DISCUSSION PAPER SERIES

IZA DP No. 7455

The Economics of Severance Pay

Tito Boeri
Pietro Garibaldi
Espen Moen

June 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Economics of Severance Pay

Tito Boeri
*Bocconi University
and IZA*

Pietro Garibaldi
*Collegio Carlo Alberto
and IZA*

Espen Moen
Norwegian Business School

Discussion Paper No. 7455
June 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7455
June 2013

ABSTRACT

The Economics of Severance Pay

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 such transfers are either ineffective or highly distortionary. In this paper we show that mandatory severance is optimal in presence of wage deferrals when there is moral hazard of employers and workers, notably when employers cannot commit not to fire a non-shirker and shirkers can also get away with it. Our model also accounts for two neglected features of EPL. The first is that dismissal costs depend not only on whether the dismissal is deemed fair or unfair, but also on the nature, economic vs. disciplinary, of the layoff. The second feature is that compensation for unfair dismissal or severance is generally increasing with tenure.

JEL Classification: J63, J65, J33

Keywords: severance, unfair dismissal, graded security

Corresponding author:

Tito Boeri
Department of Economics
Università Bocconi
Via Roentgen 1
20136 Milano
Italy
E-mail: tito.boeri@unibocconi.it

Introduction

Most OECD countries have legally mandated severance pay in case of employer initiated job separations. Such transfers from the employer to the worker are a very important component of dismissal costs as they account 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. 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.

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, 1990). Under rigid wages, severance pay generates unemployment (Garibaldi and Violante, 2005) operating de facto as a firing tax. 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, 2008).

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, provided that there are wage deferrals and the worker can decide not to invest in the productivity of the job. Under these conditions, severance deals with the moral hazard problem related to investment in job-specific productivity. 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 for incentive purposes and these tenure-related components of compensation are agreed in advance, conditional on the continuation of a job, but independent of productivity.

Our model also allows to explain two neglected features of EPL. The first relates to the role played by the nature of the dismissal. Compensation offered to workers in case of dismissals varies depending on the alleged reasons for the layoff as well as its "‘fairness ". 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 often higher than the severance for fair economic dismissals. There are also countries in which compensation is provided only for unfair 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 could concern not only the fairness of the layoff, but also its nature, economic or disciplinary (let alone discriminatory). 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 stochastic court ruling clearly affects also private settlements out of court, as such settlements will be based on the expected costs had the case gone to court.

The second neglected characteristic of EPL is the *tenure profile* of severance pay. Most countries allow for mandated severance pay to be increasing with tenure. A very few countries provide compensation to workers independently of tenure.

We show that these design features are efficient in dealing with moral hazard of workers and employers. Severance is needed as moral hazard of employers prevent them from committing not to fire workers investing in the productivity of the job. At the same time, the stochastic nature of dismissal costs and the distinction between economic and disciplinary dismissals increases the moral hazard problem of workers as they can "get away with it" even when they shirk.

Our model shows under which conditions – in terms of productivity, monitoring technologies, jurisprudence, and design of unemployment benefit systems – tenure-related severance pay increases productivity, reduces inefficient firing and induces an efficient allocation of labor. In particular, optimal severance is according to our model related to the efficiency of the judicial system: the larger the probability that a worker can get away with it, the larger is severance pay. This also rationalizes 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 relevant 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.

The plan of the paper is as follows. Part one discusses the relevance of severance pay and characterizes neglected features of EPL, notably the treatment of fair vs. unfair economic and disciplinary dismissals, and the tenure profile of severance pay. Part two develops a simple two-period model and evaluates optimal severance pay. Part three extends the model to three periods and endogenizes the detection probability. Finally, part four summarizes our key results and discusses their empirical relevance.

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 (1990) and Bertola (1990), typically treats EPL as a firing tax to be 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 (1990) 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 account for a very large share of the costs of dismissals. According to Garibaldi and Violante (2005) who carefully estimate the red tape costs of layoffs in countries like Italy, severance pay accounts for about 2/3 of total dismissal costs. Severance also explains about 50 per cent of the total cross-country variation in the OECD strictness of EPL

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 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 payment.

1.1 Stochastic severance

The severance pay levels used so far in characterizing the tenure profile of employment protection are those established by the law in case of *fair economic* individual dismissals. Fairness in the case of *economic dismissals* refers to the behaviour of the employer: she should have tried as much as possible to avoid this outcome. Although the definition of fair economic dismissal differs quite considerably from country to country, it generally implies that some "genuine and serious" exogenous shocks in firm's performance require "operational changes" in the scale, and possibly, nature of the work organization, making the worker involved redundant. Often evidence of "economic difficulties" or "technological change" is explicitly required.
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 to be "unfair", the employer in some countries is forced to reinstate the worker. Generally the reinstatement does not take place, but 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

[1] See Boeri and vanOurs (2013) [3] 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 even in case of a voluntary separation of the worker.

during the legal procedure. The decisions as to the nature of the dismissal (economic vs. disciplinary) and its fairness require some Court ruling. In practice, disputes are mostly settled before the Court ruling, 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 dependent on the evidence that the employer can provide for a disciplinary or economic dismissal. The theoretical literature on EPL, recently reviewed by Boeri and vanOurs [3] generally treats severance as a deterministic transfer from the employer to the employee. In the few cases where stochastic severance is allowed (Garibaldi [8], Malo [11]) , 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 [7] and Boeri [4]. However, Galdon-Sanchez [7] operates on a reduced form model and both Boeri [4] and Galdon-Sanchez [7] do not address the efficiency of severance pay, 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 required under unfair, fair economic and fair disciplinary dismissals 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 [6] and a study by the ILO [2].

As shown by Table 1, in all countries even fair dismissals command some compensation to the worker, either in terms of strictly severance pay or of a minimum notice period (de facto an extension of pay after the date when the worker is made redundant). The severance for unfair dismissals (first column) is, however, always higher than that provided in case of fair dismissals (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 reinstatement of the worker in the ranks of the firm. Thus, in these countries, the costs of unfair dismissals should also 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. An additional compensation should also be considered 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 the costs of unfair dismissals in countries with reinstatement include the average length of the trial period (d) plus notice period and actual compensation for unfair dismissal at 20 years as an alternative (C) to the reinstatement itself, multiplied by the likelihood that a reinstatement is actually granted (π), i.e.:

$$T_U = S + \pi(d + S) \tag{1}$$

where T_U are the costs of unfair dismissals, and S as usual includes both notice period and actual severance. Moreover, fair economic dismissals generally involve a higher compensation that fair disciplinary dismissals (fourth column displaying the difference between the two). Typically, only a relatively short notice period is required in the case of firm initiated job separations when they are due to workers' misconduct.

As detailed in Table A2 in the annex, we attribute to π the value obtained by standardizing in the unit interval the 0-3 OECD index on the likelihood of the reinstatement, where clearly 0 means never reinstatement and values close to one denote a frequent reinstatement. Finally, in the fourth and fifth columns of Table 2, we provide a summary measure of dispersion given by the weighted coefficient of variation of the compensation required for economic and disciplinary dismissals. In defining the weights, that is, the probability that a dismissal is ruled as unfair, we consider whether the burden of proof is on the employer or, at least partly, on the worker (see Table A2). In the former case, we arbitrarily assume that the probability, p that a dismissal is ruled as unfair is .5. This probability is arbitrarily reduced to .25 when the burden of proof is on the worker. An intermediate value is taken in the only case (Japan) where the burden of proof may fall either on the worker or the employer. All in all, our measure of uncertainty in the costs of dismissals (CV) is given in each country by

$$CV^J = \frac{\sqrt{p(T_U - \bar{T})^2 + (1-p)(T_F^J - \bar{T})^2}}{\bar{T}} \tag{2}$$

where the superscript J denotes either disciplinary (D) or economic (E) dismissals and $\bar{T}$ is the weighted average of the compensation when the dismissal is considered as unfair (T_U) or fair (T_F^J) respectively. The last two columns on the right-hand-side of Table 1 suggest that there is considerable uncertainty as to the actual costs of dismissal: the cross-country average for the coefficients of variation is 1.5 in case of economic dismissals and 1.9 in case of disciplinary dismissals. Moreover, uncertainty increases if the employers takes the route of the disciplinary dismissals, involving in some countries a standard deviation three times as large as the average, than that of economic dismissals.

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 and the nature, economic vs. disciplinary, of the dismissals. We will initially assume that the probability of proving the case before a Court 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. We will also show that, under some configurations of inter-temporal productivity and probability that the dismissal is treated as fair, this optimal severance is increasing with tenure. Later on, we will endogenize 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.

1.2 The elasticity of severance to tenure

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 for private sector workers at different tenure lenghts, 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 has been gathered: tenure at nine months; at one, five, ten and twenty years. At each tenure lenght, 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. The largest and the lowest of these elasticities (and their correspondent tenure levels) are displayed in the first two columns of Table 2. 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 lenght, denoting a flat severance-tenure profile. In Denmark, New Zealand and the US, there is no mandatory severance, hence the elasticity is not defined. In the other countries a flat severance-tenure profile is observed only limited to some tenure lenghts.

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 lenght of each tenure interval. Finally we normalized these overall apparent elasticities to obtain a unit value for a proportional severance scheme at all tenure lenghts (one having always a unit apparent elasticity). Formally, denoting by S the months of mandatory severance (and compulsory notice period), by τ months of tenure, and by indexing the tenure classes by subscript t, our index of *Graded Security* is given in each country by

$$GS = \sum_{t=0}^{5} \frac{\Delta S_t}{\Delta \tau_t} * \frac{\tau_t}{S_t} * \frac{(\tau_t - \tau_{t-1})}{240} \tag{3}$$

where $t = 0$ denotes the beginning of the tenured contract, $t = 1$ denotes 9 months of tenure, $t = 2$ corresponds to one year of tenure, $t = 3$ to five years, $t = 4$ to ten years and, finally, $t = 5$ to twenty years

Table 1: Severance and nature of dismissal

Country	T_U	T_F^E	T_F^D	$T_F^E - T_F^D$	CV^E	CV^D	p
Australia	11.3	3.8	1.0	2.8	1.4	2.0	0.5
Austria	25.9	4.0	4.0	0.0	1.8	1.8	0.5
Belgium	31.3	21.0	21.0	0.0	1.1	1.1	0.5
Canada (Federal)	-	4.3	2.0	2.3	-	-	0.5
Czech Republic	22.0	3.5	2.0	1.5	2.2	2.7	0.25
Denmark	19.8	9.0	6.0	3.0	1.3	1.6	0.25
Finland	20.0	6.0	6.0	0.0	1.5	1.5	0.5
France	23.2	7.4	2.0	5.4	1.6	2.7	0.25
Germany	35.4	17.0	7.0	10.0	1.2	1.7	0.5
Greece	-	12.0	4.0	8.0	-	-	0.5
Hungary	21.5	9.0	3.0	6.0	1.3	2.3	0.25
Ireland	38.7	6.0	2.0	4.0	1.8	2.1	0.5
Italy	42.9	6.0	6.0	0.0	1.8	1.8	0.5
Japan	10.3	1.0	1.0	0.0	2.2	2.2	0.375
Korea	17.0	1.0	1.0	0.0	2.9	2.9	0.25
Luxembourg	12.0	12.0	6.0	6.0	1.0	1.2	0.5
Mexico	-	-	-	-	-	-	0.5
Netherlands	14.9	4.0	4.0	0.0	1.5	1.5	0.5
New Zealand	11.8	0.5	0.5	0.0	2.1	2.1	0.5
Norway	34.0	6.0	6.0	0.0	1.7	1.7	0.5
Poland	9.0	6.0	3.0	3.0	1.1	1.4	0.5
Portugal	49.6	14.5	2.5	12.0	1.5	2.1	0.5
Slovak Republic	19.0	7.0	3.0	4.0	1.4	2.2	0.25
Spain	24.5	12.5	0.5	12.0	1.2	2.2	0.5
Sweden	38.0	6.0	6.0	0.0	1.8	1.8	0.5
Switzerland	9.0	3.0	3.0	0.0	1.5	1.5	0.25
Turkey	12.0	22.0	2.0	20.0	1.2	1.7	0.5
United Kingdom	19.3	7.6	3.0	4.6	1.3	1.8	0.5
United States	-	0	0	0.0	-	-	0.5

Sources: EPLex; OECD (2013);
See equations 1 and 2 in the main text and table A2 in annex for details.
Time is expressed in months. Reference is made to a worker with 20 years of tenure.

Figure 1: Severance payments and tenure

of tenure. The last column on the right-hand-side of Table 2 provides the value of this index for the OECD countries. Figure 1 displays the severance tenure profiles for the same set countries.

We find that 18 countries out of 29 display an index above 50 per cent. In the two countries paying the same severance at all tenure levels (Austria and Japan), the index is clearly zero and the country-specific diagram displays a flat line. Relatively low levels for GS would also be observed in case of a markedly concave severance - tenure profile, as severance at longer tenures has a larger weight in our index. Both Figure 1, and Table A1 in the Annex displaying the apparent elasticities at each tenure lenght, suggest that a number of countries starting up with a relatively high level of severance and then experiencing a mild, but steady, increase of severance with tenure. This generates a relatively low GS. Finally, in Denmark, New Zealand, and the US no index is computed as in these countries costs of dismissals can only be imposed via collective bargaining (OECD Indicators of Employment Protection, 2008). In other words, there is no legally mandated severance pay at all.

Overall, the GS index documents that most countries have mandated severance pay (and statutory notice periods) increasing with tenure. 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 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.

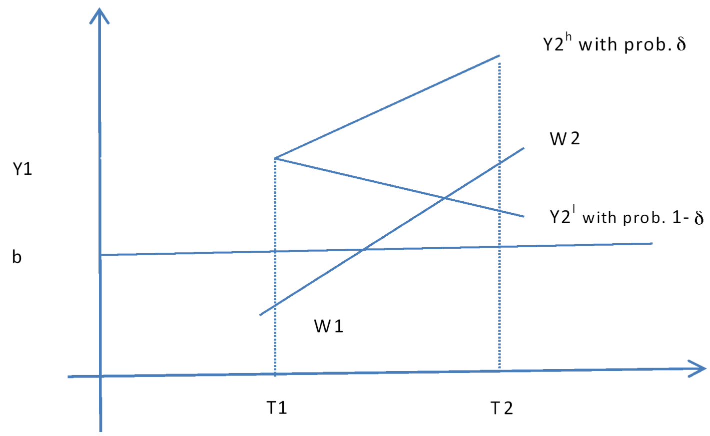
Personnel economics offers explanations for why firms offer *tenured jobs*, that is, positions that cannot

Table 2: Min and max apparent elasticity and related tenure and GS Index

	Minimum Elasticity	Related Tenure (months)	Maximum Elasticity	Related Tenure (months)	Elasticity Range	GS Index
Australia	0.0	240	3.3	12	3.3	0.31
Austria	0.0	each tenure lenght	0.0	each tenure lenght	0.0	0.00
Belgium	0.2	60	1.0	9	0.8	0.70
Canada	0.7	240	1.0	9, 60	0.3	0.84
Czech Republic	0.0	12, 60,120, 240	0.6	9	0.6	0.02
Denmark	*	*	*	*	-	-
Estonia	0.0	240	1.0	9, 12	1.0	0.26
Finland	0.6	60	1.0	9	0.4	0.77
France	0.0	9	0.9	240	0.9	0.74
Germany	0.0	9	1.4	12	1.4	0.91
Greece	0.7	60, 120	4.0	12	3.3	-
Hungary	0.0	9, 12	0.9	60	0.9	0.71
Ireland	0.7	60	2.2	12	1.5	0.79
Israel	0.8	60	3.1	12	2.3	0.95
Italy	0.0	9, 12,60	0.5	120	0.5	0.33
Japan	0.0	each tenure lenght	0.0	each tenure lenght	0.0	0.00
Korea, Rep.	0.0	9	2.0	12	2.0	0.90
Luxembourg	0.0	9, 12	0.8	60, 120	0.8	0.67
Mexico	0.1	12.0	1.0	9	0.9	0.63
Netherlands	0.0	9, 12	0.7	120	0.7	0.54
New Zealand	*	*	*	*	-	-
Norway	0.0	9, 12, 240	0.7	120	0.7	0.29
Poland	0.0	9, 12, 120, 240	0.8	60	0.8	0.17
Portugal	0.5	9, 12	1.0	9	0.5	0.82
Slovak Republic	0.0	12, 120	0.5	9	0.5	0.10
Slovenia	0.0	9	1.1	240	1.1	0.85
Spain	0.5	9	0.9	60, 120, 240	0.4	0.87
Sweden	0.0	9. 12, 240	1.0	120	1.0	0.42
Switzerland	0.0	60, 240	1.0	9	1.0	0.23
Turkey	0.0	9	2.1	12	2.1	0.87
United Kingdom	0.0	9, 12	1.1	60	1.1	0.79
United States	*	*	*	*	-	-

Source: OECD (2012), World Bank Data (2012)
Notes: * : No mandatory severance

Figure 2: An illustrative example



be severed under any set of circumstances. Tenured jobs can be rationalized as the result of learning about match quality, building on Jovanovic (1979) [9] matching model, thereby firms, after observing a sufficiently long string of positive signals on the productivity of a worker decide to retain the employee, irrespective of future negative signals. Another explanation provided for offering tenured jobs is in terms of 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, 1988)[5]. 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 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 both the "fairness" and the nature (economic vs. disciplinary) of the dismissal have to be proved before a Court.

2 The Basic Economics of Severance Payments with Unfair Dismissal

2.1 A model without specific investment

A simple motivating example, illustrated in Figure 2, delivers our key intuition as to the inefficiency of separations without severance. Consider a two periods model, where the workers' outside option is b in both periods. To keep things simple, assume that there is no discounting and there is no specific investment. Wages are given, but for some exogenous reasons, are deferred. Denote the first period and the second period wages by $w_1 < b$ and $w_2 > b$ respectively.

Productivity is y_1 in the first period while in the second period is stochastic: it can take value y_2^h with probability δ and value y_2^l with probability $1 - \delta$. Assume that

$$w_2 > y_2^l > b$$

Firms can fire in the second period conditional on the realization of the shock. Firing requires a severance payment $T \geq 0$. It follows that expected profits of the firm (the surplus of the employer) are given by:

$$\Pi = y_1 - w_1 + (1-\delta)[y_2^h - w_2] + \delta Max[y_2^l - w_2, -T]$$

The total surplus is given in each period by $y - b$. Since $y_2^l > b$ by assumption, the joint surplus is positive in the second period whatever the realization of the shock, and hence, for efficiency reasons, production should take place (both y_2^h and y_2^l are above the horizontal line at b in Figure 2. However, since $w_2 > y_2$, when $T = 0$,firms always fire conditional on a adverse shock.

The following remarks and proposition then immediately follow.

Remark 1 *A severance payment $T^* \geq w_2 - y_2^l$ prevents inefficient separation.*

Proposition 2 *When there are wage deferrals, a severance payment can prevent inefficient separation for senior workers*

2.2 Two periods model of specific investments

Consider now a more general two-period setting with endogenous wages being set in the initial period and moral hazard related to the decision of the worker to invest in firm-specific training. In our setup, workers may undertake a costly specific investment with uncertain return. We will keep our assumption that firms cannot commit "not to fire", thus workers expect that firms will always fire when -ex post- returns are too low, even when they have invested in the job. 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.

The model is 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. As in the illustrative example above, there is no discounting between periods, and the worker's outside option is b in every period. Denote the baseline productivity on the job as $y > b$ in every period. Assume further that wages are unilaterally set by the firm at the beginning of the contract. There is full commitment on the wage schedule and the firm cannot renegotiate ex post on the promised wage. The per period wage is indicated with w_i, $i = \{1, 2\}$

In period 1 the worker faces a specific investment opportunity $s = \{0, 1\}$. The investment opportunity costs to the worker C in the first period. It is only the worker who can undertake the investment opportunity. In other words, s is privately known by the worker in the first period, and revealed to the firm in the second period.

Conditional on the investment being undertaken, productivity in the second period will be $y + \varepsilon$, where ε is the specific component of productivity and it is drawn from a continuous distribution $F(\varepsilon)$ defined over the support $\varepsilon \in [\varepsilon_l, \varepsilon^u]$ with $\varepsilon_l < 0$.

The specific component of productivity is observed only by the firm. Since ε is only known to the firm, wages cannot be contingent on the specific component of productivity.

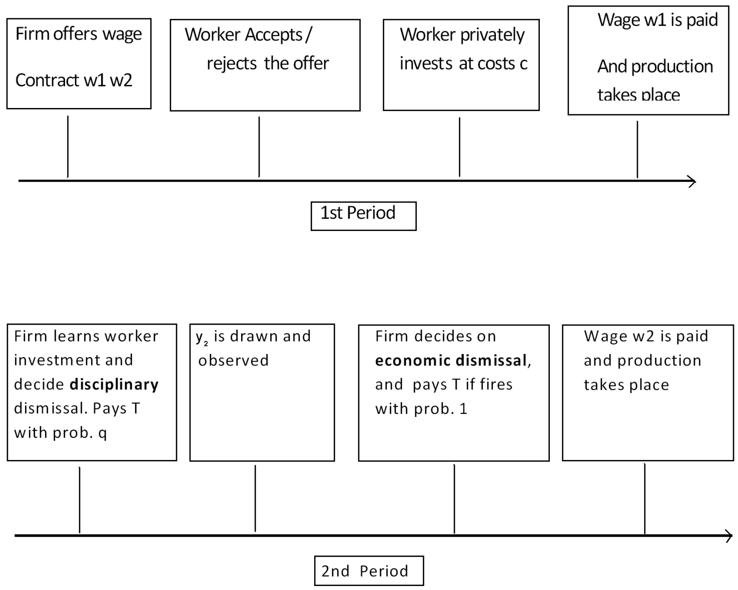
In the second period, conditional on having observed the ε draw, the firm can unilaterally fire the worker. The second period wage is therefore a compensation contingent on not having shirked in the first period.

There are two types of dismissals in the model. Economic and Disciplinary dismissals.

Definition 3 *Disciplinary Dismissal. In period 2, a firm is entitled to freely dismiss a shirking worker that did not invest in the first period.*

Definition 4 *Economic Dismissal. In period 2, when productivity is sufficiently low, a firm is entitled to dismiss a worker by paying a severance T*

Figure 3: Timeline



In terms of employment protection legislation, a worker that does not invest is fired for disciplinary reasons while a productivity related firing is a fair economic dismissal.

In our model, in the absence of publicly imposed severance payment the worker will not receive any severance payment if fired. The severance payment is a pure transfer to the worker. The severance payment T is set by the government. The firm can not commit to a severance payment because a problem of adverse selection stands on the way of a private contractual arrangement of this type. If a firm unilaterally commits to a severance payment, it would end-up selecting a group of shirkers. A mandatory severance solves this co-ordination problem. The realism of this assumption can be assessed considering that severance is either legislated or established within collective agreements at the industry, state or national level.

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 investments (hereafter shirking) in a court of law. We assume that there is a probability $1 - q$ that the court of law 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 in 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.

The timeline of the model in periods 1 and 2 is described in Figure 3.

In the case of an economic dismissal, we assume that severance payment is always due, and there is no distinction between fair and unfair dismissals.

Definition 5 *The equilibrium is obtained by a set of wages w_1, w_2, an investment decision s of the worker and a firing policy ε_d that satisfy*

- *Firm's participation constraint (non-negative expected profit in period 1)*
- *Worker's optimal investment*

- *Firm optimal firing in period 2*
- *Incentive compatible wage in period 2*
- *Worker's participation constraint*

2.3 Basic Value Functions

Let us indicate with $W_{(s=0)}$ the value of the job to the worker in case she does not invest. In this case the shirking worker gets the first period wage and will be fired in period 2 for disciplinary reasons. Since in the court of law the firing may be defined as unfair, the cumulative wage bill problem is totally trivial and the worker will be paid its outside option in both periods so that

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

where b is the flow outside option and qT is the expected severance payment in case of shirking. The worker's value function if she invests is then

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

where $F(\varepsilon_d)$ is the probability of being fired in the second period for economic reasons.

Firms expected profits in the first period if the worker invest are

$$\Pi_{1(s=1)} = y - w_1 + \int_x Max[y + x - w_2; -T]dF(x)$$

where $y - w_1$ is the operational profits in the first period while the integral in the second period is the expected profit conditional on a destruction/firing decision. Since the firing decision is clearly described by a reservation productivity, the expected profits are

$$\begin{aligned} \Pi_{1(s=1)} &= y - w_1 + \int_{\varepsilon_d}^{\varepsilon_u} [y + x - w_2]dF(x) - TF(\varepsilon_d) \\ \Pi_{1(s=1)} &= y - w_1 + \int_{\varepsilon_d}^{\varepsilon_u} xdF(x) + (y - w_2)(1 - F(\varepsilon_d) - TF(\varepsilon_d) \end{aligned}$$

Finally, if the worker does not invest, and the firm fires the worker in period 2, its expected cumulative profits are

$$\Pi_{1(s=0)} = y - w_1 - (1 - q)T$$

where $(1 - q)T$ is the expected severance payment for a fair dismissal

2.4 The reservation productivity

When the worker invests, in the second period the firm has to select a reservation productivity. Profits in the second period are conditional on a specific component of productivity and can be written as

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

Since the outside option of the firm in period 2 is an economic dismissal with full severance payment, the reservation productivity solves $\Pi_2(\varepsilon_d) = -T$ and its expression reads

$$\varepsilon_d = w_2 - y - T \tag{4}$$

Firing increases with wages while it decreases with productivity and severance payment. For a given wage, severance payments reduce firing in the second period. Substituting back into the profit functions

$$\Pi_{1(s=1)} = (y - w_1) + (1 - F(\varepsilon_d))(y - w_2) + \int_{\varepsilon_d}^{\varepsilon_u} xdF(x) - TF(\varepsilon_d)$$

Integrating by parts $\int_{\varepsilon_d}^{\varepsilon_u} x dF(x)$ and using equation4, the expected profits are

$$\Pi_{1(s=1)} = y - w_1 + (1 - F(\varepsilon_d))(y - w_2) + \varepsilon_d(1 - F(\varepsilon_d)) + \int_{\varepsilon_d}^{\varepsilon_u} (1 - F(x))dx - TF(\varepsilon_d)$$

$$\Pi_{1(s=1)} = y - w_1 + (1 - F(\varepsilon_d))[(y - w_2)+](y - w_2 + \varepsilon_d + T)] + \int_{\varepsilon_d}^{\varepsilon_u} (1 - F(x))dx - T$$

Since $y - w_2 + \varepsilon_d + T = 0$, the last expression is simply

$$\Pi_{1(s=1)} = y - w_1 + \int_{\varepsilon_d}^{\varepsilon_u} (1 - F(x))dx - T$$

. The firm participation constraints requires that $\Pi_1 > 0$

2.5 Information structure, contract space, and restrictions on parameter values

As mentioned in the introduction, the firm offers the worker a contract (w_1, w_2). The firm cannot commit to a firing policy. Hence if the firm finds it profitable to fire the worker ex post for economic reasons, it cannot commit not to do so. To ease the exposition of the model, we assume that the firm can commit to fire a worker that has not invested.

In what follows we assume that $y < b+C$ and that there exists a wage w_2 such that $(1-F(\varepsilon_d))w_2 \geq b+C$. The latter is a necessary requirement for investments to be implementable from the firm standpoint. We also assume that $(w_2 - b)[1 - F(\varepsilon_d)]$ is increasing on the relevant intervals.

2.6 Efficient Separation

Separation in the second period is efficient when it maximises the joint surplus from the job. Denote worker surplus at time $t = 2$ with $S_{w,2}$ and the firm surplus with $S_{f,2}$. Clearly

$$\begin{aligned} S_{w,2} &= w_2 - (b + T) \\ S_{f,2} &= y + \varepsilon - w_2 - (-T) \end{aligned}$$

where $b + T$ is the worker's outside option while $-T$ is the firm outside option. The joint surplus reads

$$S_2 = S_{w,2} + S_{f,2} = y + \varepsilon - b$$

where both wages and severance payments do not enter in the joint surplus, as they are a simple transfer between the two parties. Firing is efficient if and only if

$$\begin{aligned} S_2(\varepsilon^*) &= 0 \\ \varepsilon^* &= b - y \end{aligned} \tag{5}$$

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

2.7 The optimal contract

Given the reservation productivity in equation 4, the firm has to select optimally the wage in period 1 and 2. In general, the wage in period 2 will be determined by the worker's incentive compatibility constraint, while the wage in period 1 will be determined by the workers' participation constraint.

The firm must necessarily increase the second period wage in order to incentivate the worker to undertake the investment. Since the wage in period 1 will be identical for both shirker and non shirker workers, the worker will invest if and only if

$$W(s = 1) \geq W(s = 0)$$

The firm has to fire the shirker in the second period, otherwise the incentive compatibility constraint cannot be satisfied. The incentive compatibility constraint thus reads

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

or

$$(w_2 - b)[1 - F(\varepsilon_d)] \geq C + [q - F(\varepsilon_d)]T \tag{6}$$

Given our assumptions, (6) with equality defines a unique value of w_2. The general expression for the second period wage is thus

$$w_2 = b + \frac{C + [q - F(\varepsilon_d)]T}{1 - F(\varepsilon_d)} \tag{7}$$

Given the second period wage, the first period wage is set so as to ensure that the worker participation constraint is satisfied. The workers has an outside option equal to b per period, so that the worker participation constraint is

$$W(s = 1) = w_1 - C + w_2(1 - F(\varepsilon_d)) + F(\varepsilon_d)(b + T) \geq 2b \tag{8}$$

Given w_2, the solution of the participation constraint with equality gives the first period wage. The reservation productivity of equation 4 completes the solution of the model.

2.8 Wages, Dismissals and Severance Payments

The first result that we can establish concerns the firing policy of the firm. Recalling that the reservation productivity in period 2 is $\varepsilon_d = w_2 - y - T$, its equilibrium level reads

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

From the above equation, we get a first result when there are no severance payments,

Proposition 6 *If there are no severance payments* ($T = 0$), *firing is too high in the second period*

For a general value of C, there is too much firing in the second period, as

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

Recalling that the reservation productivity is $\varepsilon_d^* = b - y$, in the model there is too much firing as long as $C > 0$. This result is the essence of the firm non commitment policy vis-a-vis separations. As second period wages need to pay for the worker's investments' effort in the first period, the firm has a tendency to over dismiss in the second period a worker that did not shirk and invested in the first period.

With severance payments several features emerge. To understand firm behavior and the link between wages and severance payments, we need to study the relationship of the wage with respect to the probability q. Assume first that $q = 0$, so that shirking workers "never gets away with it", and disciplinary firing is perfectly detected. It follows that

$$w_{2(q=0)} - b = \frac{C - F(\varepsilon_d)T}{1 - F(\varepsilon_d)} \tag{9}$$

In this case, there is a perfect court monitoring and we can say that *severance payment acts as* a *discipline device*. Indeed, with perfect court monitoring, the firm needs to use less of the wage to induce workers' investment and the period 2 wage turns out to be decreasing in the severance pay, for two reasons. First, for a given ε_d, severance payments make investments more attractive, while they do not influence the pay-off for shirking workers (who get only b). Second, severance pay (and lower wages) makes firms more reluctant to fire, which also makes investing more attractive (since $w_2 > b + T$).

On the other extreme, if $q = 1$ (shirking workers always get severance pay), *severance payments are neutral*. In this case the second period wage is increasing in the severance payment

$$w_{2(q=1)} - b = \frac{C}{1 - F(\varepsilon_d)} + T \tag{10}$$

Substituting this wage in the reservation productivity we get

$$\varepsilon_{d(q=1)} = b - y + \frac{C}{1 - F(\varepsilon_d)}$$

In this case, the severance payment is neutral in terms of allocation of labor. For a given ε_d, a severance pay now increases wages by the same amount. Furthermore, when $dT = dw$, ε_d stays constant. It follows that $\frac{dw_2}{dT} = 1$.

This is a version of the Lazear (1990) neutrality result.

Proposition 7 *When severance payments are paid to the worker under any type of separation, they are neutral vis-a-vis the firms' firing allocation of labor and they simply make the wage profile steeper*

To sum up, for the general expression, by comparing (5) and (6) we derived the following set of results

Proposition 8 *a) If there are no severance payments ($T = 0$) then the firm fires workers too frequently (ε_d is too high)*

b) In case of perfect court monitoring ($q = 0$), severance payment T reduces w_2, and hence reduces firing. In other words, severance payments act as discipline device

c) If workers always get severance payments ($q = 1$), then severance increases w_2 but is neutral in terms of dismissal probabilities. In this case, $dw_2/dT = 1$ severance payments do not influence ε_d, and have no welfare implications. It only influences the wage profile by making it steeper.

d) Severance payment reduces the period 2 wage if and only if $q < F(\varepsilon_d)$

2.9 Efficient Severance Payments

Our next step is to derive the optimal severance payment T. From (4) and (5) it follows that the optimal T solves

$$\varepsilon_d = \varepsilon^*$$

which implies that

$$w_2(\varepsilon_d) - y - T = b - y$$
$$T = w_2(\varepsilon_d) - b$$

i.e., is equal to the wedge between the inside and the outside wage. Inserting from (6) with equality gives the following expression for the optimal T, denoted by T^*

$$T^* = \frac{C + [q - F(\varepsilon^*)]T^*}{1 - F(\varepsilon^*)} \qquad q < 1$$

Solving this for T^* gives

$$T^* = \frac{C}{1 - q} \qquad q < 1$$

When $q = 1$,an increase in T directly increases w_2 unit by unit. In addition, increased T (and w_2) increases ε_d, which again increases w_2.

Finally, to solve for w_1, we utilize that $w_2 = b + T$, whic inserted into (8) gives that

$$w_1 + b + T - C = 2b$$

or

$$\begin{aligned} w_1 &= b + C - T \\ &= b + C - \frac{C}{1-q} \\ &= b - \frac{qC}{1-q} \end{aligned}$$

Proposition 9 *The optimal severance pay policy can be summarized as follows:*

i) If $q = 1$ (shirkers always get severance pay) the optimal severance pay is undefined and there is no welfare loss in setting $T = 0$.

ii) For all other values of q, the optimal severance pay is strictly positive and given by

$$T^* = \frac{C}{1-q} > 0$$

3 Extensions

3.1 Three periods

Suppose now that the contract has three periods. The output of the worker in period i is $y_i + \varepsilon_i$ for $i = 2, 3$. Our set-up is robust enough to allow a general stochastic structure for the terms $\varepsilon_2, \varepsilon_3$, but for simplicity we now assume that they are iid. The outside option in all periods is b.

3.2 No investment cost in period 2

We first assume that workers do not undertake any costly investments in period 2. We also assume that if a worker is fired in period 2, he cannot be rehired in period 3. Note that the interim participation constraint of the non-shirker in period 2 is satisfied with slack and will not be discussed.

The participation constraint of a non-shirking worker reads

$$W^e = w_1 - C + (1 - F(\varepsilon_2^d))[w_2 + (1 - F(\varepsilon_3^d))w_3 + F(\varepsilon_3^d)T_3] + F(\varepsilon_2^d)T_2 \geq 3b \tag{11}$$

The value to a shirker reads

$$\begin{aligned} W^s &= w_1 + qT_2 + 2b \\ &\leq W^e \end{aligned} \tag{12}$$

The latter inequality is to avoid shirking. The joint surplus in period 3 is given by $S_3 = \max[y + \varepsilon - b, 0]$. Efficient separation in period 3 thus requires that

$$\varepsilon = \varepsilon_3^* = b - y$$

It follows that $S_3^* = \int_{\varepsilon \geq b-y} (y + \varepsilon_3 - b) d\varepsilon$. Suppose the efficient separation rate is imposed in period 3. The maximum joint surplus in period 2 is then

$$S_2^* = \max[y + \varepsilon_2 + S_3^* - b, 0]$$

Efficient separation in period 2 thus requires that

$$\varepsilon_2^* = b - y - S_3^* \tag{13}$$

Since the firm cannot re-hire the worker, there is an option value of keeping her.

Note that adding one more period gives two more instruments, T_3 and w_3. There is one more choice variable, ε_3^d. Since we have two more instruments and only one more variable, we have one degree of freedom to obtain efficiency.

In period 3, the firm retains the worker if $y + \varepsilon_3 \geq w_3 - T_3$. In order to implement the efficient separation rule $\varepsilon_3^d = b - y$, we must have that $T_3 = w_3 - b$. The resulting expected period 3 profit (with $T_3 = w_3 - b$) reads

$$\begin{aligned} E\Pi_3 &= \int_{\varepsilon \geq b-y} (y + \varepsilon - w_3) dF(\varepsilon) + F(b-y)(w_3 - b) \\ &= S^* - (w_3 + b) \end{aligned}$$

In period 2, the firm keeps the worker if and only if

$$y + \varepsilon_2 + E\Pi_3 = y + \varepsilon_2 + S_3^* - (w_3 - b) \geq w_2 + T_2$$

or

$$\varepsilon_2 \geq w_2 + w_3 - b - y - S^* - T_2$$

In order to implement (13) it follows that

$$T_2 = w_2 + w_3 - 2b \tag{14}$$

Given optimal values of T_1 and T_2, the worker is indifferent between being retained and being fired, both in period 1 and in period 2. It follows from (11) that

$$\begin{aligned} W^e &= w_1 + T + 2b - C & (15) \\ &\geq w_1 + qT + 2b & (16) \end{aligned}$$

Or that

$$T_2 \geq \frac{C}{1-q} \tag{17}$$

as before. To find the period 1 wage, plug in for (15) and (17) to get into the participation constraint (11) The incentive compatibility constraint (12) inserted thus gives that

$$w_1 + \frac{C}{1-q} - C \geq b$$

or

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

as before.

What about the wage-tenure profile and the severance pay-tenure profile? In our model this is indeterminate. The worker is indifferent between wages in period 2 and 3, and the profile does not influence T_2. However, $T_2 > T_3$ unless $w_2 < b$.

Proposition 10 *Suppose that the employment relationship involves three periods, and that workers only invest in the first period. Then the following is true*

a) The severance pay in period 2, T_2, is given by

$$T_2 \geq \frac{C}{1-q}$$

b) Optimal firing decisions in period 3 requires that $T_3 = w_3 - b$

c) Given optimal period 3 severance pay, the participation constraint gives a constraint on $w_2 + w_3$, but not on the wage-tenure profile. The severance pay T_2 is independent of the wage-tenure profile (as long as the participation constraint of the worker is satisfied).

d) If $w_2 \geq b$, then $T_2 > T_3$

3.3 Investment costs in period 2

Suppose now that the worker has to invest C in period 2 as well in order to maintain the high productivity in period 3. Again we assume that a shirker will be fired. If the worker shirks in period 2, the probability that she gets away with it is $q_2 > q$ (courts are more lenient towards senior workers). As in the previous subsection, we assume that efficiency is implemented.

Again we require that $T_3 = w_3 - b$ in order to obtain efficient separation in period 3. A period 2 non-shirker has a continuation pay-off in period 2 given by

$$\begin{aligned} W_2^e &= w_2 + F(\varepsilon_3^*)(T_2 + b) + (1 - F(\varepsilon_3^*))w_3 - C \\ &= w_2 + T_3 + b \end{aligned}$$

The second period shirker, by contrast, receives

$$W_2^s = w_2 + q_2 T_3 + b$$

Incentive compatibilty again requires that

$$T_3 \geq \frac{C}{1 - q_2}$$

Consider then period 2. The joint value of continuing the match is

$$S_2(\varepsilon) = y + \varepsilon_2 - b + S^* - C$$

The value to the firm (given efficient T_3) is given by

$$\Pi_2 = y + \varepsilon_2 - w_2 - S^* - T_3 + T_2$$

Efficiency requires that

$$\varepsilon_2 = \varepsilon_2^* = b + C - S^* - y$$

It follows that efficiency is implemented whenever

$$T_2 = (w_2 - b) + T_3 - C \tag{18}$$

It follows that

$$\begin{aligned} W^e &= w_1 - C + (1 - F(\varepsilon_2^d))[w_2 + W_3 - C] + F(\varepsilon_2^d)T_2 \\ &= w_1 - C + T_2 + 2b \end{aligned} \tag{19}$$

where W_3 is the expected pay-off to the worker if employed in period 3. By using exactly the same logic as above, it follows that IC in period 2 requires that

$$T_2 \geq \frac{C}{1 - q}$$

We have now pinned down the severance payments. Consider the wage-tenure profile starting from period 1 wage. It follows from (19) that

$$w_1 - C + T_2 + 2b = 3b$$

or

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

In period 3, we have that

$$w_3 = \frac{C}{1 - q_2} + b$$

Finally, from (18) (or just by using the period 1 participation constraint) it follows that

$$\begin{aligned} w_2 &= b + C + T_2 - T_3 \\ &= b + C + \frac{C}{1-q} - \frac{C}{1-q_2} \end{aligned}$$

Hence if $q_2 = q$, the wage in period 2 is strictly higher than the wage in period 1, but strictly lower than the wage in period 3. Furthermore, with $q_2 > q$ the difference widens.

Proposition 11 *Suppose the worker has to provide effort in both periods. Suppose further that the probability of getting away with shirking is higher for senior workers ($q_2 > q$). Then the following holds:*
a) The severance pay is increasing with tenure
b) Wages are increasing in tenure, $w_2 < w_3$. If q_2 is close to q, then we know for sure that also $w_1 < w_2$.

3.4 Burden of Proof and Endogenous "q"

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 two-periods model, but results can be readily generalized to a three-periods setting.

In particular, we assume that the investment in period 1 on the part of the worker shifts the distribution of productivity by a factor Δ. Specifically, the distribution of productivity in period 2 for a shirking worker (i.e. a worker that does not invest) is uniform between α and β so that

$$X^S \sim U[\alpha; \beta], \tag{20}$$

where X^S refers to actual productivity in period 2 for a shirking worker. Conversely, the productivity in period 2 for an investment worker is shifted to the right by a factor Δ so that

$$X^I \sim U[\alpha + \Delta; \beta + \Delta], \tag{21}$$

To make the problem interesting, we assume that there is no stochastic dominance, so that the support of the two distributions has an area of overlap:

$$\Delta < \beta - \alpha \tag{22}$$

Assume that a firm- as well as an outside court- can observe only total productivity in period 2 but does not observe whether the worker has invested or not. In other words, the outcome of the distribution X^I and X^S are not observed by the court and by the firm, even though their support is known. All that such outsiders can observe is the productivity y that- at least for some range- can potentially come from either distribution of Xs. This asymmetric information is the key assumption of this section. In this setting, the model endogenously determines the probability that a worker somehow "gets away with it", either because she is entitled to severance payment even if he shirked or because can be retained even in period 2.

The rest of the setting is analogous to the model of Section 3.

The firm's separation policy in period 2 is simply

$$Max[y - w_2; -T]$$

where y is the worker's productivity as defined above. The reservation productivity is simply

$$y\star = w_2 - T$$

Since y is observed by the firm and the court, a productivity outcome below $\alpha + \Delta$ is necessarily obtained only by a shirking worker and no severance payments are due in this case. We thus let $\tilde{y} = \alpha + \Delta$ be the productivity below which a court can confidently say that the worker did shirk.

Definition 12 *Disciplinary Dismissal. A productivity in period 2 below $\alpha+\Delta$ must necessarily be obtained only by a shirking worker and the dismissed worker is not entitled to severance payments*

Conversely, a period 2 productivity that is larger than $\tilde{y}$ but lower than $y\star$ may potentially come from either an "unlucky" non-shirker worker or from a shirking worker.

Definition 13 *Economic Dismissal. A productivity in period 2 that is below $y\star$ but larger than $\tilde{y}$ leads to a dismissal with severance payment.*

The setting implies that the firm has a burden of proof vis-a-vis the court, and that in case of uncertainty the court assumes that severance payments are due.

Before solving the model, one needs to specify where exactly the reservation productivity $y\star$ lies with respect to the upper support of the distribution of shirking workers β and the lower support of the distribution of investing workers $(\alpha+\Delta)$. Depending on the position of the productivity with respect to these supports, the model has different solutions. We here assume that

$$\alpha+\Delta < y\star < \beta; \tag{23}$$

from which the following economic implications follow

- A shirking worker can get away with it and even be employed in period 2
- A shirking worker can get away with it in the sense that he obtains severance payments even if he shirked
- An investing worker is never fired in period 2

In this section we rely on the assumption (23) and we let the other cases to the appendix. Obviously, the solution of the model must necessarily be coherent- in terms of parameters- with the assumption specified in (23). The value function for an investing worker reads

$$W^I = -C + w_1 + Pr[(\alpha+\Delta) < X^I < y\star](b+T) + Pr(X^I > y\star)w_2 \tag{24}$$

so that an investing worker is either fired for economic reasons or retained and entitled to severance payments. The endogenous corresponding probability are simply

$$Pr[(\alpha+\Delta) < X^I < y\star] = \frac{y\star-(\alpha+\Delta)}{\beta-\alpha}$$

and

$$Pr(X^I > y\star) = 1 - \frac{y\star-(\alpha+\Delta)}{\beta-\alpha}$$

where obviously $y\star = w_2 - T$ Conversely, the value function for a shirking worker reads

$$W^S = -w_1 + Pr[X^S < (\alpha+\Delta)] + Pr[(\alpha+\Delta) < X^S < y\star](b+T) + Pr(X^S > y\star)w_2 \tag{25}$$

where the corresponding probabilities are

$$Pr[X^S < (\alpha+\Delta)] = \frac{\Delta}{\beta-\alpha},$$

$$Pr[(\alpha+\Delta) < X^S < y\star] = \frac{w_2-T-\alpha}{\beta-\alpha} - \frac{\Delta}{\beta-\alpha},$$

and

$$Pr(X^S > y\star) = 1 - \frac{(w_2-T)-\alpha}{\beta-\alpha}$$

The worker ICC is simply obtained by the condition

$$W^I(w_2) = W^S(w_2) \tag{26}$$

Substituting the workers' value functions (24) and (25) in the ICC, after a few steps of simple algebra, the optimal wage reads

$$w_2 = b + c\frac{\beta - \alpha}{\Delta}; \ldots \text{ICC Wage} \tag{27}$$

Proposition 14 *The incentive compatible wage in period 2 is equal to the reservation wage plus the investment cost "augmented" by a factor $q\star = \frac{\beta-\alpha}{\Delta}$ that depends on the parameters of the distribution*

Two additional propositions characterize the endogenous premium $q\star$

Proposition 15 *If $\Delta = \beta - \alpha$ the supports of the two distributions do not overlap and the ICC wage does not involve any premium vis-a-vis the outside option of the worker and the investment cost*

Proposition 16 *If $\Delta \to 0$ the distributions of investing and shirking workers are identical and the ICC wage grows indefinitely*

To close the model of this section one just needs to obtain the first period wage from the worker's participation constraint and ensure that firm's expected profits are positive.

To be coherent with the restriction imposed at the beginning of this section, $\alpha + \Delta < w_2 - T < \beta$, making use of the incentive-compatible wage, the severance payments must be such that.

$$T < b + c\frac{\beta - \alpha}{\Delta} - \alpha - \Delta$$
$$T > b + c\frac{\beta - \alpha}{\Delta} - \beta$$

Finally, one can compare the results we just obtained with the optimal severance payment of proposition [10]. By comparing the two results we can say that an efficient severance payment requires

$$T = \frac{C}{1-q} = \frac{C(\beta - \alpha)}{\Delta} \tag{28}$$

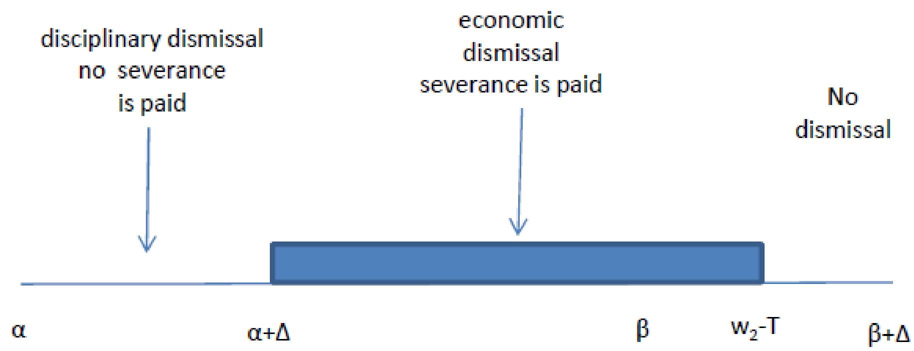
from which it follows that $q = 1 - \frac{\Delta}{\beta - \alpha}$ which in this section corresponds exactly to the probability that a shirking worker gets away with it, either because he is fired with severance payments or because he is retained in period 2.

4 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. We show that under a rather broad set of circumstances a severance scheme which is graded in tenure is privately efficient, in that it avoids inefficient separations.

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. Moreover our theory suggests that tenure-related severance is efficient

Figure 4: Court decisions



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.

A crucial issue is why severance pay has to be mandated by the Government rather than being provided by the individual firm. Adverse selection may stand on the way of these voluntary arrangements, which may end up attracting more potential shirkers to the firm unilaterally offering a severance scheme to their workers. In other words, mandatory severance acts as a coordination device across firms. In future work we plan to formally address this issue.

References

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

[2] 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.

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

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

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

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

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

[8] Garibaldi, P. 1998. Job flow dynamics and firing restrictions. International Monetary Fund, 19th Street, Washington, DC 20431, USA

[9] Jovanovic, B. 1979. Firm-specific Capital and Turnover, *Journal of Political Economy*, University of Chicago Press, vol. 87(6): 1246–60, December.

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

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

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

5 Annex

Table A1 Apparent Elasticities at different tenure lenghts

	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
Denmark	*	*	*	*	*	-
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	0.01	0.67	0.75	0.50	0.91	0.74
Germany	0.00	1.41	0.85	1.00	0.94	0.91
Greece	*	4.00	0.69	0.71	1.00	-
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	0.01	0.00	0.63	0.67	0.50	0.54
New Zealand	*	*	*	*	*	-
Norway	0.01	0.00	0.63	0.67	0.00	0.29
Poland	0.01	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	0.01	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
United States	*	*	*	*	*	-

Notes: * : No mandatory severance

Table A2. Detailed information used to produce Table 2

Country	Severance Economic at 20y, Fair	Severance Disciplinary at 20y, Fair	Typical Compensation at 20y, Unfair	Maximum Notice	Lenght of Trial τ	Prob. of Reinstatement π
Australia	2.8	0.0	6.0	1.0	6.0	0.33
Austria	0.0	0.0	6.0	4.0	5.9	1.00
Belgium	0.0	0.0	10.3	21.0	6.0[b]	0.00
Canada (Federal)	2.3	0.0	Court	2.0	3.0	0.33
Czech Republic	1.5	0.0	6.0	2.0	6.0	1.00
Denmark	3.0	0.0	6.6	6.0	9.0	0.33
Finland	0.0	0.0	14.0	6.0	9.0[b]	0.00
France	5.4	0.0	16.0	2.0	13.0[b]	0.17
Germany	10.0	0.0	15.5	7.0	3.2[b]	0.50
Greece	8.0	0.0	Court	4.0	12.0	0.67
Hungary	6.0	0.0	12.0	3.0	4.5	0.33
Ireland	4.0	0.0	24.0	2.0	12.0	0.33
Italy	0.0	0.0	21.0	6.0	20.6[b]	0.33
Japan	0.0	0.0	6.0	1.0	3.0	0.33
Korea	0.0	0.0	6.0	1.0	3.0	1.00
Luxembourg	6.0	0.0	6.0	6.0	2.3[b]	0.00
Mexico	1.0	0.0	15.0	n.a.	3.5	0.50
Netherlands	0.0	0.0	7.0	4.0	0.7[b]	0.33
New Zealand	0.0	0.0	7.6	0.5	3.0	0.33
Norway	0.0	0.0	12.0	6.0	6.0	0.67
Poland	3.0	0.0	3.0	3.0	3.0	0.33
Portugal	12.0	0.0	20.0	2.5	10.0[b]	0.83
Slovak Republic	4.0	0.0	6.0	3.0	3.0	0.83
Spain	12.0	0.0	24.0	0.5	3.5[b]	0.00
Sweden	0.0	0.0	32.0	6.0	6.0	0.00
Switzerland	0.0	0.0	6.0	3.0	2.0	0.00
Turkey	20.0	0.0	10.0	2.0	8.0[b]	0.00
United Kingdom	4.6	0.0	5.5	3.0	24.0	0.33
United States	0.0	0.0	Court	0.0	3.0	0.17

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

Notes: Data are expressed in months. 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.