Severance Pay

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

February 11, 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. According to the extensive literature on Employment Protection Legislation (EPL), such transfers are either ineffective or less efficient than unemployment benefits in providing insurance against labor market risk. In this paper we show that mandatory severance is optimal in presence of wage deferrals motivated by deterrence of opportunistic behavior of workers. Our results hold under risk neutrality and in general equilibrium. We also establish a link between optimal severance and efficiency of the legal system and we characterize the effects of shifting the burden of proof from the employer to the worker. Our model accounts for two neglected features of EPL. The first is the discretion of judges in interpreting the law, which relates not only to the decision as to whether the dismissal is deemed fair or unfair, but also to the nature, economic vs. disciplinary, of the layoff. The second feature is that compensation for dismissal is generally increasing with tenure. The model also rationalizes why severance is generally higher in countries with less efficient judicial systems and why small firms are typically exempted from the strictest EPL provisions.

"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 ***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. Such transfers from the employer to the worker are the most important component of dismissal costs. 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. 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.

According to the literature on EPL, severance pay is either neutral with respect to labor market outcomes or it is more distortionary than other institutions in providing insurance against labor market risk. It is neutral when wages are flexible and agents are risk neutral (Lazear [22]). Under rigid wages, severance pay increases unemployment (Garibaldi and Violante [16]). With risk-averse employees, severance is less efficient than other institutions – such as experience-rated unemployment benefits – in providing insurance to workers against the risk of job loss (Blanchard and Tirole [6]).

Why do we need then severance pay? In this paper we show that severance pay is efficient under risk neutrality and flexible (entry) wages, even in presence of unemployment benefits putting a floor to wages, provided that there are wage deferrals. In our model, wages are deferred in order to incentivize workers to investment in job-specific productivity. However 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 [27] 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 [26], Kotlikoff and Gokhale [20] Flabbi and Ichino [17]) 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 [21], who compared seniority-earning profiles of employees and self-employed (for which no agency problem arises) and by Barth [1], who compared the wage-tenure profile of workers paid piece-rate with that of workers receiving a flat wage.

Our model also allows to rationalize two neglected features of EPL.

The first relates to the discretion of judges in deciding upon the fairness and the nature (economic vs. disciplinary) of the dismissal. This decision deeply affects the costs of individual dismissals. 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. Shirkers may receive the compensation offered for unfair disciplinary or economic dismissals, while opportunistic employers

claiming that the dismissal is either disciplinary or due to objective economic circumstances may get away without paying the higher severance required for unfair dismissals or not paying severance at all. 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 [13]) that the organisational 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 arrangement 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 organisations 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 [11]). 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.

In this paper we show that these design features are efficient in dealing with moral hazard and adverse selection and that optimal severance is dependent on the design and efficiency of the judicial system.

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 ([38], 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 even in case of dismissals that are actually motivated on purely exogenous productivity reasons. 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). Thus, there is a non-zero probability that a shirker obtains the severance pay provided in case of economic dismissals or that an employer pays the (low if any) severance due in case of disciplinary dismissals even when the worker has invested in the productivity of the job. We do endogenize these probabilities depending on whether the burden of proof concerning the nature (economic vs. disciplinary) of the dismissal falls on the worker or on the employer.

As we dig into the legal system, our model can establish a link between the efficiency of the judicial procedures in detecting opportunistic behavior of employers or employees, and the optimal levels of severance pay for disciplinary, economic and unfair dismissals. The model also shows under which conditions – in terms of productivity, monitoring technologies, jurisprudence, and design of unemployment benefit systems – a severance pay increasing with tenure improves productivity, reduces inefficient firing and induces an efficient allocation of labor. Finally, problems in monitoring rationalize why small firms are typically exempted from the strictest EPL regulations: it is easier for employers in small firms to prove opportunistic behavior of workers before Courts as they can better monitor and document the effort made by their workers in increasing the productivity of a job.

Our results are empirically relevant. Legal rules 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 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.

The plan of the paper is as follows. Part one evaluates the relevance of severance pay in OECD countries and characterizes two neglected features of EPL, the discretion of judges in setting the level of severance pay depending on whether the individual dismissals is disciplinary, economic or unfair, and the tenure profile of severance pay. Part two develops a simple partial equilibrium, model with moral hazard of the employees, and evaluates optimal severance pay under these circumstances. It also extends these results to the general equilibrium, endogenizing the workers' outside option. 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 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. Finally, part five summarizes our key results and concludes.

1 Why Severance Matters

Employment protection legislation is one of the most widely investigated institutions in the labor market[1]. The theoretical literature, pioneered by Bentolila and Bertola [2] and Bertola [4], 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[2] is generally not framed in these models, as Lazear [22] neutrality result indicates that, with wage flexibility and risk neutrality, it only affects the tenure profile of wages leaving employment, hiring and separations unaffected.

However, severance pay accounts for a very large share of the costs of individual[3] dismissals. According to Garibaldi and Violante [16] who estimated the red tape costs of layoffs in Italy, severance pay accounts for about 2/3 of total dismissal costs. Severance also explains about 50 per cent of the cross-country variation in the OECD index of strictness of EPL. It is mandatory even in countries with negligible firing taxes to be paid to third parties.[4] When severance is not mandated by law, it is industry-level collective bargaining to force individual employers to compensate workers in case of dismissals. For instance, Kodrzycki [19] 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.

Severance pay differs from firing taxes in at least two important dimensions. The first is that its amount 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. The second distinguishing feature of severance pay is 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.

1.1 Judicial Discretion

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

[1] See Boeri and vanOurs [8] for a review of this literature.

[2] 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.

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

[4] See Postal-Vinay and Turon [35], and Boeri [7] for a theory of severance pay as a device to buy time and avoid paying firing taxes, in presence of on-the-job search.

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.

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. The theoretical literature on EPL generally treats severance as a deterministic transfer from the employer to the employee. In the few cases where stochastic severance is allowed (Garibaldi [15], Malo [25]) , 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 [14] and Boeri [9]. However, Galdon-Sanchez [14] operates on a reduced form model and both Boeri [9] and Galdon-Sanchez [14] 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.

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 [12], a study by the ILO [5] and a recent survey of Civil Justice also carried out by OECD (Palumbo [33]).

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 the reintegration of the worker is actually imposed by

[5] R*ü*hmann and *S*ü*dekum [37] 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.

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.

Consistently with these facts, the model developed in the next section will allow for both economic and disciplinary dismissals and address the moral hazard problem related to potential workers' misconduct as well as the nature, economic vs. disciplinary, of the dismissals. We will initially assume that the probability that a Court rules that a layoff is disciplinary is exogenous and show that in a model with wage deferrals (e.g., related to career concerns), a properly designed severance scheme maximizes the joint surplus from a match. Later on, in Section 3, we will explicitly frame the operation of the judicial system, by endogenizing the probability that a Court rules in favor of the employer, based on the evidence that can be provided by the firm on the productivity of the worker, and evaluate the optimal severance under this endogenous probability of "getting away with it". Finally, we will consider a case where also the employer can behave strategically: in this double moral hazard case, the employer may pretend that the dismissal is disciplinary when it is instead due to economic reasons. Compensation for unfair dismissal can be rationalized as a deterrent to this behavior of the employer.

1.2 The elasticity of severance to tenure

Figure 1 displays the severance tenure profiles in OECD countries drawing on institutional information gathered by the ILO (EPLex project) and the OECD. In 25 countries out of 30 there is evidence of severance increasing with tenure. 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 Basic Economics of Severance Pay

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

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

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

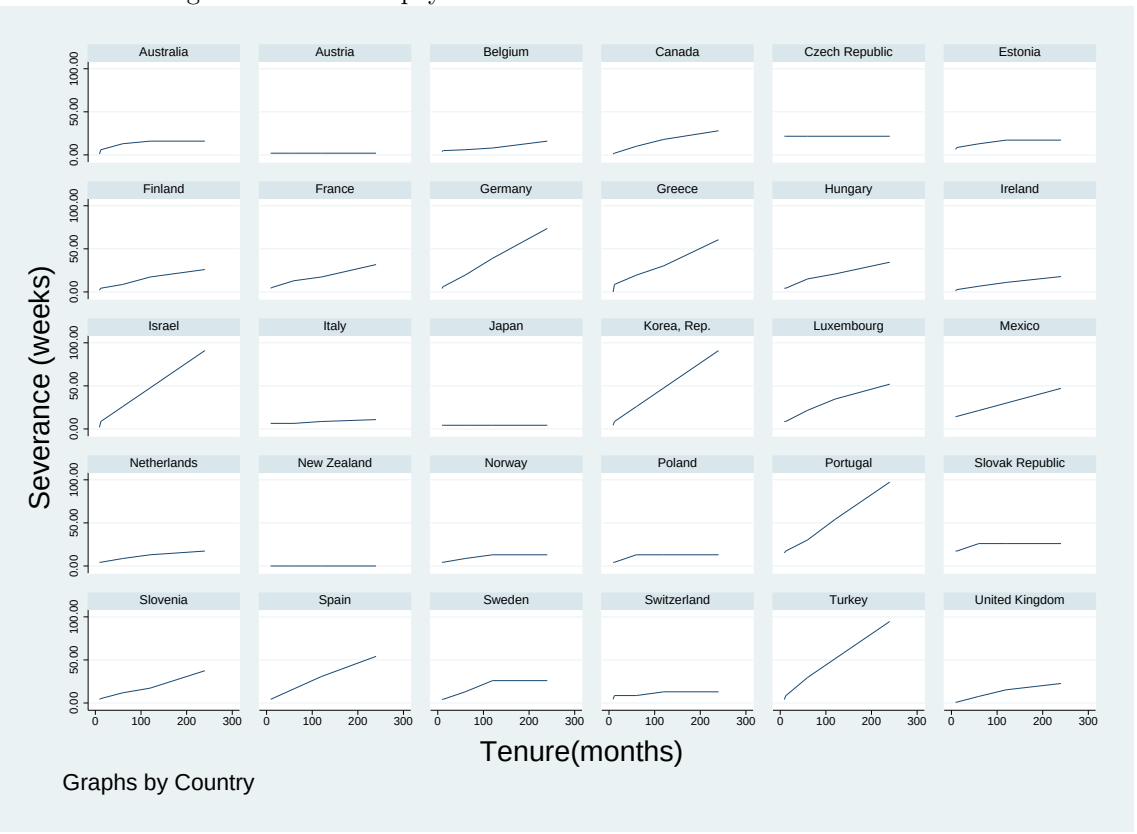
where the unit of measurement is monthly wages. As detailed in Table A1 in the annex, we attribute to π the value obtained by standardizing to the unit interval the 0-3 OECD index on the likelihood of the reinstatement, where 0 means never reinstatement and 3 denotes the case where employees can freely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

Table 1: Judicial discretion over severance pay

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

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.

Figure 1: Severance payments for fair economic dismissals and tenure



double moral hazard.

The intuition behind our model can be grasped by referring to a two-period contract with stochastic productivity realizations and exogenous wage deferrals. Suppose that the workers' outside option is constant over time, and there is no firm-specific investment in training. Productivity is known in the first period while in the second period it is stochastic: it can only take values above the worker's outside option, but possibly lower than the (non-negotiable) second period wage. In case the productivity falls below the second-period wage, the employer making (ex-post) strictly negative profits has all the incentives to layoff the worker even if the total surplus of the match is positive. Since productivity is higher than the outside option of the worker, it would be possible to split this surplus between the worker and the employer, making sure that both parties do not realize a negative surplus. However, deferred wages do not allow for this outcome, and hence even a job realizing a strictly positive (total) surplus is destroyed. Severance pay may prevent this inefficient separation for senior workers, provided that it is sufficiently large. We show below how large an optimal severance should be to prevent inefficient separations, by considering a more complex environment, notaby one with endogenous wage deferrals and firm-specific investment.

2.1 A model of specific investment and endogenous wage deferrals

Consider now a more general two-periods setting with endogenous wages being set by the firm in the initial period and moral hazard related to the decision of the worker to invest in firm-specific training. In our setup, the investment in training is costly and has uncertain returns. Wage contracts are incomplete, as firms cannot commit not to fire workers after a negative productivity realization. We are not aware of employers in the private sector signing contracts that do not allow them to layoff workers in case of exogenous shocks.

At the same time, a worker who does not invest can legitimately be dismissed as a "shirker". This firm-initiated dismissal is "disciplinary", but such a case must be proved before a Court and there is a certain probability that a disciplinary firing is deemed unfair by judges and a shirking worker can "get away with it". We thus have a moral hazard problem, since shirking workers can obtain severance payments.

We begin with partial equilibrium: one worker and one firm have a job opportunity that lasts two periods. We also impose that the worker and the firm are risk neutral, and for simplicity initially abstract from discounting. The worker's outside option is b in every period. In period 1 the worker faces a specific investment opportunity $s = \{0, 1\}$. The investment opportunity costs the worker C in the first period. The investment is private information to the worker in the first period.

If the worker does not invest, the output in both periods is non-stochastic, and denoted by $y < b$. Hence, the project cannot be profitable unless the worker invests. This restriction is imposed for expositional reasons only and will be relaxed later on. Conditional on the investment being undertaken, productivity in the second period will be $y + \varepsilon$, where ε is the specific component of productivity and is drawn from a continuous distribution $F(\varepsilon)$, defined over the support $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$. As stated above, we impose that $y + \underline{\varepsilon} < b$. The specific component of productivity is observed only by the firm, hence wages cannot be made contingent upon it. We assume that $\overline{\varepsilon}$ is sufficiently high for the firm to break even.

If the worker shirks in the first period, this is observed by the firm in the second period, either directly (it is observed that he does not have the necessary skills), or indirectly, by his low productivity (if $\underline{\varepsilon} > 0$). This assumption will be relaxed below. The firm may also want to fire the worker, even if the worker has invested, if the draw of ε is sufficiently low. We refer to this case as an economic dismissal.

Definition 1 *Disciplinary Dismissal. In period 2, a firm is entitled to freely dismiss a shirking worker who did not invest in the first period.*

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

Over and beyond economic and disciplinary dismissals, we have to distinguish between fair and unfair dismissals. In our setup, the distinction is particularly relevant for disciplinary dismissals. Whether a disciplinary dismissal can be defined as fair can only be proven in Court. The Court ruling is stochastic. In period 2 the firm observes if the worker has invested or not, but cannot necessarily prove insufficient

investment (hereafter shirking) in a Court of law. We assume that there is a probability $1-q$ that the Court observes shirking and declares the firing as fair. In such a case, the firm is exempted from paying severance payments. Hence, there is a probability q that a shirking worker gets away with it and receives severance payment. When this happens, the disciplinary dismissal is defined as unfair, and a severance payment is due. The realization of q is made after the firm has fired the worker, hence the expected severance payment for the firm when firing a shirking worker is qT. In this section, q is considered exogenous, it will be endogenized in the next section.[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 severance payment is always due, hence we abstract from moral hazard on the firm side. In principle, the firm may claim that an economic dismissal is a disciplinary dismissal in the hope of getting away with it without having to pay T. We will consider moral hazard of the employer at a later stage, and comment on its effects along the way. Our underlying assumption at this point is that the cost to the firm of being caught for unfair disciplinary dismissal is sufficiently high, so that the firm will never claim that an economic dismissal is disciplinary.

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 is a policy tool. We rationalize our assumption that the firm cannot contract upon T by alluding to an underlying, unmodelled 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 arbitrarily small, but strictly positive. Firms cannot distinguish between shirkers and ordinary workers. Consider a situation where all firms offer a contract (w_1, w_2, T), where w_1 and w_2 are period 1 and 2 wages, 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 $(w_1, w_2', T' - \varepsilon)$, where $w_2' > w_2$ and ε can be arbitrarily small. Since ordinary workers are strictly more willing to trade off severance payment for a higher period 2 wage than are shirkers, it is possible to chose w_2' 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 (for a more formal treatment of the same argument, see [29] and [36])

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.

Optimal contracts

We derive below the optimal contract of the firm, that is, wages (w_1, w_2), that maximize the profit of the firm given 1) the firm's ex post firing behavior, 2) the incentive compatibility and participation constraints of the worker, for an exogeneously given severance T. Then we solve for the optimal severance. In what follows we *assume* that the firm's participation constraint is satisfied. We will get back to this in the section on optimal severance.

Let us indicate with $W_{(s=0)}$ the value of the job to the worker in case she does not invest. When this happens, the shirking worker gets the first period wage, and will be fired in period 2 for disciplinary reasons with probability 1.

$$W_{(s=0)} = w_1 + b + qT$$

The worker's expected income if she invests is

$$W_{(s=1)} = w_1 - C + (1 - F(\varepsilon_d))w_2 + F(\varepsilon_d)[b + T]$$

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

where $F(\varepsilon_d)$ is the probability of being fired for economic reasons. The period 2 profit of the firm is given by

$$\Pi_2(\varepsilon) = Max[y + \varepsilon - w_2; -T].$$

The firm retains the worker if and only if $\varepsilon \geq \varepsilon_d$, where the threshold ε_d is given by

$$\varepsilon_d = w_2 - y - T \tag{1}$$

The worker will invest if and only if $W_{(s=1)} \geq W_{(s=0)}$. Hence the ICC reads

$$(1 - F(\varepsilon_d))w_2 + F(\varepsilon_d)(b + T) - C \geq b + qT$$

The lowest value of w_2 that satisfies the ICC is thus[8]

$$w_2 = b + \frac{C + [q - F(\varepsilon_d)]T}{1 - F(\varepsilon_d)} \tag{2}$$

Let us give some comments. First, note that if $\varepsilon_d = \varepsilon_l$, i.e., if workers who invest are never dismissed, then $w_2 = b + C + qT$. In this case the worker who invests is compensated for her outside option b, her investment cost C, and the rents qT she would get if shirking. Second, the numerator in (2) increases in T if $q > F(\varepsilon_d)$. This reflects the fact that the worker in this case is more likely to get the severance if shirking than if not shirking. If $q < F(\varepsilon_d)$, the opposite holds.

Given the second period wage, the first period wage is set so as to ensure that the worker participation constraint is satisfied. The worker has an outside option equal to b per period, so that the worker participation constraint is $W_{(s=1)} \geq 2b$ which, when strict, simplifies to

$$w_1 = b - qT \tag{3}$$

This is also intuitive. By shirking and not investing, the worker is able to achieve a rent of qT. In order to satisfy the ICC, the worker gets the same rent when investing. The exact same amount is extracted from the worker (relative to her outside option) through a low period 1 wage. Clearly, the severance does not influence the worker's lifetime income in partial equilibrium, but it makes the wage-tenure profile steeper.

Finally, in order to fully characterize the contract, we have to solve for ε_d. From (1) and (2) it follows that

$$\varepsilon_d = b - y + \frac{C - (1 - q)T}{1 - F(\varepsilon_d)} \tag{4}$$

At $\varepsilon_d = \varepsilon_l$, the left-hand side of the equation is strictly negative, while the right-hand side is positive (as long as $C \geq (1 - q)T$). Both the left-hand side and the right-hand side of the equation are increasing in ε, hence the equation may not have a solution.[9] However, if the 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 $\varepsilon = kz$, 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 and C where investments are implementable, ε_d is increasing in C and decreasing in T

Proposition 1 *Consider an arbitrary investment cost C and severance T. Then, if the investment is sufficiently productive, the investment is implementable. On intervals where the investment is implementable, ε_d is increasing in C*

[8]Suppose that we include moral hazard on the firm side, and assume that there is a probability τ that the firm does not pay severance payment at economic dismissals. One can then show that the incentive compatibility constraint reads

$$w_2 = b + \frac{C + [q - F(\varepsilon_d)(1 - \tau)]T}{1 - F(\varepsilon_d)}$$

This case is discussed in Section 3.

[9]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.

Note that if $q = 1$ (shirking workers always get severance pay), *severance payments are neutral*. From (4) it follows that ε_d is independent of T in this case, and from (2) that w_2 increases dollar for dollar with T. This is a version of the Lazear (1990) neutrality result.

Optimal severance

In this subsection we will derive the optimal severance for the firm in question. We think of the firm as a representative firm, so that our results can shed light on optimal severance more generally. We will discuss this more in detail below.

Let S_2 denote the joint period 2 surplus of the worker and the firm, defined as the joint income of the worker and the firm when together $(y+\varepsilon)$ less the sum of their incomes when separating (b). It follows that

$$S_2 = y + \varepsilon - b$$

Neither wages nor severance payments appear in the joint surplus, as they are transfers between the two parties. We say that the layoff is efficient if the cut-off level ε^* (below which a worker is fired) solves $S_2(\varepsilon^*) = 0$, i.e.,

$$\varepsilon^* = b - y. \tag{5}$$

Firing is efficient whenever the productivity from the job $(y + \varepsilon^*)$ falls below the worker's outside option b. We call ε^* the efficient reservation productivity. By assumption, $\underline{\varepsilon} + y < b$, hence $\varepsilon^* > \underline{\varepsilon}$.

The first result that we can establish concerns the firing policy of the firm. By comparing (4) and (5) at $T = 0$, it immediately follows that

$$\varepsilon_{d(T=0)} = b - y + \frac{C}{1 - F(\varepsilon_d)} > \varepsilon_d^*.$$

Hence, with no severance payments, firing is too high in the second period, as predicted by the motivating example at the beginning of this section. As second period wages need to pay for the worker's investment effort in the first period, the firm has a tendency to over dismiss in the second period a worker who did not shirk and invested in the first period.

Our next step is to derive the optimal severance payment T, that is, find the value of T such that $\varepsilon_d = \varepsilon^*$. From (1) and (5), it follows that efficient firing is obtained if and only if $T = w_2 - b$, i.e., the severance is equal to the wedge between the inside and the outside wage. Inserting this into (2) gives that the optimal T, T^*, is given (for $q < 1$) by

$$T^* = \frac{C}{1 - q} \tag{6}$$

This can also be seen directly from equation (4). Finally, to find w_1, we insert T^* into (3) to obtain

$$w_2 = b + \frac{C}{1 - q} \tag{7}$$

$$w_1 = b - \frac{q}{1 - q} C \tag{8}$$

Proposition 2 *For $q < 1$, the optimal severance T^* is given by* (6)

The expression for the optimal severance turns out to be surprisingly robust. Not only does it holds in the n-period case, it also holds when we allow for a different output structure and with endogenous court decision structure as demonstrated in the next section.

If $q = 1$, 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.

We want to point out the remarkable fact that optimal severance is independent of the distribution of ε. Optimal severance pay only depends on q, a property of the legal system, and C, the investment costs. It

seems natural to assume that q is the same for all the firms in a country. The investment cost C is probably firm-specific, however, one may think that the average value of C may vary from country to country. Hence our theory predicts that countries with a high value of q (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 3, and we evaluate its empirical relevance in Section 4. Note further that the wage tenure profile is steeper when the severance pay is higher.

The profit of the firm when implementing the project under optimal severance is

$$\begin{aligned}\Pi &= y - w_1 + \int_{\varepsilon^*}^{\overline{\varepsilon}} (\varepsilon + y - w_2) f(\varepsilon) d\varepsilon + (1 - F(\varepsilon^*))T \\ &= y - b + \int_{\varepsilon^*}^{\overline{\varepsilon}} (\varepsilon + y - b) f(\varepsilon) d\varepsilon - C\end{aligned}$$

where we insert for the participation constraint of the worker. Clearly, this is also the social value of the project. It follows that the project breaks even if and only if it is socially desirable. By construction, we know that the project is implementable. Hence we have shown the following proposition:

Proposition 3 *With socially optimal severance, $T = T^*$, the firm will implement the project if and only if this is socially desirable.*

Finally, note that with optimal severance, the firm clearly wants to fire a shirking worker. The cost of firing is qT, while the cost of keeping the worker is $w_2 - y$. Since $w_2 - y \leq w_2 - b = T$, it follows immediately that the firm fires shirking workers.

2.2 Many periods

We analyzed above how severance payments may be warranted in order to induce optimal layoff behavior of senior workers by firms. We will now extend the model to n periods, to capture the time profile of optimal severance and wages. We assume that the worker invests in every period, and shirkers are always fired in the next period. We show that in this case, the incentive compatibility constraint of the worker, as well as the optimal severance pay in period T, is determined in an analogous way as in the two-period model. In particular, the severance in period t only depends on the investment cost in the previous period, and the probability that a shirker "gets away with it" in the current period. However, with n periods, the notation as well as the analysis become more cumbersome, particularly for a general time profile of severance payments. We therefore focus our attention on deriving the optimal severance profile.

Suppose that the relationship lasts for n periods. Let C_t and b_t denote investment costs and per period outside option of the worker in the t th period of the employment relationship. Let q_t denote the probability that a worker who shirks in period $t-1$ gets away with it obtaining severance in period t. Let T_t denote the severance payment in period t. Finally, let β denote the discount factor.[10] In each period, the worker has to be incentivized.

We need first to introduce some notation. Let $W_t(\varepsilon_t)$, $\Pi_t(\varepsilon_t)$,and $S_t(\varepsilon_t)$ denote the expected NPV value of the income including period t of a non-shirking worker (net of investment costs), of the profit of the firm, and of the match surplus of the worker and the firm, as a function of ε_t (the expectation is taken with respect to future values of ε). Let B_t denote the NPV of the outside option, and ε_{dt} denote the period t cut-off for retaining the worker. Then we have that

$$\begin{aligned}W_t(\varepsilon_t) &= I(\varepsilon \geq \varepsilon_{dt})[w_t - C_t + \beta E W_{t+1}] + I(\varepsilon \leq \varepsilon_{dt})[B_t + T_t] \quad (9)\\ \Pi_t(\varepsilon_t) &= I(\varepsilon \geq \varepsilon_{dt})[y + \varepsilon_t - w_t + \beta E \Pi_{t+1}] - I(\varepsilon \leq \varepsilon_{dt}) T_t \quad (10)\\ S_t(\varepsilon_t) &= I(\varepsilon \geq \varepsilon_{dt})[y + \varepsilon_t - b_t + \beta E S_{t+1}]\\ B_t &= b_t + \beta B_{t+1}\end{aligned}$$

[10] As above, we make the assumption that the NPV value of the outside option in period t is equal to the NPV value of future b's, that is, the NPV of the outside option is $\sum_{j=t}^{n} b_j \beta^{j-t}$.

where $I()$ denotes the indicator function. Note that $ES_t = E\Pi_t + EW_t - B_t$. The firm chooses ε_d so as to maximize ex post profits, hence

$$\varepsilon_{dt} = w_t - \beta E\Pi_{t+1} - T_t - y$$

Let S_t^* denote the value of S_t contingent on efficient separation in period t and all later periods. Let ε_t^* denote the optimal threshold in period t. Efficient separation obtains if the separation threshold in period t is such that $S_t^* = 0$. It follows that for the last period,

$$\begin{aligned} \varepsilon_n^* &= b_n - y \\ ES_n^* &= \int_{\epsilon_n^*}^{\varepsilon^u} (y + \varepsilon_n - b_t) f(\varepsilon_t) d\varepsilon_t \end{aligned} \tag{11}$$

For all earlier periods

$$\begin{aligned} \varepsilon_t^* &= b_t - y + C_t - \beta ES_{t+1}^* \\ ES_t^* &= \int_{\epsilon_t^*}^{\varepsilon^u} (y + \varepsilon_t - b_t - C_t) f(\varepsilon_t) d\varepsilon_t + (1 - F(\varepsilon_t^*)) \beta ES_{t+1} \end{aligned} \tag{12}$$

Finally, it is convenient to define the *rent* associated with keeping a job in period t – prior to severance – as

$$\begin{aligned} R_t &= w_t - C_t + \beta EW_{t+1} - B_t \\ &= w_t - C_t - b_t + \beta E[W_{t+1} - B_{t+1}] \end{aligned} \tag{13}$$

R_t thus captures the value of keeping the job relative to being without a job in period t. The worker is indifferent between being retained and laid off if and only if $R_t = T_t$.

As the next lemma shows, efficiency in period t is ensured if the worker is indifferent between being separated and staying on in the firm in this and all future periods. Formally, if $T_k = R_k$ for all $k \geq t$. This is intuitive. If the worker, at any point in time, is indifferent between being fired and being retained, there are no externalities from the firm's retention decisions on workers, and a profit-maximizing firm makes the optimal decision.

Lemma 1 *Suppose the contract $w_t, , ..., w_n$ and the severance payments $T_t, ..., T_n$ are such that the worker is indifferent between being separated and not being separated. Suppose also that the worker is not shirking. Then the firm's retainment decision is optimal, i.e., $\varepsilon_j^d = \varepsilon_j^*$ for all $j \geq t$*

The proof is in the appendix. If the worker is indifferent between being retained and layed off, it follows that EW_t reads

$$\begin{aligned} EW_t &= w_t - C_t + \beta EW_{t+1} \\ &= R_t + B_t \end{aligned} \tag{14}$$

(from equation 13). Incentive compatibility in period $t-1$ requires that $C_{t-1} + \beta EW_t \leq \beta(q_t T_t + B_t)$, or, from (14), that

$$C_{t-1} + \beta R_t \leq \beta q_t T_t$$

Inserting for $R_t = T_t$ it follows that the lowest rent that is incentive compatible is

$$R_t = T_t = \frac{C_{t-1}/\beta}{1 - q_t} \tag{15}$$

Proposition 4 *The optimal severance pay in period t is given by* (15)*. 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 and the probability of being caught in any other periods.*

It follows that the severance is increasing with tenure if q_t is increasing with tenure or if C_t is increasing with tenure. Both seems reasonable, as discussed in Section 4.

We also want to study wage profiles. Inserting (14) into (13) gives $R_t = w_t - b_t - C_t + \beta R_{t+1}$. By inserting from (15) we get that

$$w_t = b_t + \frac{C_{t-1}}{\beta}\frac{1}{1-q_t} - C_t\frac{q_{t+1}}{1-q_{t+1}} \tag{16}$$

Let us look at some examples. First, suppose that $C_1 = ...C_{n-1} = C$, $q_1 = ... = q_{n-1} = q$, and $\beta = 1$. Then the wage equation (16) simplifies to

$$\begin{aligned} w_1 &= b_1 - C\frac{q}{1-q} \\ w_t &= b_t + C \qquad\qquad \text{for } t < n \\ w_n &= b_t + \frac{C}{1-q} \end{aligned}$$

In this case, the worker is compensated for the investment costs period per period. The carrot is the high wage in the last period, which prevents the worker from shirking. The worker pays for the carrot in the first period.

Suppose then that the cost in period C_t is $C_0 + t\Delta$, keeping q constant and $\beta = 1$. We start at $t = 0$ for convenience. It follows that the wage is given by

$$\begin{aligned} w_0 &= b_0 - C_0\frac{q}{1-q} \\ w_t &= b_t + C_0 + (t-1)\Delta - \frac{q}{1-q}\Delta \qquad\qquad t < n \\ w_n &= b_n + \frac{C_0 + (n-1)\Delta}{1-q} \end{aligned} \tag{17}$$

Hence wages are increasing over tenure by the same amount as the increase in per period investment costs. In the last period, the worker gets a large bonus, and this drags down wages in all earlier periods.

Finally, we assume that costs are constant and the discount factor is constant, but allow q_t to vary. More specifically, we assume that $n = 3$, and that $q_1 = 0$ and $q_2 = q > 0$. It follows that

$$\begin{aligned} w_1 &= b_1 \\ w_2 &= b_2 + C\frac{1-2q}{1-q} \\ w_3 &= b_3 + C\frac{1}{1-q} \end{aligned}$$

Wages are highest in period 3. If $q < 1/2$, wages are higher in period 2 than in period 1.

2.3 General equilibrium

In this section we endogenize the workers' outside option b. We do this by introducing a labor market with search frictions as in Mortensen and Pissarides [31], [32]. We assume that firms advertize wages as in Moen

[28]. The production technology in each firm is as described above. For notational simplicity we define period 0 as the first period.

Let $x(u,v) = Au^{\alpha}v^{1-\alpha}$, $A > 0$, $0 < \alpha < 1$ be a constant-returns-to-scale matching function mapping stocks of firms with vacancies, v, and unemployed workers, u, into a measure of new matches.[11] Let $p(\theta)$ denote the probability rate at which unemployed workers find jobs, and let $q(\theta)$ denote the probability rate at which firms with vacancies find workers. As above, b denotes the per period value of being unemployed, assumed constant, and including the gain from search. When a worker and a firm are matched, production takes place the next period. The production and separation process is as in the n-period case described above. We assume that workers live forever.

Above we defined R_t as the rent associated with continuing the relationship in period t. Ananlogously, let $R_0 = W_0 - \frac{1+r}{r}b$ denote the rents associated with getting the job. Note that R_0 is a "sufficient statistics" for workers regarding the attractiveness of a contract, as the expected income in net present value (NPV) terms is the only relevant variable when comparing wage contracts. Hence[12]

$$b = z + p(\theta)\beta R_0 \tag{18}$$

where z denotes the income during unemployment (home production or unemployment benefits).

Firms post vacancies and wages attached to them. A firm chooses the contract so as to maximize profits given the incentive compatibility constraint of the worker. However, the participation constraint of the worker never binds, as the firm in competitive search equilibrium has to offer the employee more than her outside option in order to attract applicants.

The value of a vacancy reads

$$V = -c + q(\theta)\beta(S_0 - R_0) \tag{19}$$

Hence, as in Moen and Rosen [30], the maximization problem of a firm can be decomposed into two subproblems:

1. For any given R_0, chose the contract so as to maximize profits Π_0. Since R_0 is given, this is equivalent to maximizing joint income S_0
2. Find the value of R_0 that maximizes V given by (19)

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. Hence, any allocation that can be achieved for one value of R_0 can be obtained for any value of R_0. It follows that maximum joint income $S_0^*(b)$ given by (11) can be achieved whenever (15) is satisfied, as in partial equilibrium.[13] In the following we assume that (15) is satisfied.

We turn now to the second part of the maximization problem. The optimal R_0 solves

$$V = \max_{R_0}[-c + q(\theta)\beta(S_0^*(b) - R_0)] \qquad \text{S. T. } b = z + p(\theta)\beta R_0 \tag{20}$$

Let K denote the cost of opening a vacancy. Equilibrium satisfies

$$V = K \tag{21}$$

General equilibrium is a vector (S_0^*, b, R_0, θ) satisfying (11),(18), (21) and the maximization problem in (20) 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. The following proposition thus follows:

[11] We impose that the number of matches is lower than $\min(u,v)$, and assume that the constraint does not bind.

[12] Recall that the gain from being matched is equal to the difference between NPV income when employed and when unemployed, which by definition is equal to R_0

[13] Since there is no uncertainty in the first period we skip the expectations operator before S. However, since b now is endogenous, we explicitely write S_0^* as a (decreasing) function of b.

Proposition 5 *Suppose that the severance pay is given by* (15)*. Suppose further that z reflects both the private and the social flow value of being unemployed. Then the general equilibrium allocation is efficient, independently of z and A*

Hence, even in general equilibrium, the optimal severance payment only depends on the investment cost, the discount factor, and the probability q of getting away with shirking.

Before we continue, we want to make two points. The first regards fiscal externalities. The proposition assumes that there are no fiscal externalities associated with unemployment. This is an ureasonable assumption, 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, they may find contractual arrangements that ensure that separation will indeed take place (for instance by agreeing *ex ante* that the worker returns part of the severance pay after separation). 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.

Our second point is more technical and regards the separation rate from employment. We have defined the equilibrium in terms of the job finding rate θ. 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 (12). Furthermore, $N_t = (1 - F(\epsilon_{t-1}^*))N_{t-1}$. Hence the average separation rate $\bar{s}$ in steady state reads

$$\bar{s} = \sum_{t=1}^{n+1} F(\epsilon_t^*)\Pi_{j=1}^{t}(1 - F(\epsilon_j)) \tag{22}$$

with $F(\epsilon_{n+1}^*) \equiv 1$. The unemployment rate is then $u = \bar{s}/(\bar{s} + p)$. Note that as $n \to \infty$, ϵ_t^* does not depend on t, the separation rate is independent of tenure.

3 Endogenous Court rulings

In this section we modify the setting so as to allow for the endogenous determination of the probability that a shirking worker can get away with it. To keep things simple we go back to the partial equilibrium, two-periods model, but results can be readily generalized to a n-periods setting.

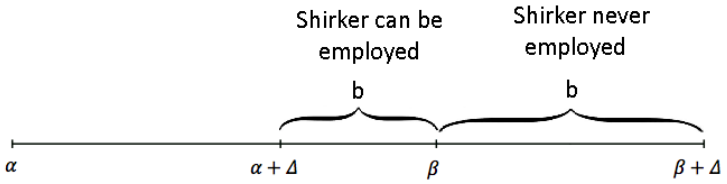
In the previous section, we assumed that the productivity distribution of a shirker was non-stochastic, and lower than b so that a shirking worker would always be fired. Here we relax these assumptions. Instead we assume that the investment in period 1 on the part of the worker shifts the distribution of productivity by a factor Δ, which is common knowledge. Specifically, the distribution of productivity in period 2 for a shirking worker is supposed to be uniformly distributed on $[\alpha, \beta]$, while the distribution of a non-shirking worker is uniformly distributed on $[\alpha + \Delta; \beta + \Delta]$. As the two distributions are identical with the exception of a shifter Δ in the case of a non shirking worker we shall indicate with $F(.)$ and F_Δ the cumulative density functions for the shirking and investing worker respectively. To make the problem interesting, we assume that the support of the two distributions has an area of overlap:

$$\Delta < \beta - \alpha \tag{23}$$

As in the baseline model, output in period 2 is only observed by the firm. The contract between the firm and the worker specifies the wage w_1, w_2 and it is fully enforceable. The firm has the right to fire the worker in period 2. The firm's decision to fire the worker reveals a low productivity level. Since the investment decision of the worker is unobserved, the issue is whether a fired worker should receive severance payment T or not.

Whether a firing is economic or disciplinary is settled by a Court ruling. We assume that the Court is able to accurately and freely establish the actual draw of the productivity in period 2, but can not directly

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



observe whether the worker did invest in period 1. All the Court can observe is the productivity y that- at least for some range- can be consistent with both investments and shirking. We return to the issue of Court monitoring in section 3.4.

As discussed in Section 1, countries differ based on whether the *burden of proof* is on the employer or partly also on the worker. We are now going to assess the implications for optimal severance of this key design feature of legal systems. When the burden of proof is on the firm, in case of ambiguity between the two types of dismissals, the Court will rule in favor of the worker and will declare the dismissal as economic. Conversely, when the burden of proof is on the employer, ambiguity is solved in favor of the employer and the case will be ruled as disciplinary. We say that a shirker is *distinguishable* from a non-shirker if her realized productivity is outside the support of the investing worker's productivity. Analogously, an investing worker is distinguishable from a shirker if her productivity is outside the support of a shirking worker's productivity. The rest of the setting is analogous to the model of Section 2.

To simplify the exposition, we shall operate only around the efficient solution, so that in the various cases we solve the reservation productivity at the level $y^*(T) = b$. In other words, T solves

$$T = w_2(T) - b$$

where $w_2(T)$ is determined by the worker ICC. Recall that at this level of severance, the worker is indifferent between being fired and retained.

As the reservation productivity b is an exogenous parameter, the model is fully solved by the triple w_1, w_2 and T. As mentioned above, w_2 is obtained by the ICC while w_1 by the participation constraint. The severance payment T is obtained by equation (6) to ensure coherence between w_2 and the efficient reservation productivity.

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 on the employee. For each burden of proof setting we shall consider two cases[14], related to the position of the parameter b to the distribution of productivity realizations (see Figure 2).

3.1 Burden of proof on the employer

When the burden of proof is on the employer, the Court has a bias toward economic dismissal. In any case there is a situation of doubt, the Court rules that the worker did invest and the severance payment is due.

Shirker never employed: $\beta < b < \beta + \Delta$

When the burden of proof is on the employer, the fim has to pay severance to the shirker as long as she is not distinguishable from an investing worker. In case of ambiguity the firm has to pay the severance payment as depicted in the top panel of Fig 3. Even though s shirking worker is never employed, she receives severance

[14]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

with probability $1-F(\alpha+\Delta)$ (the probability that her productivity is within the support of the productivity of the non-shirker). Hence

$$\begin{aligned} W^s &= w_1 + b + T(1 - F(\alpha + \Delta)) \\ &= w_1 + b + qT \end{aligned} \tag{24}$$

where $q = 1 - F(\alpha + \Delta) = 1 - \frac{\Delta}{\beta - \alpha}$ is the probability that the worker gets away with it. The pay-off to the shirker is thus as in the original version of the model. An investing worker gets

$$W^I = w_1 - C + b + T \tag{25}$$

since she is indifferent between being fired and retained. The ICC $W^I(w_2) = W^s(w_2)$ immediately gives

$$T^* = \frac{C}{1-q} = \frac{C(\beta - \alpha)}{\Delta} \tag{26}$$

The optimal contract is identical to the contract derived above, the only difference is that q is endogenously determined by the distribution F and the shifter Δ.

Shirker can be employed: $\alpha + \Delta < b < \beta$

In this case a shirking worker not only can get away with it, but also she can be employed in period 2, while an investing worker can be fired in period 2. However, both shirkers and non-shirkers are indifferent between being fired with severance and being retained. The expected utility for a shirking worker and an investing worker are still given by (24) and (25), respectively and the optimal severance by (26). Hence, also in this case the result from the previous section is replicated.

Proposition 6 *When the burden of proof is on the employer, the optimal severance is given by* $T = C/(1-q)$, *where* $q = 1 - \frac{\Delta}{\beta - \alpha}$

3.2 Burden of proof on the worker

When the burden of proof is on the worker, the behavior of the Court is different. In this case a shirker never gets severance, while an investing worker gets severance if she is distinguishable from a shirker. Hence, whenever there is a situation of doubt, the Court rules that the worker did non invest and the severance payment is not due.

The same two cases depicted in Figure 2 apply in this context.

Shirker never employed: $\beta < b < \beta + \Delta$

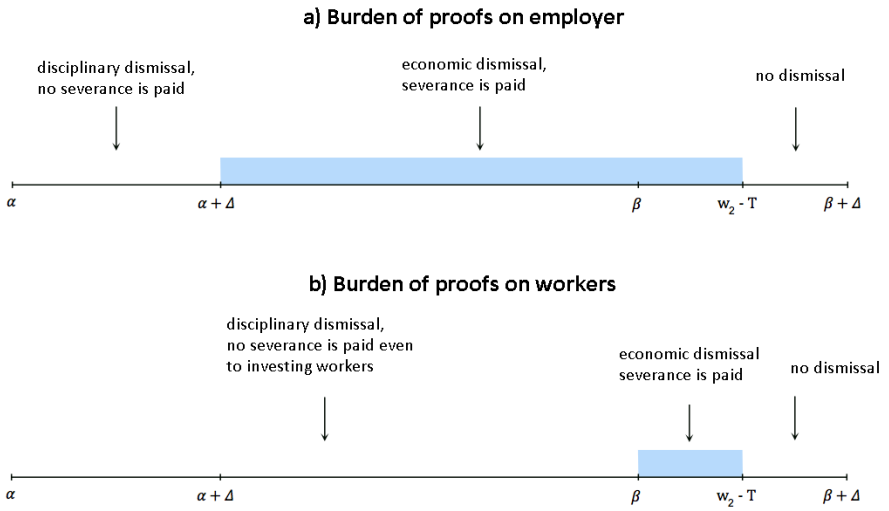
The firm will employ the worker if her productivity exceeds the period 2 wage, and the worker receives severance if her productivity is sufficiently high so that she is distinguishable from a shirker. Hence the probability that the investing worker either gets a wage or a job is $1 - F_\Delta(\beta) = \Delta/(\beta - \alpha) = 1 - q$ (recall that $q = 1 - \Delta/(\beta - \alpha)$). This is not surprising: q is the probability that the shirker's productivity falls within the support of the investor's productivity, hence $1-q$ is the probability that the shirker's productivity is outside it, and hence distinguishable from an investor's productivity. We assume that the severance, when paid, is still equal to the difference between the inside and outside payment, $T = w_2 - b$. Also in this case, there is an *unfair economic dismissal* and we address it in the next section. The value functions read in this case

$$W^s = w_1 + b$$

and

$$W^I = w_1 - C + b + (1 - q)T$$

Figure 3: Court decisions
Case where $\beta < b < \beta + \Delta$



Since $1 - F_\Delta(\beta) = q$, it follows that we can write

$$T^* = \frac{C}{1-q} \tag{27}$$

Hence, even in this case, the formula from the simple model still holds. To give intuition, note that the fact that the investing worker does not get severance when he is indistinguishable from the shirking worker reduces the pay-off of investing with qT. On the other hand, the fact that the shirking worker never gets away with it reduces the value of shirking with qT. Hence the two effects cancel out, and the ICC wage and the corresponding severance do not change. However, the reduced expected period-2 income of the investing worker is accompanied by an equally large increase in the period 1 wage so that the total value of the job remains constant.

Shirker can be employed $\alpha + \Delta < b < \beta$

This case is graphically depicted in the bottom panel of figure 3. In this case, efficient hiring cannot be achieved. If it were, the firm could fire a worker with productivity b, and since a worker who has invested cannot distinguish herself from a shirker at this productivity level, she will not get severance. Hence the firm would have an incentive to fire her.

Suppose that the ICC constraint is satisfied at wage w_2^*, and that $w_2^* < \beta$. Then the difference in the period 2 pay-off between the shirker and the non-shirker is $(w_2 - b)\Delta/(\beta - \alpha) = (w_2 - b)(1 - q)$ (the wage premium times the difference in probability of being hired), which must be greater than or equal to C for the ICC to be satisfied. The ICC wage again writes

$$w_2^* = b + \frac{C}{1-q} \tag{28}$$

If $\beta < w_2^*$, the firm can commit to pay severance if the investing worker's productivity exceeds β.

With $b < \beta$, the Government may use other instruments to achieve efficient separation. By imposing a firing tax equal to the difference between the period 2 wage and the outside option, efficient firing is restored. For instance, if $\beta > w_2^*$, a firing tax of $w_2^* - b$ 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 is always the results of a non-investment on the part of worker. In this section we still consider the case in which $b > \beta$ 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 (27) and the firm wants to fire the worker, so that the productivity is below b. Let $\hat{x}$ denote the declared productivity and x the true productivity. Assume further that the Court can audit the firm with probability λ and impose a severance payment T_U. In other words, T_U is the severance payment *for unfair economic dismissal*. If the firm reports that $\hat{x}$ is below β, even though it is above it, its expected profits read

$$\hat{\Pi}(\hat{x} < \beta | \ \ \beta < x < b) = -\lambda T_U$$

Conversely, truth telling for the firm yields

$$\Pi(\hat{x} = x | \ \ \beta < x < b) = -T$$

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

$$T_U = \frac{T}{\lambda} \tag{29}$$

where T_U 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_U and $T = T_E^F$ 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

$$T_U = \frac{C(\beta - \alpha)}{\lambda \Delta} \tag{30}$$

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

4 Discussion

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.

4.1 Severance-tenure and wage-tenure profiles

In our model severance operates as a commitment device reducing the probability that workers investing in the productivity of the match, e.g., undergoing costly training, are laid-off in case of a negative productivity realization. Optimal severance is increasing with tenure when either the costs of this investment in work capacity or the probability of getting away with it are increasing with tenure.

There is ample evidence that employees' self-assessed working capacity is declining with age [18]. This may or may not translate into a declining pattern of productivity with age, but for graded security to be

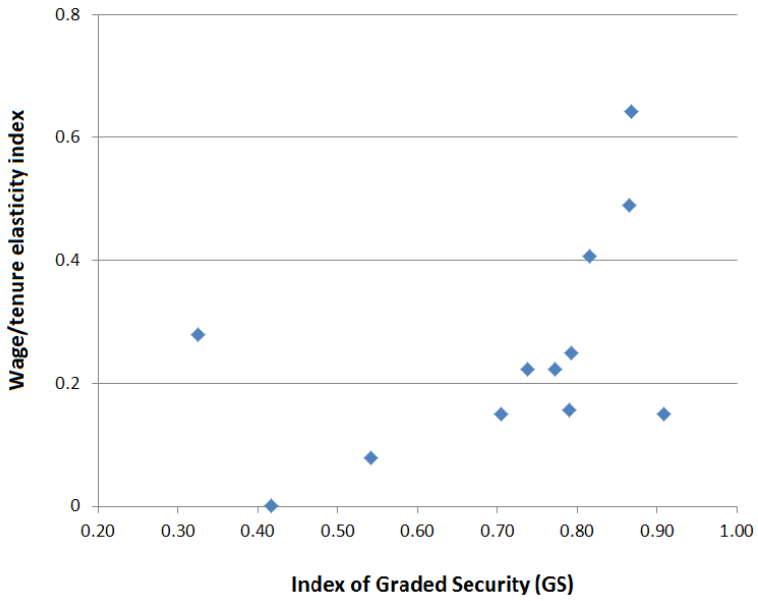
optimal what matters is that the investment in working capacity is perceived by the workers to be increasingly costly with age or tenure.

In presence of training costs 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[15] 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

$$\ln w_i = \alpha + \beta_1 \tau_i + \beta_2 \tau_i^2 + X_i \gamma + \epsilon \tag{31}$$

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

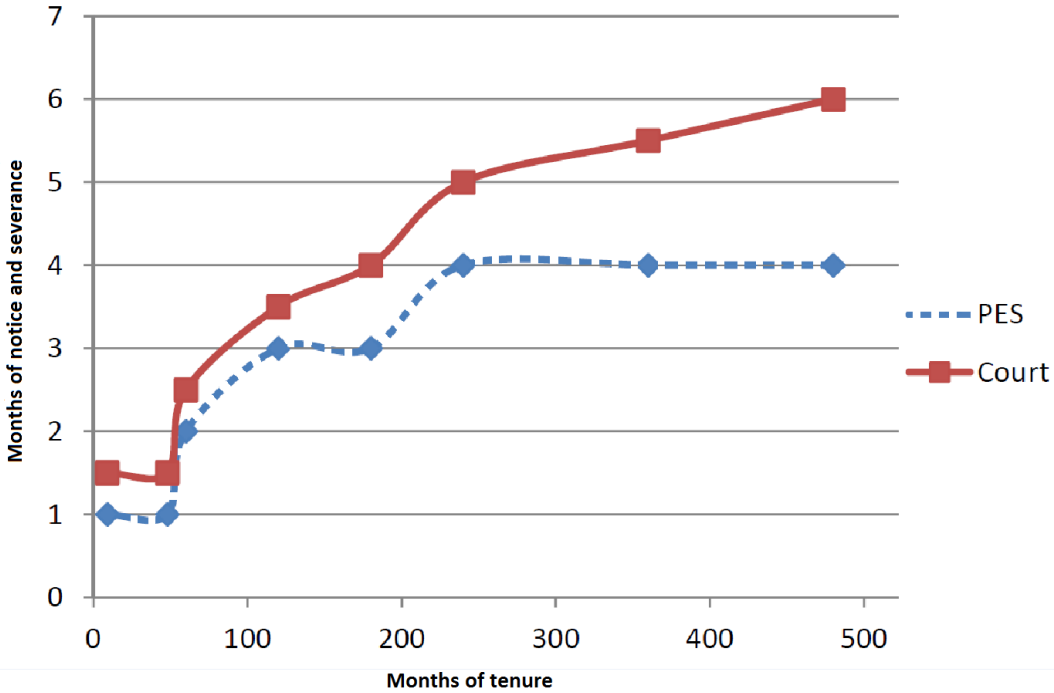
[15] No causality is involved here as both severance and wages depend on the costs of training.

being brought to Courts and the fraction of labour disputes ending with a Court ruling favourable to the worker appear to be higher during cyclical downturns and in relatively depressed labor markets [5], an indication that judges are more protective of workers under these circumstances. Insofar as senior workers, 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 implication 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 labour 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. In case of approval by the PES, the dismissal costs for the employer costs are independent of the nature of the dismissal, disciplinary or economic, provided that the length of the procedure is the same. The alternative route is represented by the involvement of Courts, which allows for a shorter procedure (about one month shorter than with the PES), but requires the payment of a Court fee, in addition to the mandatory severance in case of economic dismissals.

The obvious question is why employers do ever go to the Court given that it is always more expensive. 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. Figure 5 displays the notice and severance costs (if any) to be paid by the employer for fair economic dismissals in the two cases by tenure length, according to the legal rules and the informal guidance followed by the Courts: the procedure involving the judiciary is clearly increasing more steeply with tenure than the PES route. Pfann [34] estimates that the actual differences in the amount of the transfers between the two procedures are even higher than those displayed in the diagram, as Courts tend to use the informal guidance only as a floor in establishing the level of the severance.

Figure 5: Transfers to the worker in case of economic dismissal by tenure, Netherlands



Note: PES= Public Employment Service procedure.

4.2 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. Previous work had found that the organization of legal systems, notably the legal origin of countries played an important role in labor market outcomes [10]. 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, efficiency of judicial systems.

The first two panels at the top of the Figure 6 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. Both indicators appear to be positively correlated with the mandated months of severance in case of fair economic dismissals.

The other four panels of Figure 6 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. Our measure of the expected severance in case of unfair dismissals, T_U, is also 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, we also 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.

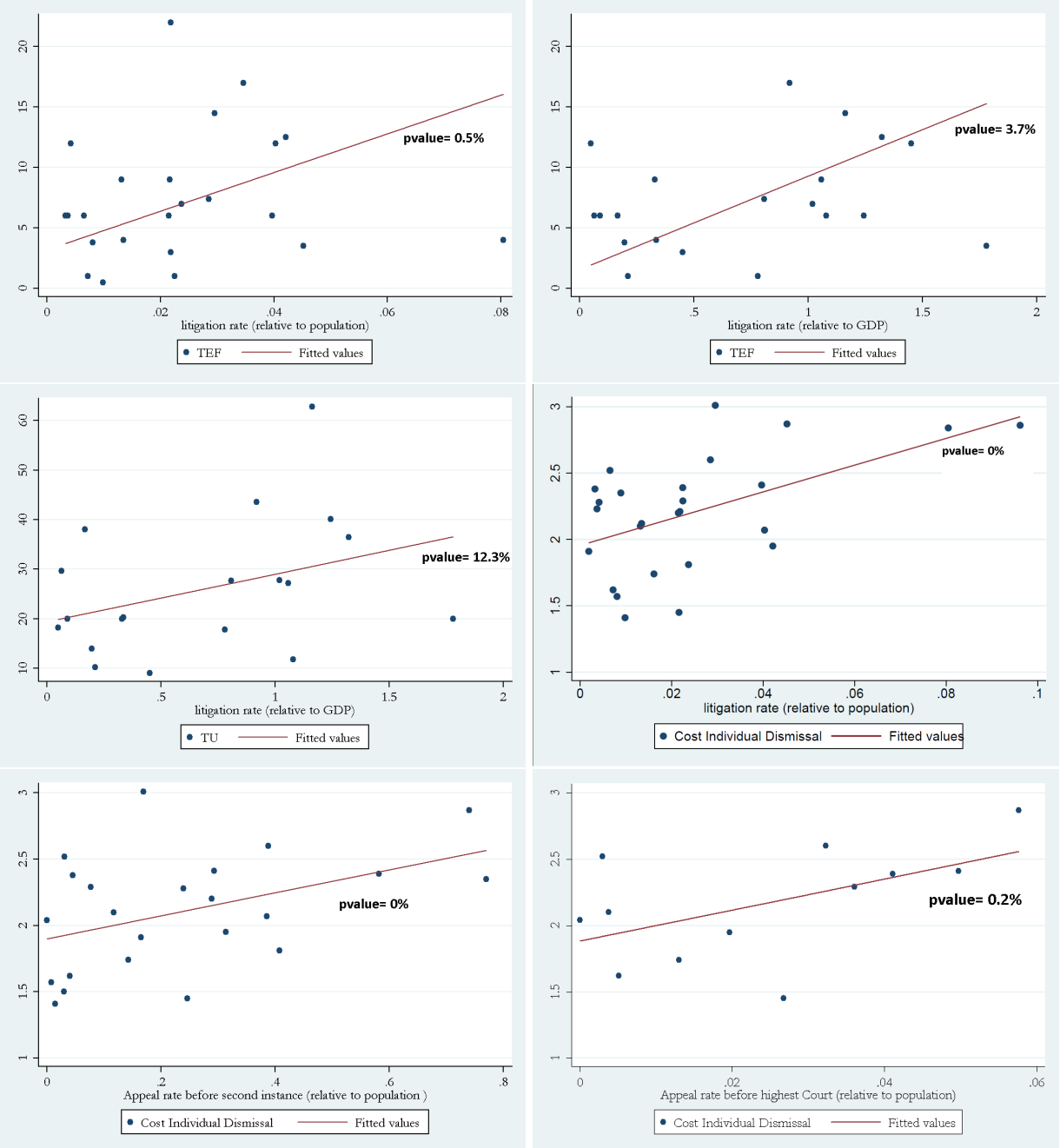
Notice finally that the OECD indicators of the efficiency of the judiciary are positively correlated with subjective evaluations of public opinion as to the quality of the legal system collected within the World Value Survey (Palumbo 2013 [33]).

5 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

Figure 6: Compensation for dismissal and judicial efficiency



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 and appeal rate from Palumbo [33].

pay, that is, for the fact that investment costs or the probability of getting away with it are increasing with tenure.

Graded employment security schemes in our model deal with the moral hazard associated with the initial investment in training. 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", that is, the coexistence of a highly protected segment of the workforce and one segregated into temporary jobs providing low, if any, employment security. The theory presented can certainly be used to assess and evaluate some of the reforms currently under discussion in various Southern European parliaments, including Italy. 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.

References

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

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

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

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

[5] 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.

[6] 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.

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

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

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

[10] 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.

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

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

[13] 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.

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

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

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

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

[18] 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.

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

[20] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[35] 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.

[36] 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.

[37] Rühmann, P. and Südekum, J. 2001. Severance Payments and Firm-Specific Human Capital, Discussion Paper n. 111 of the University of Göttingen.

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

[39] 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 [33]). 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[16].

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

$$\Sigma = \sqrt{\alpha((1-p)T_U^2 + p^2 {T_F^D}^2 + p(1-p){T_F^E}^2 - ((1-p)T_U + p^2 T_F^D + p(1-p)T_F^E)^2)} \tag{32}$$

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

countries (Austria and Japan) where the elasticity is zero throughout a 20 years tenure length, denoting a flat severance-tenure profile. [17] 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

$$GS = \sum_{t=0}^{4} \frac{\Delta T_t}{\Delta \tau_t} * \frac{\tau_{t+1}}{T_{t+1}} * \frac{\Delta \tau_t}{240} \tag{33}$$

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.

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

[17] 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.

Proof of proposition 1

For any k, the right-hand side of (4) 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 (4) 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.

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 (4) 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.

Proof of lemma 1

The proof is by induction. In period n, the worker is indifferent between being separated and not if $T_n = w_n - b$. It follows from (10) that the firm in this case retain the worker iff $\varepsilon \geq b_n - y$, which is efficient. Consider an earlier period t, and assume that the firm makes the optimal decisions in all later periods. The NPV profit of the firm by contintuing is $E\Pi_{t+1} = ES^*_{t+1} - EW_{t+1} + B_{t+1}$. The firm is thus indifferent between retaining and firing the worker whenever

$$w_t - y - \varepsilon_t + \beta(ES^*_{t+1} - EW_{t+1} + B_{t+1}) = T_t \tag{34}$$

Inserting $T_t = R_t \equiv w_t - C_t - b_t + \beta(EW_{t+1} - B_{t+1})$ into(34) gives

$$\varepsilon_t = b_t - y + C_t - \beta ES^*_{t+1} \tag{35}$$

Which is identical to (12). The proof is thus complete.