Theoretical aspects of unemployment insurance

Bertil Holmlund
Department of Economics, Uppsala University*

April 28, 2015

* E-mail: bertil.holmlund@nek.uu.se. The paper is written for the Danish commission on unemployment insurance. I thank Michael Rosholm and the commission’s secretariat for excellent comments on previous drafts.
Theoretical aspects of unemployment insurance

1. Introduction 3

2. A multitude of policy issues 4

3. How does unemployment insurance affect labor market outcomes? 7
   The microeconomics of job search 8
   Matching and equilibrium search 9
   Financing unemployment compensation 11
   Other considerations 14

4. Designing optimal unemployment insurance 15
   The optimal time profile of benefit payments 16
   Monitoring, sanctions and workfare 23
   Part-time unemployment 29
   Financing issues 30
   Estimating optimal unemployment insurance 32

5. Concluding remarks 34

References 36
1. Introduction

The economics of unemployment insurance (UI) has been an active research area in recent decades. Much of this research has been concerned with positive analysis, such as the impact of UI benefit levels on the duration of unemployment. Less research has been devoted to normative issues concerning the design of an optimal UI system. The fundamental rationale for unemployment insurance is to provide income insurance for risk-averse workers. A unified assessment of UI must consider the welfare gains associated with insurance as well as potentially adverse incentive effects. Indeed, recent years have seen a resurgence of interest in the economics of optimal UI design and provided new insights into the tradeoffs between insurance benefits and incentives.

Unemployment insurance has rarely been provided by the private sector, a phenomenon that often is attributed to informational asymmetries that are unique to this form of insurance. There are basically two kinds of asymmetric – or hidden – information: hidden actions and hidden characteristics. Hidden actions involve imperfect information on the part of the insurance provider regarding actions taken by the insured individual. Hidden characteristics pertain to the characteristics of the insured that are unobserved by the insurer. Hidden actions give rise to moral hazard problems whereas hidden characteristics may cause adverse selection.

The moral hazard problems associated with UI appear in various guises. The most frequently discussed margin of adjustment is probably job search. According to standard search theory, the properties of the UI system affect the individual unemployed worker’s search effort and reservation wage. For example, a well-known result from a prototype search model states that higher UI benefits raise the reservation wage and thereby also the expected duration of unemployment. Another potential source of moral hazard arises from decisions taken by employed workers regarding work effort. An employed worker may raise the probability of retaining the job by working rather than shirking and the level of UI benefits may conceivably influence the choice of effort since the cost of job loss is affected. Moral hazard problems may also appear through the linkage between UI benefits – and UI financing – and the design of wage and employment contracts.

Some actions taken by insured agents may be unobserved by the UI provider but not necessarily unobservable; information may be available at a cost. Monitoring of job search
and work tests are examples of devices whereby the insurer can obtain information about
search effort and availability for work among the unemployed.

The consequences for UI design of hidden characteristics among the insured depend on the
precise nature of the individual heterogeneity. One example is differences in unemployment
risks that are unobserved to the insurer but private information to the insured. These
differences can have many sources, including differences in individual productivity, search
effectiveness, or preferences for leisure. The literature on adverse selection problems in UI is
relatively small, perhaps reflecting the analytical difficulties of incorporating worker
heterogeneity in a tractable manner.¹ Adverse selection may provide a rationale for mandatory
insurance.

There are other reasons than adverse selection for public UI. Arguably most important is the
fact that unemployment risks tend to be positively correlated across workers. Private
insurance firms may not survive adverse macroeconomic shocks that lead to sharp and long-
lasting increases in overall unemployment.

This paper provides a brief survey of some theoretical aspects of unemployment insurance.²
Section 2 includes a brief overview of some of the key policy instruments in this area. In
section 3 we ask how UI policies affect various labor market outcomes, such as the duration
of unemployment spells and the overall rate of unemployment. In section 4 we consider
normative issues, i.e., how UI policies should be designed. This analysis recognizes that UI
offers welfare-improving benefits to workers while also entailing adverse incentive effects.
Most of the exposition deals with theoretical issues but some recent empirical references are
briefly mentioned.³

2. A multitude of policy issues
Unemployment insurance is a complex system that involves a large number of policy issues.
The list that follows is by no means is exhaustive.

One important policy choice concerns whether or not UI should be mandatory or voluntary.
Most advanced countries have adopted mandatory systems, probably because of some of the
reasons discussed in the introduction. In a voluntary system, benefit receipt is conditioned on

¹Jones (1986) and Chiu and Karni (1998) are two contributions of note. In both models, the absence of private
UI despite risk-averse individuals is explained by adverse selection.
²The paper draws in part heavily on Fredriksson and Holmlund (2006b). See also Tatsiramos and van Ours
(2012) and Holmlund (1998) for recent and not so recent surveys of UI issues. Chetty and Finkelstein (2013)
includes a broad discussion of social insurance issues in general.
³Policies not discussed include unemployment savings accounts and business cycle dependent benefits.
past active decisions to join the UI system. Denmark, Finland and Sweden are among the few countries that maintain voluntary UI where union-affiliated UI funds play important roles, an arrangement known as a Ghent system. See Parsons et al. (2015) for description and analysis of the Danish system.

For several reasons, the differences between mandatory and voluntary UI may not be large in practice. UI coverage – the fraction of unemployment covered by UI benefits – may be high in a voluntary system if UI participation is inexpensive as a result of substantial government subsidies (as has been the case in the aforementioned Nordic countries). Moreover, effective UI coverage may be low in a mandatory system if the qualification rules are very strict.

The distinction between mandatory and voluntary systems is also blurred if there are other forms of compensation – unemployment assistance – available for workers who have decided not to join the “regular” UI system. In fact, the Swedish system is arguably best characterized as a hybrid between a mandatory and a voluntary model since unemployed workers outside the UI funds are covered by basic unemployment assistance at a rate substantially lower than regular benefits. In addition, welfare payments have features in common with unemployment assistance, the main difference being that welfare payments usually are means tested and based on family income.

The generosity of the UI system is typically measured by two parameters, namely the benefit level and the potential duration of benefit receipt. The benefit level is usually given as a fraction of past earnings – a replacement rate – up to a cap (and sometimes also involving a floor). The potential duration is closely related to the time profile of benefits over the unemployment spell. The replacement rate (or the cap) may change over the spell, e.g. decline in a stepwise fashion.

In most countries, potential benefit duration does not vary over the business cycle. In some countries, recessions may trigger extensions of benefit duration. In Canada, these extensions are rule-based according the pre-determined formulas. In the US, severe recessions may trigger discretionary decisions on extensions. The recent Great Recession is a case in point: benefits were extended from the normal period of 26 weeks to an unprecedented length of 99 weeks.

There may be incentives for some workers, or groups of workers, to top up their benefits via individual or collective insurance arrangements. This raises the question whether there should
be government regulations of such *topping up schemes*. In Sweden, many labor unions offer topping up schemes that effectively raises the cap. The replacement rate cannot be raised beyond the statutory rate, however.

UI schemes often include a *waiting period* before benefits are paid out. Benefit receipt is conditioned on *eligibility conditions* that can be defined in terms of past employment or past earnings. The strictness of these rules is crucial for effective UI coverage. Some countries have systems where the potential duration of benefit receipt is increasing in time spent as employed. Eligibility for UI is typically restricted to workers who have been exposed to involuntary job terminations; workers who quit into unemployment are usually not eligible. In practice, the distinction between quits and layoffs is often blurred and the policy implementation may require detailed operational definitions.

Benefit eligibility is typically conditioned on some form of active job search which may or may not involve close monitoring. Registration at a public employment agency is generally required. Failure to meet search requirements (or rejections of “suitable” job offers) may lead to benefit *sanctions*, i.e., partial or complete withdrawal of benefit payments.

The *financing* of UI typically works via general income or payroll taxes. The main exception here is the US system of *experience rating* where UI taxes are levied on employers and partly based on how many workers the employer has laid off in the past. In the few countries with Ghent systems, membership fees account for part of the financing.

Existing UI systems focus mainly on workers who are fulltime unemployed after having been separated from fulltime jobs. For many workers, underemployment may be a more relevant concept than unemployment. Underemployment as measured by ILO captures “partial lack of work” as opposed to unemployment which captures “total lack of work”, both categories referring to a survey week. This raises the question of how *part-time unemployment* should be treated by the benefit system.

*Seasonal unemployment* represents another dimension of underemployment where workers are employed for a fraction of the year and temporarily laid off during the off-season. Part-time work would then be defined over the course of the year rather than pertaining to work hours during a survey week. Seasonal unemployment is largely predictable and the case for publicly provided UI therefore weaker. If seasonal unemployment were completely
predictable, there would be no uncertainty regarding employment and no reason to demand insurance as a means to smooth consumption.

Another dimension of underemployment may be thought of as cyclical, with the German short-time work schemes (Kurzarbeit) being a paramount example. These schemes cushion the recessionary impacts on unemployment by offering wage subsidies to workers on part-time work. Instead of laying workers off in times of economic crisis, hard-pressed companies are encouraged to reduce all of their employees’ hours and pay — and the government pays a portion of the workers’ unemployment benefits to help make up the difference.

UI rules for self-employed people are typically stricter than the rules for wage earners. A complete business closure may be required to guarantee that UI payments are not used as a subsidy to the business.

International migration adds new issues regarding eligibility conditions and monitoring of job search. It might seem reasonable to count work abroad as equivalent to work in the domestic country in terms of eligibility for UI. In practice, it may be difficult to get accurate information about work abroad. Moreover, it will presumably be much more difficult to monitor job search abroad than job search domestically.

UI interacts with other labor market policies, such as active labor market policy (ALMP) and employment protection legislation (EPL). ALMP may offer income support and effectively extend benefit periods. ALMP may also be an important ingredient of the monitoring system. By providing income support during joblessness, UI may reduce the political support for EPL. UI also interacts with other social insurance policies, sickness insurance being a case in point. Incentives for “benefit arbitrage” may arise if there are large differences in compensation levels between spells of sickness and unemployment.

Finally, there are issues related to whether or not family conditions should affect the UI rules. More generally, should UI be based exclusively on individual conditions or should the household be treated as the relevant unit? It is not uncommon that the UI system treats individuals with children more generously. Nor is it uncommon that individual age affects the potential duration of benefit receipt.

3. How does unemployment insurance affect labor market outcomes?
Theoretical research over the recent decades has identified various mechanisms through which UI policies can affect labor market outcomes. The early research from the 1970s
focused on how UI influenced the individual unemployed worker’s incentives to find and accept job offers. This research is closely linked to the evolution of search theory from the early 1970s and onwards. The UI policies may affect job finding via unemployed workers’ efforts to search for jobs as well as via how choosy they are when it comes to accepting job offers.

UI policies may have numerous other effects beyond those studied by means of microeconomic search models that focus on the outflow from unemployment. The policies may, for example, influence the rate at which employed workers quit into unemployment in order to search for better jobs. UI may also affect wage bargaining and thereby also firms’ decisions on hirings and firings.

This section provides a brief overview of some of the main mechanisms whereby UI may affect various labor market outcomes. The focus is theoretical although some relevant empirical evidence is sometimes alluded to.

The microeconomics of job search
The basic microeconomic theory of job search portrays an individual unemployed worker facing a known and stationary distribution of wage offers (McCall 1970). The worker searches sequentially for job offers with the objective to maximize the present value of lifetime income. It can be shown that the optimal strategy involves a decision rule such that offers exceeding a critical “reservation wage” are acceptable whereas other offers are rejected. The probability of making a transition to work is determined by the probability of finding a job offer that exceeds the reservation wage. Unemployment benefit is usually modeled as a subsidy to unemployment, i.e., as income available when unemployed. Higher unemployment benefits imply a higher reservation wage and thus a lower job finding probability and an increase in the expected duration of unemployment. It also follows that higher benefits increase the expected post-unemployment wage via a higher reservation wage.

A more elaborated microeconomic model of how UI affects job search was presented in Mortensen (1977). Key innovations were to introduce a fixed potential duration of benefit receipt as well as stochastic duration of employment spells. The model also incorporates an eligibility condition requiring a certain amount of work experience in order to qualify for UI. The most important implications derived from this model were the following:

---

4 Related contributions include Mortensen (1990) and van den Berg (1990).
1. The worker’s reservation wage declines as she approaches the date at which benefits expire; hence the exit rate increases over the spell of (insured) unemployment.
2. Search effort increases over the spell of insured unemployment, a response that also causes the job finding rate to increase over unemployment spell.
3. An increase in the benefit level makes it more attractive for presently not eligible workers to accept jobs and thereby become qualified for benefits in the future; higher benefits thus result in an increase in the exit rate from unemployment to employment for workers who are not qualified for benefits, a response known as the “entitlement effect”.
4. A rise in the benefit level causes a newly unemployed and insured worker to increase her reservation wage but induces an insured worker close to benefit exhaustion to reduce her reservation wage. The exit rate thus declines for newly unemployed insured workers but increases for workers who have come close to benefit exhaustion. The last property follows from the fact that a higher benefit level increases both the value of continued search as unemployed and the value of accepting an offer. The value of continued search is small for workers close to benefit exhaustion as they are almost in the same situation as workers not qualified for UI.
5. An increase in the potential duration of benefits results in a decline in the job finding rate, i.e., an increase in the expected duration of unemployment.

The basic search models thus focus on how UI affects job finding, i.e., the outflow from unemployment. UI may, however, also affect the inflow into unemployment, a possibility emphasized by Wang and Williamson (1996). They examine an environment where a worker’s employment status depends on the choice of search effort as well as work effort. The transition rate from unemployment to employment is increasing in search effort; analogously, the probability of remaining employed is increasing in work effort. The model thereby makes job destruction endogenous and affected by the UI policy through the worker’s choice of work effort. Higher benefits encourage shirking on the job and thus increase the probability of job loss and thereby the risk of entering unemployment.

Matching and equilibrium search
It is important to check if the results from partial equilibrium analyses remain at least qualitatively intact in general equilibrium. The most widely used model of equilibrium unemployment is the search and matching model associated with Peter Diamond, Dale
Mortensen and Chris Pissarides, sometimes referred to as DMP-models.\(^5\) This framework appears in many guises, some more elaborated than others. The model includes relationships that describe how jobs are created and destroyed, which essentially correspond to firms’ decisions on hiring and firing. Unemployed workers search for job openings in an optimal fashion, recognizing that more time and effort allocated to search increases the probability of job finding. Wages are typically determined via bargaining between the worker and the firm.

An overall increase in UI benefits has a number of predictable effects in the search and matching model. Search effort among the unemployed falls as the value of employment has declined relative to the value of unemployment; job finding thus tends to fall. There is also an increase in wage pressure as higher benefits strengthen the worker’s hand in the wage bargain. Job creation falls and job destruction increases. Unemployment is unambiguously increased when benefits are raised.

The basic DMP-model ignores many fine details of the UI system. It is, for example, assumed that UI benefits are available for all unemployed workers, a patently unrealistic assumption that ignores that a large fraction of the unemployed are not eligible for UI. In the US, more than half of the unemployed are not covered by UI. As shown by Mortensen (1977) in a partial equilibrium setting, higher benefits imply stronger search incentives for the noninsured as job finding results in future UI eligibility (the so-called entitlement effect). A recent paper by Regev (2012) has examined whether the standard result – higher benefits lead to higher unemployment – holds also in an equilibrium search model with partial UI coverage. The answer is a qualified yes. The result holds analytically in a special case of the model with a degenerate productivity distribution. In the general case with stochastic match productivity, simulations based on a calibrated model for the US suggest that higher UI benefits do increase unemployment.

The entitlement effect is observationally similar to – but distinct from – a market externality whereby actions taken by a treated group (the insured) impacts on outcomes among the untreated (the uninsured). Imagine that insured workers get higher benefits and respond by searching less actively for work. This means that competition for jobs becomes less severe for the uninsured which in turn would increase job finding in this group. This spillover effect could operate even absent any entitlement effect via search efforts among the uninsured. An

\(^5\) The framework associated with Richard Layard and Steve Nickell has much in common with the DMP-models (Layard et al. 1991). Whereas DMP focuses on flows in the labor market, the Layard-Nickell model focuses on stocks.
early empirical paper by Levine (1993) used US data to examine the impact of benefit hikes on unemployment duration among the noninsured. He found support for spillover effects; increases in UI replacement rates produced decreases in unemployment duration among workers not eligible for UI. Using Austrian data, Lalive et al. (2013) also find support for spillover effects. More research on this issue is clearly important. Ignoring spillover effects might lead to misleading (upward-biased) estimates of the equilibrium effects of benefit hikes.

**Financing unemployment compensation**

Most countries finance UI compensation through general income taxes levied on households or payroll taxes levied on employers. A notable exception is the system with *experience rating* practiced in the US where each firm’s taxes depend in part on the firm’s contribution to insured unemployment. Under complete experience rating, the firm is charged the full UI cost of a layoff; in practice, experience rating is incomplete in the sense that the firm typically covers less than the full UI cost. There is a theoretical and empirical literature, largely based on the US practice, dealing with the impact of experience rating on firms’ layoff behavior and other outcomes.

Experience rating is also practiced in Denmark through the so-called G-days. An employer who lays off a worker who is member of a UI fund has to pay unemployment benefits to that worker for the first three days of unemployment. This obligation includes unemployment that is caused by completion of a fixed-term contract. There is a cap amounting to 16 G-days per worker and calendar year. There are some notable differences between US-style and Danish-style experience rating. A US employer is charged with a UI cost that depends not only on the magnitude of layoffs but also on the duration of the subsequent insured unemployment spells. The UI cost paid by a Danish employer is targeted at entry into unemployment.

Other forms of experience rating are conceivable. For example, Canada has in the past practiced experience rating on the side of benefits by making a worker’s benefit level dependent on her unemployment history: the more frequent unemployment in the past, the lower the benefit level. In countries with Ghent systems, i.e. voluntary UI organized through union-affiliated UI funds, part of the financing comes from membership fees and the rest via general taxes. This raises issues concerning the appropriate level of government subsidies on the one hand and membership fees on the other.

---

6 The system is gradually phased out over the period 2015–2018 for employment spells that have lasted for less than three months.
Experience rating

Feldstein (1976) developed the basic implicit contract model to examine the effects of experience rating on firms. The model features a firm facing uncertain product demand and with a pool of workers “attached” to it. The firm and the workers have to agree on a contract that specifies employment, wages and perhaps work hours for every possible realization of demand (booms as well as slumps). The number of laid off workers in each state of demand is simply given by the difference between the number of attached and employed workers. Feldstein assumed that firms that lay off workers have to finance part of the UI benefits that their workers are eligible to. The model implies that a rise in the UI subsidy – a decline in experience rating – causes a reduction in employment.

The general validity of Feldstein’s result has been questioned. Burdett and Wright (1989) have relaxed the assumption of a fixed pool of workers attached to the firm and allowed the firm to choose the number of workers. This modification has important implications. It turns out that higher experience rating does reduce layoff rates but also the number of attached workers. The intuitive explanation is that higher experience rating increases labor costs, which is bound to reduce the number of workers that the firm is willing to hire. The effect on average employment is ambiguous in general, and may plausibly be negative.

The impact of experience rating has also been studied by means of other models. Mortensen (1994) makes use of a calibrated version of the Mortensen and Pissarides (1994) equilibrium search model and introduces experience rating as a tax on firing. The effect of experience rating on unemployment is in general ambiguous as it reduces layoffs but also job creation. In fact, simulations of the model indicate that full experience rating would cause a small increase in unemployment. Albrecht and Vroman (1999) study experience rating in a model where firms pay efficiency wages in order to deter shirking. Two regimes are compared, one where benefits are financed by a proportional payroll tax and the other with experience rating in the form of taxes on each firm’s separations. The numerical results indicate that experience rating leads to lower unemployment and higher output. Cahuc and Malherbet (2004) note that firing costs are generally higher in Europe than in the US and ask whether this difference weakens the case for experience rating in the European context. They develop a matching model of a labor market that includes firing costs, temporary jobs and a minimum wage. Analytical results are hard to obtain but a calibrated model suggests that experience rating is likely to reduce unemployment.
All in all, the theoretical results regarding the effects of experience rating on employment are not particularly strong. The empirical literature indicates that full experience rating would substantially reduce layoff unemployment and also the seasonality of employment. The overall impact on the level of employment is less clear. Experience rating may also have favorable effects on resource allocation by reducing cross-subsidization of industries with varying degrees of labor turnover.\(^7\)

*Unemployment insurance in unionized economies*

When UI is voluntary and organized through union-affiliated UI funds, part of the financing arises from membership fees. This raises issues concerning the impact on wages and employment of alternative ways of financing UI. Some of these matters have been analyzed by Holmlund and Lundborg (1988, 1989, 1999). They use standard models of utility maximizing labor unions. A popular model considers a trade union attempting to maximize its members’ expected utility, typically subject to the restriction that profit-maximizing firms determine employment taking the wage as given. One of the implications is that the union’s preferred wage is increasing in the benefit level, a result that follows from the fact that a higher benefit level reduces the cost of unemployment to the members. This result for the monopoly union case carries over to more realistic models where there is bargaining over wages.

Consider an industry-wide union that effectively runs its own UI fund. Benefits are financed in part by membership fees (UI premiums) and the rest by general taxes. Suppose that the subsidy scheme is such that a fraction of an increase in paid-out benefits is covered by membership fees. A cut in subsidies would thus require higher membership fees, a relationship that arguably may lead the union to exercise wage moderation. A wage hike would be more costly since union members would have to self-finance a larger part of the rise in benefit payments that is caused by a wage increase.\(^8\) It can be shown that it is the *progressivity* of the subsidy system that matters for wage bargaining. Wage moderation (and thus higher employment) is encouraged if the UI premiums are made highly responsive to benefit expenditures without having much effect on average premiums. In practice, it may be difficult to implement a reform where marginal costs of wage hikes are increased without any impact on average UI premiums.

---

\(^7\) See e.g. Topel (1983), Card and Levine (1994) and Anderson and Meyer (2000).

\(^8\) This line of argument was made already by Pigou (1933, p. 254): “If the unemployed members of a trade union have to be cared for exclusively by that union, so that heavy unemployment means a heavy drain on union funds, this fact will act as a check upon claims for higher wages. If, however, unemployed members are cared for, in the main, at the expense of other people, the union’s contribution being no larger when there are many unemployed than when there are few, this check does not operate.”
A problem associated with substantial cuts of UI subsidies is that the ensuing rise in UI fees is likely to reduce UI coverage by inducing many workers to exit from the UI funds. The Swedish experience is illustrative in this regard: a substantial cut in UI subsidies undertaken in 2007 was followed by a sharp fall in UI fund membership amounting to 10 percent of the labor force. Another problem is that the overlaps between UI funds and bargaining areas are far from perfect, a fact that will attenuate wage responses to financing reforms.

**Other considerations**

UI benefits may also affect labor force participation. Higher benefits increase the value of unemployment relative to nonparticipation and would thereby encourage labor force entry. This may in part reflect demand for benefits rather than demand for work, in which case there would be little impact on “true” labor force participation. But benefit receipt is conditioned on active search and typically involves some monitoring, facts suggesting positive effects on search and subsequent employment. The numerical results presented in the formal search model developed by Rosholm and Toomet (2005) suggest that benefit cuts have negative effects on participation. Recent empirical evidence suggest that the benefit extensions in the US during the Great Recession have had negligible effects on job finding but positive effects on participation (Rothstein 2011; Farber and Valetta 2013; Farber, Rothstein and Valetta 2015).

UI benefits may affect geographic labor mobility but it is not clear in which direction the effect goes. On the one hand, generous benefits may reduce mobility by weakening incentives to search for and accept jobs in distant localities. On the other hand, generous benefits may alleviate credit constraints and thereby make it more attractive to invest in migration. The empirical studies are few and the evidence is mixed (see e.g. Tatsiramos 2009).

The organization of UI almost certainly matters for the level of unionization. Countries with Ghent systems have markedly higher union density, reflecting the fact that union membership and UI fund membership have been closely linked. For many unions, it has been required that members also belong to the affiliated UI fund.

UI is likely to reduce income inequality since unemployment is generally higher among groups with low education and low (potential) earnings. An overall increase in benefits is therefore likely to be particularly advantageous for individuals with weak labor market prospects. This effect is amplified by the fact that most benefit schemes involve a cap which makes them progressive in the sense that effective replacement rates are higher for workers...
with relatively low previous earnings. Of course, these effects are first-order and may be partially offset if generous benefits widen the unemployment differences among groups.

Finally, UI is likely to be a useful instrument for stabilization policy by functioning as an automatic stabilizer. A recessionary chock that leads to a fall in private consumption is partly offset by an automatic increase in expenditure on unemployment benefits.

4. Designing optimal unemployment insurance
The resurgence of interest in the economics of optimal UI-design has provided new insights into the tradeoffs between insurance benefits and incentives. This section reviews some of this literature. A common assumption is that individuals have unbiased risk preferences in the sense that they make correct assessments regarding probabilities of job finding and job loss. The policy maker cares about individuals’ expected utilities.

The question of time sequencing of benefits concerns whether benefits should be paid at a fixed rate over the spell of unemployment or decline (or increase) over the spell. This issue was discussed in a few seminal contributions on optimal UI published in the late 1970s and has attracted new attention in recent research. Issues regarding monitoring and sanctions concern how much resources should be spent on checking search behavior and how sanctions, such as benefit cuts, should be implemented if prescribed search requirements are not met.

The conventional approach to the welfare analysis of UI is to contrast the consumption-smoothing role of UI benefits with the moral hazard effects that are thought to reduce employment and output. It should be noticed that some arguments suggest that UI may well increase total output over some range despite the presence of moral hazard. One idea formalized in a paper by Diamond (1981) is that prolonged job search, induced by UI, may improve matching in the labor market and therefore be socially desirable. The more recent paper by Acemoglu and Shimer (2000) makes this point in a new setting and argues that UI encourages risk-taking, i.e., encourages workers to seek higher productivity jobs and firms to create such jobs. The idea is formalized and built into a quantitative model that is calibrated on US data. The authors find that an increased generosity of the prevailing UI-system in the US would increase unemployment as the conventional wisdom predicts. However, total output and welfare would also increase as more generous UI encourages workers to hold out for more productive jobs.⁹

⁹ Several recent empirical papers have investigated the match quality aspects of UI. The results are mixed.
The optimal time profile of benefit payments

UI and search effort

The seminal papers on optimal UI appeared in the late 1970s (Baily 1977, 1978; Flemming 1978; Shavell and Weiss 1979). Shavell and Weiss focused in particular on the optimal sequencing of benefits. Consider a model of job search with identical and risk-averse individuals where the employment state is “absorbing”, i.e., once the worker has found a job he stays in it forever. The probability of job finding is partially under the control of the job searcher through the choice of effort and reservation wage. The objective of the UI provider is taken to be maximization of the newly unemployed worker’s expected utility, subject to a fixed UI budget.

Shavell and Weiss derived several results from their model. Under the most restrictive assumptions – no wealth, no borrowing and no moral hazard problem – the benefit level should be constant over the spell of unemployment. By introducing moral hazard, this result is overturned and it is found that the benefit level should decline monotonically over the spell, the reason being that a declining benefit profile provides stronger incentives to search. Shavell and Weiss were however not able to characterize the benefit profile in the general case with moral hazard and where the individual has initial wealth and can borrow.

Recently a number of papers have extended the analysis of Shavell and Weiss. One strand of the literature stays within the Shavell and Weiss framework in the sense that the focus is solely on the behavior of the worker. One example is Hopenhayn and Nicolini (1997) who enlarge the set of policy instruments by considering a wage tax after reemployment in conjunction with the sequence of benefit payments. The model is one where agents are risk averse and the unemployed worker’s probability of finding a new job depends on search effort, which is unobserved by the UI provider. The worker has no other source of income. Moreover, savings and borrowing against future income are ruled out by assumption, so income equals consumption in each state. All workers are identical, infinitely lived and receive the same (exogenous) wage while employed. Employment is treated as an absorbing state, i.e., there is no risk of reentering unemployment in the future.

The optimal UI program is taken to involve minimization of the UI budget – the expected discounted value of UI transfers – subject to a prescribed expected discounted utility to the worker. Two instruments are available that allow the insurer – the principal – to control the worker’s – the agent’s – consumption in each state, viz. unemployment benefits and a wage
tax. Benefits can vary by elapsed duration and the wage tax is allowed to depend on the worker’s unemployment history. The wage tax is constant over the reemployed worker’s (infinite) employment spell. Moreover, the tax can take negative values, in which case it acts as a subsidy to job finding.

The most important analytical results are as follows. First, unemployment benefits should decrease over the elapsed duration of unemployment. Second, the wage tax should under some (sufficient) conditions increase with the length of the previous unemployment spell. The intuition for those results is that both the declining benefit profile and the rising tax profile encourage job finding by making prolonged job search more expensive.

Hopenhayn and Nicolini also present a numerical calibration so as to gauge the welfare effects of switching from the current UI system in the United States to an optimal system. These calibrations suggest that the welfare gains can be substantial. The cost savings from an optimal system relative to the current system amount to almost 30 percent. The decline in replacement rates over elapsed duration is much smaller compared to an optimal system without a wage tax. The reason is that the “one-instrument” policy can only affect intertemporal incentives by varying benefits over time, whereas consumption possibilities during a future employment spell is beyond the principal’s control. The “two-instrument” policy has the virtue of improving intertemporal consumption smoothing as well as intertemporal incentives. The computed optimal replacement rates in the “two-instrument” case are remarkably high, ranging from 99 percent of the wage during the first weeks of unemployment to 94 percent after one year’s unemployment. The optimal time profile in the absence of a wage tax involves a decline in replacement rates from 86 percent during the first week to 13 percent after one year of unemployment. The optimal wage tax is negative for workers with short previous unemployment spells, i.e., the optimal system entails a subsidy to unemployed workers who find jobs quickly (within five weeks). The tax imposes a large penalty for long spells of unemployment. If the spell lasts for six months, the tax amounts to 2 percent of the pre-tax wage; if it lasts for 12 months, the tax is 4.5 percent.

Hopenhayn and Nicolini (2009) report further results based on a model that incorporates repeated unemployment spells. A key result is that replacement rates should decrease with current and previous unemployment spells. It is also found that UI coverage should increase with the length of previous employment spells if quits and layoffs are indistinguishable.
The two papers by Hopenhayn and Nicolini ignore the issue of how the UI system influences search behavior among those who are not qualified for UI. New entrants account for a substantial fraction of inflow into unemployment and they are typically not covered by UI since previous employment experience is required. As shown by Mortensen (1977), this feature of UI gives rise to an “entitlement effect”. This effect may well have important implications for the design of optimal UI. Another restrictive feature is the fixed wage assumption. In a search equilibrium framework, as well as in other models of equilibrium unemployment, there is a link between benefits and wages, which in turn implies a relationship between benefits and job creation.

The role of capital market imperfections
Hopenhayn and Nicolini (1997, 2009) ignore the possibility that workers may borrow and save in order to smooth consumption. Shimer and Werning (2008) address this issue using a sequential search model along the lines of McCall (1970). Whereas Hopenhayn and Nicolini study how the UI rules affect search effort, Shimer and Werning focus on incentives operating via reservation wages. The unemployed worker lives forever and gets a wage offer with some probability in each period. The optimal reservation wage is chosen so as to maximize expected lifetime income.

Shimer and Werning compare two UI policies, the first referred to as constant benefits. Under this policy, the worker gets a constant benefit while unemployed and pays a constant tax when she has got a job. The worker faces no liquidity constraints, thus being free to borrow and lend as much as she prefers. The government (insurance agency) sets the benefit level and the employment tax to minimize the cost of providing the worker with a given level of utility.

Under the second policy, referred as optimal UI, the worker has no access to the capital market and consumes her after-tax income. The insurance agency thus dictates the path of consumption by its choice of benefit path, recognizing that the choice affects reservation wages and thus the duration of unemployment. The government sets benefits so as to minimize the cost of providing the worker with a given level of utility.

Shimer and Werning show that the two UI policies are equivalent, i.e., the cost of providing a given level of utility is the same, the path of the reservation wage is the same, and the path of
consumption is the same.\textsuperscript{10} Reservation wages fall with elapsed duration of the unemployment spell. With constant benefits, this is driven by the worker’s own decisions on borrowing and lending; with optimal UI, it is implied by the benefit decision taken by the insurance agency. The bottom line is thus that the case for a declining benefit profile over the spell of unemployment hinges crucially on capital market imperfections. Constant benefits are optimal if the capital market permits unrestricted borrowing and lending.

\textit{Work effort and job loss}

Wang and Williamson (1996) call into question the result that benefits should fall monotonically over the unemployment spell. They add another source of moral hazard by examining an environment where a worker’s employment status depends on the choice of effort. The transition rate from unemployment to employment is increasing in search effort; analogously, the probability of remaining employed is increasing in work effort. The model thereby makes job destruction endogenous and affected by the UI policy through the worker’s choice of effort. Workers cannot lend or borrow, as in Hopenhayn and Nicolini.\textsuperscript{11} The wage upon reemployment is exogenous and identical across workers. The model incorporates a flow of new entrants in the labor market and the eligibility condition that employment is required in order to qualify for benefits. Unemployed workers who don’t receive UI benefits are eligible for a welfare benefit that is taken as a fixed fraction of the wage.

The optimal UI system — implied by an objective function similar to the one used by Hopenhayn and Nicolini — involves a large drop in consumption in the first period of unemployment (so as to discourage shirking), and a large reemployment bonus (so as to encourage search effort). The implied time profile of UI compensation is thus non-monotonic; compensation increases initially and then falls throughout the spell. The numerical calibrations suggest that the optimal system involves a reduction in unemployment relative to unemployment in the prevailing US system by more than three percentage points and an increase in output by more than three percent. These results are driven by declines in job destruction, reflecting a high elasticity of job retention with respect to effort on the job, and a rise in mean effort on the job.

It is difficult to assess the empirical plausibility of these results. There is virtually no empirical evidence available on the relationship between job destruction and workers’ choice

\textsuperscript{10} Strictly speaking, the equivalence of the two policies depends on preferences. The policies are exactly equivalent when preferences are characterized by constant absolute risk aversion but only approximately equivalent under constant relative risk aversion preferences.

\textsuperscript{11} Wang and Williamson (2002) is an extended version of their paper which allows for individual borrowing.
of effort. The importance attributed to on-the-job effort for job destruction and for UI analysis is probably overstated. In fact, workers who quit or who are fired for cause are typically not eligible for UI in existing systems. These institutional characteristics should temper although not eliminate the shirking incentives arising from UI.

It is noteworthy, however, that existing UI systems (and other social insurance schemes, such as health insurance) often involve a waiting period before benefits are paid out. This discourages entry into unemployment. In addition to the conceivable effect on shirking behavior, the waiting period may also affect wage and employment contracts. A large penalty to unemployment entry will reduce the attractiveness of implicit contracts with repeated use of temporary layoffs. A system with a waiting period may also be cost effective by reducing the administrative burden on the UI system.

**Search, matching and wages**

Several papers have studied optimal benefit sequencing using models with endogenously determined wages. The paper by Cahuc and Lehmann (1997) examines how the time sequencing of benefits affect equilibrium unemployment in a model with an endogenous number of jobs and union-firm bargaining over wages. A key assumption is that it is the short-term unemployed that affect wage setting. In case of disagreement in the negotiations, the “insiders” involved in the bargain become short-term unemployed and eligible for UI payments. A declining time profile of benefits thus improves the fallback position of the insiders and this tends to increase wage pressure and cause higher unemployment. Indeed, the paper finds that a constant time sequence yields a lower unemployment rate than a program with a declining time profile (taking the tax rate as exogenous).

Cahuc and Lehmann (2000) is a more recent version of the model where endogenous job search is allowed for. The model then becomes too complex to yield analytical results and the authors turn to a number of calibration exercises. One noteworthy feature of this analysis is the computation of welfare effects of alternative time profiles for both short-term and long-term unemployed as well as for employed workers. In these experiments, the tax rate is held constant so there is no attempt to characterize the optimal UI system (in which case one would choose the tax rate along with the UI parameters subject to a government budget restriction). The simulations illustrate that a declining time profile tends to encourage search effort and thereby reduce unemployment (albeit at the cost of lower welfare for the long-term unemployed). When both search and wages are endogenous, the simulations still imply that a
declining profile lowers unemployment and leads to higher aggregate welfare compared to a flat profile. However, the decline in unemployment is weaker compared to the case with exogenous wages, the reason being the rise in wage pressure.

Fredriksson and Holmlund (2001) address the question of the optimal sequencing of benefits using an equilibrium model of search unemployment along the lines of Pissarides (1990/2000). They allow for endogenous search effort among unemployed workers and, in contrast to Shavell and Weiss (1979), Hopenhayn and Nicolini (1997), Wang and Williamson (1996) and Davidson and Woodbury (1997), incorporate endogenous wage determination and free entry of new jobs. The UI program affects search effort as well as the wage bargains.

For analytical tractability, the paper mainly focuses on a two-tiered UI system, i.e., a program with two benefit levels where the first is referred to as UI and the second as unemployment assistance (UA). Workers who lose their jobs are entitled to UI. UI benefits may not be paid indefinitely, however. Workers losing their benefits are entitled to UA that has infinite duration but is potentially lower than UI payments. The paper asks whether a two-tiered system dominates, in welfare terms, a program with indefinite payments of a constant wage replacement rate. The answer to this question turns out to be an unambiguous yes, provided that time discounting is ignored. The result carries over to the case with a multi-tiered benefit structure; unemployment benefits should decline monotonically over the spell of unemployment. A feature known from models of individual worker search drives the key result: the effect of higher benefits on the individual worker’s search behavior depends on whether he is presently qualified for UI or not. A rise in benefits will in general increase search effort among those not insured, as this will bring them quicker to employment that results in eligibility for future UI payments. A two-tiered UI system exploits this entitlement effect by providing incentives for active search among workers not currently entitled to benefits.

When time discounting is accounted for, the optimality of a declining benefit sequence cannot be established analytically. The reason for the ambiguity lies in the fact that a declining sequence increases the welfare of the short-term unemployed at the expense of the long-term unemployed, which in turn induces stronger wage pressure than a flat (or increasing) sequence; this is the mechanism discussed in the papers by Cahuc and Lehmann. According to

12 Ignoring time discounting implies that future utility is given the same weight as current utility.
the numerical calibrations, however, this “wage pressure effect” is dominated by the case for exploiting the entitlement effect.

In the calibrations of the model, the optimal uniform replacement rate is around 40 percent, the exact magnitude depending, inter alia, on relative risk aversion. The optimal ratio between UI and UA varies between 1.7 and 2, with the larger number implied by higher relative risk aversion (equal to two). The welfare gain of moving from the optimal uniform to the optimal differentiated system is non-trivial. In the example with highest risk aversion, workers would be willing to pay around one percent of their consumption flow in order to move from the uniform to the differentiated system. The effect on overall unemployment of moving from the optimal uniform to the optimally differentiated system is negligible, however.

**Summary and discussion**

What conclusions can be drawn from the literature regarding the optimal time profile of benefit payments over the spell of unemployment? The case for a declining time profile seems reasonably well developed. A declining profile provides better search incentives than a flat (or increasing) profile. There are at least two caveats to this conclusion. First, as emphasized by Cahuc and Lehmann, there is a possibility that this design encourages wage pressure. However, it has not been convincingly demonstrated that this effect is so quantitatively important that it overturns the argument for restoring search incentives. A second caveat follows from the observation, made by Wang and Williamson, that it may be optimal to impose a “tax” on entry into unemployment by offering low benefits during the first week(s) of unemployment. The shirking argument for this policy is probably overstated but other mechanisms may suggest a case for low benefits during the first week(s) of unemployment. We discuss two examples.

The use of a waiting period before UI benefits is paid out is a feature of some existing systems. This effectively works as a tax on entry into unemployment and may be desirable as a means to discourage the use of temporary layoffs subsidized by UI. However, temporary layoffs can also be taken care of by experience rating provisions. That is, firms that engage in frequent layoffs can be taxed in proportion to their contribution to the layoffs, a key feature of the UI system in the United States. It is also plausible that there are economies of scale in the administration of the benefit system. There is presumably a fixed cost associated with each UI claim so that the average administration cost per week of elapsed duration declines with
increasing duration. This might provide a rationale for a system with a waiting period so as to reduce the flow of new UI claims.

A related argument revolves around self-insurance through private savings. Most of the models of optimal UI have ignored savings, the reason being that the modeling difficulties have been substantial. A reasonable conjecture is that private savings (including family transfers) can work reasonably well as a substitute to UI for very short spells of unemployment. For this additional reason, an optimal UI policy may include a waiting period before benefits are paid out.

There are a few papers that have tried to model optimal UI in general equilibrium models with endogenous savings. In general, one would expect that allowing for savings (and borrowing) would reduce the optimal replacement rate. Costain (1997) develops a model with endogenous search effort and precautionary savings and finds in numerical experiments that the optimal replacement rate should be in the range of 30 to 40 percent. He does not examine the time profile of benefits, however. The paper by Heer (2003) takes this step by means of a calibrated general equilibrium search model with endogenous savings. Heer focuses on a two-tiered benefit system, following the set-up in Fredriksson and Holmlund (2001). One important result is that the optimal UI compensation decreases over the spell of unemployment, a result in line with previous research that mostly has ignored precautionary savings. However, it is not possible to use Heer’s model to analyze the case for an initial waiting period so as to discourage entry into unemployment.

Monitoring, sanctions and workfare

In the contributions discussed above, the receipt of unemployment benefits is not affected by how hard the worker searches or how choosy she is with regard to acceptance decisions. In practice, however, the UI systems generally condition benefit payments on some performance criteria such as “availability for work” and “actively searching for work”. To make sure that these criteria are met, the benefit administration engages in some degree of monitoring of the unemployed benefit claimants. Monitoring usually takes place through public employment agencies. For example, job seekers have to show up with some regularity at the employment offices and/or they have to give evidence of job applications. A worker who fails to meet certain requirements may be exposed to a sanction, for example a temporary cut in benefits.

The economics literature dealing with monitoring and sanctions in the context of UI is small and of recent origin. There is however a literature on optimal law enforcement that is of
relevance for the analysis of optimal UI design.\textsuperscript{13} We briefly summarize some results from this literature that seem potentially relevant for the analysis of optimal UI design.

\textit{The economic theory of law enforcement}

Gary Becker’s paper on crime and punishment is the seminal contribution to this literature (Becker 1968). Becker takes individuals to be rational and risk-neutral expected utility maximizers who compare benefits and costs of violating the law. Behavior is affected by the monetary (or other) gains from crime, by the probability of being detected if choosing to violate the law, and by the severity of the punishment in case a crime is detected. A law violation is optimal as long as the benefit from the action exceeds the expected fine. The government can influence incentives primarily by affecting the probability of detection and the severity of punishment. The theory of optimal law enforcement is concerned with how the government should choose detection probabilities and measures of punishment so as to maximize a social welfare function.

Suppose that individuals are risk-neutral and contemplate an act that may be harmful to society. The social planner has two instruments at its disposal: expenditure on detection (monitoring) and a pecuniary sanction (fine). There exists a maximum feasible level of the fine, often interpreted as equal to the individual’s wealth. Assume also that the sanction can be imposed without costs. Under these assumptions (and some additional technical conditions) one can show that \textit{the optimal fine is the maximal fine}: the fine should be set to its maximum feasible level. The reason for this result is that sanctions are costless whereas monitoring is costly. Absent a bound on the feasible fine, the optimal fine would tend to go to infinity and the probability of detection to zero. However, the detection probability is strictly positive when the maximum fine is bounded. It follows that an increase in the maximal fine, for example due to an increase in wealth, may allow a reduction in the detection probability.

Much of the recent literature has been concerned with extensions of Becker’s analysis in various directions. One insight from this literature is that the maximal fine may in fact not be optimal under several plausible conditions. Indeed, existing legal systems do not seem to practice maximal fines!

One modification of Becker’s model is to allow for costly sanctions. It is possible that costs of enforcement increase, as the fine gets larger. For example, one might argue that more resources will be spent on lawyers etc., the more that is at stake in terms of sanctions. When

\textsuperscript{13} Surveys of the literature are contained in Garoupa (1997) and Polinsky and Shavell (2000).
sanctions are socially costly, the optimality of the maximum fine is no longer guaranteed. This is as should be expected since both monitoring and sanctions are costly and there is no presumption that a corner solution would be optimal. It can also be shown that the maximal fine may not be optimal when individuals are imperfectly informed about the probability of apprehension.

The maximum fine result may also be overturned when individuals are risk averse (Polinsky and Shavell 1979). Risk aversion has in itself a (costless) deterrent effect. The safe law-abiding action may be optimal even if the benefit from a crime exceeds the expected fine. The risk-averse individual requires a risk premium in order to choose a risky criminal activity. This needs to be taken into account by the social planner. A higher sanction implies a higher risk premium and it is no longer necessarily true that the sanction needs to be the maximal feasible one.

Another extension of the basic Becker model deals with the accuracy of the social planner’s information. Individuals may be falsely sanctioned (Type I error) or escape sanctions even if they in fact committed a crime (Type II error). There is a cost associated with improving the accuracy of information, i.e., reducing the probabilities of Type I and Type II errors. Under risk-neutrality it can be shown that the optimal fine is still the maximal fine. In addition, the optimal cost devoted to improving accuracy (as well as detection) is positive. Increasing the probability of detection as well as improving accuracy can raise deterrence and there is a tradeoff between the two instruments. If the social planner is strongly averse to legal errors it may however be the case that an interior solution – a sanction less than the maximal one – is optimal.

Yet another reason for why the optimal sanction may be less than the maximal one is avoidance activities. The higher the sanction, the higher presumably the resources used by offenders to avoid it. This means that a sanction is no longer costless which needs to be considered in the social welfare function. An interior solution then again appears as a possible outcome.

Theoretical modeling of monitoring and sanctions

Most theoretical models of unemployment and job search ignore monitoring and the possibility of benefit sanctions as the outcome if a worker does not comply with search requirements. There are however a few exceptions that include papers by Boone and van Ours
Boone and van Ours (2006) use a search model to explore links between the UI system and unemployment. The model here is a version of the Pissarides (1990/2000) search and matching model and has similarities with the model in Fredriksson and Holmlund (2001). Workers are risk averse, search effort among the unemployed is endogenous and wages are determined in bargaining between the firm and the individual worker. A key assumption is that the unemployed and insured worker can affect the probability of continued UI receipt by the choice of search effort; the higher the search effort, the lower the risk of being exposed to a benefit sanction. This is the monitoring system in the model. The benefit associated with additional search thus involves two components, one capturing the gain associated with a transition to employment and the other capturing the gain of not being penalized by a benefit sanction.

Boone and van Ours calibrate their model to data for the Netherlands and undertake a number of simulations to shed light on the role of monitoring and sanctions in a general equilibrium setting. They note that the effects of monitoring and sanctions involve an ex ante effect capturing deterrence as well as an ex post effect capturing higher search effort among those actually exposed to a sanction. With a low monitoring rate, any action is produced by the ex post effect; with a high monitoring rate, the main action is driven by deterrence. It is conceivable that the ex post effect – which is possible to estimate by use of micro data – can be negligible despite an overall strong effect due to deterrence. The model illustrates that the risk of being sanctioned can have strong behavioral effects even if the penalty is modest. Boone and van Ours report (utilitarian) measures of welfare for alternative scenarios and conclude that policies involving monitoring and sanctions can be welfare improving. A difficulty here is that a proper welfare analysis requires that one recognizes that monitoring, and the enforcement of sanctions, do not come without costs. Any conclusion about the optimal set of instruments – replacement rates, monitoring rates, sanctions – is bound to be crucially dependent on how costly it is to enforce stringent search requirements.

The paper by Boone et al. (2007) examines to what extent the optimal UI policy involves monitoring of search effort and benefit sanctions if observed search is deemed insufficient. The framework is an equilibrium search model along the lines of Fredriksson and Holmlund (2001) and Boone and van Ours (2006). The paper finds that introducing monitoring and
sanctions represents a welfare improvement for reasonable estimates of monitoring costs; this conclusion holds both relative to a system featuring indefinite payments of benefits and a system with a time limit on unemployment benefit receipt. The optimal sanction rates implied by the calibrated model are much higher than the sanction rates typically observed in European labor markets.

The paper by van den Berg and van der Klaauw (2006) brings attention to the fact that the unemployed may use different search channels. A policy that affects only one search method may induce a substitution among search methods. The theoretical framework is a partial equilibrium search model with endogenous search effort and two search channels, namely formal and informal search. The wage-offer distributions, possibly channel-specific, are taken as exogenous. One can think of formal search as search through the public employment service whereas informal search includes referrals through friends etc. Optimal behavior is characterized by a reservation wage rule and optimal search intensities in the usual way. Monitoring takes the form of a minimum search requirement applied to the formal search channel (since it is difficult to monitor informal search). It may involve more checks on the number of job applications, more frequent visits to the employment office etc.

Monitoring affects search to the extent that it bites, i.e., to the extent that minimum required search exceeds optimal formal search in the absence of monitoring. The effect on the transition rate to employment depends crucially on how search costs are specified. The reservation wage declines and the transition rate increases if channel-specific search costs are additive. In this case changes in formal search do not affect the marginal cost of informal search. The imposition of a binding search requirement makes it less attractive to be unemployed which implies greater willingness to accept job offers. An alternative assumption is that search costs depend on total search effort in such a way that formal and informal search becomes close or perfect substitutes. If search costs are given by a strictly convex function of total search effort, it follows that a formal search requirement will reduce effort allocated to informal search. The effect on the transition rate to employment may then be zero.

**Workfare**

Workfare – the requirement that a benefit recipient participate in some work activity in exchange for benefits – has been on the policy agenda for a very long time; indeed, examples of workfare in France and Britain can be traced centuries back (Besley and Coate 1992).

---

14 See Van den Berg and Vikström (2014) for a related paper.
Workfare has been thoroughly scrutinized in the public finance literature on poverty alleviation. In the context of UI, workfare has sometimes been discussed in conjunction with active labor market policies. One idea in this discussion is that active labor market programs (ALMP) can be useful to implement the work test of UI. ALMP as workfare – the requirement to participate in a labor market program in exchange for benefits – effectively puts a price on individuals’ time. Those who put a high value on their leisure time may self-select out of the benefit system and only those really needy would remain as benefit claimants.

The paper by Hansen and Tranæs (2005) presents a formal analysis of workfare in a model where individuals may have the same productivity but differ with respect to their preferences for leisure. There are two types of individuals, referred to as workers (with low disutility of labor) and non-workers (with high disutility of labor). The government knows the distribution of individual characteristics but not the preferences of a particular individual. Job search effort and job acceptance decisions are also private information to the individuals.

The paper examines whether workfare can be a Pareto improving welfare reform, i.e., whether it is possible to improve welfare for one type of individuals without worsening conditions for the other type. The answer is affirmative: workfare works as a welfare improving screening device if individuals are sufficiently heterogeneous with respect to their valuations of leisure. The intuition for this result is as follows. When preferences and job search behavior are private information, even non-workers (who do not search for work) may claim UI benefits. This reduces the scope for UI as income insurance for workers (who do search for work). By introducing a work requirement, the government can induce non-workers to self-select out of the UI system, the reason being that they have a strong preference for leisure. At the margin, it is possible to simultaneously raise UI benefits and introduce a work requirement so as to make non-workers indifferent between claiming and not claiming UI benefits. The rise in UI benefits represents a strictly positive welfare improvement for workers.

Empirical work on the effects of workfare is rather limited. Some research on the effects of labor market policy is of relevance. A study by Black et al. (2003) provides fairly conclusive evidence in favor of the hypothesis that the mere threat of being placed in a labor market program can reduce time spent on UI and boost job findings. Rosholm and Svarer (2008) report similar results.

Fredriksson and Holmlund (2006a) is a related paper.
Part-time unemployment

Conventional measures of unemployment do not capture all dimensions of joblessness. A case in point is underemployment among employed workers, also frequently referred to as part-time unemployment. Workers in this category are employed during a survey week but are unable to work as many hours as they wish during that week. An ILO-resolution from 1998 defines underemployment as comprising persons in employment who (i) are willing to work additional hours, (ii) are available to work additional hours, and (iii) worked less than a threshold relating to working time. Underemployment accounts for a non-negligible fraction of the labor force in most countries, although it is typically somewhat lower in magnitude than the conventional measure of unemployment. The labor force surveys show that underemployment stood at almost 5 percent of the labor force in the OECD area by the end of 2009.

Ek and Holmlund (2015) study an economy with two sectors (industries) where full-time jobs are offered in one sector and part-time jobs in the other sector. Unemployed workers prefer full-time jobs but are willing to consider part-time jobs as stepping stones to full-time jobs. The unemployed worker thus searches for part-time as well as full-time jobs and accepts the first offer that arrives. The part-timer searches only for full-time jobs and quits her part-time job as soon as a full-time offer comes along. Unemployment benefits affect search effort among the unemployed as well as among the part-timers. More generous benefits for part-timers reduce the flow from part-time work to employment as part-time status becomes more attractive relative to employment. This in turn increases the value of opening part-time jobs for firms. But benefits for part-timers also make part-time status more attractive relative to unemployment, thus increasing the outflow from unemployment to part-time work. In that respect, offering benefits to part-timers is akin to in-work benefits, a policy that has received considerable attention in both policy discussions and research.16

Ek and Holmlund (2015) examine the optimal structure of unemployment benefits focusing on benefit levels (replacement rates) and the potential duration of benefit receipt for the unemployed as well as for part-timers. The model is of the search and matching variety and is calibrated with an eye on the Swedish labor market. It is found that the optimal policy involves positive benefits also for the part-time unemployed. There are non-negligible welfare gains associated with time limits for unemployment benefits as well as for part-time benefits.

16 See Kyyra (2013) for an empirical analysis using Finnish data.
Financing issues

UI benefits are typically financed by general taxes, such as income taxes on households or payroll taxes levied on employers. An increase in benefits would thus require higher taxes and this benefit-induced tax hike might have adverse effects on employment, in addition to the “pure” benefit effect. A couple of remarks are in order on this point.

First, a general result in the theory of taxation is that statutory incidence is irrelevant for the real burden of a tax, at least in the long run. Taxing wages on the side of employers would thus have same employment effect as a tax on labor earnings on the side of workers. This holds as long as the tax base is the same, e.g. the total wage bill. Second, the employment effect of a tax on labor earnings is theoretically ambiguous. In a competitive model, the incidence depends on the interplay between the elasticity of labor demand and the elasticity of labor supply. In a model with bargaining over wages, the impact on employment of a tax on wages depend, inter alia, on the value attributed to leisure while unemployed and on the extent to which benefits are indexed to wages. Under relatively plausible assumptions, the burden of a tax on wages falls entirely on workers and labor cost and employment are not affected.17

Even if payroll taxes have negligible effects on employment, there may exist more employment-friendly ways to finance unemployment benefits. Some method of experience rating is the obvious alternative. This could involve a US-style tax on employers that depends on each employer’s layoff behavior. Or it might entail a tax on employed workers that is related to the length of the previous unemployment spell(s), along the lines of Hopenhayn and Nicolini (1997). The former model focuses on firms’ layoff incentives whereas the latter is concerned with workers’ search incentives.

There are many studies of experience rating at the firm level but few of them address normative issues. The paper by Blanchard and Tirole (2008) is an exception. The authors consider the joint design of UI and employment protection and model employment protection as a tax on layoffs (analogous to US-style experience rating). The framework is essentially a static one-period contracting model that involves risk-averse workers, risk-neutral firms and random shocks to productivity. If a worker is laid off, she receives unemployment compensation which can be financed by payroll taxes or by layoff taxes. A key benchmark result is that the optimal policy implicates that benefits should be entirely financed by layoff

17 This holds if the leisure value of unemployment is negligible and benefits are indexed to wages via a fixed replacement rate.
taxes. This strong result is modified under various alternative versions of the model but the overall gist of the analysis is that benefits should, at least in part, be financed by layoff taxes. Such taxes give firms incentives to internalize the social costs of unemployment and take efficient layoff decisions.

As noted, the wage tax upon reemployment studied by Hopenhayn and Nicolini (1997) can be viewed as a form of experience rating on workers. The optimality of using such taxes flows from the fact that prolonged search as unemployed increases the future tax cost as employed. Can such a system easily be implemented? There are informational problems that need to be addressed. One issue concerns the measurement of unemployment at the individual level. Survey-based information from the labor force surveys would obviously be useless as a basis for taxation at the individual level. Register data on unemployment must be used, such as registration at the employment offices or the UI funds. But unemployed individuals’ registration decisions will most likely be affected by properties of the UI scheme, such as benefit levels and taxes upon reemployment.

There is also the question how worker heterogeneity and distributional issues should be dealt with. Individuals who expect long unemployment spells may be inclined to avoid registration as unemployed in order to avoid future tax payments. Such behavior would, however, also imply forfeit of UI benefits. Incomplete UI coverage seems as a plausible outcome since it is highly unlikely that the policy maker will be able to perfectly identify individual unemployment risks and use that information to differentiate benefits and taxes across risk groups. This is, of course, a problem of adverse selection.

Distributional issues are relevant ex ante (before labor market status is revealed) as well as ex post (after labor market status is determined). On the ex-ante side, workers with attractive labor market characteristics would benefit from lower taxes upon reemployment. On the ex-post side, workers who are long-term unemployed simply because of bad luck will be penalized. One can imagine ways to mitigate these distributional effects but it is probably difficult to completely avoid a regressive distributional profile.

An arguably attractive feature of either form of experience rating is that less cross-subsidization of industries would improve resource allocation. Firms (or workers) in industries with marked seasonal or cyclical fluctuations in labor demand would be penalized relative to a system without any experience rating.
Estimating optimal unemployment insurance

The characteristics of an optimal UI system depend in general on many factors, such as individuals’ attitudes to risk, the importance of moral hazard (adverse incentive effects), and the functioning of the capital market. Some of these factors are very difficult to measure and it may seem as a formidable task to provide quantitative estimates of optimal replacement rates and other UI characteristics. Yet the UI literature has seen several efforts to estimate optimal UI characteristics.

One widely used approach relies on calibration and simulation; examples are discussed in previous sections. The researcher writes down a theoretical model and attempts to translate it into a quantitative model by making specific assumptions about the parameters involved. Some parameters, such as policy instruments, are often straightforward to set. Other parameters may be possible to pin down by using estimates from the empirical literature. There are in general also parameters that are largely absent from the empirical literature and therefore require guesstimates that will be surrounded by substantial uncertainty. Equipped with a quantitative model, and having chosen a social welfare function, the researcher can compute optimal values for the UI parameters of interest.

An alternative to calibration is structural estimation where key parameters are estimated rather than imposed. An example is the study by Lentz (2009) based on Danish data. One parameter of interest captures workers’ degree of risk aversion. Lentz makes use of the estimated model to compute optimal replacement rates for various types of individuals and for aggregate groups. The estimates vary quite substantially – between 40 and 80 percent – depending on the relationship between the interest rate and the subjective discount rate.

A drawback of the calibration approach is that the results are likely to be sensitive to the chosen parameter configurations. Structural estimation may narrow down the set of reasonable parameters but may require restrictive assumptions in some dimensions. Checks for robustness are common practice but there are practical limits to the amount of robustness checks. It would seem desirable to strive for more robust methods to estimate optimal UI, i.e., methods that require little economic theory and rely on parameters that can be observed empirically. Recent works by Chetty (2008) and Shimer and Werning (2007) are examples of this approach. Earlier contributions along similar lines are Baily (1978) and Gruber (1997).

Chetty (2008) notes that the impact of unemployment benefits on unemployment duration involves a “liquidity effect” in addition to the usual moral hazard effect. The liquidity effect
captures the fact that higher benefits will lessen credit constraints, thereby allowing a welfare-improving increase in duration as the worker gets more cash on hand. In US data, nearly half of job losers report no liquid wealth and Chetty finds that most of the increase in duration driven by higher benefits is attributable to the liquidity effect. Chetty also derives a formula for the optimal benefit level that depends only on reduced-form elasticities, i.e., responses that are relatively straightforward to estimate. For example, no assumptions regarding individual attitudes to risk are required. Using the formula, Chetty finds that the optimal benefit level amounts to nearly 50 percent of the wage.

Shimer and Werning (2007) make use of McCall’s sequential search model and shows that a worker’s utility while unemployed is a monotone function of her after-tax reservation wage. The goal of the policy maker is to maximize the unemployed worker’s utility subject to a budget constraint. It follows that a benefit increase is welfare improving as long as it increases the after-tax reservation wage. The authors derive a formula for optimal benefits that includes five components: the interest rate, the duration of a job, the mean duration of an unemployment spell, the elasticity of duration with respect to benefits, and the responsiveness of the reservation wage with respect to benefits. All components can in principle be estimated (and some have been estimated). Estimates of the responsiveness of the reservation wage with respect to benefits are rare but feasible to obtain. The tentative calculations presented by the authors based on US data suggest that benefit hikes may be welfare improving.

A life cycle perspective on optimal UI is offered in the recent paper by Michelacci and Ruffo (2015). The authors focus on the US labor market and argue that nontrivial welfare gains can be achieved if benefits are made age-dependent: benefits should be higher for younger workers than for older ones, an age-pattern rather different from what is common in existing UI schemes. This proposal is supported by evidence that UI is more valuable to young workers since they have less means to smooth consumption. Young workers are also found to be less responsive to benefit hikes, i.e., moral hazard is less of a problem for the young. By means of a calibrated life cycle search model, the authors find that the optimal policy involves an increase in replacement rates from 50 to 80 percent for workers in their mid-twenties and from 50 to 60 percent for workers in their thirties. For workers in their forties and fifties, replacement rates should decline to 10 percent or less.

The standard assumption in the literature on optimal UI design is that individuals have unbiased beliefs about their employment prospects so that perceived expected utility coincides
with true expected utility. There is evidence, however, that the unemployed substantially overestimate how quickly they will return to work (see Spinnewijn 2015 and references therein). As a consequence of this optimism-bias, the unemployed search and save too little and deplete their assets too rapidly. The worker maximizes perceived expected utility whereas a paternalistic social planner maximizes the worker’s true expected utility. Spinnewijn (2015) studies how the optimal policy will be affected if workers have biased risk perceptions. With too optimistic beliefs about job finding, workers underestimate the risk of becoming long-term unemployed and will be less responsive to future incentives. The planner can therefore offer more insurance to the long-term unemployed without much adverse incentive effects. Under some conditions, the optimal policy may involve increasing benefits over the unemployment spell.

5 Concluding remarks
Economic theories do rarely produce straightforward policy implications. One reason is that the theoretical models typically are based on strong simplifying assumptions, another that these models often generate ambiguous results. Policy recommendations based exclusively on theory should be taken with a grain of salt. But policy conclusions based exclusively on observed correlations without any theoretical foundation should also be treated with a dose of skepticism since correlations without theory provide little understanding of mechanisms. With these caveats in mind, the time is ripe for some broad conclusions that draw on lessons from theory and empirical work as well as my subjective assessments.

1. There is strong evidence that UI affects a variety of labor market outcomes. For example, higher benefits and, in particular, longer potential duration of benefit receipt increase the duration of unemployment among the insured and probably overall unemployment as well. UI may at the same time encourage labor force participation. Recent evidence indicates that a substantial part of the impact of benefit extensions on duration is driven by a decline in labor force exits rather than by a decline in job finding.

2. The optimal UI system is unlikely to be the system that leads to the lowest unemployment rate. There is a strong case for public UI as a welfare-improving social institution. Capital markets are imperfect and a large fraction of the unemployed lack liquid wealth.

3. Whether UI is made mandatory (as in most countries) or voluntary but subsidized (as in some Nordic countries) is of second-order importance for UI coverage and welfare. UI coverage is mainly determined by eligibility rules and government subsidies.
4. UI interacts with other policies, especially active labor market policy and employment protection legislation. This implies that the optimal generosity of UI varies depending on other institutions and policies. For example, stricter monitoring of search activity implies a higher degree of optimal UI generosity.

5. There is very little research on optimal eligibility rules, including research on whether eligibility should be achieved via past work experience or past earnings.

6. There is a reasonably strong case for a benefit path that declines over the current spell of unemployment (and possibly over the length of previous spells as well). A declining profile provides better search incentives than a flat (or increasing) profile. Existing research does not give much guidance regarding the precise shape of this profile.

7. It is probably optimal to impose a “tax” on entry into unemployment by offering low (or no) benefits during the first week(s) of unemployment. A waiting period deters excessive inflow into unemployment via temporary layoffs or shirking. Private borrowing and savings may also work reasonably well as a substitute to UI for very short spells of unemployment.

8. Existing research has little to say about how benefits should vary over the distribution of (previous) earnings. The most common model involves a constant replacement rate up to a cap. The desired policy on this point is bound to be strongly affected by distributional preferences.

9. There is a case for considering some form of experience rating at the firm level. This would improve resource allocation via less cross-subsidization across industries with very different cyclical or seasonal labor demand patterns. The Danish system with so-called G-days is an interesting example.

10. There is a case for provision of unemployment benefits for workers who are part-time unemployed. Benefits for part-timers make part-time work more attractive relative to unemployment and may therefore increase the outflow from unemployment to part-time work.

11. Voluntary UI systems of the Ghent variety, as practiced in Denmark, Finland and Sweden, are conducive to unionization. An assessment of this feature lies largely in the eyes of the beholder. However, to the extent that unions by law have been given important welfare state functions, e.g. concerning implementation of employment protection legislation, the legitimacy of such a “union friendly” welfare state model is amplified by a high unionization rate.
REFERENCES


Lalive, R, C Landais och J Zweimüller (2013), Market Externalities of Large Unemployment Insurance Programs, manuscript.


Pigou A C (1933), The Theory of Unemployment, Macmillan.


Rothstein, J (2011), Unemployment Insurance and Job Search in the Great Recession, Brookings papers on Economic Activity (Fall), 143–196.


