Why Do Firms Offer "Employment Protection"?

Christopher A. Pissarides*
Centre for Economic Performance,
London School of Economics, IZA and CEPR

September 2009

Abstract

This paper derives optimal employment contracts when workers are risk averse and there are employment and unemployment risks. Without income insurance, consumption rises during employment and falls during unemployment. Optimal employment contracts offer severance compensation and sometimes give notice before dismissal. Severance compensation smooths consumption during employment and dismissal delays insure partially against the unemployment risk because of moral hazard. During the delay consumption falls to give incentives to the worker to search for another job. No dismissal delays are optimal if exogenous unemployment compensation is sufficiently generous.

Keywords: employment risk, unemployment risk, severance compensation, notice of dismissal

JEL Classification: E21, E24, J32, J33

"Employment protection legislation" has featured prominently in the work of Layard and Nickell - see for example their *Unemployment* book or Steve Nickell's (1997) much cited paper in the *Journal of Economic Perspectives.* Nickell (1982) is one of the very first empirical papers that derived the "equilibrium" unemployment rate by equating flows in with flows out of unemployment. He discussed how "firing costs" are an influence on both inflows and outflows, with ambiguous effects on the equilibrium unemployment rate, much as subsequent work in search theory elaborated over and over again.

Layard and Nickell's work, however, like the vast majority of the literature, investigated the impact of *exogenous* employment protection legislation on labour market

*Earlier versions of this paper circulated under the title "Consumption and Savings with Unemployment Risk: Implications for Optimal Employment Contracts" and were read at the NBER Summer Institute, the LACEA conference in Montevideo, and at Princeton, di Tella and Bonn Universities. I have benefited from the comments of Chris Carol, Leonardo Felli, Raquel Fernandez, Hugo Hopenhayn, Winfried Koeniger, Rasmus Lentz, Eva Nagypal, Robert Shimer, Michael Woodford and Sevin Yeltekin, and from the research assistance of Vassileios Gkionakis.

outcomes. But employment contracts often contain voluntary firing restrictions: provisions for the payment of severance compensation to dismissed employees, or for delays in dismissals. The most common procedure that delays dismissal is the requirement to give a notice of fixed duration before dismissal. There are also other restrictions. In many countries minimum levels of severance compensation and dismissal delays are written in employment laws but private contracts contain similar, if not more, stringent requirements. The OECD (1999) reports that on average in its member states employers are required to give minimum advance notice of dismissal of 1.6 months to employees of four years standing, and to pay severance compensation of four weeks' wages.[1] The purpose of this paper is to investigate the theoretical foundations for the existence of such provisions in private employment contracts.

I study a situation in which a principal, the firm, chooses the employment contract that minimizes the cost of its job offer, which has to be acceptable to an agent, the worker. I do not attempt to justify the inclusion of severance compensation or dismissal delays in legislation but investigate whether they can be a part of an optimal employment contract. The main result of the paper is that if workers cannot insure against the risk of becoming unemployed and the risk associated with an uncertain duration of the subsequent search for another job, severance compensation and dismissal delays provide second-best alternatives that avoid some of the moral hazard problems of first-best insurance. The payment of severance compensation is a perfect substitute for insurance against the risk associated with an uncertain duration of employment (I refer to this as the *employment risk*). Giving advance notice before dismissal provides additional insurance against the uncertain duration of unemployment (the *unemployment risk*). Optimal dismissal delays, however, do not fully insure against the unemployment risk, because the failure of the firm to monitor the search strategy of the workers that it places on notice of dismissal introduces moral hazard.

Intuitively, it should not be surprising that severance payment is optimal when it is not certain how long a job will last. The severance payment replaces the need to save

[1] Provisions are more stringent in Europe than elsewhere but even in the United States, where legal provisions are virtually non-existent, similar arrangements are found in private contracts. For example, the OECD reports that in a survey conducted in 1992, it was found that between 15 and 35 percent of employees in the United States were covered by company severance plans, depending on company size. Civil rights laws and other legislation are also said to be contributing to delays in dismissals. See OECD (1999, p. 58).

for the unemployment contingency, and it is preferred because it can be less contingent on the duration of the job than accumulated private savings are. The surprising result of the paper is that even when severance payments are optimal, there is a role for a notice period before dismissal, with the worker paid a wage during the notice period but producing nothing. The reason for this result is the following.

Suppose news arrives that a job has become unproductive. If the worker is dismissed into unemployment without delay her income falls, and she starts search for another job. Search is an uncertain activity so she wants to buy insurance contingent on its outcome, but the absence of private insurance markets does not make this possible. Instead, the worker can be made better off if the firm lets her stay employed without a pay cut but she agrees to search on the job and quit when she finds another job. The total wage payments actually received by the worker are then contingent on the outcome of search, and if the firm is risk neutral it is indifferent between paying the value of marginal product and dismissing without delay, or choosing a flat wage rate with indefinite delay subject to expected total payments not exceeding the expected marginal product.

This contract is not feasible, however, when the firm cannot monitor the search strategy of the workers on notice of dismissal. The worker whose job has become unproductive will have no incentive to search if her income does not fall. As a second-best alternative, the firm can reduce moral hazard by structuring its compensation package so as to give incentives to the worker to search on the job and quit. I show that the firm can again achieve the contingency of payments by delaying dismissal but in a more subtle way. The optimal compensation package with moral hazard requires that when the job becomes unproductive the worker's utility of remaining employed should fall, and become progressively worse as the duration of employment increases. Delaying dismissal still has insurance value because by keeping the worker employed the firm can effectively monitor the outcome of search. If the firm pays employees on notice of dismissal more than the level of exogenous unemployment compensation, it knows that if the employee quits it is to take another job. Therefore, wage payments are still contingent on unsuccessful search. Workers on dismissal notice are given incentives to search by the payment of compensation to quitting workers, which is effectively a reward for successful search.

The cost of delaying dismissal is that the worker on delayed dismissal is foregoing unemployment insurance, which is subsidized by the state.[2] One of the results of this

[2]By assumption, the firm does not pay unemployment insurance to former employees. Once a

paper is that there is a trade-off between state-provided unemployment insurance and risk-shifting in private employment contracts. I consider only a simple form of state-provided unemployment insurance, the payment of a flat subsidized compensation to unemployed workers. I show that if the compensation is sufficiently high, no dismissal delays are optimal.[3]

With unlimited borrowing and lending (a maintained assumption in this paper) there is a number of ways that the firm can implement the optimal contract. One compensation package that has the required properties holds the wage rate constant for a finite length of delay time and offers severance compensation to quitting workers. The worker is fired if she is still employed at the end of the "notice" period. This is the most common structure found in employment contracts that include dismissal delays. But other compensation packages give equivalent results. One such other package is to allow wages to fall monotonically with time employed after the job becomes unproductive. I discuss the implementation of the contract only briefly in the concluding section of the paper, my main focus in this paper being the derivation and description of the optimal contract.

My results on the optimal compensation package during a delay in dismissal are similar to the results derived for the optimal time structure of unemployment compensation, especially those by Shavell and Weiss (1979), Sampson (1978) and Hopenhayn and Nicolini (1997), who show that optimal unemployment compensation declines with search duration.[4] Although the model in this paper is deliberately simplified, and ignores the aggregate implications of the employment contracts, it can easily be extended to a model of labor market equilibrium with unemployment when workers have concave utility functions and can borrow and lend.[5]

---

separation has taken place all ties between the firm and the worker are severed.

[3] A topic for future work is whether private severance compensation and delayed dismissal are still optimal under different forms of unemployment insurance. For example, limited-duration unemployment benefits may encourage delayed dismissals. Paying subsidized compensation to workers on short time may also encourage it, because the firm may place the workers on short time when it gives notice of dismissal. See Burdett and Wright (1989) on the comparison between short-time insurance and unemployment insurance.

[4] Their results are usually derived under the restriction of no borrowing or lending, so the only way consumption can be made to fall during unsuccessful search is to offer a declining compensation package. When borrowing and lending are allowed a similar result can be obtained for workers on notice of dismissal with other compensation packages, discussed briefly in the concluding section.

[5] Other papers on optimal unemployment insurance with concave utility functions address different sets of issues. For example, Hansen and Imrohoroglu (1992), Andolfatto and Gomme (1996), Costain (1997) and Wang and Williamson (1996) study computational models to derive the implications of

The features of the optimal employment contract that I derive are often collectively described in the empirical literature as "employment protection." The large literature on employment protection legislation studies partial or equilibrium models with risk neutrality in order to quantify the effect of various policy measures on employment and wages. Special emphasis is given to compulsory severance compensation and dismissal frictions.[6] The general conclusion reached in this literature is that employment protection measures do not have a significant impact on steady-state employment, but are likely to influence the dynamics of employment and wages. The main implication of the analysis of this paper for this literature is that a proper evaluation of employment protection measures should take into account the fact that they may be optimal responses to missing markets. This should influence the impact that they have on equilibrium but a proper evaluation of the empirical literature requires a model that implies the optimality of legislation, an issue that I do not take up here.[7]

Also related to the model of this paper is another strand of the literature, which studies the behavior of wealth and the unemployment hazard during search when there is risk aversion. Danforth (1979) shows that with decreasing absolute risk aversion reservation wages fall and so the probability of leaving unemployment rises. A similar result is derived by Lentz and Tranaes (2005) for a more general model of job search, with both employment and unemployment risks and borrowing and lending.

It is important for the results of this paper that the firm should be better able to insure against fluctuations in income than workers are. This property, the asymmetric access to insurance markets by firms and workers, is the key assumption behind the static implicit contract theory, and this paper can be viewed as an application of the ideas first developed in that theory to dynamic search equilibrium (see Baily, 1994, Azariadis, 1975

unemployment insurance for welfare, aggregate economic activity and the optimal level of UI benefits in calibrated economies. Acemoglu and Shimer (1999) study a model with constant absolute risk aversion to derive results on the efficiency of unemployment insurance, given that risk-averse workers accept offers too quickly. Wright (1986) is an early attempt to derive a political equilibrium with unemployment insurance, in contrast to the other papers in this footnote, which consider only welfare gains.

[6] See in particular Lazear (1990) and Nickell (1997). For recent summaries see Nickell and Layard (1999) and Bertola (1999) and for more recent contributions see Ljungqvist and Sargent (1998), Mortensen and Pissarides (1999) and Blanchard and Wolfers (2000).

[7] Recently, Bertola (2004) studied the role of arbitrary legislated severance compensation in a model with risk averse workers. He shows that it can increase welfare but he does not allow agents to write optimal contracts, which could undo the effects of the legislation. Saint-Paul (2002) considers reasons for the political support of employment-protection legislation. See also Pissarides (2001) for more discussion of the policy issues and their implications for labor market equilibrium.

and Gordon, 1994).

Section 1 outlines the framework used to study the implications of non-linear utility for consumption and job search. Section 2 studies the optimal consumption and search strategies when workers are paid their marginal product, and section 3 studies the other extreme, choices made under a full set of insurance contracts. Section 4 forms the core of the paper and studies first, the insurance implications of severance compensation and second the insurance implications of delayed dismissal. Section 4.3 shows that whereas it is always optimal to include severance compensation in employment contracts, whether dismissal delays are part of a contract or not depends on the compensation received by unemployed workers and on their risk aversion. Section 5 shows with computations that dismissal delays can be optimal for sufficiently low unemployment compensation. The concluding section 6 briefly summarizes the main findings and discusses issues in the implementation of the optimal contract. All proofs are collected in the Appendix.

1 Framework

The model is a partial one and focuses on the relation between a risk-neutral firm that owns a job and a risk-averse worker who owns a time endowment. Time is discrete and information unfolds sequentially. The results that I derive do not depend on any particular horizon length but they are more transparent in a model that has the shortest horizon length that admits variable dismissal delays. This horizon length is $T = 3$, which allows me to include both the employment and unemployment risks and dismissal delays of either one or two periods. The results that I derive, however, are not special to this horizon and can be generalized to an infinite horizon.

The worker's time endowment yields no utility but it enables her to hold a job. Utility is derived only from consumption, at the rate $u(c)$ per period, with $u'(c) > 0, u''(c) \leq 0$ and $u'(0) = \infty$, although there are also lump-sum disutilities associated with changing jobs, which are specified later. There is unlimited borrowing and lending at a safe rate of interest r, which accrues during the period, and which is also used to discount future utility. I define the discount factor $\beta = 1/(1 + r)$. The utility function, discount rates and capital structure are chosen such that under a full set of insurance markets the consumption profile is flat in all states of nature, irrespective of the income profile. This makes it easier to describe the implications of risk and optimal employment contracts.

It is assumed that there are no exits from the labor force before the end of the horizon.

My objective is to describe the features of an employment contract offered by the firm to a new employee, when there is a positive probability that the job will become unproductive and when the date of arrival of a new job (if this one ends) is uncertain. I make use of the following simplified framework (see footnote 9 for one possible rationalization of this framework).

In period 1 a worker arrives to a new job which produces output $p > 0$. The job either becomes unproductive at the end of the period, an event that occurs with fixed probability $q \in (0,1)$, or it remains productive at the same p to the end of the horizon (becomes an *absorbing state*). If the job becomes unproductive the worker may be dismissed immediately, with or without severance compensation, or dismissal may be delayed, during which time the worker produces nothing but is paid a wage. I refer to this state as being *on delayed dismissal* or *on notice of dismissal.* In the three-horizon model the delay may be either one or two periods.

Workers search on the job in period 1, because of the risk that the job will become unproductive, and may also search on the job during a period of notice. If a job that dominates the present one is found during search the worker moves to it at the beginning of the next period. Unemployed workers also search for a job and receive compensation $b < p$ in each period of unemployment. All jobs found after period 1 are absorbing states, i.e. the worker stays in them until the end of the horizon, producing output p per period.[8]

There are no search costs but before accepting an offer the worker has to pay a moving cost $x \geq 0$, which differs across jobs. The cumulative distribution of x for the best job offer available to the worker each period is denoted by $G(x)$ and has support in the positive quadrant. The mobility cost is measured in utility units and it is strongly separable from the utility of consumption. Job acceptance is governed by a reservation rule on x, such that workers accept a job at the beginning of period t if the realized acceptance cost is $x \leq R_t$. The probability that a job searcher moves to a job at the beginning of period t is therefore $G(R_t)$.

Of course, a worker on notice of dismissal can quit at any time into unemployment

[8] Employment is an absorbing state in all periods in Danforth (1979), Hopenhayn and Nicolini (1997) and Acemoglu and Shimer (1999) but not in Lentz and Tranaes (2005), who derive the effects of unemployment risk on savings (without employment contracts) in a more general environment.

and receive the subsidy b, but once she quits she cannot be rehired by the same firm. The circumstances that lead an agent to make decisions among the alternative states, the factors that influence their decisions, and whether contracts specify severance compensation or delayed dismissal, are the subject of analysis in this paper.

There are two income risks in this model, which are insurable with a full set of insurance markets. First, the risk that productive employment in the first job lasts either one period, or until the end of the worker's horizon. Second, conditional on the termination of employment in period 1, the risk that non-production (i.e., either unemployment or employment on notice of dismissal) lasts for one or two periods. The first risk is the *employment risk* and the second the *unemployment risk*, as each is associated with an uncertain duration of employment or unemployment (more precisely, job search).

The firm offers the worker an employment contract in period 1 which minimizes the cost of providing a pre-specified level of lifetime utility. I assume that the firm can monitor the worker's assets for the duration of the contract. It cannot, however, monitor the worker's search effort. Therefore, the firm can act as if it can choose a consumption profile for the worker, for the duration of the contract, subject to incentive-compatibility constraints on search effort. This is a typical principal-agent problem with moral hazard, with the firm acting as the principal who minimizes the cost of providing a consumption level to the worker.

It is convenient to set up the problem as if the contract ends either at the beginning of period 2 if the first job becomes an absorbing state, or in either period 2 or 3 if it does not. In the latter case, the contract ends either when the worker quits or when she is dismissed. For the duration of the contract the firm provides the worker with consumption. When the contract ends the firm makes a transfer of assets to the worker who then chooses her own consumption levels. There is no loss of generality if I assume that the contract ends at the beginning of period 2 if the job survives, because there is no uncertainty attached to lifetime earnings in the first job beyond period 1.[9]

[9]The following rationalization, based on Jovanovic (1979) and Wilde (1979), can be used for the assumptions about jobs and their survival made in this paper. Suppose there are two or more differentiated types of agents and jobs and the match between a worker and a job is good if they are of the same type and bad if they are of different types. Net output is p per period in all matches, irrespective of type, but mismatched workers forego a lump-sum utility cost in order to produce this output. Workers who are matched to a job of their type do not forego any utility to produce; they stay in it for ever and earn their marginal product p per period. Workers, however, do not initially know how to recognize

2 Non-contingent contracts

I derive first the lifetime consumption profile in the absence of insurance and contingent transfers from the firm to the worker. Workers receive their marginal product p when employed, and subsidy $b < p$ when unemployed. They do not receive any income if they are on notice of dismissal, making this option sub-optimal.

An equilibrium is a consumption sequence $\{c_t^s\}$ for each state and period and an acceptance rule for each period of search. The states of nature are employment $(s = j)$ or unemployment $(s = u)$, and $t = 1, 2, 3$. Agents are always employed in $t = 1$ but employed or unemployed in subsequent periods. The agent maximizes expected utility subject to a sequence of budget constraints and a value for initial assets, which is assumed to be zero, and subject to correct expectations about p, b, the distribution of costs $G(x)$ and the termination probability of the first job, q. Income and consumption flows arrive at the end of the period whereas asset values are calculated at the beginning.

I derive the optimal policy starting from the end of the horizon. A person who is unemployed in period 2 starts period 3 with initial assets A_2 and exits at the end of period 3 with zero assets. It follows that if the agent takes a job in period 3 consumption is

$$c_3^j = p + \beta^{-1} A_2. \tag{1}$$

If the agent remains unemployed she has no incentive to search, because she will exit the market at the end of the period, so she receives the subsidy b and consumes

$$c_3^u = b + \beta^{-1} A_2. \tag{2}$$

Suppose now the agent's initial assets in period 2 are A_1. If the agent is employed in period 2, either in the job that she was employed in period 1 or in a new job, income in periods 2 and 3 is p and both incomes and utilities are discounted at the common factor β. Therefore consumption in periods 2 and 3 is flat, and denoted by c_2^j :

$$c_2^j = p + \frac{1}{\beta(1+\beta)} A_1. \tag{3}$$

their job type without experiencing the job. They enter a randomly-selected job and spend the first period of their life in productive employment, learning also about their job type, and how to inspect and recognize future job types. The utility cost of a mismatch is sufficiently high that unemployment dominates production in a job of the wrong type. But the disutility is sufficiently low that all agents prefer to produce in period 1, and run the risk of mismatch, from taking leisure for ever and never learn about job types. The probability that the first job is not of the worker's type is a fixed $q \in (0, 1)$. Turnover and employment risk are higher for young individuals and at short tenures because of learning about job types and skill attributes.

If the agent is unemployed in period 2 she is also searching for another job. Job offers arrive and she can move to one at the beginning of period 3 by bearing a one-off utility cost x, which has distribution $G(x)$. By the separability of the mobility cost and the full information on $G(x)$, job acceptance satisfies the reservation property: the individual accepts a job at the beginning of period t if the realized mobility cost $x \in [0, R_t]$, where R_t is a reservation value. I denote by $\bar{x}_t$ the expected acceptance cost conditional on the reservation R_t, i.e. $\bar{x}_t = E(x|x \leq R_t)$. It follows that the utility function in the event of unemployment in $t = 2$ satisfies the Bellman equation

$$U_2^u(A_1) = \beta \max_{c_2^u, R_3, A_2} \left\{ u(c_2^u) + G(R_3) \left(U_3^j(A_2) - \bar{x}_3 \right) + (1 - G(R_3)) U_3^u(A_2) \right\}, \tag{4}$$

where $U_3^j(A_2)$ is lifetime utility when the worker moves to a job at the beginning of period 3, given by $\beta u(c_3^j)$, and $U_3^u(A_2)$ is lifetime utility when the worker is unemployed in period 3, given by $\beta u(c_3^u)$. (In general superscript u on the utility function denotes utility when the worker starts the period unemployed and superscript j when she starts in a productive job.) The end-of-period budget constraint in period 2 is

$$A_1 + \beta (b - c_2^u - A_2) \geq 0. \tag{5}$$

The first order maximization conditions yield, after application of the envelope theorem,

$$u'(c_2^u) = G(R_3) u'(c_3^j) + (1 - G(R_3)) u'(c_3^u), \tag{6}$$

and

$$R_3 = U_3^j(A_2) - U_3^u(A_2). \tag{7}$$

In period 1 the job is not an absorbing state because there is a probability q that it will end at the end of the period. The individual searches in period 1 and a job arrives at the beginning of period 2, which she can either accept or reject and join unemployment. The Bellman equation satisfied by lifetime utility at birth is

$$U_1 = \beta \max_{c_1, R_2, A_1} \left\{ u(c_1) + (1 - q) U_2^j(A_1) + q \bar{U}_2(A_1) \right\}, \tag{8}$$

where c_1 is consumption in period 1, $U_2^j(A_1)$ lifetime utility when the worker is in a job at the beginning of period 2 (given in this case by $\beta(1 + \beta) u(c_2^j)$) and $\bar{U}_2(A_1)$ is the expected lifetime utility when the job in period 1 terminates. In this event the agent either moves to another job, with (conditional) probability $G(R_2)$, or becomes unemployed. Therefore

$$\bar{U}_2(A_1) = G(R_2)(U_2^j(A_1) - \bar{x}_2) + (1 - G(R_2))U_2^u(A_1). \tag{9}$$

The budget constraint in period 1, given zero initial assets, is

$$p - c_1 - A_1 \geq 0. \tag{10}$$

The necessary and sufficient maximization conditions satisfy

$$u'(c_1) = (1 - q + qG(R_2))\, u'(c_2^j) + q(1 - G(R_2))u'(c_2^u). \tag{11}$$

$$R_2 = U^j(A_1) - U^u(A_1). \tag{12}$$

It follows from (6) and (11) that both the unemployment and the employment risk give rise to a lifetime consumption profile that is not flat. I now show (see the Appendix for proof)

Proposition 1 *The employment risk causes a rising consumption profile and the unemployment risk a falling consumption profile.*

The optimal policy is one where the agent consumes c_1 in the first period and increases her consumption permanently to a higher level if the job survives, or if she moves to another job (which is necessarily an absorbing state). If the job terminates in period 2 she reduces her consumption. During unsuccessful search consumption falls and when a job is found it rises to a permanently higher level.

The result that the employment risk yields a rise in consumption implies $A_1 > 0$ and the result that the unemployment risk yields a fall in consumption implies $A_1 > A_2$. The latter restriction also implies $c_2^j > c_3^j$, i.e., the individual is better off if she finds a job in the first period of search than in the second.[10]

[10] The property that assets fall during unsuccessful search generalizes to an infinite horizon, provided that lifetime utility is a concave function of initial assets. This in turn also implies that the probability of leaving unemployment rises during unsuccessful search. With a shorter horizon the probability of leaving unemployment may fall because of the fixed utility cost of accepting jobs. Concavity of the utility function is not guaranteed because of the influence of wealth on reservation costs but almost certain to be satisfied. See Danforth (1979), Lentz and Tranaes (2005) and Hopenhayn and Nicolini (1997) for related discussion. Lenz and Tranaes (2005) introduce explicitly lotteries to avoid non-concavities. Hopenhayn and Nicolini (1997) derive results for the range of parameters that are consistent with concavity. Danforth (1979) restricts the utility function to decreasing absolute risk aversion. Where necessary I follow Hopenhayn and Nicolini and derive results for the range of parameters that guarantee concavity.

3 Full insurance

When workers have access to actuarially fair insurance against all income risks their consumption profile becomes flat and independent of state of nature. This result emerges readily from the assumptions of constant and equal rate of interest and rate of time preference and the existence of unlimited borrowing and lending at a safe interest rate, and will not be demonstrated in full. As an illustration, consider a one-period insurance contract for workers in period 1. With a full set of insurance contracts the worker can insure against the employment risk by buying insurance that will pay her I_1 at the beginning of period 2 if she becomes unemployed. The risk of this is $q(1 - G(R_2))$, and so actuarial fairness implies that the budget constraint for period 1 changes from (10) to

$$p - c_1 - A_1 - q(1 - G(R_2))I_1 = 0. \tag{13}$$

At the end of period 2 initial assets if the agent is in a job are worth $\beta^{-1}A_1$, as before, and in the event of unemployment they are worth $\beta^{-1}(A_1 + I_1)$. Because I_1 is a choice variable, the agent can use it to transfer wealth between the states of employment and unemployment so as to maintain the same consumption level in each state. With a full set of insurance contracts the state does not influence the consumption level.

This result, however, is achieved for given transition probabilities. If the insurance company cannot monitor the search or quitting behavior of the worker, the flat consumption profile will give rise to moral hazard that will lead to the breakdown of insurance against both the employment and unemployment risks. Insurance against the unemployment risk gives rise to conventional moral hazard that prevents workers from accepting job offers, of the type commonly analyzed in the unemployment insurance literature. When there is insurance condition (12) changes to

$$R_2 = U^j(A_1) - U^u(A_1 + I_1). \tag{14}$$

With consumption equal in all states of nature both lifetime utilities are equal to $\beta(1 + \beta)u(\bar{c})$, where $\bar{c}$ is the common consumption level, giving the solution $R_2 = 0$, and the same holds in all periods during which the agent searches for another job.

Insurance against the employment risk gives rise to a different type of moral hazard, temporary layoffs. Well-matched workers and firms can gain by colluding to separate temporarily, to enable the worker to collect the contingent claim from the insurance company. The loss to the pair from separating for one period is the marginal product p and the gain is the unemployment subsidy b and the insurance payment I_1. If $b+\beta^{-1}I_1 > p$ this would be an optimal response to the contract, and if this is anticipated by the worker she might choose I_1 such that this condition is satisfied.[11]

4 Optimal employment contracts

When workers have no access to insurance markets for income risk, employment contracts can make Pareto improvements by incorporating contingent transfers between risk-neutral firms and risk-averse workers. I derive the optimal contract under the following assumptions:

1. The firm can monitor the asset position of the worker, who can borrow and lend at a safe interest rate.
2. The firm cannot monitor the worker's search strategy.
3. The firm cannot make payments to unemployed workers but can make either positive or negative payments to employed workers.
4. The firm can monitor the productivity of the job (or the quality of the match).
5. The firm cannot monitor the destination of the worker after separation.

Assumptions 1 and 2 define the typical environment analyzed in the unemployment insurance literature, with the added generalization that the worker can borrow and lend. Assumption 3 implies that there is no private unemployment insurance. The implications of the other two assumptions will be discussed shortly.

An employment contract is *optimal* if it yields at least some exogenous utility level $\bar{U}$ to the worker at minimum cost to the firm. Employment contracts are written at the beginning of period 1. Assumptions 1-5 imply that an employment contract can in general be defined as: a consumption c_1 for period 1, an asset transfer A_1 from the firm to the worker in period 2 if the job survives, a period $T \geq 1$, at the end of which

[11]The moral hazard in this connection is closely related to the one discussed in the literature on temporary layoffs in the absence of perfect experience rating. Feldstein (1977) first showed how partial experience rating leads to excessive temporary layoffs, as firms and workers collude to maximize their revenue from the government subsidy to workers on layoff.

the worker is dismissed if the job does not survive, an asset transfer A_t^j for $t < T$ if the worker quits before dismissal, an asset transfer at dismissal A_T^u, and consumption in periods 2 and 3 in the event of a delay in dismissal, c_t^n for $t = 2, 3$.

Assumption 4 ensures that the worker cannot falsely declare an unproductive job match, quit and collect A_1^j (or A_T^u if $T = 1$). This removes a potential moral hazard problem in the payment of severance compensation to dismissed employees.[12] Assumption 5 implies that the asset transfer at separation, A_T^u, is not contingent on the destination of the worker.[13]

Let now $V(U)$ be the minimum cost of the contract to the firm at the beginning of period 1. U is the lifetime utility that the contract yields to the worker. If the job survives the firm makes a transfer A_1 to the worker at the beginning of period 2 and the contract ends. If the job becomes unproductive the firm either makes a transfer A_1^u and dismisses the worker (and the contract ends) or makes transfer A_1^j and keeps the worker under notice of dismissal. A worker on notice of dismissal in period 2 either receives transfer A_2^u and leaves at the end of the period or stays on for another period of dismissal notice for a transfer A_2^j. Separation at the end of period 2 can be either because the dismissal notice was only for one period or because the worker has found another job, but since the firm cannot monitor the worker's destination the asset transfer is common. In period 3 workers still employed on notice of dismissal do nothing, because the horizon ends at the end of the period.

Thus, the superscript on the asset transfer distinguishes between the case of severance compensation and dismissal, with superscript u, and notice of dismissal and continued employment, with superscript j. In the case of a productive job the asset transfer has no superscript. I first characterize the optimal contracts for any arbitrary duration $T \geq 1$, starting with $T = 1$, before fully characterizing the optimal duration of the contract.

[12] Such conditions on the payment of severance compensation are sometimes found in practice, when the worker is paid compensation when she is fired but not when she quits against the firm's wishes.

[13] Even if the firm can monitor the worker's destination and makes transfers contingent on destination, if the transfer to workers entering unemployment is higher than the one to workers who have accepted another job, workers who find a new job can collude with the new employer to delay hiring. The worker enters in the meantime unemployment, in order to collect the severance payment. This moral hazard problem is similar to the one that gives rise to temporary layoffs and does not allow third-party insurance contracts against the employment risk.

4.1 Severance compensation

Contracts of duration $T = 1$ dismiss the worker at the end of period 1 if the job turns out to be unproductive. The contract is said to contain severance compensation if $A_1^u > A_1$, given that A_1 is the worker's savings at the beginning of period 2. The assumption that the firm can verify whether the job is good or not implies that the severance compensation acts as insurance against the risk of an unproductive job, but not as insurance against the risk attached to the destination of the worker.

The cost of the contract to the firm in period 1, $V_1(U_1)$, satisfies the simple equation

$$V_1(U_1) = \beta \min_{c_1, A_1, A_1^u, R_2} \{c_1 + (1-q)A_1 + qA_1^u\}. \tag{15}$$

The contract is acceptable to the worker if it is worth at least $\bar{U}_1$, constraining the firm to offer

$$\bar{U}_1 \leq U_1. \tag{16}$$

The utility derived by the worker from this contract satisfies the Bellman equations

$$U_1 = \beta \left(u(c_1) + (1-q)U_2^j(A_1) + q\bar{U}_2(A_1^u) \right) \tag{17}$$

$$\bar{U}_2(A_1^u) = G(R_2) \left(U_2^j(A_1^u) - \bar{x}_2 \right) + (1 - G(R_2))U_2^u(A_1^u), \tag{18}$$

The fact that there is search on the job in period 1 and the firm cannot monitor the search strategy of the worker introduces an incentive-compatibility constraint, derived from the unconditional maximization of lifetime utility with respect to the reservation R_2. Maximization of (18) with respect to R_2 yields

$$R_2 = U_2^j(A_1^u) - U_2^u(A_1^u). \tag{19}$$

The optimal contract satisfies minimization conditions for (15) subject to (17), (18) and (19).

It is easy to see that in this case the incentive-compatibility constraint is not binding, due to the fact that the firm dismisses the worker at the end of period 1 if the job is unproductive, whatever the worker's search strategy. The appendix shows that the shadow price of (19) at the optimum is $\lambda = 0$, so the optimal contract satisfies

$$u'(c_1) = U_2^{j\prime}(A_1) = G(R_2)U_2^{j\prime}(A_1^u) + (1 - G(R_2))U_2^{u\prime}(A_1^u). \tag{20}$$

Application of the envelope theorem to lifetime utilities and substitution into (20) yields

$$u'(c_1) = u'(c_1^j) = G(R_2)u'(c_2^j) + (1 - G(R_2))u'(c_2^u). \quad (21)$$

The solution is illustrated in figure 1. The fact that the incentive-compatibility constraint is not binding in this case makes severance payments a first-best insurance against the risk of the termination of the job (the employment risk). In the absence of contingent transfers, proposition 1 demonstrated that the employment risk increases consumption in a good job. (21) shows that consumption in a good job is now flat: the consumption chosen in period 1, c_1, is at the level of consumption in all future periods in a good job, c_1^j.

But the failure of the firm to monitor the destination of the worker, or make payments to her after entry into unemployment, implies that severance compensation does not insure against the unemployment risk. In the event of separation at the end of period 1, the agent's maximization problem in periods 2 and 3 is identical to the one studied in proposition 1, but now with initial assets A_1^u. Consumption increases when the worker goes to another job and falls when she joins unemployment. Equation (21) and proposition 1 yield:

$$c_2^j > c_1^j = c_1 > c_2^u. \quad (22)$$

Similarly, the optimal consumption profile in the event of a second period of unemployment also satisfies proposition 1. When the agent finds a job consumption increases permanently to a higher level but during unsuccessful search it decreases.[14]

It is straightforward to show that severance compensation is positive in this environment. Consumption from period 2 onward in the event of a productive job and consumption in another job accepted in period 2 in the event of an unproductive job both satisfy (3), with initial assets A_1 and A_1^u respectively. The first inequality in (22) then yields $A_1^u > A_1$.

Proposition 2 *The consumption profile is flat in all periods if the first-period job survives to the end of the horizon. If the job ceases to be productive and workers separate without delay, the firm pays positive severance compensation. Consumption falls in period 2 if the worker joins unemployment, or rises if she moves immediately to another job.*

[14] The only qualitative difference between the results of proposition 1 and the policy shown in figure 1 is that in the case of proposition 1 c_1^j jumps up to the level of c_2^j.

4.2 Delayed dismissal

Partial insurance against the unemployment risk can be offered by delaying dismissal in the event of an unproductive job. Delaying dismissal has insurance value because the firm can make payments contingent on the worker's state after the end of the productive relationship. The worker searches on the job during the delay period and so there is a positive probability that she will move to another job without entering unemployment. During this period the firm can effectively monitor the worker's destination, because if the worker quits, it will be to take another job. It can therefore make payments conditional on destination and so increase the insurance value of its contract. It can easily be shown that if the firm can monitor search effort the additional insurance offered by delaying dismissal can fully insure the worker against the unemployment risk but with moral hazard the insurance is second-best.

The disadvantage of delaying dismissal is that the worker cannot claim the unemployment subsidy during delay. One other potential cost and one other benefit of delayed dismissal are ignored in the analysis that follows, without loss of essential generality. If the job is costly to maintain the firm suffers losses by delaying dismissal, which can be avoided if the worker is fired. Against this, a firm may move the worker elsewhere during the delay to perform tasks that have some value to the firm.

4.2.1 One-period delay

For $T = 2$, the firm pays A_1 to workers in productive jobs at the beginning of period 2, A_1^j to workers in unproductive jobs who quit at the beginning of period 2, and guarantees utility U_2^n to workers in unproductive jobs who stay for one more period. The cost of this contract satisfies the Bellman equations

$$V_1(U_1) = \beta \min_{c_1, A_1, A_1^j, R_2, U_2^n} \left\{ c_1 + (1-q)A_1 + q\left(G(R_2)A_1^j + (1 - G(R_2))V_2(U_2^n)\right)\right\} \tag{23}$$

$$V_2(U_2^n) = \beta \min_{c_2^n, A_2^u} \left\{ c_2^n + A_2^u \right\}, \tag{24}$$

where c_2^n is the consumption level of those who remain in unproductive jobs and A_2^u the transfer to those workers at the end of the period, when by assumption they have to leave the job.

This contract yields utility

$$U_1 = \beta \left(u(c_1) + (1-q)U_2^j(A_1) + q\bar{U}_2\right) \tag{25}$$

$$\bar{U}_2 = G(R_2)\left(U_2^j(A_1^j) - \bar{x}_2\right) + (1 - G(R_2))U_2^n, \quad (26)$$

$$U_2^n = \beta\left(u(c_2^n) + G(R_3)\left(U_3^j(A_2^u) - \bar{x}_3\right) + (1 - G(R_3))U_3^u(A_2^u)\right), \quad (27)$$

The following incentive-compatibility constraint also needs to be satisfied

$$R_2 = U_2^j(A_1^j) - U_2^n. \quad (28)$$

The incentive-compatibility constraint for search in period 2, i.e., the choice of R_3, is not binding, for the same reasons that the choice of R_2 did not bind in the case of no delay.

The optimal contract satisfies minimization conditions for (23)-(24) subject to (16) and (25)-(28). As already noted (see footnote 10), the constraint set defined by these conditions may not be convex. I will again ignore the fact that with non-convexities a market for lotteries may open up and focus instead on the case where the parameters make the constraint set convex.

The important result derived in the preceding section for $T = 1$, that severance compensation fully insures the worker against the employment risk, also holds for the case of delay in dismissal, because the firm can differentiate between the asset transfer to workers in productive jobs and the asset transfer to those in unproductive jobs. In addition, because the incentive-compatibility constraint for the last period of the contract is not binding, condition (21) also holds for the last period of the contract:

$$u'(c_2^n) = G(R_3)u'(c_3^j) + (1 - G(R_3))u'(c_3^u). \quad (29)$$

As before, $c_3^j = p + A_2^u/\beta$ and $c_3^u = b + A_2^u/\beta$, and so (29) implies

$$c_3^j > c_2^n > c_3^u. \quad (30)$$

With convexity, the following optimality conditions also hold for this contract:

$$\beta q(1 - G(R_2))(u'(c_1) - u'(c_2^n)) = -\lambda_1 u'(c_1)u'(c_2^n) \quad (31)$$

$$\beta q G(R_2)(u'(c_1) - u'(c_2^j)) = \lambda_1 u'(c_1)u'(c_2^j) \quad (32)$$

$$\beta q g(R_2)(A_1^j - V(U_2^n)) + \lambda_1 = 0 \quad (33)$$

where λ_1 is the shadow price of the incentive-compatibility constraint (28).

A result that emerges readily from (31)-(33) is that if the incentive-compatibility constraints are not binding, which in this case requires monitoring of the worker's search effort, $\lambda_1 = 0$ and consumption is equalized in all states of nature. With perfect monitoring of search effort the firm can offer first-best insurance against the unemployment risk to its employees on notice of dismissal. The Appendix shows (in the proof of proposition 3) that without monitoring of search $\lambda_1 > 0$, and so, by the concavity of the utility function, (31) and (32) imply

$$c_2^j > c_1 > c_2^n. \tag{34}$$

Condition (33) implies also that $V_2(U_2^n) > A_1^j$, i.e., if the firm pays compensation to quitting employees at the beginning of period 2, this compensation is not worth as much as the compensation that it pays to those who are unsuccessful in their search. Quitting employees leave to join another firm and get paid their marginal product, which cannot be less than their current wage. This explains why their compensation is not worth as much as the compensation to those who stay. But why does the firm still pay them compensation? The reason is the incentive-compatibility constraint. The compensation induces more search effort, and counterbalances the search disincentives of the wage payments to workers on notice of dismissal.

Conditions (31) and (32) also imply the "inverted" marginal utility condition

$$\frac{1}{u'(c_1)} = G(R_2)\frac{1}{u'(c_2^j)} + (1 - G(R_2))\frac{1}{u'(c_2^u)}. \tag{35}$$

The inverted marginal utility condition appears to be the result of insurance with moral hazard, and is to be contrasted with the marginal utility condition (21) in the absence of insurance, and with the complete equality of marginal utilities when there is full insurance and no moral hazard.

4.2.2 Two-period delay

For $T = 3$, $V_1(U_1)$ satisfies the Bellman equations

$$V_1(U_1) = \beta \min_{c_1, A_1, A_1^j, R_2, U_2^n} \left\{c_1 + (1 - q)A_1 + q\left(G(R_2)A_1^j + (1 - G(R_2))V_2(U_2^n)\right)\right\} \tag{36}$$

$$V_2(U_2^n) = \beta \min_{c_2^n, A_2^j, U_3^n} \left\{c_2^n + G(R_3)A_2^j + (1 - G(R_3))V_3(U_3^n)\right\}, \tag{37}$$

$$V_3(U_3^n) = \beta \min_{c_3^n} c_3^n. \tag{38}$$

The utility level derived from the contract satisfies:

$$U_1 = \beta\left(u(c_1) + (1 - q)U_2^j(A_1) + q\bar{U}_2\right) \tag{39}$$

$$\bar{U}_2 = G(R_2)\left(U_2^j(A_1^j) - \bar{x}_2\right) + (1 - G(R_2))U_2^n, \tag{40}$$

$$U_2^n = \beta\left(u(c_2^n) + G(R_3)\left(U_3^j(A_2^j) - \bar{x}_3\right) + (1 - G(R_3))U_3^n\right), \tag{41}$$

$$U_3^n = \beta u(c_3^n). \tag{42}$$

There are now two incentive-compatibility constraints, one for search in period 1 and one for search in period 2. In period 3 there is no search. The constraints are:

$$R_t = U_t^j(A_{t-1}^j) - U_t^n, \qquad t = 2, 3. \tag{43}$$

Implicit in the specification of the minimization program is the assumption that the worker will never want to quit into unemployment during the delay in dismissal, i.e. that $U_t^n \geq U_t^u(A_{t-1}^j)$ is not binding for the duration of the contract. This follows trivially from the fact that if it were, the firm would dismiss the worker into unemployment than keep her employed, because this would reduce its costs in future periods (the financing of consumption if the worker were to remain employed).

The results for a two-period delay are natural generalizations of the results for the one-period delay. The optimality conditions yield again $c_1 = c_1^j$ and satisfy (31)-(33). In addition they satisfy:

$$\beta(1 - G(R_3))(u'(c_2^n) - u'(c_3^n)) = -\lambda_2 u'(c_2^n)u'(c_3^n) \tag{44}$$

$$\beta G(R_3)(u'(c_2^n) - u'(c_3^j)) = \lambda_2 u'(c_2^n)u'(c_3^j) \tag{45}$$

$$\beta g(R_3)(A_2^j - V_3(U_3^n)) + \lambda_2 = 0. \tag{46}$$

where λ_2 is the shadow price of (43) for $t = 2$. Once again it can be proved that without monitoring of search $\lambda_2 > 0$ and so

$$c_1 > c_2^n > c_3^n. \tag{47}$$

Consumption falls during unsuccessful search on the job. Condition (35) also holds and a similar condition in the inverted marginal utilities holds also for period 3, in place of (29).[15] Qualitatively, the consumption path for contracts with delay is similar to the one shown in figure 1:

[15] Conditions (44)-(46) hold in all periods of notice of dismissal in longer horizon models, and a condition like (35) always replaces one like (29), whenever a period of delay with on-the-job search replaces a period of unemployment with off-the-job search.

Proposition 3 *If the value function is a concave function of beginning-of-period assets and workers in unproductive jobs are given notice before dismissal, consumption is flat in all states of nature if the firm can monitor its workers' search effort. If it cannot monitor search effort consumption falls during unsuccessful search on the job, or when the worker is dismissed into unemployment, but rises if the worker is successful in her search and quits to take another job.*

The result on the declining consumption profile during unsuccessful search parallels the result first derived by Shavell and Weiss (1979) for unemployment insurance. However, whereas in their case, because of the absence of a capital market, the result required a declining level of unemployment compensation, when there is unlimited borrowing and lending it can be achieved in a variety of ways, which I discuss briefly in section 6. The key property of the optimal contract is that the lifetime utility of the worker who remains employed in an unproductive job falls over time.

Proposition 4 *If the firm's cost function is convex, in the case where search on the job cannot be monitored by the firm the lifetime utility of workers on notice of dismissal falls with the duration of employment.*

4.3 Optimal contract length

A contract of the type studied in section (4.1) dominates non-contingent wage offers because severance compensation provides insurance without causing moral hazard or increasing the firm's costs. The question I investigate here is whether a delay in dismissal can be optimal when severance compensation is allowed, i.e., whether the optimal T can be greater than 1. The optimal T is defined as the T that yields the smallest contract cost for given lifetime utility $\bar{U}_1$, or by its dual, the contract that yields highest lifetime utility for given cost $V_1(U_1)$. It is convenient in this section and the next to discuss the optimality of dismissal delays with reference to the dual, under the assumption of a given contract cost (I later assume that competition drives the cost of the contract to the expected marginal product of the worker).

A delay in dismissal ($T > 1$) may or may not be optimal, partly because of the cost of foregone unemployment compensation and partly because of moral hazard. Consider again the problem studied in section 4.2.1 for $T = 2$. Inspection of the minimization problem shows that if the worker fails to find another job at the end of period 1, the minimization problem from period 2 onward is the dual of the utility maximization problem of a worker with lifetime utility defined by U_2^n, initial assets $V_2(U_2^n)$ and zero income. Therefore, if unemployment compensation is zero, ex post the worker who fails to find a job in period 1 is indifferent between a dismissal delay of one period and immediate dismissal with severance compensation $V_2(U_2^n)$. But ex ante the worker who is offered one more period of employment is better off because of the possibility of successful search at the end of period 1. The firm's failure to monitor the worker's destination after separation implies that if it were to dismiss the worker at the end of period 1, it would have to make the same transfer $V_2(U_2^n)$ to workers who found jobs at the end of period 1 and workers who did not and joined unemployment instead. The firm that gives notice to the worker and keeps her employed for one more period can differentiate between these two payments, paying A_1^j in the event of a quit and $V_2(U_2^n)$ otherwise. I have shown that with risk aversion and moral hazard the firm chooses $V_2(U_2^n) > A_1^j$, although it could choose, if it wanted, $V_2(U_2^n) = A_1^j$. It follows that delaying dismissal always dominates immediate firing when $b = 0$.

This is a general property of the optimization problem in the last period of the contract. The trade-off faced by the firm in the last period of the contract is that it can either set $A_{T-1}^j = V_T(U_T^n)$ and dismiss the worker, so she can receive subsidy b in the event of unsuccessful search, or keep the worker for zero return and partially insure her against the risk of the outcome of search, by offering $A_{T-1}^j < V_T(U_T^n)$ to quitting workers.

Suppose now that instead of zero income, the unemployed enjoy an income which is arbitrarily close to p, their marginal product. Then, trivially (and more formally by an extension of the argument used to prove proposition 1), the worker will never prefer a delay in dismissal over unemployment. Consumption is smoothed completely when the worker can move between employment and unemployment without suffering income loss. By the continuity and monotonicity of value functions with respect to the exogenous income flow during unemployment, it follows that:

Proposition 5 *There is a unique value of unemployment income $b^* \in (0, p)$ such that for $b < b^*$ the optimal contract offers a dismissal delay but for values of $b \geq b^*$ no delay is offered.*

Intuitively, by offering a delay in dismissal the firm partially insures the worker against the unemployment risk for a "premium" b per period. It would appear from this intuition that there should be a "well behaved" relation between the subsidy b and the length of delay offered by the firm, in the sense that as b is increased from a low value, the delay in dismissal falls from two periods to one and to zero. It has not been possible to prove this as a general result, and computations confirm that although for some range of the parameters the "well behaved" relation holds, there are feasible parameter ranges that imply that the optimal contract offers either a two-period delay at low b, or no delay at high b but never a one-period delay.

5 Computations

The model is too stylized to serve as a framework for the calibration of optimal contracts in real situations. In the computations that follow I give reasonable values to the parameters and compute the solutions to demonstrate the optimality of dismissal delays in some ranges of the parameters. As far as possible, I choose the parameters of the model to match known results in consumption and job search decisions.

I choose the value of output p as the normalizing constant, and set it at $p = 1$. The utility function has constant relative risk aversion with coefficient (in the benchmark case) $\sigma = 2$. The calibration studies cited in footnote 5 use mostly a coefficient between 1 and 2 but since this is one of the key parameters of the model, I will compute the solution for a number of different values. The period of analysis in this model is defined by the rate of discount and by the probability of success in job search. Given the short horizon, the value of the rate of discount turns out not to be important in the solution. But the value of the probability of success is important. In the Current Population Survey the fraction of unemployed workers with unemployment duration less than five weeks is about 0.4. In my model this fraction is an unknown but it is influenced by the parameters of the distribution of acceptance costs $G(x)$. I assume that this distribution is exponential, which has only one parameter, and choose the value of the parameter such that in the benchmark case the probability that search is successful after one period is approximately 0.4. This parameter value is $\mu = 0.4$. The discount factor is set at $\beta = 0.995$, implying an annual discount rate of approximately 6 percent (but results are virtually identical for a large range of the discount factor).

The probability that the first job ends, q, is arbitrarily set at 0.5. It turns out that the value of this probability is not important for the results of interest, given the way that the remaining unknowns of the model are computed. Results for other values are also reported. In addition, I report results for the case where job offers in period 1 arrive with a lower probability than they do when the worker is unemployed or on notice of dismissal, to capture the idea that the worker searches with lower intensity in period 1 because she does not know with certainty that the job will end.

The choice of parameter for the distribution $G(x)$ makes the period of analysis one month. A problem faced with such a short horizon is that the wealth effects from changes in a single period's income become important. For example, in the case where unemployment income is $b = 0$, a person who experiences one period of unemployment has only two thirds the lifetime income of someone who does not experience unemployment. In practice unemployment durations are too short to have such a big impact on lifetime wealth, although evidence shows that at least for the unemployed, the impact of current income on consumption is more than the impact implied by a permanent-income model (Hansen and Imrohoroglu, 1992, Gruber, 1997). In some computations that I did with the infinite horizon version of this model (not reported), the impact of unemployment on consumption with the assumed capital market structure was negligible. Gruber (1997) reports that in the PSID, when a worker who is not entitled to unemployment compensation becomes unemployed, consumption drops by 22 to 25 percent and that if the replacement rate (the ratio of unemployment income to the wage rate) is increased by ten percentage points the drop falls by about 3 percent.

Because the impact of unemployment income on consumption is crucial for the results of this paper, I attempted to match the facts reported by Gruber in the benchmark case. There are two ways in which the facts can be matched. The horizon can be extended or shortened until the wealth effect from becoming unemployed is sufficiently big to give the required fall in consumption. Or, the value of b can be given a more general interpretation as the sum of non-market income (home production) and unemployment compensation and a value for it chosen such that in the three-period horizon case the drop of consumption in the benchmark case is in the range 22-25 per cent. This value of b can then be interpreted as the one corresponding to zero unemployment compensation.[16]

[16]Note that I have chosen the normalization $p = 1$ but the assumptions that initial assets and total income in the absence of unemployment insurance are both zero are also normalizations. The argument made in this paragraph is that to match Gruber's facts one of these normalizations has to be relaxed.

I have experimented with both methods and the results for the key question of the model, the optimality of dismissal delays, turned out to be virtually identical. I report only results for the three-period horizon and various values of b, but take $b = 0.2$ as the benchmark case because of its implications for the drop in consumption in the first month of unemployment. It also turns out in this case that the marginal effect of increasing b reduces this drop by an amount of the order of magnitude reported by Gruber. So in what follows $b = 0.2$ should be interpreted as the value of home production consistent with the consumption facts in the three-period model.

The computations were carried out under the assumption that competition ensures that the expected discounted value of earnings is equal to the expected discounted value of marginal product. Table 1 gives some results for different parameter values. I give the percentage consumption drop in period 2 in the event of unemployment, the probability that the worker becomes unemployed when the first job is unproductive (i.e., the product $q(1 - G(R_2))$, and the unemployment hazard after one period of unemployment, $G(R_3)$, for the case where only severance compensation is offered by the contract. It turns out that the benchmark case of $\sigma = 2$ and $b = 0.2$ implies that either a contract with severance compensation only or one that also includes a one-period delay before dismissal give virtually identical lifetime utilities (see below) and very similar results. The numbers in brackets in the table are for the optimal contract when this is not the one with severance compensation and no delay in dismissal.

The results are as expected. More risk averse individuals take action that reduces their consumption drop in the event of unemployment, and are more prepared to move in order to reduce both the probability of entering unemployment and the duration of unemployment. Unemployment income has a big impact on the drop in consumption, because of income effects, and also a big impact on unemployment experience, because of moral hazard. A higher probability that the first job becomes unproductive in the second period (higher q) increases the consumption drop because the higher q reduces the expected lifetime contribution of the worker to the firm, and so the firm offers less insurance in the form of severance compensation. Not surprisingly, the higher q also increases the probability that the worker will enter unemployment, so the worker will want to save more to finance consumption in the event of unemployment. The lower expected income, however, dominates these two conflicting influences. To check that this is the reason for the larger consumption drop, I also computed the solution for the case $q = 0.5$ but with a lower probability of moving directly to another job if the first job turns out to be unproductive in period 2. This effect was obtained by multiplying the probability of moving to another job, $G(R_2)$, by a constant a, and computing the solution for $a = 0.5$. It can be justified by the assumption that on-the-job search is less efficient than off-the-job search. The result shows a much smaller drop in consumption than in the benchmark. The firm increases the severance compensation in this case because of the higher probability of entering unemployment.

Finally, the results show that a more efficient search, in the form of lower mobility costs (higher μ in the table), increases the consumption drop in the event of unemployment because of lower probability of entering unemployment and lower probability of remaining unemployed.

The numerical solutions were next used to illustrate the paper's main results, that dismissal delays are more likely to be offered to more risk averse workers and to workers with low income during unemployment. Figure 2 plots the value of each contract type to the worker for the benchmark case of $\sigma = 2$ against unemployment income. The value of the contract with no severance and no dismissal delay is not shown because it is always dominated by the contract with severance compensation only. The figure confirms proposition 5, that at $b = 0$ the contract with the longest delay is optimal, and at sufficiently high b no delay in dismissal is optimal. Of course, the contract with a two-period delay is independent of unemployment income, because the worker never enters unemployment, whereas the contract with no delay shows the strongest dependence on unemployment income because there is a positive probability that the worker will experience up to two periods of unemployment. Figure 2 confirms that in the benchmark case no delay, or a one-period delay, give virtually identical utilities. It also satisfies the intuitive result that could not be proved, that at intermediate values of unemployment income, a delay of one period is optimal. This result, however, is reversed at some parameter values; i.e., it is possible to find feasible ranges of the parameters which imply that the line for the one-period delay contract is entirely below the other two lines (for example, at large values of μ, when the efficiency of search is high).

Figure 3 plots the regions in a space of risk aversion against unemployment income for which each contract type is optimal. The results conform with intuition and with the propositions proved in this paper: more risk averse workers are offered dismissal delays and the delay is less when exogenous unemployment income is higher. But it also shows that more risk aversion is needed than the values normally used in calibrations of labor-market equilibrium models to obtain delays in dismissals at non-trivial unemployment compensation levels.

In addition to the influence of risk aversion and unemployment income, the results in brackets in table 1 also show that a delay is optimal when the period 1 job is more likely to prove a success and when the probability of success in job search is higher. The intuition for these results is also clear. In the first case, the worker is more likely to be productive in periods 2 and 3, so the firm can afford to offer more insurance in the form of longer dismissal delays in the event of a mismatch. In the second case the worker is more likely to quit if she is unproductive, so the delay is less expensive to the firm. The results in the table also show that the search disincentive effects of the contract are stronger when a delay is offered than when the worker is dismissed with severance compensation only. The probability of leaving unemployment (last row of table 1) is always higher in the case of severance compensation only than in the bracketed terms of delayed dismissals.

6 Conclusion

This paper has established that in the absence of insurance against income risk, a firm that cannot monitor the search strategy of its workers will offer employment contracts that include severance payments to dismissed employees and, under certain conditions, also delays in dismissal. An important result is that a delay in dismissal is less likely to be offered the more generous is the compensation to unemployed workers offered by an exogenous unemployment insurance system. Another important result is that the firm offers incentives to employees in unproductive jobs, whose dismissal is delayed for insurance reasons, to search on the job and quit. These incentives take the form of severance compensation in the event of a quit and income transfers that induce a fall in consumption during unsuccessful search.

An important issue for future work is the joint optimality of employer-provided insurance and government-provided insurance. There are two aspects to this issue. The first is whether there is scope for both employer-provided insurance (in the form of severance compensation and dismissal delays) and optimally-designed unemployment insurance by the state. The results of the paper hint that severance compensation may still be optimal but dismissal delays may not be, given that even modest levels of unemployment compensation reduce the optimal contract to one of severance compensation only. But it may turn out that when employment contracts can offer the dismissal delays the optimal level of unemployment insurance is one that is sufficiently low to make these delays optimal.

The second aspect is whether severance compensation and dismissal delays should be in legislation and not in private contracts, and jointly designed with unemployment insurance. Firms in our framework have an incentive to renege on the contract once the job proves to be unproductive, because they are required to make payments to employees who are quitting. We did not examine the conditions under which this incentive can give rise to legislated "employment protection." It is, however, interesting that there appears to be a negative correlation across countries between legislated employment protection measures and the generosity of unemployment compensation (See Boeri et al. 2001). Such correlation may reflect the model's prediction that firms offer less protection when the unemployment compensation offered by the government is more generous and suggest that there is scope for further research on the joint determination of optimal employment contracts and unemployment insurance.

The method used to derive the results is that of a principal, the firm, minimizing the cost of offering a contract worth a pre-determined utility level to an agent, the worker. The implementation of the optimal contract was not discussed but it is straightforward to describe some of its main features. The key to the results is twofold. First the worker should be made progressively worse off during a dismissal delay. Second, the contract should be able to differentiate between the asset transfer to workers who quit and the asset transfer to those who stay with the firm.

Two features of labor contracts, commonly found in practice, have both these implications. The first, is the "up or out" feature. The firm pays a worker initially a wage which is less than marginal product and if the job proves productive it raises the wage in subsequent periods. If the match is unproductive it dismisses the worker, with or without notice, but with some severance compensation. The "promotion" implied by the pay rise makes those who fail to get it worse off than their expectation when starting a job, giving incentives for search on the job. The severance compensation paid to those who leave differentiates the asset transfer of those who leave from the asset transfer of those who stay.

Numerical computations confirm that more risk averse workers and workers with less entitlement to unemployment compensation are more likely to be offered contracts with dismissal delays. More efficient search increases this likelihood, despite the fact that in all cases there are more disincentives to search during a delay in dismissal than during unemployment.

References

[1] Acemoglu, D. and R. Shimer (1999). "Efficient Unemployment Insurance." *Journal of Political Economy* 107, 893-928.

[2] Andolfatto, D. and P. Gomme (1996). "Unemployment Insurance and Labor-Market Activity in Canada." *Carnegie-Rochester Conference Series on Public Policy* 44, 47-82.

[3] Azariadis, C. (1975). "Implicit Contracts and Underemployment Equilibria." *Journal of Political Economy* 83, 1183-1201.

[4] Baily, N. (1974). "Wages and Employment under Uncertain Demand." *Review of Economic Studies* 42, 37-50.

[5] Bertola, G. (1999). "Microeconomic Perspectives on Aggregate Labor Markets." In *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card. Amsterdam: North-Holland.

[6] Bertola, G. (2004). "A Pure Theory of Job Security and Labour Income Risk." *Review of Economic Studies* 71: 43-61.

[7] Blanchard, O. J. and J. Wolfers (2000). "The Role of Shocks and Institutions in the Rise of European Unemployment: the Aggregate Evidence." *Economic Journal* 110, C1-C33.

[8] Boeri, T., A. Borsch-Supan and G. Tabellini (2001). "Welfare State Reform." *Economic Policy* 32, 7-44.

[9] Burdett, K. and R. Wright (1989). "Unemployment Insurance and Short-Time Compensation: The Effects of Layoffs, Hours per Worker, and Wages." *Journal of Political Economy* 97, 1479-1496.

[10] Costain, J. S. (1997). "Unemployment Insurance with Endogenous Search Intensity and Precautionary Saving." University of Pompeu Fabra Economics Working Paper no. 243.

[11] Danforth, J. P. (1979). "On the Role of Consumption and Decreasing Absolute Risk Aversion in the Theory of Job Search." In *Studies in the Economics of Search*, edited by S. A. Lippman and J. J. McCall. Amsterdam: North-Holland.

[12] Feldstein, M. S. (1977). "Temporary Layoffs in the Theory of Unemployment." *Journal of Political Economy* 84, 937-957.

[13] Gordon, D. F. (1974). "A Neoclassical Theory of Keynesian Unemployment." *Economic Inquiry* 12, 431-459.

[14] Gruber, J. (1997). "The Consumption Smoothing Benefits of Unemployment Insurance." *American Economic Review* 87, 192-205.

[15] Hansen, G. D. and A. Imrohoroglu (1992). "The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard." *Journal of Political Economy* 100, 118-142.

[16] Hopenhayn, H. A. and J. P. Nicolini (1997). "Optimal Unemployment Insurance." *Journal of Political Economy* 105, 412-438.

[17] Jovanovic, B. (1979). "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87, 972-990.

[18] Lazear, E. (1990). "Job Security Provisions and Unemployment." *Quarterly Journal of Economics* 55, 699–726.

[19] Lentz, R. and T. Tranaes (2005). "Job Search and Savings: Wealth Effects and Duration Dependence." *Journal of Labor Economics* 23, 467-90.

[20] Ljungqvist, L. and T. J. Sargent (1998). "The European Unemployment Dilemma." *Journal of Political Economy* 106, 514-550.

[21] Mortensen, D. T. and C. A. Pissarides (1999). "Unemployment Responses to 'Skill-Biased' Shocks: The Role of Labour Market Policy." *Economic Journal* 109, 242-265.

[22] Nickell, S. J. (1982). "The Determinants of Equilibrium Unemployment in Britain." *Economic Journal* 92: 555-575.

[23] Nickell, S. J. (1997). "Unemployment and Labor Market Rigidities: Europe versus North America." *Journal of Economic Perspectives* : 55-74.

[24] Nickell, S. J. and R. Layard (1999). "Labor Market Institutions and Economic Performance." In *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card. Amsterdam: North-Holland.

[25] OECD (1999). "Employment Protection and Labour Market Performance." In *Employment Outlook*, edited by OECD. Paris: OECD.

[26] Pissarides, C. A. (2000). "Employment Protection." *Labour Economics* 8,131-159.

[27] Saint-Paul, G. (2002). "The Political Economy of Political Protection." *Journal of Political Economy* 110, 672-704.

[28] Sampson, A. A. (1978). "Optimal Redundancy Compensation." *Review of Economic Studies* 45, 447-452.

[29] Shavell, S. and L. Weiss (1979). "The Optimal Payment of Unemployment Insurance Payments over Time." *Journal of Political Economy* 87, 1347-1362.

[30] Wang, C. and S. Williamson (1996). "Unemployment Insurance with Moral Hazard in a Dynamic Economy." *Carnegie-Rochester Conference Series on Public Policy* 44, 1-41.

[31] Wilde, L. L. (1979). "An Information-Theoretic Approach to Job Quits." In *Studies in the Economics of Search*, edited by S. A. Lippman and J. J. McCall. Amsterdam: North-Holland.

[32] Wright, R. (1986). "The Distributional Roles of Unemployment Insurance and the Dynamics of Voting." *Journal of Public Economics* 31, 377-399.

7 Appendix

7.1 Proofs of Propositions

Proof of Proposition 1. It follows immediately from (1), (2) and the restriction $b < p$ that $c_3^j > c_3^u$. Therefore, by the concavity of the utility function and (6), $c_3^j > c_2^u > c_3^u$. Suppose $c_2^u > c_2^j$. Then from (3) and (5) I obtain

$$A_2 - \frac{1}{1+\beta} A_1 + p - b < 0. \tag{48}$$

But $c_2^u > c_2^j$ implies $c_3^j > c_2^j$ because $c_3^j > c_2^j$, and so (3) and (1) yield

$$A_2 - \frac{1}{1+\beta} A_1 > 0, \tag{49}$$

which contradicts (48). Therefore, $c_2^j > c_2^u$ and by the concavity of the utility function and (11), $c_2^j > c_1 > c_2^u$.

Proof of Proposition 2. Let $\mu > 0$ be the Lagrangian multiplier associated with the constraint in (16), given the definition of U in (17), and λ_1 the one for (19). The condition that minimizes (15) with respect to R_2 subject to (16) and (19) is

$$\mu q g(R_2)[U_2^j(A_1^u) - R_2 - U_2^u(A_1^u)] + \lambda_1 = 0,$$

which, in view of (19) immediately yields $\lambda_1 = 0$. The conditions that minimize (15) with respect to c_1, A_1 and A_1^u are

$$1 - \mu u'(c_1) = 0 \tag{50}$$

$$1 - \mu U_2^{j\prime}(A_1) = 0 \tag{51}$$

$$1 - \mu[G(R_2)U_2^{j\prime}(A_1^u) + (1 - G(R_2))U_2^{u\prime}(A_1^u)] = 0. \tag{52}$$

Application of the envelope theorem to (51)-(52) proves the proposition. Equality between c_1 and c_1^j follows immediately and the other results follow from proposition 1.

Proof of Proposition 3. Consider the firm's choice problem for $T = 2$, that for $T = 3$ following immediately. The firm minimizes (23) subject to (16) and (25)-(28). Let $\bar{U}_2^n$ be the inherited (state) value of the contract to the worker at the beginning of period 2, given the choices made by the firm in period 1. Let μ as before be the shadow price of (16), μ_2 the shadow price of the constraint $\bar{U}_2^n \leq U_2^n$ and λ_1 the Lagrangian for the constraint in (28). As before, minimization with respect to the controls c_1 and A_1 yield (50) and (51), so $c_1 = c_1^j$. Minimization with respect to c_2^n, U_2^n and A_1^j yields

$$1 - \mu_2 u'(c_2^n) = 0 \tag{53}$$

$$\beta q(1 - G(R_2))(V_2'(U_2^n) - \mu) + \lambda_1 = 0 \tag{54}$$

$$\beta G(R_2)(1 - \mu U_2^{j\prime}(A_1^j)) - \lambda_1 U_2^{j\prime}(A_1^j) = 0. \tag{55}$$

The envelope theorem gives $V_2'(U_2^n) = 1/u'(c_2^n)$ and $U_2^{j\prime}(A_1^j) = u'(c_2^j)$, which, upon substitution into (54)-(55) yields (31)-(32). Minimization with respect to A_2^u yields

$$1 - \mu_2[G(R_3)U_2^{j\prime}(A_2^u) + (1 - G(R_3))U_2^{u\prime}(A_2^u)] = 0, \tag{56}$$

which can be used in conjunction with (53) and the envelope theorem to derive (29). Applying the results of proposition 1 gives (30). Finally, minimization with respect to R_2 immediately yields (33).

If the firm can monitor search effort the incentive compatibility constraints are not binding and so $\lambda_1 = 0$. It then follows that $c_2^j = c_2^n = c_1$. If the firm cannot monitor search the proposition's result hinges on the sign of λ_1. I now show that if the incentive compatibility constraints are binding, $\lambda_1 > 0$.

Suppose $\lambda_1 < 0$ and so by (33), $A_1^j > V_2(U_2^n)$. From (31) and (32), $c_2^n > c_1 > c_2^j$ and so by (30) $c_3^j > c_2^j$. Given that $c_3^j = p + A_2^u/\beta$ and $c_2^j = p + A_1^j/\beta(1+\beta)$, it follows that $A_2^u > A_1^j/(1+\beta)$. To get a contradiction, note that (24) and $A_1^j > V_2(U_2^n)$, derived from (33) when $\lambda_1 < 1$, yield

$$\begin{aligned} A_1^j &> \beta(c_2^n + A_2^u) \\ &> \beta(c_2^j + A_2^u) \\ &= \beta(p + A_1^j/\beta(1+\beta) + A_2^u), \end{aligned}$$

which yields the contradiction $A_2^u < A_1^j/(1+\beta)$. Therefore $\lambda_1 > 0$. By an analogous argument it can be shown that $\lambda_2 > 0$ in the case $T = 3$.

Proof of Proposition 4 From the optimization conditions (53)-(54), and given the envelope property $V_1'(U_1) = 1/u'(c_1) = \mu$,

$$\beta q(1 - G(R_2))(V_2'(U_2^n) - V_2'(U_2)) = -\lambda_1 < 0.$$

Therefore, if $V_2(.)$ is convex, $U_2 > U_2^n$. A similar property holds for $T = 3$, which yields $U_2^n > U_3^n$.

Proof of Proposition 5 The proof follows immediately from the argument made in the text. Let $b = 0$ and compare the maximization problem when there is a one-period delay and a two-period delay. The income flow in the two cases is the same $(p, 0, 0)$. The optimal allocations obtained from (23)-(27) are identical to the optimal allocations that can be obtained from (36)-(43) by imposing the additional constraint $A_2^j = \beta c_3^n \equiv A_2^u$. Therefore, a two-period delay always dominates a one-period delay when $b = 0$. Working backwards a similar argument shows that a one-period delay dominates no delay.

Consider next $b = p$. With no delay the income flow in the event of an unproductive job is (p, p, p), with a one-period delay it is $(p, 0, p)$ and with a two-period delay $(p, 0, 0)$. The firm will never delay because in the former case allocations replicate the full-insurance equilibrium. Since the maximization in (36)-(43) is independent of b but utility is increasing in b in all other contract types a two-period delay is always dominated as b increases towards p. A similar argument shows that a one-period delay is also dominated as b increases towards p.

Table 1
Numerical Solutions

	bench mark	σ		b		q		a	μ	
		1	4	0	.5	.2	.7	.5	.2	.6
cons drop	.23	.26	(.17) .22	(.24) .35	.11	.17	(.29) .28	.09	.11	(.27) .37
prob enter unempl	.24	.26	(.22) .17	(.26) .17	.37	.12	(.36) .28	.27	.36	(.17) .11
prob exit unempl	.37	.32	(.57) .61	(.35) .57	.20	.27	(.43) .47	.31	.16	(.61) .82

Notes.

1. The solutions without brackets are for the contract with severance compensation but no delay in dismissal. When this contract is not optimal the optimal solution is given in brackets. There are four columns with bracketed terms. When σ=4 or b=0 or μ=.6 the optimal contract requires a two-period delay and when q=.7 it requires a one-period delay. Comparing for example the two entries in row 1 of the column headed b=0, we find that if the firm offers only severance compensation the consumption of dismissed employees drops by 35 per cent, but when the optimal contract of a two-period delay is offered, the consumption drop is only 24 per cent.
2. The benchmark case is σ=2, b=.2, q=.5, a=1, μ=.4, p=1.
3. The consumption drop is the percentage drop in consumption from period 1 to period 2 when the job is revealed to be unproductive. The probability of entering unemployment is the product $q(1\text{-}G(R_2))$, which gives the probability of entering unemployment (or a period of delayed dismissal) when the job in period 1 turns out to be unproductive. The probability of exiting unemployment is $G(R_3)$, the probability of becoming employed in a productive job in period 3.

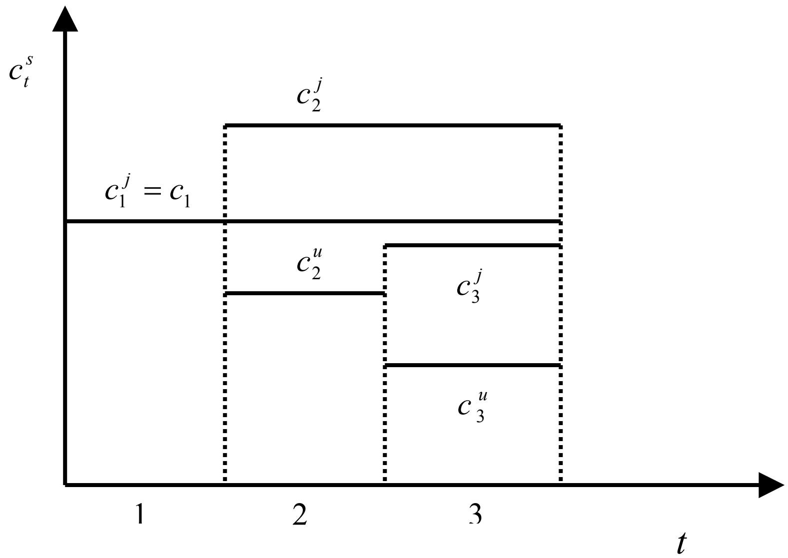

Figure 1
The consumption profile with severance compensation

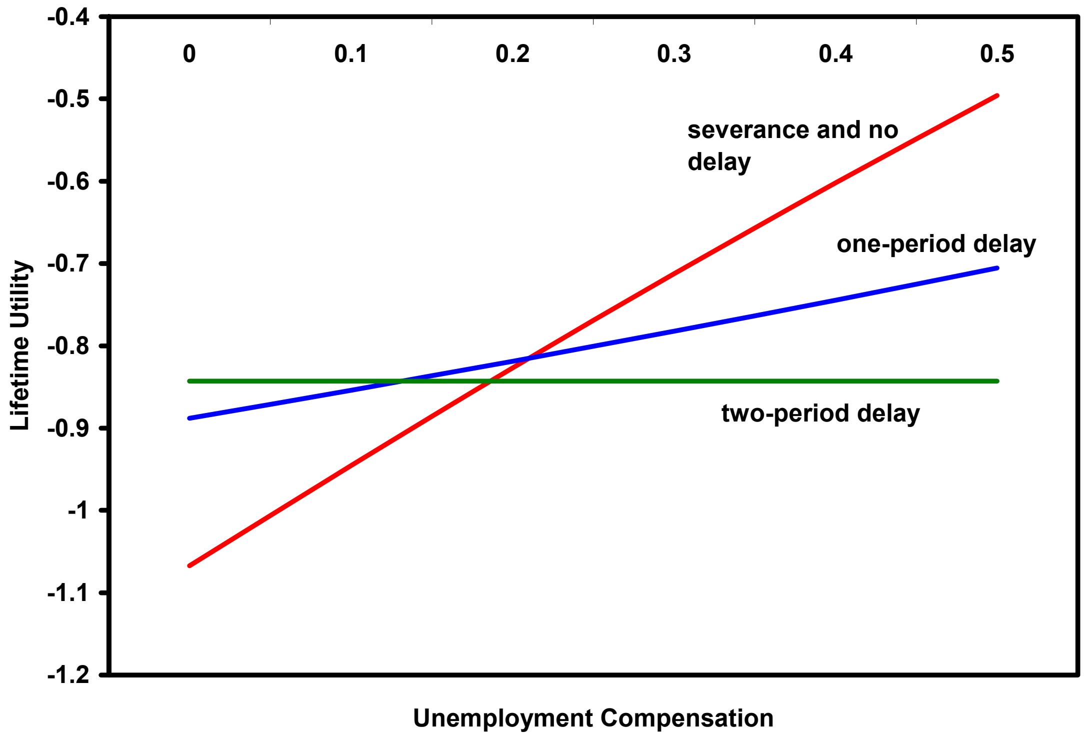

Figure 2

Unemployment compensation and the lifetime utility of three contract types (risk aversion $\sigma = 2$)

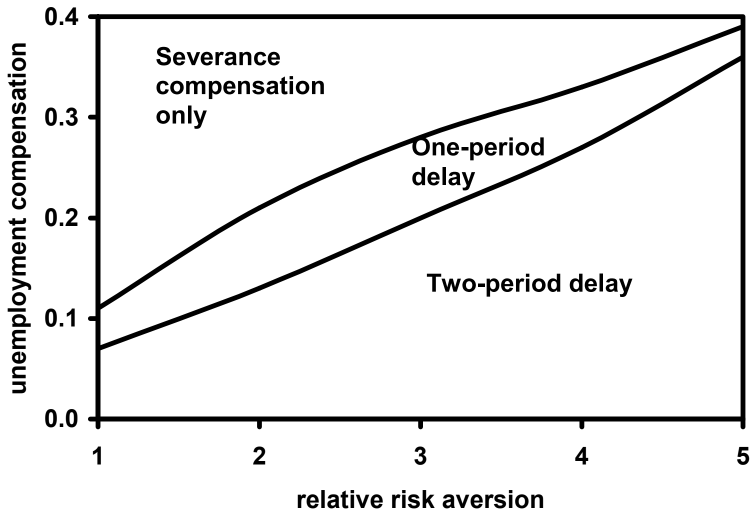

Figure 3
Optimal contract regions: the influence of risk aversion and unemployment compensation