that removed existing severance payments would constitute a windfall loss for workers in existing jobs and a windfall gain for their employers and would result in inefficiently high destruction of existing jobs. This provides an additional argument, over and above political economy considerations (see, e.g., Saint-Paul, 2002), for a reform to apply only to newly created jobs.

The paper is obviously related to the broad literature applying Lazear’s result to study the effect of firing costs in flexible wage models and to a number of papers that study the equilibrium effect of severance payments when wages are downward rigid.

Within the second class of models, Saint-Paul (1995), Fella (2000) and Alvarez and Veracierto (2001) have shown that severance payments reduce job destruction and, over a range, increase job creation and efficiency. Severance payments that are too large, however, reduce job creation and efficiency in the environments studied in those papers. Severance payments reduce job creation and efficiency in the environments studied in those papers. Severance payments reduce job creation and efficiency in the environments studied in those papers. Severance payments reduce job creation and efficiency in the environments studied in those papers.

Severance payments reduce job creation and efficiency in Galdón-Sánchez and Güell’s (2003) shirking model, as long as there is a strictly positive probability that they are also paid in the case of a disciplinary dismissal. Garibaldi and Violante (2005) show that severance payments are neutral if wages for ongoing matches are endogenously determined by a monopoly union and entry wages are flexible. Finally, Postel-Vinay and Turon (2011) find the firing taxes increase employment in a version of Mortensen and Pissarides’s (1994) with on-the-job search and rigid exogenous wages.

The paper is structured as follow. Section 2 introduces the economic environment. Section 3 studies the equilibrium with exogenous, rigid wages in ongoing matches. Section 4 studies the equilibrium under two micro-foundations for wage rigidity: right-to-manage bargaining and efficiency wages. Section 5 discusses some of the assumptions underlying the results, while Section 6 concludes.

2 The model

2.1 Environment

The economic environment is effectively the same as in Mortensen and Pissarides (1994). The economy comprises a unit mass of infinitely-lived workers and an endogenous measure of establishments (or firms). Time is discrete and there is no aggregate uncertainty. All agents have linear preferences and discount the future at the constant rate $r > 0$. Workers are endowed with one indivisible unit of leisure whose utility is $z$. There is a unique good which is also the numeraire.

Each establishment requires one worker in order to produce, and can either have a vacancy and be looking for a worker or be matched to a worker. Firms with vacant positions and unemployed workers are brought together by a random matching process according to a constant returns to scale, strictly increasing and concave matching technology mapping the stocks of unemployed workers $u$ and of vacancies $v$ into a flow $m(u, v)$ of new matches. With constant returns, matching probabilities depend only on market tightness.

---


6 I discuss the sources of the difference between these results and those in this paper in the main text.
\( \theta = v/u \). They are denoted by \( q(\theta) = m(u,v)/v \), for vacant firms, and \( p(\theta) = m(u,v)/u \), for unemployed workers. Keeping an open vacancy entails a flow cost \( c > 0 \) and there is free entry in vacancy posting.

Firm-worker matches (or jobs) are ex ante identical, but subject to idiosyncratic productivity shocks. Following standard practice, it is assumed, without loss of generality, that all new matches begin at the top of the productivity support, normalized to 1, and that there are positive gains from forming a new match.\(^7\) Shocks hit all jobs with probability \( \lambda \). Following a shock, the match productivity \( y \) takes a new value, i.i.d. across time and drawn from a continuous cumulative density function \( G(y) \), density \( g(y) \) and support \([0,1]\).

When a new match is formed, the parties negotiate a first-period (entry) wage \( w_e \) and a state-independent (continuation) wage \( w_c \) that applies throughout the remaining duration of the match.\(^8\) The pair \((w_e, w_c)\) is assumed to maximize the Nash bargaining solution with weight \( \beta \in (0,1) \) on the worker’s gain. The continuation wage \( w_c \) has to satisfy a lower bound \( w_l \), which captures, possibly binding, wage rigidities. The binding lower bound \( w_l \) is assumed to be exogenous in Section 3 and is endogenized in Section 4. All flows are measured in units of end-of-period output.

Employment protection legislation (EPL) mandates that firms have to transfer a severance payment \( F \) to workers if a separation is labeled a layoff but not if it is a quit or a voluntary redundancy. Namely, separation payments are contingent on which party takes verifiable steps to end the relationship. A separation is deemed a layoff if the firm gives written notice that the worker’s services are no longer required. On the other hand, a separation is deemed a quit, and no payment is due, if the worker gives written notice that he or she no longer intends to continue in employment (or simply stops showing up for work without obtaining leave). Any claim by one party that the other has unilaterally severed the relationship must be supported by documentation. This is consistent with existing practices in most industrialized countries.

I allow for the possibility that at the time of separation the parties “renegotiate” the mandated layoff payment \( F \), and agree to separate with a lower transfer, if doing so yields a Pareto improvement.\(^9\) Indeed, the parties can implement this outcome in two equivalent ways. They can agree to label the separation a quit (or a voluntary redundancy) rather than a layoff, in which case transfers between the parties are unconstrained by legislation. Alternatively, they can label the separation a layoff, with the worker rebating to the firm, on the spot, the difference between the legislated and the Pareto optimal severance transfer.

### 2.2 Analysis

The timing is standard. At the beginning of each period, the stochastic state is revealed and the relevant decisions are taken after observing it. As noted above, to simplify notation we assume that all output/income flows accrue at the end of the period, while value functions are expressed in units of output at the beginning of the period.

\(^7\)This last assumption ensures the existence of a non-trivial equilibrium with positive employment.  
\(^8\)The bargaining mechanism is similar—in fact isomorphic—to that in Rocheteau (2002).  
\(^9\)The assumption satisfies the natural, joint-rationality constraint that the firm and worker do not leave money on the table. Section 5 discusses empirical evidence consistent with the assumption.
2.2.1 Bellman equations

The joint return $Z(y,w)$ from production in a match with current productivity $y$ and wage $w$ satisfies

$$(1 + r)Z(y,w) = y + \lambda \left[ \int_R Z(y',w') \, dG + G(R)(U + V) \right] + (1 - \lambda)Z(y,w').$$

(1)

When a shock hits the match, the latter survives if the new productivity realization $y'$ is above some reservation value $R$ and is destroyed otherwise. In the latter case, the joint payoff from separation is the sum of the worker's expected utility from unemployment $U$ and the value of a new vacancy $V$. Note that equation (1) implies that $Z(y,w)$ depends on $w$ to the extent that the latter affects the next period's wage $w'$ and through it the reservation productivity $R$.

The lifetime utility of a worker in a match with wage $w$ satisfies

$$(1 + r)W(w) = w + \lambda \left[ [1 - G(R)]W(w') + \int_0^R [U + Q(y',w')] \, dG \right] + (1 - \lambda)W(w').$$

(2)

The current wage $w$ equals $w_e$ for new hires and $w_c$ for all workers in ongoing matches. Next period, workers receive the continuation wage $w' = w_c$, unless the match is hit by a shock that results in separation. In the latter case, the worker's payoff equals the return to search $U$ plus the equilibrium separation transfer $Q(y',w')$, conditional on the productivity draw $y'$ and the ruling wage $w'$. The equilibrium transfer $Q(y',w')$ differs from the mandated payment $F$ whenever the latter is renegotiated. Note that $Q(y',w')$, and therefore $W(w)$, are functions of the mandated payment $F$. Yet, since $F$ is a time-invariant, common policy parameter it does not enter the state vector.

Since $w' = w_c$ for all matches and $w = w'$ for an ongoing match, it follows from equation (2) that the value of employment in an ongoing match satisfies

$$rW(w_e) = w_e + \lambda \int_0^R [U + Q(y',w_e) - W(w_e)] \, dG.$$

(3)

and that

$$W(w_e) = W(w_c) + \frac{w_e - w_c}{1 + r}.$$  

(4)

The corresponding value of an ongoing job with productivity $y$ is

$$J(y,w_c) = Z(y,w_c) - W(w_c).$$

(5)

Finally, the value of search for an unemployed job seeker satisfies

$$rU = z + p(\theta) \left( W(w_e) + \frac{w_e - w_c}{1 + r} - U \right),$$

(6)

while the corresponding value for a vacancy is

$$rV = -c + q(\theta) \left( Z(1,w_c) - W(w_c) - \frac{w_e - w_c}{1 + r} - V \right).$$

(7)

\textsuperscript{10}Equation (1) assumes the reservation property in order to simplify notation. Since $Z(y,w)$ is increasing in $y$, it will become clear later in the analysis that the assumption is indeed satisfied.
Before continuing, a few remarks on the above modeling framework are appropriate. First, as in Mortensen and Pissarides (1999) and Garibaldi and Violante (2005), the two-tier wage structure captures the fact that severance payments are due to terminate a match, but not for failing to form a match upon contacting a worker. Secondly, the two-tier wage structure allows to evaluate the effect of severance payments, usually associated with quantity flexibility, in isolation from the conceptually different issue of the flexibility of (the present value of) wages for new hires. Finally, since the model features no aggregate uncertainty, the assumption that the continuation wage \( w_c \) does not depend on the productivity realization buys tractability with no loss of generality.\(^{11}\) It is well known\(^{12}\) that, if labor services are not traded in spot auction markets, the mapping between productivity realizations and wages is allocationally irrelevant and, given risk-neutrality, the same is true for welfare. In the present set-up, the only allocationally-relevant payoffs are the payoff to the marginal worker and the ex ante payoff to a new hire. The combination of the entry and continuation wages allows decoupling of the two payoffs in the same way as in the, standard, instantaneous-bargaining set up.

### 2.2.2 Job destruction

Given the payoff from employment continuation \( W(w_c) \), a worker is indifferent between continuing employment and separating with a transfer equal to \( W(w_c) - U \). Similarly, if the match productivity is \( y \), the firm is indifferent between continuing the match and separating with a transfer equal to \( V - J(y, w_c) \). Since the firm has also the option to unilaterally lay the worker off by paying her the mandated severance payment \( F \), the firm reservation severance payment is equal to \( \min\{V - J(y, w_c), F\} \).

It follows that separation with a transfer between the two reservation severance payments is a Pareto improvement relative to continuing the match; or more formally

**Definition 1** For given \( y \), \( W(w_c) \) and \( R \), the interval

\[
\Omega(y, w_c) = [W(w_c) - U, \min\{V - J(y, w_c), F\}]
\]

is the set of Pareto-improving and individually-rational separation transfers.

Therefore, if the parties bargain efficiently over separation payments, they will separate in all states in which \( \Omega(y, w_c) \) is non-empty. If \( \Omega(y, w_c) \) is non-empty and does not contain \( F \), they will agree to separate with a transfer strictly smaller than \( F \).

If

\[
W^r(w_c) = \frac{w_c + \lambda G(R)(U + F)}{r + \lambda G(R)}
\]

denotes the value of \( W(w_c) \), when the severance payment is never renegotiated—\( Q(y, w_c) = F \) for all \( y \leq R \)—the following proposition holds.

\(^{11}\)In fact, Section 3.1 shows that if the lower bound \( \bar{w} \) on the continuation wage \( w_c \) is not binding the equilibrium allocation coincides with the one in Mortensen and Pissarides (1994), that features period-by-period wage bargaining. In the case in which \( \bar{w} \) is binding, the equilibrium is isomorphic to one in which the wage is bargained in every period subject to a lower bound that implies a positive rent for the marginal worker.

\(^{12}\)See, e.g., Barro (1977) and Pissarides (2009).
Proposition 1 For given \( w_c \):

1. If \( w_c \geq r(U + F) \), the equilibrium separation transfer \( Q(y, w_c) \) equals \( F \) for all \( y \leq R \). The reservation productivity satisfies \( Z(R, w) = W^r(w_c) + V - F \geq U + V \) with strict inequality if and only if \( w_c > r(U + F) \).

2. If \( w_c < r(U + F) \) and bargaining over separation is efficient, the equilibrium separation transfer \( Q(y, w_c) \) is strictly smaller than \( F \) for some \( y \leq R \). The reservation productivity satisfies \( Z(R, w_c) = U + V \).

Proof. See Appendix. \( \blacksquare \)

Note that \( w_c \geq r(U + F) \) is equivalent to \( W(w_c) \geq r(U + F) \). Therefore, if \( w_c \geq r(U + F) \), the worker’s reservation severance payment \( W(w_c) - U \) is (weakly) larger than the mandated payment \( F \) and \( \Omega(y, w_c) \) is either empty—if \( w_c > r(U + F) \)—or contains the single point \( F \). The mandated severance payment is never renegotiated—\( W(w_c) = W^r(w_c) \)—and the firm lays the worker off whenever the payoff from continuation \( J(y, w_c) = Z(y, w_c) - W(w_c) \) falls below the return from firing \( V - F \). If \( w_c > r(U + F) \) the joint payoff from employing the marginal worker \( Z(R, w_c) \) exceeds the joint payoff from separation \( U + V \) and job loss is involuntary. The worker would strictly prefer to accept a wage cut to being laid off. Conversely, if \( w_c = r(U + F) \), the separation rule maximizes joint wealth—\( Z(R, w_c) = U + V \).

Instead, \( w_c < r(U + F) \) implies \( W(w_c) < U + F \) and \( Z(y, w_c) - W(w_c) > V - F \) when the productivity \( y \) is marginally below the joint-wealth-maximizing reservation value \( Z(R, w_c) = U + V \). Yet, privately inefficient labor hoarding does not follow despite the fact that it is not optimal for the firm to unilaterally fire the worker. Since wage rigidity does not constrain spot wealth transfers upon separation and \( \Omega(y, w_c) \) is non-empty when \( Z(y, w_c) < U + V \), it is jointly optimal for the parties to agree to separate with a separation payment below the mandated one\(^{13} \) whenever \( W(w_c) + V - F < Z(y, w_c) < U + V \).

It follows that, for given \( U \), the separation rate is either higher than—if \( w_c > r(U + F) \)—or the same as under flexible wages. Mandated severance payments can never induce privately inefficient labor hoarding and the reservation productivity satisfies

\[
Z(R, w_c) = U + V + \max\{W^r(w_c) - F - U, 0\}. \tag{10}
\]

This is the first main result in the paper.\(^{14} \) It contrasts with results in Bentolila and Bertola (1990), Fella (2000), Alvarez and Veracierto (2001) and Garibaldi and Violante (2005) that imply that in the presence of wage rigidities large enough severance payments induce privately (and socially) inefficient labor hoarding up to the point where all separations other than bankruptcies are prevented. The difference is that those papers assume that a firm-worker pair always exchange the legislated severance payment upon separation. This implies that the parties leave money on the table, despite the fact that, whenever \( Z(y, w_c) < U + V \), they can be both better off by agreeing to separate with a lower consensual severance payment in the non-empty interval \( \Omega(y, w_c) \).

\(^{13}\)Appendix A.2 derives the value of the equilibrium transfer \( Q(y, w_c) \) under one particular (efficient) bargaining solution.

\(^{14}\)An early, static, version of this result was first circulated as Proposition 4 in Fella (1999), the working paper on some of whose early ideas this paper is based.
2.2.3 Bargaining with a new hire

The pair \((w_e, w_c)\) solves the standard Nash bargaining problem

\[
\max_{w_e, w_c} \left( W(w_c) + \frac{w_e - w_c}{1 + r} - U \right)^\beta \left( Z(1, w_c) - W(w_e) - \frac{w_e - w_c}{1 + r} - V \right)^{1-\beta},
\]

s.t. \(w_c \geq w\),

\[(11)\]

where it is assumed, without loss of generality, that \(w \geq rU\), so that it is not necessary to impose a separate participation constraint for the worker.

The first-order necessary and sufficient conditions for an optimal choice of \(w_e\) and \(w_c\) can be written as

\[
\frac{W(w_c) + \frac{w_e - w_c}{1 + r} - U}{Z(1, w_c) - W(w_e) - \frac{w_e - w_c}{1 + r} - V} = \frac{\beta [Z(1, w_c) - U - V]}{1 - \beta},
\]

\[
\frac{\partial Z(1, w_c)}{\partial w_c} = \frac{\lambda g(R)}{r + \lambda} \left[ U + V - Z(R, w_c) \right] \frac{\partial R}{\partial w_c} \leq 0, \quad w_c \geq w
\]

\[(13)\]

\[(14)\]

where the two inequalities in (14) hold with complementary slackness.

Equation (13) determines the present value of wages and implies that the entry wage \(w_e\) ensures that the firm and worker’s surpluses from a new match are proportional to each other and to the joint match surplus \(Z(1, w_c) - U - V\). The parties’ respective shares equal their bargaining weights.

One can use equation (13) to solve for the entry wage

\[
w_e = w_c + (1 + r) \{ \beta [Z(1, w_c) - U - V] - [W(w_c) - U] \}.
\]

\[(15)\]

The entry wage \(w_e\) is below the continuation wage \(w_c\) if and only if the marginal worker’s rent \(W(w_c) - U\) next period exceeds a new hire’s quasi-rent \(\beta [Z(1, w_c) - U - V]\); namely if the lower bound on the wage in ongoing matches implies a rent which exceeds the quasi-rent due to the existence of matching frictions. Given a flexible entry wage \(w_e\), mandated severance pay does not affect the ex ante division of the surplus, as workers are not entitled to it if they are not hired. To the extent that severance pay increases the worker’s future rent \(W(w_c) - U\), a lower entry wage reestablishes the ex ante surplus division.

Equation (14), instead, makes clear that the role of the continuation wage \(w_c\) is to ensure that the reservation productivity maximizes the ex post—hence the ex ante \(Z(1, w_c)\)—joint payoff subject to the wage rigidity constraint. It also implies that the ex ante joint return \(Z(1, w_c)\) is maximized if and only if the separation decision is jointly optimal; i.e., it satisfies \(Z(R, w_c) = U + V\).

If \(w > r(U + F)\), it follows from Proposition 1 and equation (10), that \(\partial R/\partial w_c > 0\)—separation is not jointly optimal—and the square bracket is negative. Therefore, the first inequality is strict for any \(w_c > w\) and the optimal wage satisfies \(w_c = w\).

Conversely, Proposition 1 implies that the first inequality holds as an equality, and there is a continuum of equilibria indexed by \(w_c\), for any \(w_c \in [w, r(U + F)]\). Yet all such equilibria are allocationally equivalent as they imply the same, jointly optimal, separation rule \(Z(R, w_c) = U + V\) and, from equation (13), the same division of the surplus from a new match that would prevail under flexible wages.
Therefore from an allocational perspective one can restrict attention, without loss of generality, to equilibria in which \( w \geq r(U + F) \); i.e., equilibria for which equation (14) reduces to

\[
w_c = w
\]

and \( Q(y, w_c) = F \) whenever separation takes place.

### 3 Equilibrium with exogenous wage rigidity

#### 3.1 Decentralized equilibrium

We can now define a stationary equilibrium for an economy in which the value of \( w \) is exogenous. To keep the exposition focused, we restrict the formal definition alone to the case in which \( w \geq r(U + F) \).

**Definition 2** Given \( w \geq r(U + F) \) and \( F \), a stationary equilibrium is a set of value functions \( \{Z(y, w_c), W(w_c), J(y, w_c), U, V\} \), payments \( \{w_e, w_c, Q(y, w_c)\} \), a market tightness \( \theta \), a reservation productivity \( R \) and an unemployment rate \( u \) such that: 1) free entry of vacancies implies \( V = 0 \); 2) the value functions \( \{Z(y, w_c), W(w_c), J(w_c), U\} \) satisfy (1), (3), (5) and (6); 3) \( \{w_e, w_c, Q(y, w_c)\} \) are determined by (13), (16) and \( Q(y, w_c) = F \); 4) market tightness satisfies (7); 5) the reservation productivity is determined by (10); 6) the unemployment rate satisfies the flow equilibrium condition

\[
u = \frac{\lambda G(R)}{p(\theta) + \lambda G(R)}
\]

As usual in this class of models, the system of equations in the previous section can be collapsed into two, job destruction and job creation, equations that, together, determine the equilibrium values of the pair \((\theta, R)\).

One can use the ex ante surplus sharing condition (13) and free entry to replace in (7) to obtain

\[
\frac{c}{q(\theta)(1 - \beta)} = Z(1, w) - U = \frac{1 - R}{r + \lambda} + Z(R, w) - U,
\]

where \( w_c = w \) follows from equation (16) and where the last equality follows from (1). Similarly, substitution of (17) and (13) into (6) yields the standard equation

\[
rU = z + \frac{\beta}{1 - \beta} \theta
\]

for the value function of an unemployed worker.

The job destruction locus can be derived by integrating equation (1) by parts and evaluating it at the reservation productivity \( R \) to obtain

\[
Z(R, w) - U = \frac{h(R, \theta)}{r + \lambda G(R)}.
\]

\[15\]As discussed in the previous section, the restriction is without loss of generality from the allocational perspective. It just implies that the mandated severance payment is never renegotiated and allows relegating details of bargaining over separation to Appendix A.2. The equilibrium allocation, though, is characterized for the general case in what follows.
where

\[ h(R, \theta) = R + \lambda \int_{R}^{1} \frac{1 - G(y)}{r + \lambda} dy - \left( z + \frac{\beta}{1 - \beta} c\theta \right), \] (20)

and where the term in bracket is the permanent income from unemployment from (18).

Using (19) to replace for \( Z(R, w) - U \) in (17) implies the job creation condition

\[ \frac{c}{q(\theta)(1 - \beta)} = \frac{1 - R}{r + \lambda} + \frac{h(R, \theta)}{r + \lambda G(R)}. \] (21)

Finally, substituting \( Z(R, w) - U = \max\{W_r(w) - U(\theta + F); 0\} \) in (19) and using (9) yields, after some rearrangement, the job destruction condition

\[ \max\{w - rU(\theta) - rF, 0\} = h(R, \theta). \] (22)

For given \( w \), equations (19) and (21) form a system of two equations in the two unknowns \( \theta \) and \( R \).

Before studying the general case, it is useful to consider first the case in which \( w \) is so low that \( w \leq r(U(\theta + F)) \) for any value of \( \theta \). It follows from Proposition 1 that equation (19) reduces to \( h(R, \theta) = 0 \) which is the same job destruction equation as in Mortensen and Pissarides (1994) and, from (20), describes an upward sloping locus \( JD^f(\theta) \) in the \((R, \theta)\) space. This is the curve \( JD^f \) in Figure 1, where the superscript \( f \) stands for flexible wages. It is only slightly more involved to show that the job creation condition (21) implies the \( JC \) curve in Figure 1. The curve is positively sloped, and lies below \( JD^f \), for values of \( R \) such that \( Z(R, w) < U \); i.e., \( h(R, \theta) < 0 \) from equation (19). Conversely \( JC \) is negatively sloped (vertical for \( R \) high enough) and lies above \( JD^f \), for values of \( R \) such that \( Z(R, w) > U \). The flexible-wage equilibrium solution \((\hat{\theta}^f, \hat{R}^f)\) lies at the unique intersection A of the two curves in Figure 1. It is straightforward to verify that it is the same solution as in Mortensen and Pissarides (1994) as equation (21) collapses to their job creation equation when one imposes \( h(R, \theta) = 0 \).

\[^{16}\text{The advantage of writing the job creation as in equation (21) is that it is unaffected by the rigid wage \( w \) and severance pay. These two enter only the job destruction condition (19).}\]
Consider now the case in which for \( \theta \) low enough it is \( w > r(U(\theta) + F) \). The term \( rU(\theta) \) cancels out in equation (22); the reservation productivity is a function of \( w - rF \) alone and the JD curve is horizontal. Yet, since the left hand side of equation (22) is decreasing in \( \theta \), there exists a value of \( \theta \) above which it is zero and the JD curve coincides with \( JD^f \). The thick curve in Figure 1 is one possible such JD curve.

If, as in Figure 1, the horizontal section of the JD curve intersects the JC above point A the following result holds.

**Proposition 2 (Involuntary layoff)** If \( w > r(U(\hat{\theta}^f) + F) \), job creation is lower and job destruction is higher than under flexible wages. An increase in \( F \), reduces job destruction and increases job creation.

If the firm shadow cost of labor \( w - rF \) exceeds the worker’s reservation wage \( rU(\hat{\theta}^f) \) in the flexible-wage economy, wage rigidity affects the allocation of labor. For given \( w \), an increase in \( F \) reduces the firm shadow cost of labor and the reservation productivity, shifting the horizontal section of the JD curve downward.

If instead the wage constraint \( w \) is such that the intersection is below point A, the equilibrium is the same as under flexible wages. Shifting the horizontal section of the JD curve further down does not change its intersection with the JC curve.

**Remark 1 (Privately eff. separation)** If \( w \leq r(U(\hat{\theta}^f) + F) \), the equilibrium allocation is the same as under flexible wages. Any increase in \( F \) has no allocational effect.

Remark 1 is the equilibrium counterpart of Proposition 1. The following corollary then follows directly from equation (24).

**Corollary 1** If \( d(w - rF)/dF < 0 \), large enough severance payments induce the same equilibrium allocation which obtains in the absence of wage rigidities.

When \( w \) is exogenous the condition is trivially satisfied. In general, though, \( w \) is an endogenous function on \( F \) and the condition has to be verified case-by-case. Furthermore, since \( w \) may also be a function of market tightness \( \theta \), the \( JD^r \) locus may intersect the JC curve more than once. Even if the condition is satisfied, there may be an equilibrium with perverse comparative statics. Yet, \( A \) will be the unique equilibrium for \( F \) large enough.

Section 4 establishes the extent to which the condition in Corollary 1 is satisfied under two main micro-foundations for wage rigidities and the associated workers’ rents: inefficient union bargaining and efficiency wages.

### 3.2 Socially optimal job destruction

The previous section has established that, to the extent that they reduce the reservation productivity, large enough severance payments increase job creation and employment.

It is well known, though, that matching frictions imply that the decentralized equilibrium is not in general constrained-efficient, even in the absence of binding wage rigidities; i.e., when \( Z(R, w) = U \). Efficiency obtains only if Hosios’s (1990) condition that the workers’ bargaining weight \( \beta \) equals the elasticity of the probability of filling a vacancy holds. In such a second-best environment, one cannot conclude that a reduction in job
destruction and an increase in employment is associated with an increase in social welfare without solving explicitly the social planner problem. This section solves such a problem.

Given risk-neutrality, the natural social welfare metric is the present value of output, net of vacancy posting costs, discounted at rate $r$. A constrained-efficient allocation maximizes social welfare subject to the relevant constraints.

Let $(\hat{\theta}^f, \hat{R}^f)$ denote the equilibrium values of market tightness and reservation productivity in the flexible-wage equilibrium associated with point $A$ in Figure 1. One can prove the following result.

**Proposition 3** In a stationary decentralized equilibrium in which the job creation equation (21) is satisfied, a marginal policy change that reduces the equilibrium job destruction rate increases efficiency if $R > \hat{R}^f$ and reduces efficiency if $R < \hat{R}^f$.

**Proof.** See Appendix.

Proposition 3 is the second main result in the paper. It implies that, even if the social planner cannot choose market tightness—i.e., for any value of $\beta$—severance pay improves efficiency, to the extent that it reduces equilibrium job destruction relative to the laissez-faire equilibrium with involuntary layoffs. The result is particularly surprising in the case in which $\beta$ is smaller than the elasticity of the probability of filling a vacancy and the decentralized equilibrium with flexible wages features inefficiently high job creation and low job destruction (see, for example, Pissarides, 2000, p. 190). A priori one could expect that a distortion that increases job destruction—thus also reducing the ex ante joint surplus and job creation—relative to the decentralized flexible wage equilibrium, could be welfare improving in such a case. Proposition 3 rules this out. It implies that privately inefficient separation—$Z(R, w) \not= U$—always reduces the aggregate net flow of consumable resources relative to the flexible wage equilibrium, even in a second best world in which the Hosios condition is not satisfied.

Intuitively, the flexible-wage, decentralized equilibrium pair $(\hat{\theta}^f, \hat{R}^f)$ coincides with point $A$ on the $JC$ curve where market tightness $\theta$ is maximized, as a function of the reservation productivity. As the economy moves along the $JC$ curve towards point $A$, the expected utility of unemployed workers in equation (18) increases with $\theta$. It is well known\(^\text{17}\) that the expected utility of the unemployed is maximized at the socially efficient workers’ share $\beta$. The above result implies that the maximum of the unemployed expected utility and social welfare coincide also in a second-best world in which $\beta$ is suboptimal and job creation is decentralized.

### 4 Endogenous wage rigidity

This section endogenizes the lower bound on the wage $w$ using two common micro-foundations for wage rigidity: inefficient union bargaining and efficiency wages.

#### 4.1 Unionized wage setting

Fully dynamic theories of inefficient unionized wage setting are hard to come by. In a static setting, privately inefficient job destruction due to inefficient bargaining over

\(^{17}\)See, for example, Pissarides (2000) p. 187.
wages alone—right to manage—obtains only if the marginal revenue product of labor is decreasing.\textsuperscript{18} A decreasing marginal revenue product implies that under right-to-manage marginally increasing the wage above the marginal worker’s reservation wage increases an infra-marginal worker’s wage but does not affect her employment. Therefore, right-to-manage bargaining results in inefficient job destruction if the representative union member is infra-marginal.\textsuperscript{19}

In the present constant-returns setup, bargaining over state-dependent (i.e., productivity-dependent) wages would be privately efficient. One way to introduce inefficient unionized wage setting is to assume, following Garibaldi and Violante (2005), that bargaining is over a unique, state-independent wage that applies to all union members.

Bargaining takes place at the sectoral level. There is a large number (possibly a continuum) of sectors with identical productivity distribution $G(y)$. Within each sector the wage is the outcome of right-to-manage bargaining between the workers’ and employers’ unions. That is bargaining is over the (productivity-independent) wage alone and then employers choose whether to end or not each specific match.

Both unions represent only their employed members and have a utilitarian objective function. As long as $w \geq r(U + F)$, it follows from Proposition 1 that the equilibrium severance payment always equals $F$, and the workers’ union objective function is

$$\bar{W}(w) = [1 - G(R)]W(w) + G(R)(U + F) = U + F + [1 - G(R)]W(w) - U - F. \quad (23)$$

Correspondingly, the employers’ union objective function is

$$\bar{J}(w) = \int_R^1 J(y, w)dG - G(R)F = -F + \int_R^1 \frac{y - R}{r + \lambda}dG. \quad (24)$$

Since each sector is atomistic, the parties take the value of market tightness $\theta$ as given. Over the range for which $w \geq r(U + F)$, the wage solves the bargaining problem

$$\max_{W(w)} \left( \bar{W}(w) - W \right)^{\beta_u} \left( \bar{J}(w) - J \right)^{1 - \beta_u}$$

s.t. $h(R, \theta) - [r + \lambda G(R)] \max\{W(w) - F - U, 0\} = 0, \quad (26)$

where $\beta_u$ indexes the bargaining power of the (employed workers’) union and does not necessarily coincide with the bargaining power $\beta$ of new hires in the previous sections.

The parties choose $w$ to maximize the Nash product subject to the ex post job destruction condition (26). The threat points $W$ and $J$ are respectively the workers’ and the firms’ no-trade, fallback utilities. Through the appropriate choice of threat points the Nash maximand (25) can encompass most strategic and axiomatic bargaining solutions. It can also accommodate the monopoly union model as a special case when $\beta_u = 1$.

Let $R(\theta)$ denote the solution to problem (25) as a function of $\theta$ and $R^f(\theta)$ the flexible-wage, reservation productivity as a function of $\theta$. The following result holds.

\textsuperscript{18}See Booth (1995) p. 53 and references therein.

\textsuperscript{19}In this respect, as long as the representative union member is infra-marginal, it makes little difference whether she is the median—if individual preferences are aggregated by majority voting—or the average member—in the utilitarian-union case considered here.
Proposition 4  Suppose: (1) $\beta_u < 1$, and (2) either $\partial W/\partial F < 1$ or $\partial J/\partial F > -1$. For given $\theta$, a higher severance payment $F$ reduces the reservation productivity $R(\theta)$ as long as $R(\theta) > R'(\theta)$.

Proof. See Appendix. ■

The result is best understood by remembering from Section 2.2.2 that, as long as $w > r(U + F)$, it is is $W(w) > U + F$ which implies that the value of the curly bracket in equation (26) is $W(w) - F - U$. Replacing for $W(w)$ using equation (23) and the job destruction constraint (26) yields

$$\max_{R} N = \left( U + F + \frac{[1 - G(R)]h(R,\theta)}{r + \lambda G(R)} - W \right)^{\beta_u} \left( -F + \int_{R}^{1} \frac{y - R}{r + \lambda} dG - J \right)^{1-\beta_u} \quad (27)$$

It follows from equation (27) that a higher $F$ reduces the the workers’ payoff and increases the firms’ at given $R$. As long as $\partial W/\partial F < 1$ or $\partial J/\partial F > -1$, this redistribution results in a higher workers’ and lower firms’ surplus at given $R$, relative to the respective threat points. The Nash solution calls for the wage to partly offset this redistribution so that the shadow cost of labor and $R$ fall at an internal maximum. Therefore, $R$ falls with $F$.

It is instructive to compare Proposition 4 with the following result.

Proposition 5  For given $\theta$ the reservation productivity $R(\theta)$ is independent of the severance payment $F$ if either

1. $\beta_u = 1$;
2. $W = U + F$ and $J = -F$.

Case 1 corresponds to the monopoly-union model analyzed in Garibaldi and Violante (2005). Since the optimal $R$ maximizes the first bracket in (27), $F$ does not enter the first order condition. In case 2 the severance payment drops out of the Nash maximand in equation (27). In both cases, severance payments raise the worker’s payoff one-for-one and are neutral even if separation is inefficient.

The two cases covered by Proposition 5, while appealing for their analytical convenience, are restrictive. In general, unions do not set wages unilaterally but instead bargain over them with firms or firms’ representatives, as argued for example in Booth (1995). Furthermore, while the threat points in case 2 are standard in the matching literature, they can be justified on the basis of convenience only if they do not constrain the equilibrium outcome in a qualitatively important way. This is not the case here.

As first pointed out by Binmore, Rubinstein, and Wolinsky (1986), whenever there is ambiguity about the appropriate threat points in bargaining one should get guidance by modeling the bargaining process strategically. They show that, when this is done, threat points are usually different from outside options. Outside options are the payoffs that the parties obtain by irreversibly breaking the match to trade outside. Since an agent’s threat to abandon the match is not credible unless her outside payoff exceeds what she would receive under bilateral monopoly, outside options only act as a constraint for the
bargaining outcome, and not as a threat point. The appropriate threat points are instead the parties’ expected payoffs in case of perpetual disagreement.\textsuperscript{20} Depending on whether the parties search or not while bargaining these are either the expected returns to search as in Wolinsky (1987) or the present value of income flows while no trade takes place as in Coles and Hildreth (2000) and Hall and Milgrom (2008). In the latter case, they are obviously independent of severance payments. If instead the parties searched while bargaining, because of search frictions they would find a new partner with probability strictly less than one in finite time. With positive discounting, even if the firm were obliged to pay the severance payment regardless of which party abandoned bargaining for a new match, $F$ would increase the workers’ and decrease the firms’ threat point less than one-for-one.

Finally, the following proposition partially characterizes the reservation productivity $R(\theta)$ as a function of $\theta$.

**Proposition 6** The reservation productivity satisfies either $R(\theta) = R^f(\theta)$ for any $\theta \geq 0$ or there exists a unique $\theta_1 > 0$ such that $R(\theta) > R^f(\theta)$ for $\theta < \theta_1$ and $R(\theta) = R^f(\theta)$ otherwise.

**Proof.** See Appendix. \qed

The proposition states that if the rigid-wage JD curve lies above its flexible-wage counterpart for low enough $\theta$, it eventually intersects the flexible-wage $JD^f$ curve and coincides with it to the right of such intersection. For given $\theta$ the reservation productivity is inefficiently high to the left of the intersection. Figure 2 draws the $JC$ and $JD$ curves in the case in which the latter lies above its flexible-wage counterpart $JD^f$ at the flexible-price equilibrium point A. The job destruction locus $JD$ is given by the thicker curve and the portion of the $JD^f$ locus to the right of point Z.

Since one cannot rule out that the job creation and job destruction curves both slope down over the relevant range, multiple equilibria are a possibility. The intuition for this is apparent from the Nash maximand (27). A given value of $W$ can be achieved with either a high $U$ (high $\theta$) and low rents (low $R$), or low $U$ and high rents. If the loss of joint surplus associated with workers’ rents and the separation rate is inefficiently high and large enough, the optimal fall in the rent in response to a relatively small increase in $U$ may be large and the $JD$ locus be sufficiently steep.

It follows from Proposition 5 that severance payments shift down the job destruction curve over the range where it lies above its flexible wage counterpart. If the equilibrium is unique, the $JD$ curve cuts the $JC$ locus from below. Provided wage rigidity is binding at equilibrium—$\hat{R} > \hat{R}^f$—job destruction falls and job creation and efficiency increase. If there are multiple equilibria, at least one of these (e.g., point C in Figure 2) behaves perversely. Severance payments increase job destruction, thus reducing job creation and efficiency. However, Proposition 6 implies that for large enough $F$, the $JD$ curve in Figure 2 lies everywhere below $JC$ to the right of point A. Therefore, the unique equilibrium coincides with the constrained-efficient equilibrium under flexible wages corresponding to point A.

\textsuperscript{20}Furthermore, since workers are not entitled to severance payments if they unilaterally abandon the match, even in an axiomatic bargaining framework it is unclear why the severance payment $F$ should enter their threat point.
4.2 Efficiency wages

Suppose that a moral hazard problem requires firms to pay efficiency wages along the lines of Galdón-Sánchez and Güell’s (2003) version of Shapiro and Stiglitz (1984). Workers enjoy zero utility from leisure when working, but enjoy a flow utility $z$ if they shirk on the job. A shirking worker is detected with probability $q$ in each period.

Courts can only imperfectly distinguish between disciplinary and economic dismissals and no severance payment is due when a dismissal is deemed disciplinary. Let $F$ denote the expected severance payment conditionally on the actual cause of separation being economic and the worker not having shirked.\footnote{Namely, $F$ is the product of the mandated severance payment for economic redundancy times the probability that courts correctly identify an economic redundancy. Denoting the expected payment by $F$ allows exploiting the equations derived in the previous sections.} Let $\pi \leq 1$ denote the ratio between the probability that courts wrongly deem a disciplinary dismissal to be economic and award severance pay, and the probability that courts correctly award severance pay for an economic redundancy if the worker does not shirk. It follows that the expected severance payments conditional on being caught shirking and dismissed is $F^d = \pi F$.

As shown in Appendix A.3 incentive compatibility requires a worker’s expected continuation rent to satisfy

$$q\{[1 - \lambda G(R)][W(w^*) - U - F^d] + \lambda G(R)(F - F^d)\} = z,$$

for the expected cost of being caught shirking to offset the benefit from shirking $z$. The expected cost equals the probability $q$ of being caught and dismissed times the expected rent the worker would have enjoyed had she not shirked. The latter is the weighted average of the rent $W(w^*) - U - F^d$ in case the match would have continued at wage $w^*$ and the separation rent $F - F^d$ stemming from the difference between the expected severance payment in case of economic redundancy and of disciplinary dismissal.
Replacing for $F^d$ and rearranging, equation (28) can be rewritten as

$$W(w') - U = \pi F + \frac{z/q + \lambda G(R)(\pi - 1)F}{1 - \lambda G(R)},$$

(29)

which implies that, to the extent that $\pi > \lambda G(R)$, severance pay increase the workers’ rent by reducing her expected loss from shirking.

Importantly, equations (28) and (29) also make clear that incentives come from expected future, not current, gains. This aspect is often lost in continuous-time versions of the shirking model. Incentive compatibility determines the future wage $w'$, not the current one. It follows that efficiency wage constraints do not pin down the entry wage $w_e$ and the division of the surplus from a new match.\footnote{On the other hand, the solution to the efficiency wage constraint (29) pins down the lowest future wage consistent with incentive compatibility in the current period.} Consistently with the rest of the analysis, the Nash bargaining surplus-sharing condition (13) is still necessary to close the model and the job creation condition (21) still applies.

Turning to job destruction, using equation (29) in the job destruction condition (10) yields

$$Z(R) = U + \max \left\{ \frac{z/q + (\pi - 1)F}{1 - \lambda G(R)}, 0 \right\},$$

(30)

or, replacing in (19),

$$\max \left\{ \frac{z/q + (\pi - 1)F}{1 - \lambda G(R)}, 0 \right\} = \frac{h(R, \theta)}{r + \lambda G(R)}.$$  

(31)

It follows that as long as $\pi < 1$, severance payments reduce the wedge between the firm shadow cost of labor $W - F$ and the worker’s return from unemployment $U$ on the left hand side of equation (31). The condition $\pi < 1$ holds if the probability that courts wrongly award severance payments for a disciplinary dismissal is less than the probability that they correctly award the payment for an economic redundancy. The condition just requires courts to perform better than a coin toss and is most likely to be satisfied.

If $\pi < 1$, the effect of severance pay is basically the same as under inefficient union bargaining.

**Proposition 7** If $\pi < 1$:

1. there exists an equilibrium in which severance payments reduce equilibrium job destruction and increase job creation and efficiency, as long as $F < z/[q(1 - \pi)]$;

2. large enough severance payments induce a unique equilibrium with the same allocation as under flexible wages.

**Proof.** See Appendix. ■

Point 2 follows directly from equation (31). Higher severance pay shifts (and stretches) the rigid wage job destruction curve $JD$ in Figure 3 to the right until is coincides with $JD'$. 

\footnote{22On the other hand, the solution to the efficiency wage constraint (29) pins down the lowest future wage consistent with incentive compatibility in the current period.}
Figure 3: Equilibrium market tightness and reservation productivity - efficiency wages

Point 1 states that, if $F < z/[q(1 - \pi)]$, there always exists one equilibrium, such as point B, in which severance payments efficiently reduce job destruction locally as well. Yet, one cannot exclude that $JD$ has the shape illustrated in Figure 3. The intuition is the following; a higher $R$ increases the expected flow payoff $h(R, \theta)$ from the marginal job but also reduces the expected job duration $\lambda G(R)$. If the second effect prevails, the present value of the joint payoff on the right hand side of equation (31) is decreasing in $R$ and a fall in the firm shadow cost of labor increases, rather than reduces $R$, at given market tightness. This is the case in all equilibria, such as point C in Figure 3, in which the $JD$ curve is negatively-sloped and flatter that the $JC$ curve.\(^{23}\)

Proposition 7 is at odds with Galdón-Sánchez and Güell’s (2003) finding that severance pay reduces job creation and efficiency. This is due to two main differences between the two models. First, the separation rate is exogenous, hence unaffected by severance pay, in their model. Second, here an increase in severance pay would be neutral, not welfare reducing, if the job destruction rate were unaffected. Intuitively, the worker would prepay for any increase in the continuation rent through a lower entry wage, according to equation (15). Conversely, Galdón-Sánchez and Güell (2003) assume continuous time with the implication that the entry and continuation wages collapse into one and workers cannot prepay without introducing unrealistic sign up fees. Note, however, that even if one were to impose $w_e = w_c$ in the present environment, Proposition 7 would still obtain for $\pi$ small enough. Saint-Paul (1995) and Fella (2000) obtain that severance payments reduce job destruction and increase job creation and efficiency in the case in which $w_e = w_c$ and $\pi = 0$; i.e., severance payments are paid only for economic dismissals. By continuity there exists some value $\hat{\pi}$ strictly between zero and one such that severance payments are welfare-improving even in the absence of bonding.

\(^{23}\)The perverse comparative statics applies not only to severance payments but also to the utility of leisure $z$ and the detection probability $q$.\)
5 Discussion

The findings in this paper hinge on three main assumptions that deserve to be discussed here.

The ability to negotiate around privately inefficient employment protection measures is crucial for the result that mandated severance pay never induce privately inefficient labor hoarding. There is considerable evidence that negotiation of Pareto-optimal transfers upon separation is more than a theoretical construct. One example is the frequency with which one learns of voluntary redundancy packages or early retirement incentives offered by downsizing firms. By revealed preference, these must be jointly optimal if workers accept them. Furthermore, if contracting firms make the effort to negotiate such packages, the associated cost must be smaller than the (possibly shadow) cost of unilaterally laying off workers. Also, in Germany firms cannot legally carry out mass redundancies (i.e., the mandated layoff cost is infinite) unless they agree with workers' representatives on a social plan covering layoff procedures and compensation packages. For Italy, a country usually associated with extreme levels of employment protection, IDS (2000) reports that employers often negotiate incentive payments to induce employees to take voluntary redundancy and sign agreements waiving their right to take legal action. Finally, Toharia and Ojeda (1999) document that it is common for Spanish firms to agree with workers to label economic dismissals as disciplinary ones to economize on advance notice and procedural costs. Over the 1987-97 period, between 60 and 70 percent of all layoffs took this form and involved bargaining over the size of termination payments.

Efficient renegotiation also requires efficient bargaining. In the standard symmetric-information, risk-neutral framework, Merlo and Wilson (1995) have shown that the equilibrium of bilateral bargaining games is efficient under quite general circumstance.

Independently from the possibility of negotiating side-payments upon separation, some form of wage rigidity is necessary for mandated severance payments to be non-neutral, as shown by Lazear (1990) and Burda (1992). From an empirical perspective, a number of institutional features, such as minimum wages and a significant union presence and coverage of union contracts, constrain wage bargaining at the individual level in Continental European countries. This paper emphasizes wage rigidity at the firing margin by assuming: 1) binding downward wage rigidity for ongoing matches; but 2) flexible wages for new hires.

Concerning the first assumption, one may argue that there is an unnatural asymmetry between the assumption that the parties can negotiate around inefficiently high severance pay, but not around wage rigidities. It would indeed be in workers' interests to rebate a fraction of their wage to firms as long as separation is involuntary and inefficient. A

\[24\] For example, Allied Irish Banks has recently negotiated a voluntary redundancy deal providing for a one-off payment of EUR17,000, the payment of a service-related lump sum of up to one year's salary and a EUR10,000 grant for children in full-time education to employees above the age of 50. (“Redundancy Scheme Set to cost AIB Euro26m,” The Irish Times, 27 March 2007)

\[25\] The same source reports a total cost for individual redundancy of 10-12 months of wages for a worker paid around 2 million ITL a month.

\[26\] An exception are repeated-game models, like Haller and Holden (1990) and Fernandez and Glazer (1991), in which an agent ability to “burn money” implies the existence of inefficient equilibria with delay. Some authors, e.g., MacLeod and Malcolmson (1995), argue that such equilibria are not very reasonable as both parties would be better off by playing one of the efficient equilibria.
crucial difference is that, while renegotiation of separation payments involve a one-off transfer, payments aimed at undoing a binding wage rigidity would have to be made in each period in which productivity is low. In fact, if wage rigidity is due to a binding efficiency-wage constraint of the kind explored in Section 4.2, any expectation of a partial rebate of the future efficiency wage would reduce a worker’s expected future surplus below its incentive-compatible level (see footnote 22). A second difference is that voluntary redundancy agreements are legal and often explicitly contracted ex post, as attested by the evidence discussed above. Conversely, to the extent that wage rigidity is stemming from statutory wage minima, violations of such minima are illegal and expose employers to sanctions. More importantly, from an empirical perspective, several recent studies have found evidence of considerable downward real wage rigidity for job stayers in a number of OECD countries (see Barwell and Schweitzer, 2007, Bauer, Bonin, and Sunde, 2007, Christofides and Nearchou, 2007, Devicienti, Maida, and Sestito, 2007, Dickens et al., 2007, Fehr and Goette, 2005). Dickens et al. (2007) have also found that, across countries, real wage rigidity is positively and significantly correlated with union density. Of course, the fact that real wages for job stayers are downward-rigid does not imply that the same is true for the marginal job stayer and that, therefore, layoffs are privately inefficient. Yet, the evidence in Barwell and Schweitzer (2007), Bauer et al. (2007), Devicienti et al. (2007) and Fehr and Goette (2005) lends support to the view that downward real wage rigidity is associated with excess labor turnover or higher unemployment rates, at least in the countries they consider; i.e., the UK, Germany, Italy and Switzerland.

Finally, concerning the assumptions that wages in new jobs are flexible, a number of studies (see Carneiro, Guimaraes, and Portugal, 2012, Holden and Wulfsberg, 2009, Martins, Solon, and Thomas, 2012) have found evidence of substantial flexibility of hiring wages. Although the assumption that entry wages are fully flexible may be extreme, it is nonetheless useful to isolate the effect of severance payments on quantity flexibility from that of entry wage inflexibility. The point that constraints on entry wages may imply that pure severance payments reduce job creation has already been made by Bertola (1990) and more recently Garibaldi and Violante (2005). Yet, if minimum wage constraints or union coverage of new hires imply downward-rigid entry wages the same is likely be true, a fortiori, for wages in marginal ongoing matches. We have shown that if this is the case, severance payments increase ex post efficiency. The appropriate policy response to reestablish ex ante efficiency is either to relax legislative constraints on the flexibility of entry wages, or to subsidize job creation. While this paper does not put forth any role for mandated severance pay, as with flexible entry wages it is ex ante optimal to incorporate severance payments in private contracts, a reform that removed existing severance payments would constitute a windfall loss for workers in existing jobs, and a windfall gain for their employers, and would result in inefficiently excessive destruction of existing jobs.

27For Portugal, Carneiro et al. (2012) find that also wages for existing workers are significantly flexible, though less than hiring wages.
6 Conclusion

Firms and workers have an incentive to negotiate around privately inefficient employment protection legislation. This paper characterizes the implications of mandated severance pay in the presence of matching frictions and wage rigidities when renegotiation, by means of spot side payments upon termination, is allowed for.

Severance payments have real effects when wage rigidity implies privately inefficient separation under employment at will. If this is the case, severance payments increase job creation and efficiency as long as they reduce job destruction. While their marginal impact on job destruction is ambiguous if multiple equilibria are possible, large enough severance payments always reduce job destruction relative to employment at will. Their maximum relevant size is bounded and their marginal effect is zero when their size exceeds the one that induces the same allocation that prevails under flexible wages. In the latter case, Pareto optimal renegotiation by means of spot payments upon separation implies that any increase in severance pay is neutral.

A Appendix

A.1 Proofs

Proof of Proposition 1. Since the firm can fire the worker at cost $F$, it is $Q(y, w_c) \leq F$.

1. Suppose $w \geq r(U + F)$. Assume by contradiction that the mandated payment $F$ is renegotiated, that is $Q(y, w_c) < F$, for some $y < R$. Since renegotiation is consensual, it has to be $W(w_c) - U \leq Q(y, w_c)$, which implies that the square bracket in (2) is always non-negative and $W(w_c) \geq U + F$. It follows that $\Omega(y, w_c)$ is either empty or contains only the point $F$, a contradiction. Therefore, all separations are layoffs—$Q(y, w_c) = F$ for all $y \leq R$—and $W(w_c) = W^r(w_c)$. It ensues that $R$ satisfies $Z(R) = W^r(w_c) + V - F \geq U + V$, where the last inequality holds as an equality iff $w = r(U + F)$.

2. Suppose $w < r(U + F)$. Since $Q(y, w_c) \leq F$, it follows from (2) that $W(w_c) < U + F$. For all $y > R$, with $Z(R) = U + V$, the match continues, as $J(y, w_c) = Z(y) - W(w_c) > V - F$ and $\Omega(y, w_c)$ is empty. For $y \leq R$, $\Omega(y, w_c)$ is not empty and ending the match with a payment $Q(y, w_c) \in \Omega(y, w_c)$ is jointly optimal. The mandated severance payment is renegotiated downward for all $y \leq R$ and such that $J(y, w_c) > V - F$. ■

Proof of Proposition 3. Suppose the social planner can optimally choose the reservation productivity $R$ but is otherwise constrained by the evolution of unemployment described by

$$u' = (1 - p(\theta))u + \lambda G(R)(1 - u),$$

and the decentralized job creation condition (21). The social maximand is the present value of the aggregate net flow of consumable resources $Y + uz - c\theta u$, where aggregate output $Y$ satisfies the difference equation

$$Y' = (1 - \lambda)Y + p(\theta) u + \lambda (1 - u) \int_{R}^{1} ydG.$$
Denoting by $\mu$ the Lagrange multiplier associated with (21) the social planner problem can be written in recursive form as

$$Z(Y, u) = \max_{\theta, R} \frac{Y + uz - c\theta u + \mu \left[ \frac{c}{q(\theta)(1-\beta)} - \frac{1-R}{r + \lambda} - \frac{h(R, \theta)}{r + \lambda G(R)} \right]}{1 + r} + Z(Y', u')$$  \hspace{1cm} (A.3)

s.t. (A.1), (A.2)

where $Z(Y, u)$ is the value function and $h(R, \theta)$ is given by (20) in the main text. The first order conditions for

$$-cu - \frac{\mu}{1-\beta} \left[ \frac{q'(\theta)c}{q(\theta)^2} - \frac{\beta c}{r + \lambda G(R)} \right] + (Z_{Y'} - Z_u')p'(\theta)u = 0$$  \hspace{1cm} (A.4)

and for $R$

$$\mu \frac{h(R, \theta)}{[r + \lambda G(R)]^2} + (1 - u) (Z_u' - Z_Y'R) = 0,$$  \hspace{1cm} (A.5)

together with the envelope conditions

$$Z_Y = \frac{1 + Z_{Y'}(1 - \lambda)}{1 + r},$$  \hspace{1cm} (A.6)

$$Z_u = \frac{z - c\theta + Z_{Y'} \left( p(\theta) - \lambda \int_R^1 ydG \right) + Z_u' (1 - p(\theta) - \lambda G(R))}{1 + r}.$$  \hspace{1cm} (A.7)

have to hold along an optimal path.

Solving for the steady state values of $Z_Y$ and $Z_u$ one can show that

$$Z_u - Z_Y R = -\frac{h(R, \theta)}{r + \lambda G(R)}$$  \hspace{1cm} (A.8)

which can be used to replace on the right hand side of (A.5) to obtain

$$\frac{h(R, \theta)}{r + \lambda G(R)} \left[ \frac{\mu}{r + \lambda G(R)} - (1 - u) \right] = 0.$$  \hspace{1cm} (A.9)

Finally, solving for $\mu$ using (A.4), $Z_Y = (r + \lambda)^{-1}$ and (A.8) and replacing in (A.9) one obtains

$$-\frac{h(R, \theta)(1 - u)}{r + \lambda G(R)} \times \frac{r \eta + \beta \theta q(\theta) + \lambda G(R)}{\eta [r + \lambda G(R)] + \beta \theta q(\theta)} = 0,$$  \hspace{1cm} (A.10)

with $\eta = -q'(\theta)/q(\theta) > 0$. It follows that the sign of the right hand side of (A.10) is the same as $-\text{sgn} h(R, \theta)$, where the last equality follows from equation (19) in the main text. Therefore, optimality requires that the planner reduce $R$ until $h(R, \theta) = 0$.

**Proof of Proposition 4.** If $R(\theta) > R^f(\theta)$, the FOC for problem (27) is

$$\frac{\partial N}{\partial R} = \frac{\beta_u}{W(w_c) - W} \left[ \frac{-g(R)(r + \lambda)h(R, \theta)}{[r + \lambda G(R)]^2} + \frac{1 - G(R)}{r + \lambda} \right] - \frac{1 - \beta_u}{J(w_c) - J} \frac{1 - G(R)}{r + \lambda} = 0.$$  \hspace{1cm} (A.11)
The result can be proved by monotone comparative statics. Assumption (2), together with (27), implies that \( \partial (\tilde{W}(w_c) - W) \partial F \geq 0 \) and \( \partial (\tilde{J}(w_c) - J) \partial F \leq 0 \) with at least one strict inequality. It follows from (A.11) that \( \partial^2 N / \partial R \partial F < 0 \) at an optimum. Since \( \partial N / \partial R \) is strictly increasing in \(-F\), Corollary 1 in Edlin and Shannon (1998) implies that the solution to (A.11) is strictly increasing in \(-F\); i.e., strictly decreasing in \( F \).

**Lemma 1** \( R(\theta) = R^f(\theta) \) for some \( \theta_1 \) implies \( R(\theta) = R^f(\theta) \) for any \( \theta > \theta_1 \).

**Proof.** Rewrite the partial derivative in (A.11) as

\[
\frac{\partial N}{\partial R} = \frac{\beta_u}{W(w_c) - \bar{W}} \left[ -g(R)(r + \lambda)h(R, \theta) + \frac{1 - G(R)}{r + \lambda} \left( 1 - \frac{1 - \beta_u}{\beta_u} \tilde{W}(w_c) - \bar{W} \right) \right].
\]

Since \( \tilde{W}(w_c) - \bar{W} > 0 \) at an optimum, the sign of \( \partial N / \partial R \) is the same as the sign of the square bracket in (A.12). For given \( \theta \), the square bracket is largest at \( R = R^f(\theta) \), as \( G(R) \) and \( J(w_c) \) are respectively increasing and decreasing in \( R \) and \( \tilde{W}(w_c) \) is lowest when \( h(R, \theta) = 0 \). Hence, if (A.12) is negative at \( R^f(\theta_1) \) for \( \theta = \theta_1 \) it is negative for any \( R > R^f(\theta_1) \). It follows that the optimal reservation productivity satisfies \( R(\theta_1) = R^f(\theta_1) \) and \( w_c \leq r(U + F) \). Finally, since \( U \) is increasing in \( \theta \), (A.12) is declining in \( \theta \) along \( R^f(\theta) \) and it is \( R(\theta) = R^f(\theta) \) for any \( \theta > \theta_1 \).

**Proof of Proposition 6.** It follows from the proof of Lemma 1 that if \( \partial N / \partial R \leq 0 \) at \( \theta = 0 \) it is \( R(\theta) = R^f(\theta) \) for all \( \theta \). Suppose, instead, that \( \partial N / \partial R > 0 \) at \( \theta = 0 \) and therefore \( R(0) > R^f(0) \). The square bracket in (A.12) evaluated along \( R^f(\theta) \) is strictly decreasing in \( \theta \). It follows that there exists some \( \theta_1 \) for which \( R(\theta_1) = R^f(\theta_1) \). Lemma 1 implies that the same holds for all \( \theta > \theta_1 \).

**Proof of Proposition 7.**

1. A non-degenerate equilibrium with positive employment exists by assumption. Hence, the \( JD \) locus characterized by (31) intersects the \( JC \) locus in Figure 1 at least once. Equation (31) implies \( JD \) lies everywhere to the left of \( JD^f \) and does not admit turning points as a function of \( \theta \). It follows that the slope of \( JD \) has to be larger than that of \( JC \) at its vertically lowest intersection with it. If \( \pi < 1 \), an increase in \( F \) shifts \( JD \) down and locally decreases job destruction, and increases job creation and efficiency, at any intersection at which \( JD \) is steeper than \( JC \).

2. A large enough \( F \) drives \( z/q + (\pi - 1)F \) to zero.

**A.2 Bargaining over separation**

This section derives one efficient bargaining solution to determine the renegotiated severance payment \( Q(y, w_c) \) in Section 3 and shows that, in general, higher mandated severance pay increases a worker’s payoff from an ongoing match \( W(w_c) \).

Remember from Section 2.2.2 that the set of Pareto-improving and individually-rational separation transfers

\[
\Omega(y, w_c) = \{ W(w_c) - U, \min\{ V - J(y, w_c), F \} \}
\]

(A.13)
is non-empty—and mandated severance pay is renegotiated down with positive probability—only if \( W(w_c) - U < F \), that is if case 2. in Proposition 1 applies. We assume this in what follows.

Suppose that bargaining over separation consists of a take-it-or-leave-it offer of a separation transfer. If the offer is rejected trade takes place at the the current wage contract. Let \( \beta \) denote the probability that the worker is the proposer, the firm proposing with the complementary probability.

Since \( W(w_c) - U < F \), when proposing, the firm offers the worker reservation severance payment \( W(w_c) - U \) that the worker accepts. Similarly, when proposing, the worker offers the firm reservation severance payment \( \min\{V - J(y, w_c), F\} \) that the firm accepts. Since the two proposals are made respectively with probability \( \beta \) and \( 1 - \beta \), it follows that the parties would be willing to agree on the expected severance payment

\[
Q(y, w_c) = (1 - \beta)(W(w_c) - U) + \beta \min\{V - J(y, w_c), F\}. \tag{A.14}
\]

Noticing that \( J(y, w_c) = Z(y) - W(w_c) \), equation (A.14) can be rewritten as\(^{28}\)

\[
Q(y, w_c) = W(w_c) - U + \beta \min\{U + V - Z(y), F - W(w_c) + U\}. \tag{A.15}
\]

Note that, as stated in Proposition 1, the worker is better off separating—\( Q(y, w_c) > W(w_c) - U \)—if and only if separation maximizes joint wealth; i.e., for \( y < R \) where \( Z(R) = U + V \). The same is true for the firm. Also the mandated payment \( F \) is renegotiated down with positive probability for \( R < y < R \), where \( R \) is the firm firing threshold satisfying \( Z(R) - W(w_c) = V - F \).

### A.3 Efficiency wage determination

This section derives the incentive compatibility constrain in equation (28) in the main text under the assumption that the efficiency wage constraint is binding or, equivalently, that the severance payment is not renegotiated.

Let \( W(w) \) denote the value function for a worker receiving wage \( w \) in the current period before choosing whether to shirk or not. It satisfies the Bellman equation

\[
(1 + r)W(w) = w + \max \left\{ \left[ 1 - \lambda G(R) \right] W(w') + \lambda G(R) (U + F), \right. \\
\left. z + (1 - q) \left[ 1 - \lambda G(R) \right] W(w') + \lambda G(R) (U + F) \right\} + q(U + \pi F). \tag{A.16}
\]

The first term inside the maximum operator is the expected continuation utility for a worker who does not shirk, as in equation (2). The second term is the expected utility from shirking. The worker enjoys the utility from leisure \( z \), but is detected and loses her job with probability \( q \). In the latter case, her payoff equals the utility from unemployment

\(^{28}\)As shown by Fella (2005) the alternative—but qualitatively similar—efficient solution \( Q(y, w_c) = \min\{W(w_c) - U + \beta(U + V - Z(y)), F\} \) obtains if the parties bargain over the separation payment according to the alternating offer bargaining game of MacLeod and Malcomson (1993). The equilibrium severance payment shares the surplus from separation \( U + V - Z(y) \) as long as the latter is positive and the firm outside option \( V - F \) is not binding.
U plus the expected severance payment conditional on the true cause of dismissal being disciplinary.

For the worker not to shirk the first term inside the maximum operator has to be larger than the second one. Since a higher wage inefficiently increases the job destruction rate, the optimal efficiency wage $w'$ equates the two terms and satisfies equation (28) in the main text.

References


