SciencesPo
LABORATOIRE INTERDISCIPLINAIRE
D'ÉVALUATION DES POLITIQUES PUBLIQUES

LIEPP Working Paper

February 2016, nº47

Inside Severance Pay

Tito Boeri
IGIER-Bocconi University
tito.boeri@unibocconi.it

Pietro Garibaldi
Collegio Carlo Alberto, University of Torino
pietro.garibaldi@carloalberto.org

Espen R. Moen
Norwegian Business School
espen.r.moen@bi.no

www.sciencespo.fr/liepp

Inside Severance Pay

Tito Boeri*, Pietro Garibaldi**, Espen R. Moen****

October 30, 2015

Abstract

All OECD countries have either legally mandated severance pay or compensations imposed by industry-level bargaining in case of employer initiated job separations. In the literature such transfers are either ineffective or less efficient than unemployment benefits in providing insurance against labor market risk. The paper shows that mandatory severance is optimal in presence of wage deferrals induced by workers' moral hazard. We also establish a link between optimal severance and efficiency of the legal system and characterize the effects of shifting the burden of proof from the employer to the worker. Quantitatively, the welfare effects of suboptimal severance payments vary in general equilibrium between 1 and 3 percent. The model accounts also for two neglected features of the legislation. The first is the discretion of judges in declaring the nature, economic vs. disciplinary, of the layoff. The second feature is that compensation for dismissal is generally increasing with tenure.

"You should be aware that tribunals can be unpredictable in their decisions" (Understanding Employment Tribunals,Citizens Advice Bureau, UK)

"The firm does not have a clue about the actual costs of the layoffs. There is a range of costs and then substantial discretion of judges in deciding which cost to apply" (Lucia Zorza, HR Manager, Sirap Group)

"Judges retain substantial discretion over individual dismissal norms. For instance, the concept of manifest unfairness is very poorly defined and it is very important when deciding upon the reintegration of the worker.(....) Another source of uncertainty is related to the length of the judicial procedures which may last several years and involve costs for the employer much different than those initially envisaged" (Stefano Franchi and Fabio Storchi, Italian Federation of Metal Working Employers)

Keywords: Severance, Unfair Dismissal, Graded Security, Legal Systems.
JEL codes: J63,J65,J33.

**IGIER-Bocconi University **Collegio Carlo Alberto, University of Torino ***Norwegian Business School.

Introduction

Most OECD countries have legally mandated severance pay in case of employer initiated job separations. When rules for compensations to workers are not specified by the law, it is collective bargaining at the industry or national level to mandate severance to individual employers. For instance, Kodrzycki [24] reports that 86% of workers in Massuchettes are covered by a severance pay agreement, involving one week's wage per year of service, that is, about 50% of the maximum unemployment insurance for the workers in the sample. Such transfers from the employer to the worker are the most important component of individual dismissal costs[1]. The average compensation for unfair dismissals is about two years of pay in case of a worker with at least 20 years of tenure. There are countries in which severance may involve up to 5 years of pay. According to Garibaldi and Violante [20] who estimated the red tape costs of layoffs in Italy, severance pay accounts for about 2/3 of total dismissal costs. Severance pay also accounts for almost 50% per cent of the cross-country variation in the OECD index of the strictness of employment protection legislation (EPL) for regular workers, the reference measure of EPL in the literature. It is mandatory even in countries with negligible firing taxes to be paid to third parties.[2]

A fundamental difference between severance payment and firing taxes is that the latter is paid to a third party, while the former is a pure transfer. Severance pay differs from firing taxes also in that it is generally dependent of tenure, while firing taxes are independent of tenure and are indeed modeled by the literature as a flat cost for the employer. Furthermore, the amount of severance pay depends both on the nature – disciplinary vs. economic – of the dismissal, and on whether it is deemed fair or unfair by a Court ruling. This role of the legal system that can only imperfectly monitor the performance of workers and employers is very important in assessing the incentives associated with the provision of severance pay.

Reforms of these regulations are high on the policy agenda and have been explicitly requested by IFI (International Financial Institutions) to the so-called program countries during the Eurozone public debt crisis. Distinguished economists active in the US policy debate have been suggesting that severance should be increased at least during downturns, while several labor economists in Europe have been advocating the introduction of a tenure-related security contract involving severance being gradually increasing with tenure.

In spite of the very high relevance of severance pay, we still lack a proper framework to evaluate it. The extensive literature on Employment Protection Legislation fails to characterise some of the key features of mandatory severance. In particular, it does not take into account that the size of these transfers typically varies depending on the seniority of the worker, on the nature – economic or disciplinary – of the dismissal and on its legitimacy (fair or unfair), as established in a Court ruling.

The purpose of this paper is twofold. First, we provide a normative framework, aimed at extending and systematizing earlier results as to the efficiency of severance pay. Second, we model the relationship between tenure and severance, and open the blackbox of judicial systems, notably characterizing how the efficiency of the legal system and the rules concerning the burden of proof affect the optimal design of severance. We show that mandatory severance is optimal even in the absence of risk aversion and when there are no ex-ante rents to be split between the worker and the firm, which are posited in the efficiency wage literature. Therefore, the 'bonding critique'does not apply. What is sufficient to make severance efficient is the presence of constraints to wage renegotiations either in terms of wage deferrals motivated by deterrence of opportunistic behavior of workers or of downward nominal wage rigidity. We provide a formal argument of why severance should be enforced by a co-ordinating mechanism outside the single firm, which is based on the fact that adverse selection stands on the way of severance schemes introduced by individual employers. Our results hold also in general equilibrium, and we can assess quantitatively the impact of removing severance pay. We demonstrate that under reasonable parameter values, the welfare effects of removing severance pay are between 1 and 3 percent, depending on the size of the initial severance pay. Our results suggest also that severance should be increasing in the inefficiency of the legal system. We also find that, under general circumstances, optimal severance is increasing with tenure. Our results are empirically relevant. Legal rules

[1] Group layoffs, that is, collective dismissals involving a discrete number of workers of the same firm, are not considered in this paper

[2] See Postal-Vinay and Turon [40], and Boeri [9] for a theory of severance pay as a device to buy time and avoid paying firing taxes, in presence of on-the-job search.

about the severance-tenure profile appear to be positively correlated with the wage-tenure profile that we estimate drawing on longitudinal data. We also find that OECD measures of efficiency of judicial systems are correlated with severance pay for individual economic and unfair dismissals in a way which is consistent with the implications of the model.

The plan of the paper is as follows. Part one reviews in some details our contribution to the extended literature on EPL. Next, it characterizes two neglected features of EPL, the discretion of judges in setting the level of severance pay depending on whether the individual dismissal is disciplinary, economic or unfair, and the tenure profile of severance pay. Part two presents the model with moral hazard of the employees, and evaluates optimal severance pay under these circumstances. It also provides a formal argument of why severance should be mandated to individual employers. Part three extends the model looking into Court rulings as to the nature of dismissals, endogenizing the probability that not investing workers get severance pay for economic dismissal and that the dismissal is considered unfair. Part four extends the results to the general equilibrium, endogenizing the workers' outside option, and provides numerical simulations of the effects of different levels of severance pay on welfare and unemployment. Part five goes back to the data investigating the correlation between severance and efficiency of judicial systems, and the severance tenure profile under different regimes as to wage deferrals and involvement of Courts in layoff procedures. The final section summarizes our key results and concludes.

1 Severance pay: literature and neglected features

1.1 Our contribution to the literature

Employment protection legislation is one of the most widely investigated institutions in the labor market[3]. The theoretical literature, pioneered by Bentolila and Bertola [4], Bertola [6] and Lazear [27], typically treats EPL as a firing tax to be dissipated or paid to a third party by the employer in case of a layoff.

Severance pay, that is, a transfer from the employer to the worker contingent on employer initiated separations[4] is generally not framed in these models, as Lazear [27] neutrality result indicates that, with wage flexibility and risk neutrality, it only affects the tenure profile of wages leaving employment, hiring and separations unaffected. When instead, wages are rigid, severance pay increases unemployment (Garibaldi and Violante [20]).

Why do we need then severance pay then? There are three key rationales for severance pay according to the literature.

The first draws on moral hazard and adopts the standard setup of the efficiency wage models a la Shapiro-Stiglitz [44]. A severance to be paid to the workers who are fired without having been shirking, acts like a commitment to an employment policy that does not strongly react to negative shocks. By playing this role, severance reduces labor costs. Without the severance, wages would have to be increased to deter shirking. This is because the penalty to shirkers is provided by the expected rent associated to the fact of having a job with respect to the value of unemployment, and this ex-ante rent falls when there is an exogenous probability of being laid-off even if a worker is not shirking. Fella ([15] and [16]) draws on this initial intuition by Saint-Paul ([43]) to show that an optimal severance can be as high as to equalize wages across all possible productivity realizations. These results have been extended by Baumann [3] to the case where even some shirkers can receive severance pay and to double moral hazard (of employers choosing over projects having different levels of risk, in addition to employees deciding as to whether to put effort). This extension draws on the work by Galdon-Sanchez and Guell ([18]) who introduced the possibility that shirkers 'can get away with it 'in a standard model of employment protection, but did not evaluate the efficiency properties of severance schemes.

A problem with this explanation of severance is that it is subject to the 'bonding critique 'which challenges the efficiency wage literature. A cheaper deterrent to opportunistic behavior is for firms to commit to a wage schedule offering initially lower wages (even below the marginal product) and higher wages if confirmed in the

[3]See Boeri and vanOurs [10] for a review of this literature.

[4]Our definition of severance clearly does not encompass deferred compensation schemes, such as private pension arrangements, which are paid at retirement or at any separation, including voluntary quits.

firm. In other words, severance pay is not needed when there are wage deferrals allowing for wages increasing with tenure as those documented by the empirical literature estimating Mincer-type wage equations.

The second argument for severance pay rests on risk aversion. Severance pay protects workers against uninsurable labor market risk, just like unemployment benefits do. Under full insurance (when the value of employment and unemployment is equalized), Blanchard and Tirole ([8]) show that severance pay has to be preferred to unemployment benefits because it internalizes the costs of layoffs. By the same token, employers could pay themselves the unemployment benefits or, at least, contribute more to the funding of unemployment benefits when they make use of them. There is, in other words, a full substitutability between severance pay and unemployment benefits when the latter can be experience-rated. When full insurance is not feasible, there is no longer full substitutability between the two institutions. There is always a role for the state in the provision of unemployment insurance and severance is not fully crowded out by unemployment benefits.

A problem with this explanation is that the borders between severance and unemployment are not very well defined. An unemployment benefit can be transformed into a severance scheme by forcing employers laying off workers to contribute more to the funding of unemployment insurance than other firms. It is debatable whether this is a theory of severance or a theory of unemployment benefits. What is clear is that it involves a substitutability between the two institutions.

A third rationale for severance pay provided by the literature is related to wage rigidities. With constant wages, search frictions and idyosincratic shocks to productivity, severance prevents excess job reallocation, as in the model by Alvarez and Veracierto [1]. A problem with this explanation is that it needs once more to posit a counterfactual wage rigidity – the fact that wages are constant over the lifetime – to attribute a welfare enhancing role to severance pay.

In our model, severance pay is efficient even when workers are risk-neutral and entry wages are fully flexible. Still wage deferrals needed to incentivize workers create a rationale for severance pay; once the firm is committed to a wage schedule, it may find it ex-post optimal to layoff a worker even when the job is generating some surplus over the value of unemployment. This wage schedule is not exogenous, but it is optimized at the time of job creation, and cannot be made contingent on a random component of output. Furthermore, there is no substitutability between unemployment benefits and severance pay. The two institutions are complementary as unemployment benefits, under a broad set of circumstances, make the wage tenure profile steeper, inducing more inefficient separations when severance pay is absent.

In our model, wages are deferred in order to incentivize workers to investment in job-specific productivity. We argue that our mechanism is relevant in all situations in which wages are deferred. The underlying assumption is that firms can commit to a future wage schedule (not contingent on individual productivity), but not on the employment relationship. Hence a firm cannot commit not to fire a worker if that is in the firm's interest *ex post*, or to any payments to workers who are fired. This is a standard assumption in the literature (see Menzio and Moen [32] and the references therein). Under these conditions, severance deals with the moral hazard problem associated with firms firing too frequently senior workers receiving deferred wages. The result is general as wage deferrals are a common feature of labor markets: most firms allow for a significant component of remuneration to be postponed to avoid agency problems and to motivate workers and these tenure-related components of compensation are agreed in advance, conditional on the continuation of a job, but independently of productivity realizations. Moreover, studies measuring both wages and productivity (e.g., Medoff and Abraham [31], Kotlikoff and Gokhale [25] Flabbi and Ichino [22]) suggest that the effects of seniority on wage profiles observed by a large body of empirical literature can be attributed mainly to incentive reasons, and are not necessarily associated with a higher productivity of senior workers. There is also a large body of indirect evidence of deferred compensation. For instance, it is consistent with the findings by Lazear and Moore [26], who compared seniority-earning profiles of employees and self-employed (for which no agency problem arises) and by Barth [2], who compared the wage-tenure profile of workers paid piece-rate with that of workers receiving a flat wage.

We also relate to specific features of severance pay. The theoretical literature on EPL typically treats severance as a deterministic transfer from the employer to the employee. In the few cases where stochastic severance is allowed (Garibaldi [19], Malo [30]), it is modeled more as an option to fire (a firing permission) than as a distribution of alternative costs of dismissals. Moreover, no reference is made by this literature to the moral hazard problem related to the distinction between economic and disciplinary dismissals. Two partial exceptions are Galdon-Sanchez [18] and Boeri [11]. However, Galdon-Sanchez [18] operates on a reduced form model and both Boeri [11] and Galdon-Sanchez [18] do not address the efficiency of severance pay[5], but only consider its effects on unemployment and the layoff behavior of firms of different size.

1.2 Neglected features

Our model also allows to rationalize two neglected features of Severance Pay that affect the cost of individual dismissal.

The first relates to the discretion of judges in deciding upon the fairness and the nature (economic vs. disciplinary) of the dismissal. Compensation is generally not offered to workers being fired for disciplinary reasons unless a Court ruling declares that the dismissal is unfair. When the individual layoff is instead motivated by the economic conditions of the firm, that is, it occurs independently of the behavior of the worker, compensation is typically offered also for fair dismissals, that is, cases where there is no evidence of opportunistic behavior of the employer. In the case of unfair dismissals, however, compensation is higher than the severance for fair economic dismissals. There are also countries in which compensation is provided only for unfair dismissals and fair economic dismissals do not involve mandated severance to the workers. Due to these wide differences in the levels of compensation related to the nature of dismissals, there are strong incentives for the employee or the employer to bring the case before a Court. Involvement of judges in the determination of the level of severance cannot be avoided by state contingent contracts, and since workers' effort and employers' investments in the duration of the job are not perfectly observable, the decisions of the judges will tend to be imperfect. The judicial discretion clearly affects also private settlements out of Court, as such settlements will be based on the expected costs had the case gone to Court. These relevant interactions between EPL and the efficiency of judicial systems have been neglected to date by the theoretical literature on EPL although there is evidence (Fraisse, Kramarz and Prost [17]) that the organizational structure of judicial systems does affect significantly labor market outcomes.

The second neglected characteristic of EPL is the *tenure profile* of severance pay. As documented in this paper, most countries allow for mandated severance pay to be increasing with tenure. We are not aware of any theory rationalizing these arrangements on the basis of purely efficiency considerations. Personnel economics offers explanations for why firms offer *tenured jobs*, that is, positions that cannot be severed under any set of circumstances. Tenured jobs can be rationalized as the result of learning about match quality or hiring incentives in organizations where incumbents have control over hirings, e.g., in academic institutions. Tenure prevents the strategic choice of incumbents of hiring only low quality workers in order to reduce competition with outsiders (Carmichael [13]). These theories explain why employers may decide to commit not to layoff some workers, but do not explain why a *mandated* profile of severance increasing with tenure is chosen for potentially *all* private firms, irrespective of whether incumbents in these organizations play any role in hiring decisions or there is substantial heterogeneity in the quality of applicants. Moreover, these models do not address problems of commitment: private firms generally cannot credibly commit not to layoff some workers, irrespective of their performance.

Judicial discretion and burden of proof

Statutory severance pay levels depend on the nature, economic vs. disciplinary, and on the fairness of dismissals. Fairness in the case of *economic dismissals* refers to the behavior of the employer: she should have tried as much as possible to avoid this outcome. Although the definition of fair economic dismissal differs quite considerably from country to country, it generally implies that some "genuine and serious" exogenous shocks in firm's performance require "operational changes" in the scale, and possibly, nature of the work organization, making the worker involved redundant. Often evidence of "economic difficulties" or "technological change" is explicitly required.

[5] R*ü*hmann and S*ü*dekum [42] consider the efficiency of severance payments in terms of human capital investment, but do not address the moral hazard problems associated with Court involvement and the severance-tenure profile.

In the case of *disciplinary dismissals*, the fairness refers to the behavior of the worker. Fair disciplinary dismissals are those for which there is evidence of misconduct on the part of the worker, where "misconduct" is often not defined, and the burden of proof typically falls onto the employer. When the economic or disciplinary dismissal is found by a Court to be "unfair", the employer in some countries is forced to reinstate the worker. Generally the reinstatement does not take place, but the compensation paid to the worker increases. Everywhere, the costs of unfair dismissals are significantly higher than those of fair economic dismissals. Moreover, the employer, in addition to providing severance pay, typically has to pay the legal costs of the employee and compensate for the foregone months of pay during the legal procedure. The decisions as to the nature of the dismissal and its fairness require some Court ruling. In practice, disputes are mostly settled before the Court decision, taking in consideration the nature of the dismissal, the probability that is considered fair and the severance and additional compensations envisaged under the different circumstances. Thus, in practice the level of severance ultimately depends on decisions made by third parties having limited information on the behavior of workers and employers. For all of these reasons the actual costs of layoffs are stochastic, and generally depend on the evidence that the employer can provide for a disciplinary or economic dismissal.

Uncertainty as to the actual costs of the dismissal is increasing, inter alia, in differences in the level of mandatory compensation required under the three types of dismissals discussed above, that is, fair economic, fair disciplinary, and unfair dismissals. Table 1 displays the maximum compensation (severance pay plus notice period) required in these three cases in OECD countries. The table is based on the analysis of the country files used by the OECD in building up the summary measure of strictness of EPL, a report prepared for a European conference of labor lawyers [14], a study by the ILO [7] and a recent survey of Civil Justice also carried out by OECD (Palumbo [38]).

As shown by Table 1, in all countries even fair dismissals command some compensation to the worker, either in terms strictly of severance pay or of a minimum notice period (de facto an extension of pay after the date when the worker is made redundant). The compensation for unfair dismissals (T_U, first column) is, however, always higher than that provided in case of fair dismissals (either economic, T_F^E, or disciplinary, T_F^D, second and third columns). One of the reasons why unfair dismissals cost more than fair dismissals is that in several countries (see Table A2 in the Annex), in addition to a monetary compensation, an unfair dismissal may also be sanctioned with the compulsory reinstatement of the worker in the ranks of the firm. Thus, in these countries, the costs of unfair dismissals should include the duration of the trial period, as reinstated workers should be back paid the full wage between the date of the dismissal and that of the Court ruling, and an additional compensation, as the worker and the employer generally agree on a monetary transaction in lieu of an actual reinstatement after the Court ruling. This compensation will be clearly related to the protection provided to job-holders, that is, to the severance in case of unfair dismissals in that specific country. Thus, we estimate the costs of unfair dismissals as given by the statutory notice period (N) and severance (S), plus, limited to the countries with reinstatement, the average length of the trial period (d) and the compensation for unfair dismissal (S), which is a proxy for the cost of the reinstatement, the latter two terms multiplied by the likelihood that a reintegration of the worker is actually imposed by the Court.[6] As shown by the fourth and fifth columns of Table 1, unfair dismissals are significantly more expensive than fair economic dismissals, while the latter are more expensive than fair disciplinary dismissals, which typically involve only a relatively short notice period.

The above suggests that there is substantial uncertainty as to costs of dismissals for an employer. This uncertainty is summarized in two measures provided in the Annex (Table A2), notably a measure of dispersion, and a measure of judicial discretion. Another fact highlighted by the last column on the right-hand-side of Table 1, is the presence of a significant cross-country variation in legal rules concerning the burden of proof. Most countries put the burden on the employer, but there are cases even outside Eastern Europe (e.g.,

[6] Denoting by (π) the probability that a reintegration of the worker is imposed by the Court, we have that:

$$T_U = N + S + \pi(d + S)$$

where the unit of measurement is monthly wages. As detailed in Table A1 in the Annex, we attribute to π the value obtained by standardizing to the unit interval the 0-3 OECD index on the likelihood of the reinstatement, where 0 means never reinstatement and 3 denotes the case where employees can freely decide upon the reinstatement in the case where the dismissal is ruled to be unfair.

France, Denmark, Switzerland, and Japan) where the burden of proof falls partly on the worker according to the OECD.

The elasticity of severance to tenure

In 25 countries out of 30 there is evidence of severance increasing with tenure (Figure 6 in the Annex). If we add the notice period (de facto an extension of the contract after the notification of the dismissal giving to the worker time to find alternative employment, see Table A3 in the Annex), only two countries pay the same compensation at all tenure levels, notably Austria and Japan.

Why do regulations in so many countries allow for severance graded with tenure? Is this profile efficient from the standpoint of the individual worker and firm involved? There may be social efficiency considerations for having employment protection increasing with tenure, e.g. related to the fiscal externalities associated to layoffs in presence of tenure-related unemployment benefit systems and/or job finding rates declining with age. There can also be equity considerations for offering stronger protection against layoffs to older workers, but we are not aware of theories rationalizing these arrangements from the standpoint of purely private efficiency.

In the model presented in Section 2, a privately efficient and positive severance-tenure profile emerges as a result of moral hazard related to the stochastic nature of severance pay and the difference between disciplinary and economic dismissals. The stochastic nature of severance is due to the fact that the nature (economic vs. disciplinary) of the dismissal has to be proved before a Court, and there is an exogenous probability that disciplinary dismissals are treated as economic dismissals. In Section 3, this probability is endogenized by looking into the Court ruling in relation to different possible productivity realizations. Opportunistic behavior of the employer is also considered in this section, allowing for unfair dismissals.

2 The model

In this section we develop a model of (one-sided) moral hazard of the worker, endogenous wage deferrals and exogenous Court ruling. This model is extended in Section 3 to allow for endogenous Court decisions, and double moral hazard, and in Section 4 to allow for endogenous outside option of the workers in general equilibrium.

One worker and one firm have a job opportunity that lasts n periods. The worker and the firm are risk neutral, and both discount the future at rate β, with $\beta < 1$. For simplicity of calculation, we assume that the worker is infinitely lived. Time is discrete. The worker's outside option in every period is U, which denotes the lifetime value of unemployment, and will be obtained in general equilibrium, but we take it as given for the time being. In period $\{i\}_{i=0}^{n-1}$ the worker faces a specific investment opportunity $s_i = \{0, 1\}$. The investment opportunity costs $\{C_i\}_{i=0}^{n-1}$ to the worker in each period. The investment is private information to the worker in each period.

The firm contract- to be discussed below- ensures that the worker invests in every period. Conditional on the investment being undertaken at time $i = 0, 1, n-1$, productivity in the following period will be ϵ_i, where ϵ_i is drawn from a continuous distribution $F(\epsilon_i)$, defined over the support $Z = [\underline{\epsilon_i}, \overline{\epsilon_i}]$. The distribution is time invariant. Note that there is no investment in period n. Further, productivity in period 0 is deterministic and fixed at $\tilde{\epsilon}_0$. Productivity is observed only by the firm, hence wages cannot be made contingent upon it. We assume that $\overline{\epsilon_i}$ is sufficiently large for the firm to break even, and for investment to be profitable in all periods (unless the worker is fired).

If the worker shirks in period i, this is observed by the firm in the following period, either directly or indirectly because his productivity is low. In this case, the firm will want to undertake a disciplinary dismissal. In the next section, where we focus on the burden of proof, we will be very specific about the productivity of the shirking worker. At this stage, we just assume that the firm initiates a disciplinary dismissal when it finds out about a non-investing worker. Note that the firm may also want to fire the worker, even if the worker has invested, if the draw of ϵ_i is sufficiently low. We refer to this case as an economic dismissal. The legislation applies different rules to these different types of dismissal.

Table 1: Judicial discretion over severance pay

Country	T_U	T_E^F	T_D^F	$T_U - T_E^F$	$T_E^F - T_D^F$	Burden of Proof
Australia	13.90	3.80	1.00	10.10	2.80	Employer
Austria	20.29	4.00	4.00	16.29	0.00	Employer
Belgium	31.30	11.15	11.15	20.15	0.00	Employer
Canada(Federal)	-	4.3	2.00	-	2.30	Employer
CzechRepublic	19.99	3.50	2.00	16.49	1.50	Worker
Denmark	19.97	9.00	6.00	10.97	3.00	Worker
Finland	20.00	6.00	6.00	14.00	0.00	Employer
France	27.67	7.40	2.00	20.27	5.40	Worker
Germany	43.58	17.00	7.00	26.58	10.00	Employer
Greece	-	12.00	4.00	-	8.00	Employer
Hungary	27.16	9.00	3.00	18.16	6.00	Worker
Ireland	40.90	6.00	2.00	34.90	4.00	Employer
Italy	40.14	6.00	6.00	34.14	0.00	Employer
Japan	10.16	1.00	1.00	9.16	0.00	Both
Korea	17.81	1.00	1.00	16.81	0.00	Worker
Luxembourg	18.20	12.00	6.00	6.20	6.00	Employer
Mexico	-	-	-	-	-	Employer
Netherlands	16.67	4.00	4.00	12.67	0.00	Employer
NewZealand	12.49	0.50	0.50	11.99	0.00	Employer
Norway	29.61	6.00	6.00	23.61	0.00	Employer
Poland	11.82	6.00	3.00	5.82	3.00	Employer
Portugal	62.85	14.50	2.50	48.35	12.00	Employer
SlovakRepublic	27.79	7.00	3.00	20.79	4.00	Worker
Spain	36.50	12.50	0.50	24.00	12.00	Employer
Sweden	38.00	6.00	6.00	32.00	0.00	Employer
Switzerland	9.00	3.00	3.00	6.00	0.00	Worker
Turkey	32.00	22.00	2.00	10.00	20.00	Employer
UnitedKingdom	17.67	7.60	3.00	10.07	4.60	Employer
United States	-	0.00	0.00	-	0.00	Employer

Notes: All magnitudes are expressed in monthly wages.
Reference is made to a worker with 20 years of tenure.
T_U is compensation for unfair dismissal; T_E^F is compensation for fair economic dismissal, and T_D^F is compensation for fair disciplinary dismissal.
Sources: EPLex; OECD (2013);
See the main text and Table A1 in the Annex for details.

Definition 1 *Disciplinary Dismissal. In period i, a firm is entitled to freely dismiss a shirking worker who did not invest in the period $i-1$.*

Definition 2 *Economic Dismissal. In period i, when productivity is sufficiently low, a firm is entitled to dismiss a worker by paying a severance T_i.*

Note that disciplinary dismissals, which we refer to as worker's shirking, need a third party intervention to certify the case. In the paper, we refer to the Court of law. From the match standpoint, the Court ruling is stochastic. We assume that there is a probability $1-q_i$ that the Court observes shirking and declares the firing as fair and pure disciplinary. In such a case, the firm is exempted from paying severance payments. Hence, there is a probability q_i that a shirking worker gets away with it and receives the severance payment. Since the realization of q_i is made after the firm has fired the worker, the expected severance payment for the firm when firing a shirking worker is $q_i T_i$. In this section, q_i is considered exogenous. It will be endogenized in section 3.[7] In what follows we assume that a firm always finds it in its interest to fire shirking workers, and then demonstrate that this is always the case when the severance is optimally set.

In the case of an economic dismissal, we assume that the severance payment is always due, hence we abstract from moral hazard on the firm side. In the next section we show that this is a valid assumption if the burden of proof is on the firm. Furthermore, if the burden of proof is on the worker, and there is moral hazard on the firm side, the optimal severance is unchanged.

Before we continue and derive the optimal contract, let us comment on our driving assumption that a firm cannot commit to a severance payment, and hence that the severance payment T_i is a policy tool. We rationalize our assumption that the firm cannot contract upon T_i by alluding to an underlying, not-modeled problem of adverse selection that stands on the way of a private contractual arrangement. If a firm unilaterally commits to a severance payment, it would be a victim of negative selection, and would end up hiring less favorable workers.

Let us be more specific. Suppose that there are two types of workers; ordinary workers as described above and shirkers, with $C=\infty$. Hence the shirkers always shirk. The fraction of the "shirkers" may be small, but strictly positive. Firms cannot distinguish between shirkers and ordinary workers. Consider a two periods situation where $n=1$, and all firms offer a contract (w_0, w_1, T_1), where w_0 and w_1 are wages in two periods, respectively, and $T>0$ is a privately imposed severance. We will argue that this cannot be an equilibrium. Consider a firm that deviates and offers a contract $(w_0, w_1', T_1-\epsilon)$, where $w_1' > w_1$ and ϵ can be arbitrarily small. Since ordinary workers are strictly more willing to trade off severance payment for a higher period 1 wage than are the shirkers, it is possible to chose w_1' so that ordinary workers strictly prefer the new contract and shirkers strictly prefer the old contract. Hence the deviator only attracts the more profitable ordinary workers, and the equilibrium unravels. This argument can be used for any equilibrium candidate in which also ordinary workers receive severance pay. Thus, an arbitrarily small fraction of shirkers drives out severance pay for ordinary workers altogether. A formal treatment of this argument is provided in the Annex (see also [34] and [41]).

A mandatory severance solves this co-ordination problem. The realism of this assumption can be assessed considering that severance, wherever it exists, is either legislated or established within collective agreements at the industry, state or national level.

2.1 Optimal contract

A wage contract Ω specifies a sequence of wages $\{w_i\}_{i=0}^{i=n}$ from period 0 to period n. The sequence of severance payments $\{T_i\}_{i=0}^{i=n}$ is taken by the firm as given. In each period, the firm will fire the worker if it is in its interest. We will refer to this as the firm's firing constraint.

The optimal contract $\Omega^{\xi} = \{w_i^{\xi}\}_{i=0}^{i=n}$ maximizes the firm's profit given the following constraints

1. worker's incentive compatibility constraint (ICC);

2. worker's participation constraint (PC) ;

[7]In terms of the definitions used in Section 1, we have that T_F^D=0 and T_F^E=T_i in this setting. T_U is introduced in section 3 below. Here we focus only on moral hazard of the employee.

3. firm's firing constraint (FC).

In what follows we discuss these three constraints in some details. After deriving the optimal contract Ω^{ξ}, we will solve for the optimal severance sequence $\{T_i^*\}$. We will solve the model by backward induction starting in period n. To begin with, we define the basic value functions and derive the workers' incentive compatibility constraints from period $i = 1$ to period $i = n-1$. We then proceed- by backward induction- to derive the optimal firing policy. The participation constraint is also illustrated. It will be made endogenous in the general equilibrium part of the paper.

The incentive compatibility constraint

The value of a job for an investing workers at period $i-1$ is

$$W_{i-1} = w_{i-1} - C_{i-1} + \beta\left[(1 - F(\epsilon_i^d))W_i + F(\epsilon_i^d)(U + T_i)\right] \tag{1}$$

where w_{i-1} is the wage and C_{i-1} is the investment cost in period $i-1$ that will affect productivity in period i. W_i is the value of the job in the following period. Note that an investing worker is retained the following period with probability $(1 - F(\epsilon_i^d))$, where ϵ_i^d will be chosen optimally by the firm so as to maximize ex post profits. With the complementary probability $F(\epsilon_i^d)$ the worker will get the outside option U augmented by the severance payment T_i. Suppose that the incentive constraints holds from period i onward. The value of employment for a shirker in period $i-1$ is $\tilde{W}_{i-1}$ and reads

$$\tilde{W}_{i-1} = w_{i-1} + \beta[q_i(U + T) + (1 - q_i)U] \tag{2}$$

The shirker does not invest C_i, and is fired in period i. Once fired, he will be certainly entitled to the outside option U, but with probability q_i he will get also a severance payment T_i.

The incentive compatibility constraint implies that the wage w_{i-1} be set so that $W_{i-1} \geq \tilde{W}_{i-1}$. In the Annex we show the following lemma

Lemma 1 *Suppose $T_i \leq T_i^*$ for all i. Then the worker's incentive compatibility constraint binds in all periods*

A rough intuition goes as follows: suppose that the contract gives rents to the worker in period i. Suppose that the firm reduces the wage in period i down to the incentive compatibility constraint, and increases the wage in period $i-1$ so that the worker's expected income W_{i-1} (and hence also all earlier periods) stays constant. This lowers the firing threshold of the worker in period i, increasing the joint income of the worker and the firm in that period, and hence the joint income of the match. Since the worker is on his participation constraint, this increase in joint income accrues to the firm.

The ICC condition thus writes

$$-C_{i-1} + \beta\left[(1 - F(\epsilon_i^d))W_i + F(\epsilon_i^d)(U + T_i)\right] = \beta(U + q_i T_i)$$

After simple algebra, the incentive compatibiliy constraint can be written as follows:

$$W_i - U = \frac{C_{i-1} + T_i\beta[q_i - F(\epsilon_i^d)]}{\beta[1 - F(\epsilon_i^d)]} \tag{3}$$

The rent $R_i = W_i - U$ has to be given to the worker in order to induce her to invest at time $i-1$.

Let us give some comments on the optimal rent indicated by equation (3). First, note that if $\epsilon_i^d = \underline{\epsilon}_i$, i.e. if workers who invest are never dismissed, then $R_i = C_{i-1}/\beta + q_i T$. In this case the worker who invests is compensated for her outside option U, her investment cost C_i, and the rents $q_i T_i$ she would get if shirking. Second, the numerator in (3) increases in T_i if $q_i > F(\varepsilon_d)$. This stems from the fact that the worker in this case is more likely to get the severance if shirking than if not shirking. If $q_i < F(\epsilon_d^i)$, the opposite holds.

The worker's participation constraint

Given the sequence of value functions for an investing worker $\{W_i\}_{i=1}^{i=n}$, the investing worker in period 0 invest C_0 to affect its productivity in period 1. The wage w_0 is set so as to satisfy the participation constraint U so that

$$W_0 = w_0 + \beta \left[(1 - F(\epsilon_1^d))W_1 + F(\epsilon_1^d)(U + T_1)\right] \geq U \quad (4)$$

For the optimal contract Ω^ξ, the participation constraint (PC) binds and uniquely determines w_0.

The firing constraint in period n

Given the ICC of equation (3) and the PC of equation (4), we can now move to the firm problem, starting in period n In the last period of the relationship, the job comes to an end and the firm continuation value is zero. The period n firm problem is static and, conditional on observing the idiosyncratic productivity ϵ, its profits are $\Pi_n(\epsilon) = \epsilon - w_n$ As there is no investment in period n and no firing in period $n+1$, the worker value function is $W_n = w_n + \beta U$, where w_n is period n wage and U is the participation constraint of the worker. Using the the ICC of equation (3) and the $n-1$ investment, the wage reads

$$w_n = b + \frac{C_{n-1} + T_n\beta[q_n - F(\epsilon_n^d)]}{\beta[1 - F(\epsilon_n^d)]} \quad (5)$$

where $b(U) = U(1-\beta)$ is the flow value of unemployment. In what follows, to ease notation, we will simply indicate $b(U) = b$. The firm profit is then

$$\Pi_n(\epsilon) = \epsilon - b - \frac{C_{n-1} + T_n\beta[q_n - F(\epsilon_n^d)]}{\beta[1 - F(\epsilon_n^d)]}$$

The firm's continuation policy satisfies the reservation rule and we call ϵ_n^d the reservation productivity in period n. The reservation rule solves $\Pi_n(\epsilon_n^d) = -T_i$ and the corresponding productivity reads

$$\epsilon_n^d = b + \frac{C_{n-1} - (1 - q_n)\beta T_n}{\beta[1 - F(\epsilon_n)]} \quad (6)$$

This equation uniquely determines the reservation productivity in period n and represents the firm's firing constraint (FC). We will use the previous equation below, as we derive by backward induction the (identical) productivity for the general period i in equation (11). We can thus proceed to derive the full sequence for wages, rents and reservation productivity for the entire contract. $\{w_i, W_i, \epsilon_i^d\}_{i=1}^n$.

The firing constraint in earlier periods

Having solved the wage and the reservation productivity in period n, we can proceed by backward induction to solve the firm problem from $i = n-1$ to period 1. Given the optimal behaviour in all other periods, profits at period i for a firm that operates at productivity ϵ are given by

$$\Pi_i(\epsilon) = \epsilon - w_i + \beta \left\{ \int_{z \in Z} Max[\Pi_{i+1}(z); -T_i] dF(z) \right\} \quad (7)$$

The max operator reflects the fact that the firm has always the option to severe the relationship and pay T_i. Firm's continuation policy satisfies the reservation rule, and, at each i, we have that ϵ_i^d solves

$$\Pi_i(\epsilon_i^d) = -T_i \quad (8)$$

While $\Pi_i(\epsilon)$ refers to to the profit function conditional on a given ϵ, we need to introduce the expected profit at period i conditional on firing. The latter is simply

$$\Pi_{i+1}^e = \int_{\epsilon_{i+1}^d}^{\bar{\epsilon}_i} \Pi(z) dF(z)$$

where we stress that expected profits are a costant vis-á-vis a particular value of ϵ. We can thus get rid of the max operator in equation (7), and write profits in period i as

$$\Pi_i(\epsilon) = \epsilon - w_i + \beta\left[\Pi^e_{i+1} - F(\epsilon^d_{i+1})T_{i+1}\right] \tag{9}$$

Define the surplus at time i as

$$S_i(\epsilon) = W_i(\epsilon) - U + \Pi_i(\epsilon)$$

Since severance payments are a pure transfer, they do not enter the definition of (joint) surplus. The expected surplus in period $i+1$ is simply

$$S^e_{i+1} = \Pi^e_{i+1} + [W_{i+1} - U]\,(1 - F(\epsilon^d_{i+1}))$$

where again, S^e_{i+1} does not depend on the actual productivity in period i. Using the ICC wage from equations (1) and (3) to get rid of the flow wage value w_i, the period i wage reads

$$w_i = b + \frac{C_{i-1} + T_i\beta[q - F(\epsilon^d_i)]}{\beta[1 - F(\epsilon^d_i)]} + C_i - \beta\left[(1 - F(\epsilon^d_{i+1}))(W_{i+1} - U)\right] - \beta F(\epsilon^d_{i+1})T_i$$

The expression of the profits conditional on a productivity ϵ, from equation (9), can be compactly written as

$$\Pi_i(\epsilon_i) = \epsilon - b - \frac{C_{i-1} + T\beta[q_i - F(\epsilon^d_i)]}{\beta[1 - F(\epsilon^d_i)]} - C_i + \beta S^e_{i+1} \tag{10}$$

The reservation productivity is defined as in equation (8) and by simple algebra and substitution reads

$$\epsilon^d_i = b + \frac{C_{i-1} - T_i\beta(1 - q_i)}{\beta[1 - F(\epsilon^d_i)]} + C_i - \beta S^e_{i+1} \tag{11}$$

This equation is the generic FC expression, and generalizes the reservation productivity at period n of equation (6). Note that both the left-hand side and the right-hand side of the equation are increasing in ϵ^d_i, hence the equation may not have a solution.[8] In addition, we require that investment is better than closing down the firm at the end of the period, i.e., that $S^e_{i+1} \geq C_i$ for all $i < n$. If investments are sufficiently productive, in a well defined sense, the equation has a solution, and we say that the investment is implementable. To be more precise, suppose that the distribution can be written as $\epsilon = kz$, where z is a stochastic variable on $[0, 1]$ with median value of $z^m > 0$ and expected value of $\bar{z}$. The scalar k is a measure of the productivity of the investment. Furthermore, on intervals of T_i and C_{i-1} where investments are implementable, ϵ^d_i is increasing in C_i and decreasing in T_i

Lemma 2 *Consider an arbitrary investment cost C_i and severance T_i. Then, if the investment is sufficiently productive, the investment is implementable. On intervals where the investment is implementable, ϵ^d_i is increasing in C_i*

Proof. See the Annex Note that if $q_i = 1$ (shirking workers always get severance pay), *severance payments are neutral.* From equation (11) it follows that ϵ^d_i is independent of T_i. This is a version of the Lazear (1990) neutrality result.

2.2 Efficient separation and optimal severance pay

Efficient separation in the last period n of the contract is defined simply as a productivity ϵ^*_n that ensures zero surplus, so that $S_n(\epsilon^*_n) = 0$. Given, ϵ^*_n, we can define the expected efficient surplus in the last period, S^{*e}_n. The surplus in the last period is simply $\epsilon_n - b$. [9]. We thus have that

[8] Or it may have multiple solutions, in which case the lowest solution is the relevant one since the firm chooses the lowest possible incentive compatible wage.

[9] To see this, recall that $S_n = W_n + \Pi_n - U = \epsilon_n + (1-\beta)U = \epsilon_n - b$

$$\epsilon_n^* = b \quad (12)$$
$$S_n^{*e} = \int_{\epsilon_n^*}^{\bar{\epsilon}} (z - \epsilon_n^*) dF(z)$$

For earlier periods, efficient firing and efficient surplus conditional on future efficient firing can be defined recursively. Given S_{i+1}^{*e}, the joint surplus in period i as a function of ϵ_i reads[10]

$$S_i^*(\epsilon) = \epsilon_i - b - C_i + \beta S_{i+1}^{e*}$$

Optimal firing in period i (defined as $S_i^*(\epsilon_i^*) = 0$) and S_i^{*e} requires then

$$\epsilon_i^* = b - C_i + \beta S_{i+1}^{e*} \quad (13)$$
$$S_i^{*e} = \int_{\epsilon_i^*}^{\bar{\epsilon}} (z - b - C_i + \beta S_{i+1}^{e*}) dF(z) \quad (14)$$

Note that, as we already noted, neither wages nor severance payments appear in the joint surplus, as they are transfers between the two parties. Equation (13) suggests that firing is efficient whenever the productivity from the job (ϵ_i^*) falls below the worker's outside option $b = (1-\beta)U$ augmented by the investment cost C_i, and the continuation surplus S_{i+1}^{*e} .

Definition 3 *The severance payment sequence* $\{T_i^*\}_{i=1}^n$ *is efficient if it is such that* $\{\epsilon_i^d = \epsilon_i^*\}_{i=1}^n$*, i.e.*

$$\epsilon_i^* = \epsilon_i^d(T_i^*) \quad (15)$$

Hence efficient severance payments is obtained when the reservation productivity $\epsilon_i^d(T_i^*)$ is identical to its efficient counterpart level. Before we derive the optimal severance, note that from equation (6), and (12) we have that

$$\epsilon_n^* - \epsilon_n^d = \frac{C_{n-1} - (1-q_n)\beta T_n}{\beta[1 - F(\epsilon_n)]}$$

The previous equation immediately suggests that at $T_n = 0$, $\epsilon_n^* > \epsilon_n^d$ so that firing is too much in the last period when severance payments are zero. As wages need to pay for the worker's investment effort in earlier periods, the firm has a tendency to over dismiss a worker who did not shirk, and invested in the previous period. To derive the optimal severance, we again do this recursively. Consider any period i and suppose that $T_j = T_j^*$ for all $j > i$ (if any), so that $S_j^e = S_j^{*e}$ for all $j > i$. From equation (11) and (13), it follows that efficient firing is obtained if and only if

$$\frac{C_{i-1} - T_i \beta (1-q_i)}{\beta[1 - F(\epsilon_i^d)]} = 0 \quad (16)$$

Or

$$T_i^* = \frac{C_{i-1}}{\beta(1-q_i)} \quad \forall i = 1, ...n \quad (17)$$

Recall that the rent $R_i = W_i - U$ that accrues to the worker is given by (3). By inserting T_i^* it follows imediately that the optimal severance is such that the worker is indifferent between being fired and being retained. In this case the firm's firing decision has no external effects on the worker.

The expression for the optimal severance turns out to be surprisingly robust. It holds with endogenous Court decisions (endogenous q_i) as shown in the next two sections. In addition, it holds in general equilibrium, as we show in section 4. If $q_i = 1$, the severance does not influence the firm's hiring decision, and is then useless as a policy tool for inducing optimal retention by the firm.

[10]Note that $S_i = W_i + \Pi_i - U = \epsilon_i - b + \beta E(W_{i+1} + \Pi_{i+1} - U) = \epsilon_i - C_i - b + \beta S_{i+1}^e$

Proposition 1 *For $q_i < 1$, the optimal severance T_i^* is given by (17). It is increasing in the investment cost in the previous period, and in the probability of getting away with it if shirking. It does not depend on investment costs, or the probability of being caught in any other periods.*

It follows that the severance is increasing with tenure if q_i is increasing with tenure or if C_i is increasing with tenure. Both seems reasonable, as discussed in Section 5.

Corollary 2 Upward Sloping Severance Payments
Suppose that q_i (C_i) is increasing in i and that C_i and (q_i) are not decreasing in tenure. Then the optimal severance payment is increasing with tenure

Proof The proof is a direct consequence of equation (17)

We want to point out the remarkable fact that optimal severance is independent of the distribution of ε_i. Optimal severance pay only depends on q_i, a property of the legal system, and C_{i-1}, the investment costs. It seems natural to assume that q_i is the same for all the firms in a country. The investment cost C_i is probably firm-specific. However, one may think that the average value of C_i may vary from country to country. Hence our theory predicts that countries with a high value of q_i (inefficient judicial system), and where workers tend to have high investment costs, the optimal severance pay is high. We return to this prediction in Section 4, and we evaluate its empirical relevance in Section 5. Note further that the wage tenure profile is steeper when the severance pay is higher.

Before we continue, we want to make a point regarding fiscal externalities. Although the optimal severance maximizes the joint income of the worker and the firm, there may still be fiscal externalities associated with firing, as unemployed workers receive unemployment benefits while employed workers pay taxes and firms payroll taxes. This fiscal externality implies that the private value of unemployment exceeds the social value, while the opposite is true for employment. One may think that this is an argument for increasing the severance pay even further. However, this needs not work. When it is privately optimal for the worker and the firm to separate, the optimal contract will prescribe a wage that is higher than w_i^{ξ} derived above, so that the (privately) efficient separation rates will be realized. In this case, the ICC constraint of the worker will not bind in periods $1, ..., n$, and the worker will compensate the firm for the slack through a lower wage in period 0. In this case, a more direct and effective policy tool will be a firing tax, paid by the firm to the Government, reflecting the fiscal externalities associated with firing the worker.

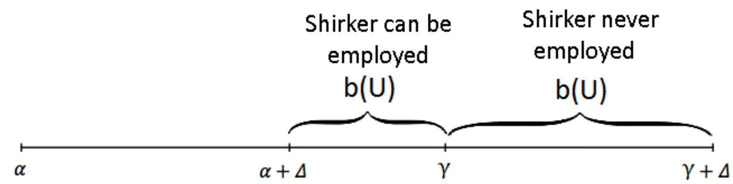
3 Burden of proof and endogenous Court rulings

In the previous section we assumed that a worker would get away with shirking with an exogenous probability, while the firm always had to pay the severance payment if laying off the worker, and hence by assumption never would get away with it. We also assumed that a shirking worker always would be fired, and we were silent about the productivity of a shirking worker. In this section we explicitly model the decisions made by workers and firms. We also dig deeper into the legal system and obtain endogenously the probability that a shirking worker gets away with it. We also define under which circumstances the firm may get away with it.

By opening up the black box of the legal system, we are able to show that the party that can get away with is the party that has the burden of proof. However, and somewhat surprisingly, the allocation of the burden of proof does not affect the optimal level of severance. Finally, we also link the value of q_i to the stochastic properties of the decisions made by the Courts. To keep things simple, we focus on a two-periods model, so that the relevant workers value functions are W_0 and W_1. Results can be readily generalized to a generic n-periods setting as well as to general equilibrium.

We assume that both shirker and non shirker workers draw a productivity level ϵ in the first period. Yet, the investment in period 0 on the part of the worker shifts the distribution of productivity by a factor Δ, which is common knowledge. Specifically, the distribution of productivity $\tilde{F}(\epsilon)$ in period 1 for a shirking worker is supposed to be uniformly distributed over $[\alpha, \gamma]$, while the distribution of a non-shirking worker $F(\epsilon)$ is uniformly distributed over $[\alpha + \Delta; \gamma + \Delta]$. Note that the two distributions are identical with the exception

Figure 1: Outside option of the worker and productivity: 2 cases



of a shifter Δ. To make the problem interesting, we assume that the support of the two distributions has an area of overlap (Figure 1):

$$\Delta < \gamma - \alpha \tag{18}$$

We also assume that the investment cost C_0 is sufficiently small so that the firm always wants to induce the worker to invest.

Whether a firing is economic or disciplinary is settled by a Court ruling. We consider a Court able to accurately and freely establish the actual draw of the productivity in period 1, but that can not directly observe whether the worker did invest in period 0. All what the Court can observe is the productivity ϵ that- at least for some range- can be consistent with both investments and shirking.[11] We return to the issue of Court monitoring in section 3.4.

Total surplus is $S_1(\epsilon) = \epsilon - b$. Efficient separation requires $\epsilon_1^* = b$. To simplify the exposition, we shall operate only around the efficient solution, so that the reservation productivity set by the firm will always be $\epsilon^d(T_1^*) = b$. Hence

$$T_1^* = w_1(T_1^*) - b \tag{19}$$

where w_1 is the wage in period 1 determined by the ICC. Recall that the worker's rent in period 1 is $W_1 - U = w_1 - b$. Equation (19) thus implies that the first-period rent is equal to the severance payment- $W_1 - U = T_1$- , a result that we already know from the previous section.

We will consider first the case where the burden of proof is on the employer, and subsequently model a case where the burden of proof is on the employee.[12]

3.1 Burden of proof on the employer

When the burden of proof is on the employer, the Court has a bias toward economic dismissals. In any case where there is a situation of doubt, the Court rules that the worker did invest and the severance payment is due. Using the fact that the worker rent in period 1 is $W_1 - U = T_1^*$, W_0 writes

$$W_0 = w_0 - C_0 + \beta[U + T_1^*]$$

A shirking worker will get severance if her productivity is above $\alpha + \Delta$, and may even be hired if $\gamma > b$. A worker is equally well off being retained and being fired with severance, the income in both cases is $U + T$. Hence, a shirking worker gets

[11]We retain the assumption that the wage cannot be made contingent on productivity. If they were, the Courts would have to intervene in each period in order to ensure that the wage contract was honored, and this may be prohibitively costly. In principle, when the Court can observe a noisy signal of the worker's productivity, the employer may offer a wage contract that depends on this noisy signal. However, even in this case there is a difference between letting the Court decide on wages and on severance pay in the case of firing, as wages are set every period, while firing takes place less frequently. Going to court every period to receive the signal may be prohibitively costly. That being said, the threat from the worker of going to court may discipline the firm, and make it willing to offer higher wages when output is high. Analyzing this more complicated game is on our agenda for future research.

[12]The case where $\alpha < b < (\alpha + \Delta)$ is not interesting. In such an environment, an investing worker is never fired, hence severance is never paid in equilibrium

$$\tilde{W}_0 = w_0 + \beta\left[\tilde{F}(\alpha+\Delta)U + (1-\tilde{F}(\alpha+\Delta))(U+T_1^*)\right]$$

Since $\tilde{F}(\alpha+\Delta) = \frac{\Delta}{\gamma-\alpha}$, the ICC ($W_0 = \tilde{W}_0$) implies that

$$T_1^* = \frac{C_0}{\beta F(\alpha+\Delta)} = \frac{C_0(\gamma-\alpha)}{\beta\Delta} \tag{20}$$

The shirker gets away with it with probability $q^e = 1 - F(\alpha+\Delta)$, hence the expression corresponds to our earlier results, with $q_1 = q_1^e$ given by

$$q_1^e = 1 - \frac{\Delta}{\gamma-\alpha}$$

where the topscript e indicates that q_1 is endogenous.

To complete the model, and for comparison with the specification on the following section, the period zero wage satisfies the participation constraint and solves $W_0 \geq U$ so that

$$w_0 = b + C_0 - \beta T_1^* \tag{21}$$

where it is clear that the severance payment is prepaid by the worker in period 0.

3.2 Burden of proof on the worker

When the burden of proof is on the worker, the behavior of the Court is different. Whenever there is a situation of doubt, the Court rules that the worker did non invest, and the severance payment is not due. A shirker never gets severance, while an investing worker who loses her job gets severance if she is distinguishable from a shirker. Hence, it is now the firm that can get away with it by firing a worker for economic reasons without paying severance, i.e., carry out an *unfair economic dismissal.* This will be discussed further in the next section.The same two cases depicted in Figure 1 apply in this context.

Never efficient to retain shirker: $\gamma < b$

An investing worker will be employed if her productivity exceeds the period 1 wage, and severed if her productivity is below the wage, but above γ, so that she is distinguishable from a shirker. Hence, she will get severance or wage if her productivity exceeds γ, and the probability of this event is $1 - F(\gamma) = \Delta/(\gamma-\alpha) = 1 - q_1^e$ (recall that $q_1 = 1 - \Delta/(\gamma-\alpha)$). This is not surprising: q_1^e is the probability that the investor's productivity falls within the support of the shirker's productivity, hence $1 - q_1^e$ is the probability that the investor's productivity falls outside of it, in which case she gets the wage or the severance. We assume that the severance, when paid, is still equal to the difference between the inside and outside payment, $T_1 = w_1 - b$.

The shirker is never employed in period 1 nor receiving severance. Hence her value function reads

$$\tilde{W}_0 = w_0 + \beta U$$

The corresponding value for a non-shirker is

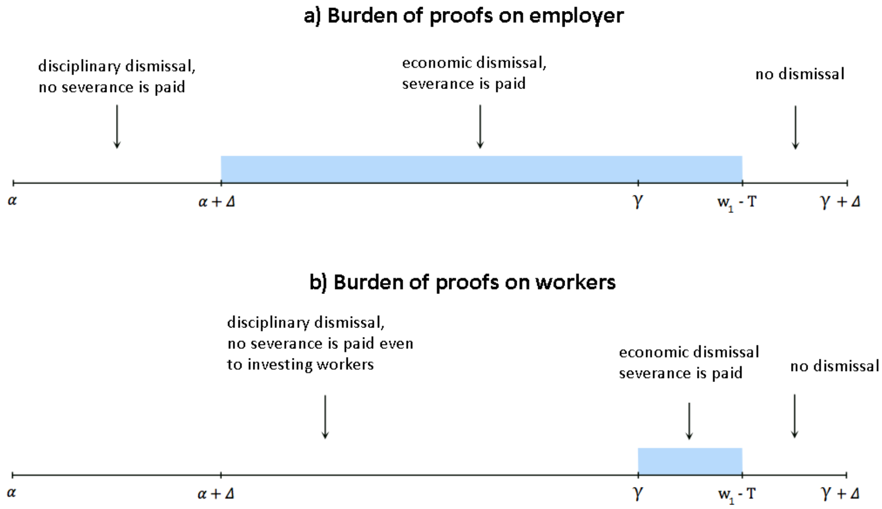
$$W_0 = w_0 - C_0 + \beta[U + (1-F(\gamma))T_1^*]$$

Since $1 - F(\gamma) = q_1^e$, it follows that we can write

$$T_1^* = \frac{C_0}{\beta(1-q_1^e)} \tag{22}$$

Hence, even in this case, the result from the model in section 2 still applies. To give intuition, note that the fact that the investing worker does not get severance whens he is indistinguishable from the shirking worker reduces the pay-off of investing with $q_1 T_1^*$. However, the fact that the shirking worker never gets away with

Figure 2: Court decisions
Case where $\gamma < b(U) < \gamma + \Delta$



it reduces the value of shirking with $q_1 T_1^*$. Hence the two effects cancel out, and the ICC wage and the corresponding severance do not change. However, the reduced expected period 1 income of the investing worker is accompanied by an equally large increase in the period 0 wage so that the total value of the job remains constant. The participation constraint reads

$$W_0 = w_0 - C_0 + \beta[U + (1 - F(\gamma))T_1^* = U$$

so that the first period wage is

$$w_0 = b + C_0 - \beta T_1^* + \beta F(\gamma) T_1^*$$

This finding suggests that the wage tenure profile is flatter when the burden of proof is on the worker. Note also that the firing is still efficient. This follows from the assumption that b is outside the support of the productivity of the shirker. Hence, around the efficient firing threshold, the firm has to pay severance, and therefore makes the efficient decision. Investing workers with a lower productivity may not receive severance, but these workers would be fired anyway.

May be efficient to retain worker: $b < \gamma$

In this case, the worker can fire a worker with productivity b and get away with it. Since the ICC requires that $w_1^* > b$, it follows that it is impossible to use severance to induce the firm to retain a worker with productivity in the interval $[b, \min[w_1^*, \gamma]]$.

Note that if $\gamma < w_1^*$, shirking workers will never be retained, while investing workers will be retained or severed if and only if $\epsilon \geq \gamma$. The probability of this event is $\frac{\Delta}{\gamma - \alpha} = 1 - q_1^e$. The gain from investing is thus $(w_1^* - b)(1 - q_1^e)$. If, on the contrary, $\gamma > w_1^*$, shirking workers are hired with positive probability, and severance is never paid. The difference in the probability of receiving a wage between the shirker and the non-shirker is $\frac{\Delta}{\gamma - \alpha} = 1 - q_1^e$. Thus, the gain from investing is again $(w_1^* - b)(1 - q_1^e)$. Hence, in both cases the ICC wage again writes

$$w_1^* = b + \beta \frac{C_0}{1 - q_1^e} \tag{23}$$

With $b < \gamma$, the Government may use other instruments to achieve efficient separations. By imposing a firing tax equal to the difference between the period 1 wage and the outside option, efficient firing is restored. For instance, if $\gamma > w_1^*$, a firing tax of $w_1^* - b$ will restore efficiency along the firing margin. However, a firing tax will, in contrast with severance, reduce the profitability of the firm, as the firing tax cannot be offset by a lower wage in the first period. Hence, a firing tax will lead to inefficient entry.

3.3 Double moral hazard and unfair economic dismissals

When the burden of proof is on the worker, the firm has an incentive to declare that a productivity level below $b = (1-\beta)U$ is always the results of a non-investment on the part of the worker. In this section we still consider the case in which $b > \gamma$ so that shirking workers are never employed, but introduce the possibility that the Court monitors the declared productivity level of the firm.

To tackle the issue in a formal way, let us assume that we are in the efficient solution of equation (22), and the firm wants to fire the worker, so that the productivity is below b. Let $\hat{x}$ denote the declared productivity and x be the true productivity. Assume further that the Court can audit the firm with probability λ and impose a severance payment T_{U1}^*. In other words, T_{U1}^* is the severance payment for an *unfair economic dismissal*. If the firm reports that $\hat{x}$ is below β, even though it is above it, its expected profits read

$$\hat{\Pi}(\hat{x} < \gamma| \ \ \gamma < x < b) = -\lambda T_{U1}^*$$

Conversely, truth telling for the firm yields

$$\Pi(\hat{x} = x| \ \ \gamma < x < b = -T_1^*$$

The incentive compatibility constraint for the firm requires that $\hat{\Pi} = \Pi$ so that

$$T_{U1}^* = \frac{T_1^*}{\lambda} \tag{24}$$

where T_{U1}^* is the severance in case of unfair dismissal. This means that severance payments for unfair economic dismissals should be higher than severance payments for economic reasons, a property that holds in all countries (see Table 1). Furthermore, we expect T_{U1}^* and $T_1^* = T_E^F$ to be positively correlated, another property which is in the data (the cross-country correlation of T_U and T_E^F is .6 which is statistically significant at conventional levels). More interestingly, combining the above condition with the no-shirking (and participation) condition for the worker, we have that

$$T_{U1}^* = \frac{C_0}{\beta\lambda(1-q_1^e)} \tag{25}$$

which establishes a relationship between optimal severance and efficiency of the judicial system, here defined more precisely in terms of the audit technology. A higher "λ" implies a more efficient legal system and a lower severance payment.

Proposition 3 *In a more efficient legal system, severance payments for unfair dismissals are lower.*

4 Severance pay in general equilibrium

In partial equilibrium, the worker and the firm take the worker's outside option U as exogeneous. In general equilibrium, U is determined endogenously. In this section, we close the model by embedding the partial equilibrium contracting problem into a Diamond-Mortensen-Pissarides framework with directed search. Then we calibrate severance pay and numerically solve the model, looking at the welfare and unemployment effects of suboptimal severance.

We follow the competitive search approach á-la-Moen, and assume that firms post period 0 rents to maximize the value of a vacancy. The matching function is a standard Cobb-Douglas with constant returns to scale. There is a cost K of opening a vacancy. The vacancy operates until it is filled. However, if a worker and an employee separate, the vacancy is lost. In line with the main literature, we shall indicate with θ the aggregate vacancy unemployment ratio and with $q(\theta)$ the probability of finding a worker. The value of a firm with a vacancy can be written as

$$V = -c + \beta \left[q(\theta)\Pi_o^e + (1 - q(\theta))V \right] \tag{26}$$

where c is the flow cost of maintaining a vacancy. Let z denote the pure flow utility value of unemployment and $\theta q(\theta)$ is the worker's probability of finding a new job. The net present income of an unemployed worker reads

$$\begin{aligned} U &= z + \beta \left[\theta q(\theta) W_o + (1 - \theta q(\theta) U) \right] \\ &= \frac{z + \theta q(\theta) R_0}{1 - \beta} \end{aligned} \tag{27}$$

where $R_0 \equiv W_0 - U$ is the *search rent* associated with employment. Firms advertize wage contracts, and workers can figure out the expected discounted income W_0 (or equivalently, the rent R_0) associated with the contract. For workers, W_0 is a sufficient statistics for the attractiveness of the contract.

The model's general equilibrium thus features an imperfect labor market with a search rent. This has to be distinguished from the information rent that the worker receives in later periods. We basically, need to be precise on the worker participation constraint and the worker's rent from the contract. The previous section showed that the worker will get an information rent $R_i = W_i - U$ from period $i = 1$ until period $i = n$. At the time of job creation in period 0, within an imperfect labor market, the firm will offer the worker a search rent R_0. The following lemma clarifies the relationship between R_0 and the optimal contract.

Lemma 3 *Suppose that the firm promises the worker an expected income $W_0 > U$, or equivalently a search rent $R_0 > 0$. Then the optimal contract Ω^R is identical to the contract Ω^ξ for $w_1, ...w_n$, while $w_0^R = w_0^\xi + R_0$.*

The proof is given in the Annex.

Definition 4 General Equilibrium
General equilibrium is a wage contract Ω^g, a npv income U^g for unemployed workers, and a search rent R_0^g for a worker becoming employed such that

1. *$w_i^g = w_i^\xi$ for all $i \geq 1$ and $w_0^g = w_0^\xi + R_0^g$.*
2. *The search rent R_0^g maximizes V given by (26) subject to (27) for $U = U^g$.*
3. *Zero profits, $V = K$.*

Total surplus at the time of job creation is $S_0 = \Pi_0 + R_0 - K$. In the Annex we show that the equilibrium can be characterized by the following equations:

$$S_0 = y - C_0 + \bar{\epsilon} + \beta[F(\epsilon_1)U + (1 - F(\epsilon_1))W_1 + \Pi_1^e] - U - K \tag{28}$$

$$K = \frac{-c + \beta q(\theta)(1 - \eta(\theta))S_0}{1 - \beta} \tag{29}$$

$$b = z + \beta \left[\theta q(\theta) \eta(\theta) S_0 \right] \tag{30}$$

To sum up, equations (29) and (30) (28) solve for S_o, U and θ Furthermore, we know from the last section that the separation decision within each firm is efficient, and that search frictions do not create inefficiencies in competitive search equilibrium.[13] The following proposition thus follows:

[13]We have not given general conditions for existence of equilibrium, as we consider this outside the scope of this paper. One can show that the equilibrium exists and is unique if $E\epsilon$ is sufficiently high, if $\bar{\epsilon}_0 < b$, and if the sequence C_i is non-increasing or not increasing too rapidly. Note also that existence for specific parameters are implicitly shown through the simulations.

Proposition 4 Efficient General Equilibrium
Suppose that the sequence of severance pay is given by a set of T_i that satisfies equation (17). Suppose further that z reflects both the private and the social flow value of being unemployed. Then the general equilibrium allocation is efficient.

Proof: This is just an application of Moen (1997) with endogenous destruction.

In order to find the unemployment rate, we have to characterize the separation rate. This is conceptually simple but somewhat tedious. Let N_t be the stock of workers employed in firms with tenure t. Then the separation rate in that period is $F(\epsilon_t^*)$, where ϵ_t^* is given by equation (13). Furthermore, $N_t = (1 - F(\epsilon_{t-1}^*))N_{t-1}$. Hence, the average separation rate $\bar{s}$ in steady state reads

$$\bar{s} = \sum_{t=1}^{n+1} F(\epsilon_t^*)\Pi_{j=1}^{t}(1 - F(\epsilon_j)) \tag{31}$$

with $F(\epsilon_{n+1}^*) \equiv 1$. The unemployment rate is then $u = \bar{s}/(\bar{s} + p)$.

Numerical simulations

This section performs a set of simulations of the general equilibrium of the model. The basic idea is to provide a numerical sense of the results of the previous section. We perform three exercises. The first one is to solve numerically the general equilibrium with optimal severance payments. In order to get some discipline on our choice of parameters, we assume that the observed severance levels are indeed optimal. We thus calibrate values for the optimal severance payments using the evidence provided in Table 1, and provide values for a "high" and "medium" severance pay economies. Finally, we ask what are the effects of removing severance pay in the two economies. The simulations show that with reasonable parameters the welfare loss varies between 1 percent in the "medium" severance pay economy to a 3 percent loss in the "high" severance pay economy.

The time period is one year and the discount rate is 0.98. The time period of the labor contract is 20 years, so as to be coherent with the severance pay values available in Section 1, and in the Annex. The rest of the labor market values are reported in Table 2. The most important parameters of the exercise of this section are the sequence $\{C_i\}$ and $\{q_i\}$. The logic of the calibration of $\{T_i\}$ is the following. In the top part of Table 3 we report values for the "high" severance pay economy in terms of monthly wages, in a way coherent with the evidence provided.[14] With respect to the severance pay reported in Table 1, we aim at capturing the size of T_E^F, i.e. the pay for a fair economic dismissal. Given the productivity distribution specified in Table 2, we calculate the average productivity of our economy, and transform the value of the monthy wage in terms of monthly productivity.[15] These average productivity levels are reported in the column labeled *Data* in Table 3 and are the values that we are trying to match.[16] Note that in the column *Data* severance payments increase with tenure.

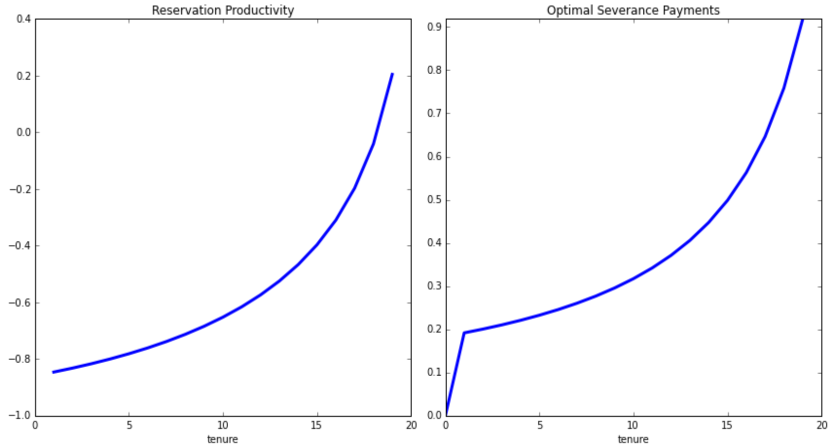
The logic of the choice of $\{C_i\}$ and $\{q_i\}$ to back out $\{T_i\}$ and matching the data is the following. We work with a time invariant value of C_i, so as to approximately match the average level of the severance payments. We then use the sequence q_i to match the slope. While our calibration is not perfect, results in Table 3 show that we are able to obtain reasonable values for both the level and the steepness of severance pay. The overall fitting of the ''high" severance pay economy is better than the corresponding fitting of the "medium" economy. Figure 3 plots the optimal severance payment for the "high" economy as well as the reservation productivity, with respect to tenure. The severance payments has an initial spike in the early years, and thereafter grows smoothly with tenure, driven by the sequence of $\{q_i\}$. From Figure 3, also the

[14] The highest severance pay economies in Table 1 are Turkey and Portugal. The medium economy has a severance pay level coherent with that of Germany

[15] The average productivity level is $\overline{y} = \sum_i^n \frac{\int_{\epsilon_i^d}^{\overline{\epsilon}} zf(z)}{1-F(\epsilon_i^d)}$, where the summation is over the tenure periods

[16] Consider, for example, the 5 monthly wage at five years in the 'high' severance economy reported in Table 3. Since the average yearly effective productivity in the model is 0.6, 5 months of productivity correspond to the value of 0.25 reported in Table 3

Figure 3: Severance Payments by tenure in the High Severance Payment Economy



reservation productivity increases with tenure. This is the well known labor hoarding effect, and reflects the option value of delaying dismissal typical of dynamic models of labor demand.

The policy experiment is reported in Table 4. Starting from the parameters specified in Table 2, we change only the level of severance payments. Specifically, the severance payments move from the value reported in Table 3 to zero. In Table 4 we report the equilibrium unemployment, and market tightness. In addition, the table reports the equilibrium value of unemployment U, a natural measure of welfare in our competitive search economy. The exercise shows that the equilibrium unemployment rises and market tightness falls. The key question is the effect on welfare of removing severance payments. The table shows that the welfare loss is approximately 1 percent in the case of the ''medium" severance economy and close to 3 percent in the ''high" severance economy.

Table 2: Baseline Parameters in two Economies

Parameter	Notation	High T_i	Medium T_i
Pure Discount Rate	β	0.980	0.980
Stochastic productivity: mean	$\overline{\epsilon}$	0.400	0.400
Stochastic productivity: variance.	$\sigma\{\epsilon\}$	0.820	0.650
unemployed income	z	0.160	0.100
matching function elasticity	α	0.400	0.400
matching function parameter	A	1.000	1.000
search cost parameter	c	0.100	0.100
entry cost	k	0.900	1.770

Source: Authors' calculation

Table 3: Calibrating Severance Payments

High Severance Pay Economy			
Tenure/Years	Monthly Wage	Data [a]	Model [b]
5 years	5	0.251	0.221
10 years	10	0.503	0.296
20 years	20	1.006	0.918
Training Technology	Notation		Model
Training Costs:	C_0		0.180
Prob. of Getting away with it: min	q_0		0.000
Prob. of Getting away with it: max	q_n		0.800
Medium Severance Pay Economy			
Tenure/Years	Monthly Wage	Data [a]	Model [b]
5 years	2	0.077	0.134
10 years	5	0.193	0.203
20 years	10	0.386	0.267
Training Technology	Notation		Model
Training Costs:	C_0		0.131
Prob. of Getting away with it: min	q_0		0.200
Prob. of Getting away with it: max	q_n		0.500

[a] Monthly productivity in the model corresponding to $\{T_i\}$ in terms of monthly wage in the data.
[b] Optimal severance pay in the model in terms of monthly productivity

Table 4: Effects of Suboptimal Severance Payments

Sev. Payments	Welfare U	Unemp. u	Tightness θ	Welfare Loss [a]
High optimal:	10.238	0.066	0.558	
High no SP:	9.958	0.071	0.488	-0.027
Medium optimal:	7.413	0.069	0.524	
Medium no SP:	7.338	0.070	0.508	-0.010

[a] Percentage changes in unemployment value U from reducing severance pay to zero.

5 Discussion

In our theory, severance is needed to deter opportunistic behavior of workers. It has to be mandated by Governments as adverse selection prevents individual employers from committing not to fire workers investing in the productivity of the job. Incentive reasons, notably deterrence of shirking ([44]), also explain why severance for economic dismissals is higher than for disciplinary dismissals. At the same time, this difference, especially when at least part of the burden of proof falls on the worker, induces employers to play strategically. Severance in case of unfair dismissals should be set at even higher levels to deter firms from taking the disciplinary dismissal route in case of dismissals that are actually motivated on purely exogenous productivity grounds. These differences in severance pay levels by nature of individual dismissals, and the associated informational asymmetries enhance the discretion of judges, hence the unpredictability of the costs of dismissals stressed by many employers (see the quotes at the beginning of this paper).

Our model has a number of testable assumptions and implications as to the time profile of severance and the relationship between T_U, T_E^F and the efficiency of the legal system.

5.1 Severance-tenure and wage-tenure profiles

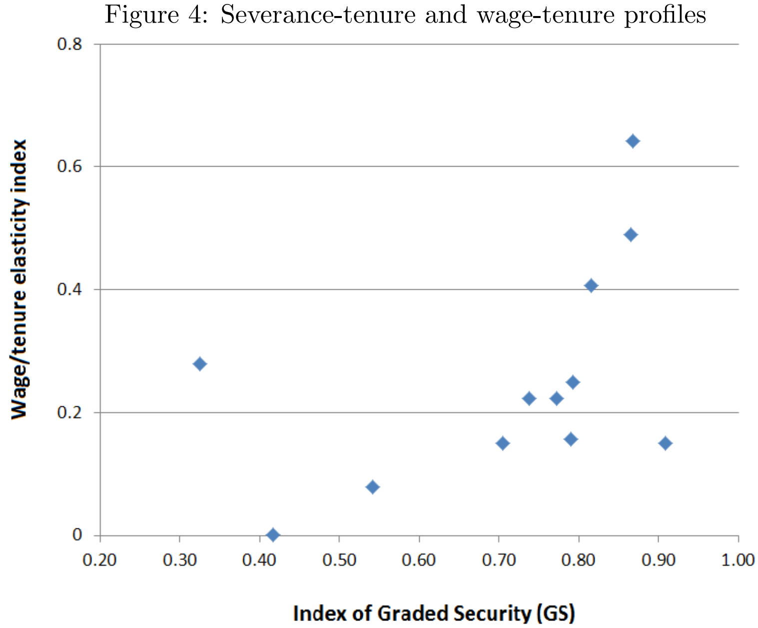
Graded security is optimal according to our model when investment costs are increasing with age and when the probability of getting away with it is not decreasing in tenure. There is ample evidence that employees' self-assessed working capacity is declining with age [23]. This suggests that investment in maintaining working capacity, e.g. training, is perceived by the workers to be increasingly costly with tenure.

When wages are deferred and training costs are increasing with tenure, also wages should be increasing with tenure according to our model. Thus, indications as to its empirical relevance may come from the correlation[17] between the severance-tenure and the wage-tenure profiles. Figure 4 displays the apparent elasticities of severance and wages with respect to tenure in all countries for which data are available. In particular, we recover the severance-tenure elasticities from the legal rules as to the mandatory notice period and redundancy payment in the case of fair economic dismissals in the different countries. Cross-country comparable data on EPL specify the level of severance at discrete tenure intervals. Based on this information, we could compute apparent elasticities at different tenure lengths and then aggregate them in the GS index presented in the Annex. For each country of the European Community Household Panel (ECHP), which is particularly suited to this as it is a long panel allowing to identify separately age and tenure, we estimated the following augmented Mincer-type wage equation against micro data on workers' earnings

$$\ln w_i = \alpha + \beta_1 \tau_i + \beta_2 \tau_i^2 + X_i \gamma + \epsilon_i \qquad (32)$$

where w denotes hourly wages, τ years of tenure, and the vector X includes educational attainment dummies (tertiary and secondary education), age and gender. The correlation is positive and statistically significant at conventional (90 per cent) levels. We also looked at the within country correlation as some countries have different rules for the level of severance for blue-collar and white-collar workers. This is the case of Austria, Denmark and Greece. In all of these cases, white collar workers command steeper severance-tenure profiles than blue-collar workers, and the same pattern is observed for wage-tenure profiles, perhaps because of the greater role played by training in the case of white collar workers.

Figure 4: Severance-tenure and wage-tenure profiles



Note: The Index of Graded Security is a weighted average of apparent elasticities of severance to tenure at different tenure lengths. The Wage/tenure elasticity index is computed in the same way, and at the same durations, based on the empirical estimates of the wage-tenure profile. See the Annex for details.

Our model also implies that severance should be increasing with tenure when q is higher. We are not aware of data on Court rulings by tenure of workers. There is some empirical support to the view that judges internalize the re-employment probabilities of workers being laid-off: both the percentage of cases being brought to Courts and the fraction of labor disputes ending with a Court ruling favourable to the worker appear to be higher during cyclical downturns and in relatively depressed labor markets [7], an indication that judges are more protective of workers under these circumstances. Insofar as senior workers

[17]No causality is involved here as both severance and wages depend on the costs of training.

face lower re-employment probabilities than junior workers, the legal system may turn out to allow for a greater probability of getting away with it as tenure progresses.

The case of the Netherlands is particularly interesting in evaluating the implications of the model as to the relationship between the severance-tenure profile and the probability of getting away with it. All firms in the Dutch labor market have to seek administrative authorisation to layoff a worker, and can follow two alternative routes in obtaining this permission to fire. The first route is represented by the request of an authorisation to the Public Employment Service (PES). This involves a relatively long procedure, but, in case the PES acknowledges the fairness of the dismissal, the severance pay is not due. Hence, the costs for the employer consist only of the notice period and the wage paid to the worker during the procedure.

The obvious question is why employers do ever go to the Court given that it is always more expensive [39]. The reason is that a PES procedure may be brought to a Court by one of the two parties. Thus, in all cases where q is high, employers prefer to go directly to Courts in order to avoid to pay for the PES procedure plus the Court one. Typically it is indeed small firms (better monitoring, low q) to go for the PES route, while large firms fill for the Court ruling. Incidentally our model also provides an explanation as to why in most countries severance pay is lower for small firms than for large firms: the presence of threshold scales below which the strictest employment protection provisions do not apply can be rationalized in terms of the better monitoring of workers in small organizations.

5.2 Severance and the legal system

Our model has predictions about the relationship between employment protection and the efficiency of the legal system. In particular, it suggests that we should expect to observe higher levels of severance in the countries where the judicial system is less efficient and where the burden of proof is entirely on the employer. Previous work had found that the organization of legal systems, notably the legal origin of countries played an important role in labor market outcomes [12]. Our model can provide an indirect explanation for this, which is based on the effects of the legal system on employment protection regulation. Moreover, we can directly evaluate the empirical relevance of the link between severance and the judicial system implied by the model. Based on recent work done by the OECD in creating cross-country comparable data on legal systems, we can indeed analyse the cross-country correlation between, on the one hand, compensation for fair and unfair dismissals, and, on the other hand, characteristics of judicial systems.

The first two panels at the top of the Figure 5 display the correlation between, on the one hand, T_F^E, and, on the other hand, the litigation rate, that is, the number of the new civil cases commenced in any given year normalized by the population or GDP. This indicator captures congestion, and, per given supply of services, a longer duration and lower quality of judicial services. We allow for separate intercepts for countries where the burden of proof is entirely on the employer, and for countries where it is also partly on the worker. The litigation rate, normalized either by population or GDP, is positively correlated with the mandated months of severance in case of fair economic dismissals. Furthermore, as suggested by the top diagram on the right-hand-side of Figure 5, the level of compensation for fair economic dismissals is higher in countries where the burden of proof is entirely on the employer than in countries where it also falls on the worker (all the countries of this second group are below the regression line for the other group and the intercept is statistically significant at conventional levels).

The other four panels of Figure 5 look at the compensation in case of unfair dismissals (T_U) as well as to a broader measure of the compensation to employees in the case of fair and unfair dismissals. They show that T_U is positively correlated with the litigation rate. Alternative measures of the efficiency of legal systems are trial lenght, and appeal rates before the second instance or higher Courts. These indicators of the efficiency of the judiciary are positively correlated with subjective evaluations of public opinion as to the quality of the legal system collected within the World Value Survey (Palumbo 2013 [38]). We find that severance in case of unfair dismissals, T_U, is strongly and positively correlated with trial length, as well as appeal rates before the second instance or higher Courts, but it can be a spurious correlation as appeal rates and trial length appear in our measure of the costs of unfair dismissals. Thus, in the remaining three diagrams we consider a global measure of the compensation for fair and unfair dismissals from regular contracts produced by the OECD, which does not draw on information on trial length and appeal rates. The correlation is once more positive and statistically significant.

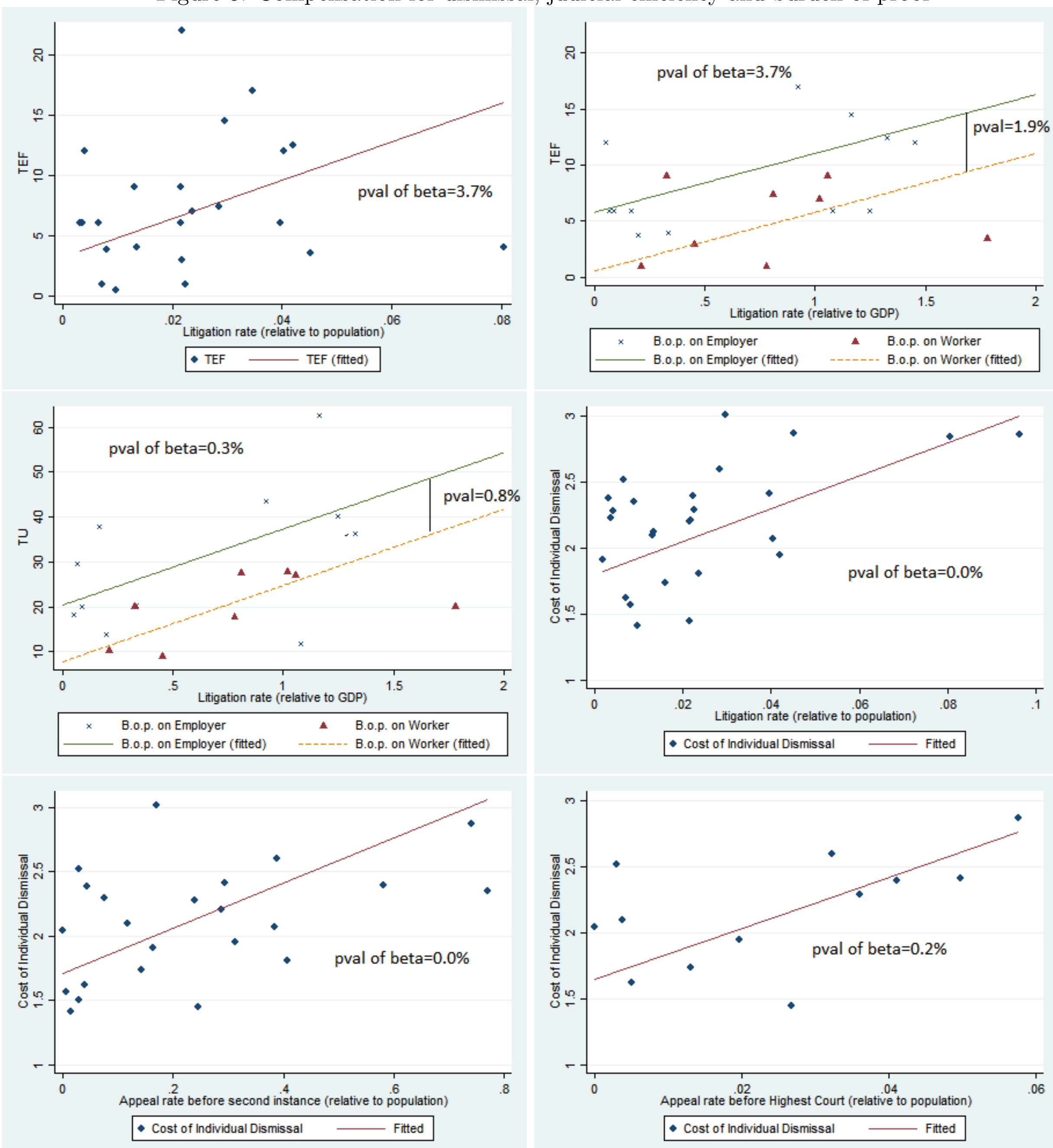
6 Final remarks

Research on employment protection fails to account for the relevance of mandatory severance pay in OECD countries. It also neglects two critical features of EPL: the tenure profile of severance pay and the fact that dismissal costs are not only stochastic, but also vary depending on whether they are motivated by economic or disciplinary reasons. In this paper we provide a normative theory of tenure-related severance pay which draws on the involvement of third parties in the decision about the nature, fair or unfair as well as disciplinary or economic, of dismissals. In our model severance pay has to be mandated by the Government (or industry-level collective bargaining) rather than being provided by the individual firm. This is because adverse selection stands on the way of these voluntary arrangements, potentially attracting more shirkers to the firm unilaterally offering a severance scheme. In other words, mandatory severance acts as a coordination device across firms.

We show that under a rather broad set of circumstances, and without having to impose any restriction on the distribution of productivity shocks, a severance scheme which is increasing in firm-specific investment costs and in the inefficiency of the legal system is privately efficient in that it avoids separations of jobs that are still originating a positive surplus. This result, which is new for the literature on employment protection, is in line with the reported correlation between, on the one hand, mandatory severance pay, and, on the other hand, OECD indicators of the inefficiency of the legal systems. It implies that reforms of the judiciary can be more effective than labor market reforms in reducing the level of employment protection. We also find empirical support for the key rationalization provided by the paper for a positive tenure profile of severance pay, that is, for the fact that investment costs or the probability of getting away with it are increasing with tenure.

The results of this paper are important in evaluating proposals to introduce mandatory compensation increasing steadily with tenure in countries characterized by "contractual dualism", that is, the coexistence of a highly protected segment of the workforce and one segregated into temporary jobs providing low, if any, employment protection. It is also informative as to the optimal slope of the severance tenure profile, depending on the way in which Courts typically protect senior workers and on the costs of training for older workers. Our theory is therefore particularly useful in assessing the scope for "insertion contracts", involving mandatory compensation increasing steadily with tenure. Such "unifying" contracts have been advocated in a number of countries as a measure to reduce "contractual dualism". The theory presented can certainly be used to assess and evaluate some of the reforms currently undertaken or under discussion in various Southern European parliaments. Moreover our theory suggests that tenure-related severance is efficient even under the typical conditions faced by "temporary workers", that is, under flexible wages, provided that agreed compensation is deferred and that the employer cannot commit not to layoff the worker who has invested in training.

Figure 5: Compensation for dismissal, judicial efficiency and burden of proof



Sources: T_F^E and T_U are as in Table 1.
Cost of Individual Dismissal: Index of Compensation in case of individual dismissal produced by the OECD.
Data on Litigation rate, trial length, appeal rate and burden of proof from Palumbo [38].

References

[1] Alvarez, F. and Veracierto, M. 2000. Labor Market Policies in an Equilibrium Search Model, *NBER Macroeconomics Annual*, 14, 265-304.

[2] Barth, E. 1997, Firm-Specific Seniority and Wages, *Journal of Labor Economics*, 15 495-506.

[3] Baumann, F. 2010. Severance Payments as a Commitment Device, *Journal of Institutional and Theoretical Economics*, 166, 715-734.

[4] Bentolila, S., and Bertola, G. 1990. Firing Costs and Labor Demand: How Bad is Eurosclerosis?, Review of Economics and Statistics,57, 381-402.

[5] Bentolila, S. 1997. Las decisiones judiciales sobre despidos en Espaa: una perspectiva econmica, Cuadernos de Informacin Econmica, 128-129.

[6] Bertola, G. 1990. Flexibility, Investment and Growth. CEPR Discussion Papers 422.

[7] Bertola, G., Boeri, T., and S. Cazes. 2000. Employment protection in industrialized countries: The case for new indicators. *International Labour Review*, 139(1):57-72.

[8] Blanchard O.J. and Tirole J. 2008. The Joint Design of Unemployment Insurance and Employment Protection: A First Pass. *Journal of the European Economic Association*, MIT Press, vol. 6(1), pages 45-77, 03.

[9] Boeri, T. 1999. Enforcement of employment security regulations, on-the-job search and unemployment duration, European Economic Review, 43, 65-89.

[10] Boeri, T. and vanOurs, J. 2013. The Economics of Imperfect Labor Markets, Princeton University Press.

[11] Boeri, T. and Jimeno, J. 2005. The effects of employment protection: Learning from variable enforcement, *European Economic Review*, 49 (8), 2057-77.

[12] Botero, S, Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. 2004. The Regulation of Labor, *Quarterly Journal of Economics 119*, 1339-1382.

[13] Carmichael, H.L. 1988. Incentives in Academics: Why Is There Tenure?, *Journal of Political Economy*, 96(3): 453-72, June

[14] European Labour Law Network. 2011. Dismissal - particularly for business reasons - and Employment Protection. Thematic Report, Contract No. VC/2010/1636.

[15] Fella, G. 2000. Efficiency wage and efficient redundancy pay, *European Economic Review*, 44, 1473-1490.

[16] Fella, G. 2012. Matching, wage rigidities and efficient severance pay, *The B.E. Journal of Economic Analysis and Policy* 12.

[17] Fraisse, H, Kramarz, F. and Prost, C. 2009. Labor Court Inputs, Judicial Cases Outcomes and Labor Flows: Identifying Real EPL. Working Paper Banque de France, n.404.

[18] Galdon-Sanchez, E. J. and M. Guell. 2003. Dismissal conflicts and unemployment. *European Economic Review*, Elsevier, 47(2):323–335.

[19] Garibaldi, P. 1998. Job flow dynamics and firing restrictions. *European Economic Review* 42, No. 2: 245-75.

[20] Garibaldi, P. and Violante G.L. 2005, The Employment Effects of Severance Payments with Wage Rigidities. *The Economic Journal*, 115: 799832

[21] Guerrieri, V., Shimer, R., and Wright, R. (2010). Adverse selection in competitive search equilibrium. Econometrica, 78(6), 1823-1862.

[22] Flabbi,L. and A.Ichino. 2001. Productivity, seniority and wages: new evidence from personnel data, *Labour*, Elsevier, vol 8(3), pp 359-387.

[23] Ilmakunnas, P., vanOurs, J. and V., Skirbekk (2010), Age and Productivity, in Garibaldi, P., Oliveira-Martins, J. and vanOurs, J. (eds.) *Ageing, Health and Productivity*, Oxford University Press.

[24] Kodrzycki, Y. K. (1998), Effects of employer-provided severance benefits on reemployment outcomes,*New England Economic Review*, pp.41-68.

[25] Kotlikoff, L. and Jagadeesh G. 1992. Estimating a Firm's Age-Productivity Profile Using the Present Value of Workers' Earnings. *Quarterly Journal of Economics*, 107, 1215-1242.

[26] Lazear, E.P. and Moore R.L. 1994. Incentives, Productivity and Labor Contracts. *Quarterly Journal of Economics*, 99, 275-296.

[27] Lazear, E.P. 1990. Job security provisions and employment. *Quarterly Journal of Economics*, 105:699–726.

[28] Malo, M.A. 1998. Las indemnizaciones por despido, un problema de negociacin. Editorial ACARL, Madrid.

[29] Malo, M.A., and Toharia, L. 1999. Costes de despido y creacin de empleo. Ministerio de Trabajo y Asuntos Sociales, Madrid

[30] Malo, M.A. 2000. A simple model of severance pay determination: The case of individual dismissals in Spain. *Labour*, 14(2):269–290.

[31] Medoff, J.L. and Katharine G. A. 1980, Experience, Performance and Earnings, *Quarterly Journal of Economics*, 95, 186-216.

[32] Menzio, G., and Moen, E.R. 2010. "Worker Replacement". *Journal of Monetary Economics*, 57, 623-636

[33] Moen, E.R. (1997). "Competitive Search Equilibrium". *Journal of Political Economy* 1997, vol 105, No.2, 385-411.

[34] Moen, E. R. and Rosen, Å. 2005. Incentive Contracts and Adverse Selection, *Scandinavian Journal of Economics*, 107(2), 279 - 298.

[35] Moen, E.R. and Rosen, Å(2011). Incentives in Compettitve Search Equilibrium (with sa Rosen). *Review of Economic Studies* 2011 vol 78 (2) 733-761.

[36] Mortensen, D., and C. Pissarides, 1994, Job Creation and Job Destruction in the Theory of Unemployment. *Review of Economic Studies*

[37] Mortensen, D., and C. Pissarides, 1999, New development in models of search in the labor market, Elsevier.

[38] Palumbo, G. et al. 2013. The Economics of Civil Justice: New Cross-country Data and Empirics. OECD *Economics Department Working Papers* 1060, OECD Publishing.

[39] Pfann, G. and Frenk, M. (2014) Firing Costs and Labor Demand in Recessions, mimeo.

[40] Postel-Vinay, F. and Turon, H. 2013. The Impact of Firing Restrictions on Labor Market Equilibrium in the Presence of On-the-job Search, Economic Journal, 124, 31-61.

[41] Rothschild, M. and Stiglitz, J. 1976. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics 17*, 557-66.

[42] Rủhmann, P. and Sủdekum, J. 2001. Severance Payments and Firm-Specific Human Capital, Discussion Paper n. 111 of the University of Gȯttingen.

[43] Saint-Paul, G. 1995. Efficiency Wage, Commitment and Hysteresis, Annales d'Economie et de Statistique, 37-38, 39-53.

[44] Shapiro, C. and Stiglitz, J. 1984. Equilibrium Unemployment as a Worker Discipline Device, *American Economic Review 74*, 433-444.

[45] Venn, D. 2009. Legislation, Collective Bargaining and Enforcement: Updating the OECD Employment Protection Indicators. OECD *Social, Employment and Migration Working Papers* 89, OECD Publishing.

Annex

An index of judicial discretion

Table A1 provides the legal information used to obtain measures of severance payment in different countries, dependent on the nature of the dismissal. Table A2 characterizes the dispersion of these measures, offering the weighted standard deviation of these dismissal costs. Weights are proxies for the probabilities of the three different outcomes (fair economic, fair disciplinary, unfair) under the country-specific rules concerning the burden of proof. Section 3, provides support to the assumption that the burden of proof affects the probability that a dismissal is considered fair economic, fair disciplinary or unfair by a Court of justice. In particular, we assume that the (unconditional) probability that a dismissal is considered unfair, $(1-p)$ takes the value .75 when the burden of proof is on the employer, .25 when the burden of proof is on the worker and .5 in the intermediate case where it can be on both parties. Higher up in the decision tree, the employer chooses whether to notify a disciplinary dismissal or take the economic dismissal route, internalizing the probability that the dismissal is considered unfair in the two circumstances. When the burden of proof is on the employer, it is more difficult that a firms takes the, less costly and hence most preferred, disciplinary dismissal route. The conditional probability that a disciplinary dismissal is ruled as unfair is then $p(1-p)$ and that an economic dismissal is ruled unfair is $(1-p)^2$. Table A2 also provides a measure of judicial discretion, Σ. The latter is obtained by simply multiplying the weighted standard deviation of dismissal costs by the appeal rates before the second instance as a percentage of population (α, see Table A1), a measure of uncertainty of Court rulings. The rationale for using appeal rates as a factor scaling up or down (in relative terms) the standard deviations is that the probability that a case is brought to a higher instance is increasing in the uncertainty as to the expected outcome of the litigation. This proxy is also used by the OECD in its review of the efficiency of legal systems (Palumbo 2013 [38]). When the judicial outcome is more certain, the parties would find an agreement extra-judicially, saving on legal costs. Overall, our measure of judicial discretion is given by

$$\Sigma = \sqrt{\alpha(E[T^2] - E[T]^2)}$$

where $E[.]$ denotes the expectation operator[18].

[18]Based on the conditional probabilities defined above, Σ is computed as follows

$$\Sigma = \sqrt{\alpha((1-p)T_U^2 + p^2 {T_F^D}^2 + p(1-p){T_F^E}^2 - ((1-p)T_U + p^2 T_F^D + p(1-p)T_F^E)^2)} \quad (33)$$

Table A1. Detailed information used to produce Table 1

Country	Severance Economic at20y,Fair	Severance Disciplinary at20y,Fair	Typical Compensation at20y,Unfair	Maximum Notice	Reinstatement Option	Appeal rate (per 100 citizens)	length of trial, first instance, months	Burden of Proof
Australia	2.80	0.00	6.00	1.00	1	0.77	6.42	e
Austria	0.00	0.00	6.00	4.00	3	0.39	4.29	e
Belgium	0.00	0.00	10.30	20	0	-	7.78	e
Canada	2.30	0.00	Court	2.00	1	-	-	e
CzechRepublic	1.50	0.00	6.00	2.00	3	74.02	4.49	w
Denmark	3.00	0.00	6.60	6.00	1	11.71	6.63	w
Finland	0.00	0.00	14.00	6.00	0	4.47	7.31	e
France	5.40	0.00	16.00	2.00	0.5	38.75	9.12	w
Germany	10.00	0.00	15.50	7.00	1.5	3.81	6.65	e
Greece	8.00	0.00	Court	4.00	2	38.49	5.17	e
Hungary	6.00	0.00	12.00	3.00	1	24.55	6.67	w
Ireland	4.00	0.00	24.00	2.00	1	2.96	9.02	e
Italy	0.00	0.00	21.00	6.00	1	29.31	18.81	e
Japan	0.00	0.00	6.00	1.00	1	3.99	3.57	m
Korea	0.00	0.00	6.00	1.00	3	7.66	4.81	w
Luxembourg	6.00	0.00	6.00	6.00	0	23.89	8.75	e
Mexico	1.00	0.00	15.00	n.a.	1.5	16.43	11.40	e
Netherlands	0.00	0.00	7.00	4.00	1	-	10.17	e
NewZealand	0.00	0.00	7.60	0.50	1	1.42	5.70	e
Norway	0.00	0.00	12.00	6.00	2	-	5.33	e
Poland	3.00	0.00	3.00	3.00	1	28.86	5.55	e
Portugal	12.00	0.00	20.00	2.50	2.5	16.93	14.15	e
SlovakRepublic	4.00	0.00	6.00	3.00	2.5	40.75	11.81	w
Spain	12.00	0.00	24.00	0.50	0	31.38	9.07	e
Sweden	0.00	0.00	32.00	6.00	0	3.05	6.20	e
Switzerland	0.00	0.00	6.00	3.00	0	10.43	4.36	w
Turkey	20.00	0.00	10.00	2.00	0	-	7.05	e
UnitedKingdom	4.60	0.00	5.50	3.00	1	5.47	8.34	e
United States	0.00	0.00	Court	0.00	0.5	-	-	e

Sources: EPLex; OECD (2013); [b]CEPEJ (2012)

Notes: Data are expressed in monthly wages. When notice period differs between categories of workers (e.g. white and blue collars) or between reasons of dismissal (e.g. personal and redundancy), the longest period is chosen; Court: Free determination by Court. Fair dismissal: severance pay at 20 years of tenure; Unfair dismissal: typical compensation at 20 years of tenure; Length of trial: Data from CEPEJ (2012) represent the average length of proceedings for employment dismissal cases at first instance Courts for the latest year available; the other data on length of trial period (OECD, 2013), represent the maximum legal length for this type of proceeding. π: probability (0-1) that, in case of unfair dismissal, the judge opts for the reinstatement of the worker. It is based on the 0-3 measure of the likelihood of the reinstatement provided by OECD (2013): 0= no right or practice; 1= rarely or sometimes made available, 2= fairly often made available, 3= almost always made available. For Netherlands, data refer to PES procedure.

Table A2. Judicial discretion

Country	st.dev	Σ
Australia	4.71	0.41
Austria	7.06	0.44
Belgium	4.46	-
CzechRepublic	7.65	6.58
Denmark	5.83	2.00
Finland	6.06	1.28
France	10.72	6.68
Germany	12.77	2.49
Hungary	10.07	4.99
Ireland	15.57	2.68
Italy	14.78	8.00
Japan	4.58	0.91
Korea	7.28	2.02
Luxembourg	3.58	1.75
Netherlands	5.48	-
NewZealand	5.19	0.62
Norway	10.22	-
Poland	2.92	1.57
Portugal	22.39	9.21
SlovakRepublic	10.41	6.64
Spain	11.98	6.71
Sweden	13.86	2.42
Switzerland	2.60	0.84
Turkey	7.81	-
UnitedKingdom	4.96	1.16

Notes: Reference is made to a worker with 20 years of tenure.
Sources: EPLex; OECD (2013);
See the main text for details.

Table A2 suggests that there is substantial cross-country heterogeneity on both, dispersion of potential costs of dismissals, and judicial discretion. Countries that are typically pointed out by employers of multinational corporations as having particularly costly dismissal procedures (e.g., France, Italy, Portugal, and Spain) display high values of the two indicators. Notice that they are not monotonically increasing in dismissal costs. For instance, countries with relatively large costs of unfair dismissals, like Sweden, display a much lower index of judicial discretion than countries, such as the Czech Republic, where unfair dismissal costs are about 50% lower than in Sweden.

An index of graded security

In order to characterize the severance-tenure profile of EPL in different countries, we developed a simple measure of graded security for regular workers, that is workers with open-ended contracts. The index is obtained by adding up mandatory severance and notice periods in case of fair economic dismissals for private sector workers at different tenure lengths, drawing on institutional information gathered by the ILO (EPLex project) and the OECD. In particular, we considered the following tenure classes for which cross-country comparable information was available: tenure at nine months; at one, five, ten and twenty years. At each tenure length, we computed an apparent elasticity of severance to tenure (plus notice) in between any two consecutive tenure levels and the ratio of tenure to the number of months in that interval. This suggests that there is significant cross-country variation in the slope of the severance-tenure profile, but only two

Figure 6: Severance payments for fair economic dismissals and tenure

countries (Austria and Japan) where the elasticity is zero throughout a 20 years tenure length, denoting a flat severance-tenure profile. [19] In the other countries, a flat severance-tenure profile is observed only limited to some tenure lengths.

As there is an apparent elasticity per period, we also developed a summary measure of graded security, by adding up the elasticities using weights proportional to the length of each tenure interval. Finally we normalized these overall apparent elasticities to obtain a unit value for a proportional severance scheme at all tenure lengths (one having always a unit apparent elasticity). Formally, denoting by $S + N = T$ the months of mandatory severance and compulsory notice period, by τ months of tenure, and indexing the tenure classes by subscript t, our index of *Graded Security* is given in each country by

$$GS = \sum_{t=0}^{4} \frac{\Delta T_t}{\Delta \tau_t} * \frac{\tau_{t+1}}{T_{t+1}} * \frac{\Delta \tau_t}{240} \tag{34}$$

where t indexes the tenure length classes.

We find that in 27 countries out of 29, the index is positive and in most of them (18) it is above 50 per cent. In the two countries paying the same severance at all tenure levels (Austria and Japan), the index is clearly zero.

[19] In Denmark, New Zealand and the US, there is no national mandatory severance, hence the elasticity is not defined. Therefore, these countries are not included in Table 2.

Table A3 Apparent Elasticities at different tenure lengths and overall measure of Graded Security

	Apparent Elasticities					GS Index
	at 9 months	at 12 months	at 60 months	at 120 months	at 240 months	
Australia	1.00	3.33	0.67	0.38	0.00	0.31
Austria	0.00	0.00	0.00	0.00	0.00	0.00
Belgium	0.75	0.80	0.21	0.50	1.00	0.70
Canada	1.00	2.00	1.00	0.89	0.71	0.84
Czech Republic	0.60	0.00	0.00	0.00	0.00	0.02
Estonia	1.00	0.99	0.42	0.50	0.00	0.26
Finland	1.00	2.15	0.63	1.00	0.67	0.77
France	0.00	0.67	0.75	0.50	0.91	0.74
Germany	0.00	1.41	0.85	1.00	0.94	0.91
Greece	0.00	4.00	0.69	0.71	1.00	0.87
Hungary	0.00	0.00	0.90	0.55	0.79	0.71
Ireland	1.00	2.18	0.74	0.80	0.76	0.79
Israel	1.00	3.11	0.83	0.91	0.95	0.95
Italy	0.00	0.00	0.00	0.50	0.40	0.33
Japan	0.00	0.00	0.00	0.00	0.00	0.00
Korea, Rep.	0.00	2.01	0.83	0.91	0.95	0.90
Luxembourg	0.00	0.00	0.75	0.75	0.67	0.67
Mexico	1.00	0.12	0.40	0.57	0.73	0.63
Netherlands	0.00	0.00	0.63	0.67	0.50	0.54
Norway	0.00	0.00	0.63	0.67	0.00	0.29
Poland	0.00	0.00	0.83	0.00	0.00	0.17
Portugal	1.00	0.50	0.54	0.88	0.89	0.82
Slovak Republic	0.50	0.00	0.42	0.00	0.00	0.10
Slovenia	0.00	0.67	0.71	0.63	1.08	0.85
Spain	0.51	0.57	0.87	0.93	0.87	0.87
Sweden	0.00	0.00	0.83	1.00	0.00	0.42
Switzerland	1.00	2.00	0.00	0.67	0.00	0.23
Turkey	0.00	2.08	0.90	0.84	0.92	0.87
United Kingdom	0.00	0.00	1.09	1.00	0.65	0.79

Non-existence of severance in equilibrium with adverse selection

The argument we put forward in order to unravel a pooling equilibrium with strictly positive severance is similar to the arguments in Rothschild and Stiglitz [41] that a pooling equilibrium cannot exist under asymmetric information. We assume that there exist two types of workers, regular workers as described above, and shirkers, who always shirk (effort cost infinitely large). A fraction α of the workers are shirkers. We work with a two periods model, and contrary the assumption in the main text we assume that the worker also lives for two periods only. We label the periods as 0 and 1.The period 1 outside option b_1 (for ordinary workers) and b_1^s for shirkers are taken as exogenous. Trivially, as long as the firms strictly prefer to incentivize ordinary workers, they prefer to hire ordinary workers rather than shirkers. However, firms cannot tell shirkers and non-shirkers apart. We first study a frictionless economy. The value of not working is normalized to zero for both worker types.

A (pooling) equilibrium is a triple (w_0, w_1, T_1). Consider a candidate equilibrium with T_1 in $(0, T_1^*]$, where T_1^* is the efficient severance. The wage w_1 is set such that the incentive compatibility constraint is satisfied (possibly with slack). In addition we require a zero profit condition on firms (this is not necessary for the main argument). The period 0 profit of a firm that hires an ordinary worker reads

$$\Pi_0 = -w_0 + E\max[z - w_1, -T] \tag{35}$$

The profit if hiring a shirker is $\Pi^s = -w_0 - q_1 T_1$. Hence the zero profit condition reads

$$\alpha[-w_0 - q_1 T_1] + (1-\alpha)[-w_0 + E\max[z - w_1, -T_1]] = 0 \tag{36}$$

which determines w_0. Clearly, w_0 is decreasing in α. We assume that E is sufficiently high so that w_0 is strictly positive for relevant values of w_1 and T_1.

We can now calculate the value of participating in the market for shirkers and non-shirkers, U and U^s. It follows that

$$U = w_0 + (1 - F(w_1 - T_1))w_1 + F(w_1 - T_1)(T_1 + b_1) - C \tag{37}$$

$$U^s = w_1 + q_1(T_1 + b_1^s) \tag{38}$$

where F denotes the cumulative distribution function of an investing worker's period 2 productivity, as in the main text.

Now consider a firm that deviates and offers $(w_0, w_1 + \epsilon_w, T_1 - \epsilon_T)$, where ϵ_w, ϵ_T are small numbers and chosen such that the non-shirking workers still receive U in expectation. It follows from the equation for U^s that the deviating firm does not attract any shirkers, as their participation constraint is not satisfied. In addition the incentives for workers are strengthened, hence the change preserves incentive compatibility. It follows that the deviating firm attracts regular workers but not shirkers, and hence the deviation is strictly profitable. As already hinted at, the result is analogous to the no-pooling equilibrium result in Rothschild and Stiglit.

Under the reasonable assumption that $w_0 \geq 0$ and $T_1 \geq 0$, a separating equilibrium with one wage contract $(0, w_1, 0)$ may exist. Now w_1 solves the zero profit condition (36), and the incentive compatibility constraint is satisfied with slack. The shirkers do not work (or, equivalently, work and receive nothing) while the non-shirkers work and receive no severance.[20]

As in the model of Rothschild and Stiglitz, a separating equilibrium may not exist. However, if we allow for search frictions, an equilibrium does exist, as shown by Guerrieri et al (2010) [21]. By using exactly the same argument as above, it follows that an equilibrium, if it exists, must be one where firms offer wage contracts $(0, w_1, 0)$. The incentive compatibility constraint is satisfied with slack. Shirkers are indifferent between searching for a job and not, and are assumed not to search. Shirkers receive zero.

Consider now a small measure of firms that deviate and introduce a strictly positive severance T_1. Independently of w_1, this will give a shirker a strictly positive pay-off. Hence the queue length in the corresponding submarket is infinite, $\theta = 0$. Therefore, no ordinary workers apply. It follows that the deviating firm gets a negative profit. Hence, with search frictions, the separating equilibrium exists.

Proof of lemma 1 and 3

The proof is by induction. Suppose the the optimal contract $\Omega' = \{w'\}_{i=0}^{n}$ is such that the incentive compatibility constraint does not bind for $i = n-1$, i.e., that $w'_n > w^\xi_n$. It follows that $\epsilon'_n > \epsilon^\xi_n \geq \epsilon^*_n$, where ϵ'_n denotes the period n firing cutoff with the wage contract Ω', ϵ^ξ_n the optimal cut-off with the contract Ω^ξ, and ϵ^*_n the efficient cut-off. We will show that there exists another contract $\Omega^d = \{w^d\}_{i=0}^{n}$ that satisfies all the constraints and and is more profitable. To this end, set $w^d_n = w^\xi_n$ and $w^d_{n-1} = w'_{n-1} + (1 - F(\epsilon'_n))(w'_n - w^\xi_n)/\beta$. For all other periods, $w'_i = w^d_i$.

It follows that

$$W^d_{n-1} = w^d_{n-1} - C_{n-1} + \beta \int_{\epsilon^\xi_n} (w^d_n + \beta U) dF + (1 - F(\epsilon^\xi_n))(T_n + U) \tag{39}$$

$$= W'_{n-1} + \beta \int_{\epsilon^\xi_n}^{\epsilon'_n} (w^d_n - (T_n + b)) dF \tag{40}$$

$$\geq W'_{n-1} \tag{41}$$

since T_n is at or below the efficient level so that the worker prefers to be hired. Hence the ICC of the worker is satisified in period $n-1$. For the firm,

$$\Pi^d_{n-1} = \epsilon_{n-1} - w^d_{n-1} + \beta \int_{\epsilon^\xi_n} (\epsilon_n - w^d_n) dF - (1 - F(\epsilon^\xi_n))T_n \tag{42}$$

$$= \Pi'_{n-1} + \beta \int_{\epsilon^\xi_n}^{\epsilon'_n} (\epsilon_n - w^d_n + T_n) dF \tag{43}$$

$$> \Pi'_{n-1} \tag{44}$$

[20]Suppose output in the first period is $y_1 > 0$ for both shirkers and no-shirkers, and that the minimum period 1 wage is $\underline{w}$, $0 \leq \underline{w} \leq y_1$. Then a separating equilibrium may exist in which the shirkers are offered a contract $(y_1, 0, 0)$ and the non-shirkers are offered a contract $\underline{w}, w_2, T$, where $T = (y_1 - \underline{w})/q$. The shirkers are then just indifferent between the contract "designed" for them and the contract designed for the non-shirkers. In this case the good workers may get some severance, but it may well be far below the efficient level.

For a given cut-of in period n, the profit is the same. However, since the firm choses to lower the threshold when the period n wage is lower, this will strictly increase the profit.

Since the contracts are identical for all other periods, the ICC will hold for all earlier periods, and $\Pi_i^d > \Pi_i'$ for all periods. Hence the contract Ω' cannot be optimal.

Suppose now that the incentive compatibility constraint does not bind in period $i' < n-1$, but binds thereafter. It follows that W_j, Π_j and S_j are the same with the contract Ω' as with the wage contract Ω^ξ for all $j > i'$, while $w_i' > w^\xi$, $\epsilon_{i'}^\xi < \epsilon_{i'}'$. Again we consider a deviation. As above, let the wage contract Ω^d be equal to the wage contract Ω' in all periods except period i' and $i'-1$. Furthermore, let $w_{i'} = w_i^\xi$ and $w_{i'-1}^d = w_{i'-1}' + (1 - F(\epsilon_{i'}'))(w_{i'}' - w_{i'}^\xi)/\beta$. It follows that

$$
\begin{aligned}
W_{i'-1}^d &= w_{i'-1}^d - C_{i'-1} + \beta \int_{\epsilon_{i'}^\xi} (w_{i'} + \beta W_{i'+1}^d) dF + (1 - F(\epsilon_{i'}^\xi))(T_{i'} + U) && (45) \\
&= W_{i'-1}' + \beta \int_{\epsilon_{i'}^\xi}^{\epsilon_{i'}'} (w_{i'}^d - b - T_{i'} + \beta(W_{i'+1} - U)) dF && (46) \\
&\geq W_{i'-1}' && (47)
\end{aligned}
$$

since the worker prefers hiring to firing. Hence the incentive compatibility constraint is satisfied in period $i'-1$, and since the contracts are identical for all earlier periods, also for all earlier periods.

For the firm, it follows that

$$
\begin{aligned}
\Pi_{i'-1}^d &= \epsilon_{i'-1} - w_{i'-1}^d + \beta \int_{\epsilon_{i'}^\xi} (\epsilon_{i'} - w_{i'}^d + \beta E \Pi_{i'+1}) dF - (1 - F(\epsilon_{i'}^\xi)) T_{i'} && (48) \\
&= \Pi_{i'-1}' + \beta \int_{\epsilon_{i'}^\xi}^{\epsilon_{i'}'} (\epsilon_{i'} - w_{i'}^d + T_{i'} + \beta E \Pi_{i'+1}) dF && (49) \\
&\geq \Pi_{i'-1}' && (50)
\end{aligned}
$$

The integral in the second equation reflects the gain from reducing the cut-off, and since the firm maximizies profit this is positive. Hence the profit increases. Since $w_i^d = w_i'$ for all $i < i'$, it follows that ICC is satisfied for all earlier periods and that $\Pi_0^d > \Pi_0'$, and the contract Ω' cannot be optimal.

Note that the argument is independent of the value of R, the search rent given to the worker. Hence the search rent should be given to the worker in the first period, where it will not distort the firing decision of firms.

Proof of lemma 2

We first show that (11) has a solution for k sufficiently high.Consider period n. For any k, the right-hand side of (11) goes to infinity as $\varepsilon_d \to \varepsilon^u \equiv \varepsilon_l + k$. Hence it is sufficient to show that the left-hand side is greater than the right-hand side for some ε^* in the support of ε. To this end, consider the median $\varepsilon^m = \varepsilon_l + kz^m$. At this value, $F(\varepsilon^*) = 1/2$. Hence, as k increases, the right-hand side of (11) stays constant while the left hand-side increases to infinity with k. Hence, for a sufficiently high value of k, the equation has a solution.

Profits at period n can be compactely written as

$$
\Pi_i(\epsilon_i) = y + \epsilon - b(U) - \frac{C_{i-1} + T\beta[q - F(\epsilon_i^d)]}{\beta[1 - F(\epsilon_i^d)]} - C_i + \beta S_{i+1}^e \tag{51}
$$

The profit of the firm is given by

$$
\begin{aligned}
\Pi &= 2y - 2b - qT + \frac{1}{2} E^{\varepsilon \geq \varepsilon_d} \varepsilon \\
&\geq 2y - 2b - qT + \frac{1}{2}[\varepsilon_l + k\bar{z}]
\end{aligned}
$$

The right-hand side goes to infinity with k, hence the proposition follows.

An increase in C shifts the right-hand side of (11) up, and the left-hand side down. Since the left-hand side crosses the right-hand side from below, it follows that ε_d increases in C and decreases in T.

Second, as k goes to infinity, it follows that S_{i+1}^e goes to infinity, and hence certainly is greater than C_i for a sufficiently high value of k.

Deriving the competitive search equilibrium

Using the definition of surplus at time 0 and the associated rent R_0, the firm problem is

$$V = Max_{R_0}\{-c + \beta q(\theta)[S_0 - R_0]\} \tag{52}$$

subject to

$$b = z + \beta\left[\theta q(\theta) R_0\right] \tag{53}$$

where $b = (1-\beta)U$ as before

A binding participation constraint pins down the first period wage w_o. We defined the period 0 worker rent as $R_0 = W_0 - U$ and $S_0 = \Pi_0 + R_o$, where Π_0 is given by equation (54). This last set of equalities highlight an important feature of our model and wage setting, namely the fact that we have $\frac{d\Pi_0}{dR_0} = 0$, and higher period 0 rents do not lead to higher output. In other words, there are no efficiency wages.

The contract starts in period 0 and the worker must invest for period 1, so that the value function is

$$W_0 = w_0 - C_0 + \beta[F(\epsilon_1^d)U + (1 - F(\epsilon_1^d)W_1]$$

Since the first period productivity is fixed at $\bar{\epsilon}$, profits are

$$\Pi_0(\bar{\epsilon}) = y + \bar{\epsilon} - w_0 + \beta\Pi_1^e \tag{54}$$

The first thing to note is that the solution to the first maximization problem is independent of R_0. This follows from the fact that the firm can always scale up or down R_0 by increasing or decreasing w_0, without influencing the incentive compatibility constraint. The first problem is thus the basic problem that we did solve in partial equilibrium. We turn now to the second part of the maximization problem, obtaining the optimal R_0. The worker indifference curve acts as a constraint to the firm and it implies

$$\frac{\partial\theta}{\partial R_0} = -\frac{\theta}{(1-\eta(\theta)R)};$$

The first-order condition of the firm implies

$$\frac{\partial\theta}{\partial R_0}[S_0 - R_0]\frac{q'(\theta)}{q(\theta)} = 1$$

Substituting the worker indifference condition into the firm first-order condition yields

$$R_0 = \eta(\theta)S_0; \quad S_0 - R_0 = (1-\eta(\theta))S_0 \tag{55}$$

The previous condition is the traditional efficient rent splitting of competitive search equilibria. The optimal value of a vacancy in equilibrium must be equal to entry costs, or $V = K$. The equilibrium values of S_0, U and θ can then be easily obtained as in the text.

Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) est un laboratoire d'excellence (Labex). Ce projet est distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR). Il est financé dans le cadre des investissements d'avenir.

(ANR-11-LABX-0091, ANR-11-IDEX-0005-02)

www.sciencespo.fr/liepp

Directeurs de publication:
Bruno Palier et Etienne Wasmer

Sciences Po - LIEPP
27 rue Saint Guillaume
75007 Paris - France
+33(0)1.45.49.83.61
liepp@sciencespo.fr