Universität St.Gallen

Optimal Use of Labour Market Policies

Conny Wunsch

July 2007 Discussion Paper no. 2007-26

Department of Economics University of St. Gallen

Editor: Prof. Jörg Baumberger
University of St. Gallen
Department of Economics
Bodanstr. 1
CH-9000 St. Gallen
Phone +41 71 224 22 41
Fax +41 71 224 28 85
Email joerg.baumberger@unisg.ch
Publisher: Department of Economics
University of St. Gallen
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 25
Fax +41 71 224 22 98
Electronic Publication: http://www.vwa.unisg.ch

Optimal Use of Labour Market Policies[1]

Conny Wunsch

Author's address: Conny Wunsch
Swiss Institute for International Economics and Applied Economic Research
Bodanstrasse 8
CH - 9000 St. Gallen
Tel. +41 71 224 2474
Fax +41 71 224 2298
Email conny.wunsch@unisg.ch
Website www.siaw.unisg.ch/lechner/wunsch

---

[1] Swiss Institute for International Economics and Applied Economic Research, University of St. Gallen, Bodanstrasse 8, CH - 9000 St. Gallen, Switzerland, e-mail: conny.wunsch@unisg.ch. I particularly thank Stephan Wiehler, Michael Lechner, Blaise Melly and Fabio Trojani as well as seminar participants at the University of St.\ Gallen for helpful discussions and comments. I also thank Michael Lechner for providing some of his GAUSS procedures, adapted versions of which I used for some of the estimations.

Abstract

Labour market policies for the unemployed combine passive income support with active measures that aim at improving jobseekers' employment prospects. This paper extends the theoretical framework developed by Pavoni and Violante (2005a) for the optimal choice between different active and passive policies for the unemployed to a setting which allows for the use of a job search assistance programme that affects the exit rate to employment by raising search effectiveness but not productivity in the job. These programmes are one of the most widely used activation measures in OECD countries and should, therefore, be taken into account when considering the optimal design of labour market policies. The enriched model allows to answer a wide range of interesting policy questions. It is used to assess the optimality of the West German policy in the period 2000-2002 as well as the benefits from introducing tight monitoring. It is shown that sizeable budget savings could have been realised by switching to the optimal scheme, but that the net gains from monitoring are only small. In addition, some interesting results on the optimal use of job search assistance and training are derived. It is shown that existing policies already share some but not all features of the optimal scheme.

Keywords

Unemployment insurance, active labour market policies, recursive contracts, job search, human capital

JEL Classification

D82, J24, J64, J68

1 Introduction

In most countries labour market policies for the unemployed rely on two sets of instruments. On the one hand, income support during unemployment provides consumption insurance and gives jobseekers the opportunity to look for an appropriate job (so-called passive measures). On the other hand, so-called active labour market policies (ALMPs) aim at improving jobseekers' employment prospects, e.g. by increasing search effectiveness or skills. Expenditures on such policies are substantial, ranging from 1% to 5% of GDP in OECD countries (OECD, 2007).

Traditionally, passive and active measures for the unemployed have been treated separately in the literature. Starting with the seminal work of Shavell and Weiss (1979), there is a large literature on the optimal design of unemployment insurance payments. Here, the main issue is how to provide consumption insurance while ensuring that jobseekers provide an appropriate level of search effort, which is usually unobserved by the insurer. Different aspects of this question have been analysed such as the joint optimisation of benefits and employment taxes, both with and without repeated unemployment experience (Hopenhayn and Nicolini, 1997, 2005), or the impact of human capital depreciation during unemployment (Pavoni, 2003), among many others.

In contrast, the question of the optimal design of ALMPs, i.e. their optimal duration, timing within the unemployment spell and allocation to jobseekers, has been addressed only relatively recently. So far, the main interest of this literature is in finding the optimal (social welfare maximising) assignment mechanism for jobseekers to different programmes (e.g. Manski, 2000, 2001, 2002, 2004, 2005; Dehejia, 2005; Hirano and Porter, 2006). One part of this still rather small literature is empirical and tries to find the most effective programme for each individual on the basis of estimates of individual treatment response from historical data (e.g. Frölich, Lechner, and Steiger, 2003; Lechner and Smith, 2007; Frölich, 2007). A first attempt to empirically address timing issues has been made recently by Lechner and Wiehler (2007).

Yet, there are obviously important interactions between providing search incentives through unemployment insurance, and active measures. On the one hand, ALMPs try to increase exit rates to employment, e.g. by improving search effectiveness or skills, in order to reduce unemployment insurance payments. If successful, this would also ease the provision of search incentives because the returns to search are increased. On the other hand, these programmes are costly and participation reduces the time available for active job search. Empirical studies for different countries show that these so-called lock-in effects can be substantial depending on programme duration (e.g. van Ours, 2004; Lechner, Miquel, and Wunsch, 2006; Sianesi, 2004).

So far, there has been only one attempt to jointly optimise search incentives, unemployment insurance payments and the use of ALMPs. Pavoni and Violante (2005a, 2006) develop a theoretical framework for analysing the optimal choice between different active and passive policy instruments during the unemployment spell. In their model, workers differ with respect to human capital, which affects both the job finding rate and wages and which depreciates during unemployment. Pavoni and Violante (2005a) consider four policies: unmonitored search, job search monitoring, training which increases human capital, and social assistance.[1] The latter does not require job search. They derive the optimal path of benefits during unemployment and income taxes once employed under each policy, as well as the optimal sequence of policies over the unemployment spell. They show that human capital dynamics are necessary for policy transitions to occur. Moreover, in the absence of training, the typical sequence of policies would be unemployment insurance with unmonitored search followed by monitored search and then social assistance, which is shown to be an absorbing policy of "last resort". Optimal benefits turn out to be generally decreasing or constant during unemployment, but must increase after a successful spell of training in order to incentivise the worker to provide positive effort. As an illustration, they calibrate the model to the U.S. labour market and show that the optimal welfare-to-work scheme would yield sizeable welfare gains compared with the current system.

One important drawback of the model by Pavoni and Violante (2005a, 2006) is that it does not allow for active policies that affect the job finding rate but not productivity in the job. The most prominent example of such a policy are job search assistance programmes, which are one of the most widely used activation measures in OECD countries (OECD, 2007). They aim at increasing individual job finding rates by raising search effectiveness. The objective is to help those unemployed without any severe skill deficits to find a job as quickly as possible,

[1] Pavoni and Violante (2006) is a condensed version of Pavoni and Violante (2005a) where training is not considered.

e.g. by teaching them how to locate job vacancies and how to formulate a job application, or by simulating and practising job interviews. The programmes are usually short and, hence, relatively inexpensive. Moreover, they have proven to be quite effective in many countries (e.g. Fay, 1996; Heckman, LaLonde, and Smith, 1999; Martin and Grubb, 2001; Kluve and Schmidt, 2002) which explains their popularity and importance.

This paper tries to overcome this drawback by extending the theoretical framework of Pavoni and Violante (2005a) to a setting which allows for the use of a job-search-assistance technology that can raise search effectiveness, which is introduced as an additional state variable, hereby increasing the exit rate to employment but not productivity in the job because the latter only depends on human capital.

The theoretical results that can be derived from the enriched model confirm that the presence of the information asymmetry with respect to the worker's effort requires benefits to increase upon success of his activity and to fall upon failure in order to incentivise the worker, thus underlining the generality of this result first derived by Shavell and Weiss (1979). It is shown that social assistance remains an optimal policy of last resort. Furthermore, it is confirmed that it is never optimal to switch from monitored to unmonitored search. However, human capital dynamics are no longer necessary for policy transitions to occur because of the dynamics implied by search effectiveness as additional state variable.

To derive more detailed insights on the optimal sequence of policies the model is calibrated to the West German economy in the period 2000-2002. West Germany is an interesting case because it is comparable to most industrialised OECD countries. Moreover, in the course of substantial reforms of the German unemployment insurance system there have been heated debates on its optimal design in recent years. Finally, the availability of exceptionally rich administrative data allows most parameters of the model to be estimated nonparametrically from the same data and for the same sample, which is a large improvement compared to most existing calibrations of economic models.

The calibrated model is used to assess the optimality of the West German policy in the period 2000-2002 and to compare the features of the optimal scheme with the recent reforms of the

German system. It turns out that about 32% or 28 billion EUR of government expenditures could have been saved by switching to the optimal scheme. Moreover, the optimal duration of an unemployment insurance claim and the optimal level of social assistance for a worker with median characteristics are strikingly close to what has been introduced recently in Germany.

The simulation is also used to assess whether a currently unavailable tight monitoring technology would be beneficial and worth introducing in West Germany. It is found that it would be optimal to use monitoring for almost all levels of human capital and even when monitoring costs are high. Moreover, the optimal timing and duration of monitoring within the unemployment spell depends on the worker's search effectiveness. Thus, there is an important role for search effectiveness in the model even in the absence of a job-search-assistance technology. However, the net gains from monitoring are too small to justify its introduction.

Another salient feature of the simulation is that by varying the parameters of the job-search-assistance and training technologies it is possible to obtain more general insights on their optimal use. It is found that existing ALMPs already share some but not all features of the optimal scheme. In line with existing ALMPs it is shown that participation prolongs the period for which unemployment insurance payments rather than social assistance are optimal. Moreover, for job search assistance (JA) it is usually optimal to use it at the beginning of unemployment as well as after depreciation of search effectiveness following the successful use of JA later in the unemployment spell. Furthermore, the optimal duration of JA is usually short. However, it is shown that the accumulation rate of search effectiveness must exceed 3% per half-month in order for JA to be used at all.

For training (TR) it turns out that a wage accumulation rate of at least 1.8% per half-month is necessary for TR to be used at all, which means that the effect of TR must be relatively large. Then, it is optimal to use TR for a range of intermediate values of human capital, which increases the larger the effect of TR. The optimal duration of TR can be long for intermediate accumulation rates while it should be relatively short otherwise. Because TR affects both the job finding rate and wages, it usually dominates JA. However, if JA is sufficiently effective, it is optimal to use TR only after a short period of JA at the beginning of the unemployment spell.

The remainder of the paper is organised as follows. The next section describes the details of the model. In Section 3 the basic mechanisms of the model are discussed and the theoretical results on the optimal benefit scheme and sequence of policies are derived. In Section 4 the model is calibrated to the West German economy in the period 2000-2002. The optimal benefit scheme and sequence of policies is derived and the optimal use of job search monitoring, job search assistance and training is discussed in detail. A sensitivity analysis complements this section. The last section concludes. An appendix contains the main proofs of the theoretical results as well as further details on the calibration. Additional proofs as well as descriptive statistics of the data are provided in a supplementary technical appendix (Wunsch, 2007).

2 The model

2.1 The baseline setup

The baseline economy considered by Pavoni and Violante (2005a) is characterised by infinitely lived workers who have time-separable preferences over consumption $c \geq 0$ and effort a. The latter is either high ($a = e > 0$) or low ($a = 0$), which underlines the role of fixed costs and the extensive margin of participation decisions. Agents discount the future at rate $\beta \in (0, 1)$, and period utility is given by $u(c_t) - a_t$, where $u(\cdot)$ is strictly increasing, strictly concave and smooth with $\lim_{c\to\infty} u'(c) = 0$. Moreover, the first derivative of u^{-1} is assumed to be convex.[2]

Workers can either be employed ($z = z^e$) or unemployed ($z = z^u$). They differ in their human-capital or skill endowment h, which may accumulate during employment or after a successful spell of training and depreciates at rate δ^h during unemployment without training or if training fails.[3] Human capital evolves according to

$$h^s = g(h), \qquad g'(h) > 0, \quad g''(h) \leq 0, \tag{1}$$

$$h^f = (1 - \delta^h)h, \tag{2}$$

[2] As Pavoni and Violante (2005a, 2006) argue, a wide range of utility functions, including the CARA class and a large part of the CRRA class, satisfies this condition.

[3] In the revised and condensed version of Pavoni and Violante (2005a), the authors abstract from the accumulation of human capital and training (Pavoni and Violante, 2006).

where the superscripts s are used for success and f for failure of the worker's search or training activity. Pavoni and Violante (2005a) assume that $g(h) = Ah^{\alpha} + (1 - \delta^{h})h,\ \alpha \in [0,1],\ A \geq 0$. Here, no additional assumptions are made with respect to the function $g(\cdot)$.

During employment a worker of type h produces output $w(h) \in [0, w_{max}]$ with $w(0) = 0$ and $w(\cdot)$ being a continuous and increasing function. Following Hopenhayn and Nicolini (1997) employment is assumed to be absorbing in order to concentrate on the (current) unemployment experience.[4] The disutility of effort during employment is generally not restricted to be the same as during search or any of the active labour market programmes. However, it must be ensured that accepting a job offer always dominates staying unemployed.[5] Human capital is assumed to accumulate at no cost (learning by doing) to ensure that employment remains dominant.

During unemployment, human capital h can be raised by use of a training technology at cost $\kappa^{TR} > 0$. Traditionally, it is argued that this increases productivity in a job, hereby increasing the number of jobs a worker qualifies for and potential wages upon reemployment (e.g. Mortensen, 1970). This, in turn, should raise the job finding probability and the value of employment, thus, increasing the returns to search. In the model, training is successful with probability $\theta^{TR}(a)$, which is only positive if effort is also positive, i.e. $\theta^{TR}(e) > \theta^{TR}(0) = 0$.[6] In this case, human capital accumulates according to (1).

For simplification it is assumed that workers do not have access to storage, insurance or credit markets. In particular, it is assumed that workers cannot self-insure against the random outcome $y \in \{s, f\}$ of their search or training activity, e.g. by saving. Pavoni and Violante (2005a) show that when workers can save through credit markets but still face a no-borrowing constraint, which is a reasonable assumption for unemployed workers, the same optimal contract can be implemented by introducing a linear, time-invariant interest tax.

[4] Qualitative results for the same unemployment spell do not change as long as the job separation rate is exogenous. Optimal contracts with endogenous job separation are studied by Zhao (2000) and Hopenhayn and Nicolini (2005) who show that in this case the optimal contract has to take into account the worker's full employment history.

[5] For example, assuming that effort is the same during employment and search would imply that the worker searches for a job with the same intensity and time input as he works.

[6] This incorporates the idea that the (expected) effect of training is increasing in training effort.

The basic contractual relationships are as follows. There is a risk-neutral planner who offers a risk-averse agent, the unemployed worker, an insurance contract at time $t = 0$. This contract maximises expected discounted (at rate β) net fiscal revenues,[7] subject to providing the agent with at least an expected discounted utility level of U_0. The latter is exogenously given, e.g. the outcome of voting, and can be regarded as a measure of the generosity of the welfare system.

The planner can observe the worker's employment status $z \in \{z^u, z^e\}$, his activity (search or programme participation) and the outcome $y \in \{s, f\}$ of this activity. But he cannot observe the worker's effort choice a, so that he faces a moral hazard problem. However, there is a monitoring technology available where, upon payment of a cost $\kappa^{JM} > 0$, the planner can observe the worker's search effort during job search.[8]

The contract specifies, for each period t and contingent on all observable histories up to t, the transfers to the worker, the policy option to be used and the corresponding recommended effort choice of the worker. With the exception of monitored search where there is no informational asymmetry anymore, all recommendations of high effort must be incentive compatible. Moreover, Pavoni and Violante (2005a) allow the planner to specify the contract contingent on the observable realisation $x_t \in [0, 1]$ of a uniform random variable X_t. This randomisation is a device to convexify the planner's problem (see also Phelan and Stacchetti, 2001; Phelan and Townsend, 1991, as well as the technical appendix).

2.2 Extensions of the baseline model

One important limitation of the model by Pavoni and Violante (2005a) is that it only allows for active policies that affect both the job finding rate and productivity in the job. Although this covers both formal training and programmes that provide work experience, it excludes job-search-assistance programmes as one of the most widely used activation measure for the unemployed in OECD countries (OECD, 2007). These programmes aim at increasing individual job finding rates by raising search effectiveness without affecting productivity in the job. The

---

[7] In other words, the planner minimises expected discounted net expenditure of the unemployment insurance system.

[8] The planner may, for example, pay a caseworker who monitors closely the workers search activities. Such technology is, however, unavailable during programme participation because the worker's learning effort, in particular his attention and concentration, are very hard to verify.

objective is to help those unemployed without any severe skill deficits to find a job as quickly as possible, e.g. by teaching them how to locate job vacancies and how to formulate a job application, or by simulating and practising job interviews. Usually, these programmes are short and, thus, less costly than traditional training or wage subsidies. This, together with the fact that they have proven to be quite effective in many countries (e.g. Fay, 1996; Heckman, LaLonde, and Smith, 1999; Martin and Grubb, 2001; Kluve and Schmidt, 2002), explains the popularity and importance of job-search-assistance programmes. Therefore, these programmes should be taken into account when considering the optimal design of labour market policies.

For this reason, the economy considered by Pavoni and Violante (2005a) is enriched by introducing search effectiveness as another source of worker heterogeneity. It is assumed that the worker is endowed with an initial level of search effectiveness $\hat{p}$. This level can be raised by use of a newly introduced job-search-assistance technology at cost $\kappa^{JA} > 0$, but the effect depreciates over time. However, in contrast to human capital h it is assumed that search effectiveness p cannot fall below its original level $\hat{p}$. Similar to the training technology, the outcome of job search assistance is stochastic with success probability $\theta^{JA}(a)$ where $\theta^{JA}(e) > \theta^{JA}(0) = 0$. Depending on whether job search assistance was successful or failed, the law of motion for p is given by

$$p^s = f(p), \qquad f(0) = 0, \qquad f'(p) > 0, \quad f''(p) \leq 0, \tag{3}$$

$$p^f = \max\{\hat{p}, (1-\delta^p)p\} \tag{4}$$

with no additional assumptions being made with respect to the function $f(\cdot)$.

In the enriched setting, the job finding probability of an unemployed worker is not only affected by search effort $a \in \{0, e\}$ and human capital h, but also by search effectiveness p. The corresponding hazard rate is denoted by $\pi(h, p, a)$ and it is assumed that $\pi(h, p, 0) \equiv 0$ and that $\pi(h, p, e) \equiv \pi(h, p) \in (0, 1)$ is continuous and increasing in both elements. Note that the monotonicity of π in h together with human capital depreciation induces negative duration dependence in π, which is an empirically well established fact in labour economics (see e.g. the survey of the relevant literature by Machin and Manning, 1999).

In total, there are eight policy options in the enriched economy: unmonitored and monitored search as well as job search assistance and training, each with high (positive) and low (zero) effort recommendation. However, the planner will never combine monitoring with low effort because any deviation from no search can be observed at no additional cost since $\pi(h, p, 0) = 0$. Moreover, job search assistance and training will always be combined with high effort, because otherwise the cost of the programme would not be compensated by a return since $\theta^{JM}(0) = \theta^{TR}(0) = 0$. Thus, there is a set of five instruments from which the planner will choose: unmonitored search with positive effort (unemployment insurance, UI) or zero effort (social assistance, SA), and, each with positive effort, job search monitoring (JM), job search assistance (JA), and training (TR). In addition, because the insurer can control consumption during employment as well, he has the opportunity to impose a wage tax or to pay a wage subsidy.

2.3 The planner's problem in the enriched setting

The optimisation problem of the planner is formulated in the recursive form proposed by Pavoni and Violante (2005a) with two differences. First, job search assistance is added to the set of policies. Second, search effectiveness is introduced as another state variable, so that a state is defined by the promised continuation utility U, the level of human capital h and search effectiveness p, i.e. by the triple (U, h, p).

At the beginning of each period, the planner chooses the optimal policy instrument $i(U, h, p)$ for an unemployed worker who enters the period with state (U, h, p) by solving

$$V(U, h, p) = \max_{i \in \{JA, JM, SA, UI, TR\}} V^i(U, h, p) \tag{5}$$

where V is the upper envelope of the values associated to the different policies which are described in detail below. If the planner decides to use randomisation through X, the planner solves

$$\begin{aligned} \mathbf{V}(U, h, p) \quad &= \quad \int_0^1 \max_{U(x) \in D} V\big(U(x), h, p\big) dx \\ s.t. \quad U &= \int_0^1 U(x) dx, \end{aligned} \tag{6}$$

where the second equation is the promise-keeping constraint which obliges the planner to deliver continuation utility U to the agent in expected value (with respect to the shock x) terms.

In the following the value functions for employment and each of the five different policy instruments are described in detail. In each case, the planner chooses an effort recommendation $a(U,h,p) \in \{0,e\}$, the transfer $c(U,h,p)$ and the continuation utilities $U^y(U,h,p)$ conditional on the outcome $y \in \{f,s\}$ of the agent's activity.

To concentrate on unemployment, *employment* is assumed to be an absorbing state without informational asymmetries. For an employed worker and state (U,h,p) the planner solves

$$\begin{aligned} W(U,h,p) \quad &= \quad \max_{c,U^s} \; w(h) - c + \beta W(U^s,h^s,p^f) \\ &s.t. \quad U = u(c) + \beta U^s. \end{aligned} \tag{7}$$

It is easy to show that W is continuous, increasing in h, and decreasing, concave and continuously differentiable in U given the properties of w and u. Moreover, W is independent of p because employment is absorbing, as well as separable in U and h:

$$W(U,h) = \frac{\Omega(h)}{1-\beta} - \frac{u^{-1}\big((1-\beta)U\big)}{1-\beta}, \tag{8}$$

with $\Omega(h) = (1-\beta)w(h) + \beta\Omega(h^s)$ being the discounted stream of wages. Thus, consumption of the agent will be smoothed fully and promised utility will remain constant over time. The promise-keeping constraint then implies that the optimal transfer is constant over time and given by $c^W(U) = u^{-1}\big((1-\beta)U\big)$. The implicit wage tax (if positive) or subsidy (if negative) that is imposed on employed workers is, thus, given by $\tau(U,h) = w(h) - c^W(U)$ and does not depend on search effectiveness p for obvious reasons.

The planner's problem under the *unemployment insurance (UI)* scheme is given by

$$\begin{aligned} V^{UI}(U,h,p) \quad &= \quad \max_{c,U^f,U^s} \; -c + \beta\big[\pi(h,p)W(U^s,h^f,p^f) + \big(1-\pi(h,p)\big)\mathbf{V}(U^f,h^f,p^f)\big] \\ &s.t. \quad U = u(c) - e + \beta\big[\pi(h,p)U^s + \big(1-\pi(h,p)\big)U^f\big] \\ &\qquad\; U \geq u(c) + \beta U^f, \end{aligned} \tag{9}$$

where h^f is generated by (2), and W and $\mathbf{V}$ are given by (7) and (6), respectively. The first constraint is the promise-keeping (PK) constraint and the second one the incentive compatibility (IC) constraint.

If the planner chooses to monitor the search activities of the agent by using *job search monitoring (JM)*, he solves the problem

$$\begin{aligned} V^{JM}(U,h,p) &= \max_{c,U^f,U^s} -c-\kappa^{JM}+\beta\big[\pi(h,p)W(U^s,h^f,p^f)+\big(1-\pi(h,p)\big)\mathbf{V}(U^f,h^f,p^f)\big] \\ & \quad s.t. \quad U=u(c)-e+\beta\big[\pi(h,p)U^s+\big(1-\pi(h,p)\big)U^f\big]. \end{aligned} \tag{10}$$

Note the absence of the IC constraint because in this case the planner can observe search effort (against payment of the cost κ^{JM}) so that there is no incentive problem during JM.

If the agent is enrolled in a programme i, which is either *training (TR)* or *job search assistance (JA)*, the planner solves

$$\begin{aligned} V^i(U,h,p) &= \max_{c,U^f,U^s} -c-\kappa^i+\beta\big[\theta^i\mathbf{V}(U^s,h^i,p^i)+(1-\theta^i)\mathbf{V}(U^f,h^f,p^f)\big] \\ & \quad s.t. \quad U=u(c)-e+\beta\big[\theta^i U^s+(1-\theta^i)U^f\big] \\ & \qquad\quad\;\; U\geq u(c)+\beta U^f, \end{aligned} \tag{11}$$

where $i\in\{JA,TR\}$ and $(h^{TR},p^{TR})=(h^s,p^f)$, and $(h^{JA},p^{JA})=(h^f,p^s)$. It is assumed that search and programme participation are mutually exclusive activities within a period, implying that participants in i cannot exit to employment directly from the programme.[9] But since the length of a period can be arbitrarily small, this assumption is not restrictive. With probability θ^i, i raises human capital if $i=TR$, and search effectiveness if $i=JA$, at the cost κ^i. Since positive effort is required ($a=e>0$), both the PK and the IC constraint are needed.

Under the *social assistance (SA)* scheme the worker is not required to search ($a=0$) but receives some transfer that ensures delivery of promised utility U. Since effort is zero, no IC

[9] This corresponds to a so-called lock-in effect of programme participation, which has been documented in many empirical studies (e.g. Gerfin and Lechner, 2002; van Ours, 2004; Lechner, Miquel, and Wunsch, 2006; Sianesi, 2004, 2007; Jespersen, Munch, and Skipper, 2004).

constraint is needed. The planner's problem is, therefore, given by

$$V^{SA}(U,h,p) \quad = \quad \max_{c,U^f,U^s} -c + \beta \mathbf{V}(U^f, h^f, p^f) \qquad (12)$$
$$s.t. \quad U = u(c) + \beta U^f.$$

Proposition 0 in the appendix states the properties of the value functions $\mathbf{V}$, W and V^i. In particular, they are continuous functions which are increasing in h and p, and decreasing, concave and continuously differentiable in U.

3 Theoretical results

3.1 Optimal sequence of payments under each policy

Proposition 1 characterises the optimal sequence of benefits under the five policies and provides some results on the optimal wage tax/subsidy levied on the worker upon reemployment.

Proposition 1: *(i) Benefits are constant during JM and SA. (ii) During UI, JA and TR benefits are constant if the incentive compatibility constraint does not bind. If it binds, benefits decrease if the worker's activity fails and increase if it succeeds. (iii) During JM the wage tax decreases.*

The proof in Appendix A shows that results (i) and (ii) are independent of the presence of human capital and search effectiveness as state variables in the model, because they are driven by the presence or relevance of the IC constraint only. Thus, they are completely general.

Intuitively, under JM and SA there is no incentive problem because under JM search effort can be verified, and SA releases the worker from any search activity. The same holds if the IC constraint is not binding during UI, JA and TR. Thus, full insurance with constant benefit payments is provided. In contrast, the necessity to incentivise agents to provide positive effort during UI, JA and TR when the IC constraint binds requires the planner to 'punish' the agent upon failure and to 'reward' him upon success of his activity.

Result (iii) follows from wages $w(h)$ decreasing during unemployment because of human capital depreciation, which implies that the wage tax $\tau(U,h) = w(h) - c^W(U)$ decreases during JM.

In contrast, the behaviour of $\tau(U,h)$ during UI is a quantitative issue because both $w(h)$ and $c(U,h,p)$ decline. If consumption declines faster than wages, then $\tau(U,h)$ increases. Otherwise it decreases, or remains constant if both decrease at the same speed.

3.2 Optimal sequence of policies

In the following, both the economic forces at work and the theoretical results on the optimal sequence of policies that can be derived from the model are discussed.

Whenever positive effort e is required - as is the case for all policies except SA - the planner must compensate the agent for this effort (so-called effort compensation cost). Since the disutility of providing e is fixed while, because of the concavity of $u(\cdot)$, the marginal utility of consumption falls with increasing U, the effort compensation in terms of (utility from) transfer payments must increase with U. Thus, the effort compensation cost will be prohibitively large for high levels of U, making SA most attractive. Since promised utility U does not fall during SA (see below), SA should remain optimal.

Another case where SA is most attractive is when the depreciation of human capital and (potentially) search effectiveness has reduced the returns to search to a prohibitively low level. Since human capital depreciates further while search effectiveness does not increase during SA, any policy other than SA becomes even less attractive after one period of SA implying that SA should remain optimal in this case as well.

Proposition 2: *(i) Whenever choosing SA in period t is optimal, it is also optimal thereafter. (ii) If in some period t the optimal programme is JM, then the next period it is never optimal to switch to UI if the conditions of Lemma A1 and A2 in Appendix A are satisfied. (iii) If $f(p)$ is strictly concave, JA is not absorbing. (iv) If $g(h)$ is strictly concave, TR is not absorbing.*

Proposition 2 (i) shows that SA is always absorbing. As a result and because $\pi(h,p,0)=0$, the equilibrium value of SA does not depend on h and p and is given by $\hat{V}^{SA}(U) = -c^{SA}(U)/(1-\beta)$ where the benefit paid during SA is constant and given by $c^{SA}(U) = u^{-1}\big((1-\beta)U\big)$.

Another important cost component are the incentive costs which arise from having to obey an IC constraint during JA, UI and TR. Using the PK constraints, the IC constraints in these three cases can be rewritten as

$$U_i^s - U_i^f \geq \frac{e}{\beta\theta^i}, \qquad i \in \{JA, TR\}, \qquad U_{UI}^s - U_{UI}^f \geq \frac{e}{\beta\pi(h,p)}, \tag{13}$$

which is independent of the transfer c. Since larger differences in the utility upon success and failure of the worker's activity correspond to larger differences in the respective consumption levels, the risk-averse agent has to be compensated with a larger average transfer for a given level of promised utility U. If the IC constraint is binding, (13) implies that incentive costs remain constant for JA and TR but increase for UI when h and potentially p fall during unemployment because $\pi(\cdot)$ declines in this case.

For $i \in \{JM, JA, TR\}$ the are also direct cost κ^i which have to be incurred. In case of JA and TR this adds to the incentive cost, while for JM it replaces the incentive cost that has to be incurred during UI. The latter implies, that UI should only be used before JM as long as the incentive cost is lower than the fixed cost κ^{JM}. Moreover, since the incentive cost increases during UI, a switch to JM becomes more likely with increasing unemployment duration. Once κ^{JM} is lower than the incentive cost during UI, a switch back to UI should only occur after successful use of JA or TR because, otherwise, the incentive cost keeps rising for UI because of the depreciation of h and potentially p. Result (ii) of Proposition 2 shows that this is indeed the case under very general conditions.

Whether and if so when in the unemployment spell JA and TR will be used depends on the returns of these programmes relative to their costs. The former are bounded above because both the job finding rate and wages are bounded above, making the use of these policies less likely the higher h and p. Furthermore, Proposition 2 shows that JA (TR) will not be absorbing if the increase in p (h) upon success of JA (TR) declines with p (h) while the cost of the programme, κ^{JA} (κ^{TR}), remains constant, since in this case, the programme will become unprofitable at some point after its successful use. Note, moreover, that during both programmes effort compensation, incentive and direct costs have to be incurred. Thus, at least for one period this cost has to be offset by the (expected) returns in order for them to be used at all.

To complete the predictions of the model, it is interesting to see what additional results can be derived in the absence of human capital dynamics. Pavoni and Violante (2005a, 2006) show that in the baseline model in this case all policies are absorbing, i.e. that human capital dynamics are necessary for policy transitions to occur. The results in Proposition 3 imply that this is generally not the case in the enriched setting.[10]

Proposition 3: *Consider the case where human capital is fixed over time. If either $p = \hat{p}$ or $\delta^p = 0$ the following results hold. (i) If JM is optimal at (U, h, p), JM will always be optimal thereafter. (ii) If JA is not optimal at (U, h, p) and $\theta^{JA} \leq \pi(h, p)$, JA will never be optimal thereafter. (iii) If UI is optimal at (U, h, p) and $\theta^{JA} \leq \pi(h, p)$, UI will always be optimal thereafter.*

Proposition 3, which is proven in Appendix A, implies that if JA is not used in period 0, i.e. at $(U_0, h, \hat{p})$, it will never be used if $\theta^{JA} \leq \pi(h, \hat{p})$. Because of the latter condition this is more likely the larger h and the original level of search effectiveness, $\hat{p}$, and the lower the success rate of JA, θ^{JA}. Moreover, in this case the other policies are absorbing. On the other hand, if JM is optimal in period 0, i.e. at $(U_0, h, \hat{p})$, then it is used forever even if $\theta^{JA} > \pi(h, \hat{p})$.

However, in all other cases JM and UI are not absorbing because whenever JA is used and search effectiveness changes, either because it accumulates during JA or because it depreciates after use of JA, the returns to search and the incentive costs during UI change.

4 Simulation for West Germany

In order to derive more detailed insights on the optimal sequence of policies the model is calibrated to the West German economy in the period 2000-2002. West Germany is an interesting case to study because it is comparable to most industrialised OECD countries and, in the course of substantial reforms of the German unemployment insurance system, there have been heated debates on its optimal design in recent years. The calibrated model is used to assess the optimality of the policy implemented in West Germany in the period 2000-2002 by comparing it with the optimal scheme that results when the same initial utility $\tilde{U}_0(h_0, p_0)$ is delivered to

[10] Note that in this case the optimal choice of policies is driven by the state variables U and p, only. In the absence of human capital dynamics there is no role for a training programme.

a worker with characteristics (h_0, p_0) as implied by the actual system. Second, the potential benefits of introducing a - currently unavailable - tight monitoring technology in West Germany are evaluated. Third, to draw some general conclusions on the characteristics and optimal use of beneficial job-search-assistance and training technologies, ranges of parameter values of these policies for which these programmes are optimally used as well as their optimal duration and timing within the unemployment spell are identified.

4.1 Unemployment insurance in West Germany

In Germany, unemployment insurance is mandatory and employees who have contributed for at least 12 months within the 3 years before entering unemployment are eligible for unemployment benefits (UB) if they register with the public employment service (PES). The minimum UB entitlement is 6 months. In the period under consideration, the maximum claim increased stepwise with total contribution time in the 7 years before becoming unemployed, and age, up to a maximum of 32 months at age 54 or above with previous contributions of at least 64 months. Since 1994, the replacement rate is 67% of previous average net earnings from insured employment with dependent children, and 60% without.

Until 2005, unemployed could become eligible for unemployment assistance (UA) after exhaustion of UB. In contrast to UB, UA was means tested and potentially indefinite. However, like UB, UA was proportional to previous earnings but with lower replacement rates than UB (57% and 53% with and without dependent children, respectively). Unemployed who were ineligible for UB and UA could receive social assistance, which was a fixed monthly payment unrelated to previous earnings, means-tested and administered by local authorities.

In 2005, one of the largest reforms of German unemployment insurance became effective. The maximum UB entitlement has been cut to 12 and 18 months for unemployed below and above age 55, respectively. Moreover, UA and social assistance have been combined to a new fixed means-tested welfare payment of normally 345 EUR that is called unemployment benefits II.

Actual payment of benefits is conditional on active job search, regular show-up at the PES, and participation in labour market programmes. In case of noncompliance with benefit conditions sanctions, i.e. reductions in or suspensions of benefits, can be imposed. However, search activities are monitored not very strictly at the moment due to capacity problems within the PES.

Table 1 provides numbers on the use of and expenditures on the most important active and passive policy measures in West Germany. Given an unemployment rate of about 9% of the civilian labour force with a relatively large share of long-term unemployed (about 30%) in the period 2000-2004 (BA, 2001-2005), expenditures on UB and UA are relatively large.

Table 1: Use of alternative policy instruments in West Germany

	Entries/recipients in 1000					Expenditure in million EUR				
	2000	2001	2002	2003	2004	2000	2001	2002	2003	2004
Job search assistance	286	319	513	690	789	183	196	289	364	312
Training	338	242	259	161	124	4226	4412	3998	3024	2342
Subsidised employment	180	143	158	135	136	1849	1608	1464	1118	765
Support of self-employment	62	65	89	178	249	534	576	768	1271	2045
Unemployment benefits	998	1045	1240	1326	1288	15342	16403	18426	20944	21231
Unemployment assistance	786	761	877	1055	1213	8067	7519	8011	9257	11024

Note: The numbers from 2005 onwards are not comparable due to a complete change in legislation.
Source: BA (2001-2005)

Expenditures on training are substantial given the number of participants for two reasons. First, with durations of up to two years, training programmes are relatively long in Germany compared to most OECD countries. Second, participants usually receive a special form of benefit (so-called maintenance allowance) while in the programme, which is of the same amount as UB or UA. Until 2005 these benefits did not affect remaining UB claims and, hence, increased total benefit claims, thus providing strong incentives to participate in training.

In terms of the number of participants, job search assistance has become the most important activation measure, by far, in recent years. Expenditures are moderate because durations are short (up to three months but usually no more than one month). Support of self-employment also has gained importance in recent years while the use of subsidised employment is declining, both in the number of entries and in durations and expenditure.[11]

[11] Besides the usual counselling and placement services, there are also special instruments for youth, elderly unemployed and the disabled in Germany. For further details on these measures and German unemployment insurance see Wunsch (2006).

4.2 Data and population of interest

With the exception of the preference and programme cost parameters, all parameters of the model are calibrated using an administrative database which has been built up by the German Institute for Employment Research. The database is a 2% random sample from all individuals who have been subject to German social insurance at least once since 1990. It covers the period 1990-2005 and combines spell information from social insurance records, programme participation records and the benefit payment and jobseeker registers of the PES. Thus, it provides detailed information on employment status (employed, unemployed with or without programme participation) on a daily basis for the period 1990-2005.

The database comprises very detailed information in several dimensions. Personal characteristics include education, age, gender, marital status, number of children, profession, nationality and health. The benefit payment register provides information on type and amount of benefits received, remaining benefit claims and imposition of sanctions. The jobseeker register includes information on the desired form of employment, compliance with benefit conditions and the number of placement propositions by the PES. Moreover, the data comprise information on employments including form of employment, industry, occupational status and wages. With respect to programme participation the data cover type of the programme as well as planned and actual duration. Detailed regional information, which include federal state, local unemployment rate, migration, demographic and industry structure, infrastructure and urbanity, complement the database (see the technical appendix for a full list of variables).

For the simulation, the model is calibrated to the population of West German workers who registered unemployed with the PES between January 2000 and December 2002 and received benefits from the unemployment insurance system (UB or UA).[12] To concentrate on the main body of the workforce, apprentices, young men doing their civilian or military service as well as elderly workers in special forms of employment are excluded. This reduced sample is referred to as the reference population of interest in the following.

[12] Note that this excludes recipients of social assistance.

4.3 Calibration

In the simulation, time units are defined by half-months. This is the smallest level of aggregation which is reasonable with the data. Human capital is measured in terms of wages and calibrated using the wage from the last employment spell before entering unemployment.[13] The definition of search effectiveness is more complicated since it is not observed in the data. Here, it is estimated by the contribution of variables related to search effectiveness to the probability of finding employment within 3 months after entering unemployment. Concretely, search effectiveness of individual j is estimated as

$$\begin{aligned} \hat{p}_j &= P(Y=1|X=x_j, Z=z_j) - P(Y=1|X=x_j) \\ &= P(Y=1|X=x_j, Z=z_j) - E_Z[P(Y=1|X=x_j, Z=z_j)] \end{aligned} \tag{14}$$

where $Y = 1$ if the worker succeeded in finding employment within 3 months after entering unemployment, and where Z is a vector of variables that are related to search effectiveness while X contains all other variables that affect the job finding probability. A probit model is estimated with Z including - inter alia - measurements of labour market attachment, duration of past unemployment spells, compliance with benefit conditions, number of placement propositions by the PES in the past, local labour market conditions, health, foreigner status and presence of (young) children. See Appendix B for more details.[14]

Preferences are parametrised as follows: The half-monthly discount factor β is chosen to match an interest rate of 4% per annum, which prevailed in the EURO area in the period of interest. Period-utility over consumption is assumed to be logarithmic, i.e. $u(c) \equiv \ln(c)$, but specifications with intertemporal elasticities of substitution below and above one are tested as well. To calibrate the disutility of effort, e, an approach similar to the one suggested by Pavoni and

[13] Most individuals enter directly from employment (71%). For all other persons, the wage is depreciated according to the calibrated wage depreciation rate for the time between leaving employment and entering unemployment. Moreover, for part-time-workers the wage is scaled up to obtain an approximate fulltime-equivalent. Figure 8 in Appendix E displays the corresponding wages. See Appendix B for more details.

[14] In the numerical solution h and p are discretised, while U is treated as continuous variables. For h and p the grids are geometrically spaced at rate 10% with 20 grid points chosen for h and 10 for p. The grid chosen for U has 500 equidistant points in the interval $[50, 1500]$. For each combination of h and p, the value functions for all policies with respect to U are computed using Chebychev polynomials up to the 20th order.

Violante (2005a, 2006), which originates from common practice in calibrating macroeconomic models, is used. Let the disutility of time n spent working be logarithmic as well, and denote by ϕ the relative weight on leisure versus consumption. Assuming a standard Cobb-Douglas production function, the static optimality condition of the worker yields a value of $\phi = 2.35$ given a labour share of 0.73 (BMAS, 2003), a consumption-income ratio of 0.72 (Statistisches Bundesamt, 2000-2002) and a fraction of time spent working of $n = 0.3$. This implies a value for the disutility of work effort of $e = \phi[\ln(1) - \ln(1-n)] = 0.84$ (Chari, Christiano, and Kenhoe, 1995). For the baseline calibration it is assumed that the disutility of effort during unemployment equals the one during work. Yet, as a sensitivity check the case where only half of time spent on work is spent on search and programme participation is considered as well.

Nonparametric estimates of the baseline exit rates to employment as well as the wage depreciation and accumulation rates without participation in JA or TR can be obtained from nonparticipants who are matched to the reference population of interest using the radius matching estimator proposed and applied by Lechner and Wunsch (2006). The matching is performed because interest is in the values of these parameters for the reference population which also includes participants and, therefore, differs systematically from actual nonparticipants in characteristics that also affect the exit rate to employment as can be seen from Table **??** in the technical appendix. However, as is argued in Appendix D, the data are sufficiently rich to control for the potential sources of selection bias with respect to the reference population of interest by use of matching techniques. See Appendix C for all details.

It turns out that the job finding rate is increasing in both h and p. However, the impact of p is much stronger, and for intermediate to high values of h the job finding rate stays almost constant. The estimate for the wage depreciation rate is $\hat{\delta}^h = 0.985$. For the wage accumulation rate during employment a value of 1.003 is obtained.[15]

In order to calibrate the job-search-assistance and training technologies, estimates of the effects of JA and TR of different durations on, respectively, p and h for the reference population of interest are required, because each period the planner has to decide whether the worker

[15] In the simulation, depreciation and accumulation of h and p are stochastic. The respective depreciation and accumulation probabilities are calculated from the respective depreciation and accumulation and the geometric rate at which the grids for h and p are spaced.

should be assigned to a programmes or, if he already attends a programme, whether he should stay for another period. However, when trying to estimate these effects two kinds of selection problems arise. First, as can be seen from Table **??** in the technical appendix, the respective programme participants differ systematically from the population of interest in ways that are also related to the respective outcome variables of interest. Second, actual programme durations are potentially endogenous. In the reference population of interest actual durations differ from planned durations by more than 15% in 11% and 21% of the cases for JA and TR, respectively. Thus, actual programme durations cannot be regarded as exogenous.

Given sufficiently rich data, both selection problems can be solved using a so-called dynamic (or sequential) treatment evaluation approach as suggested by Lechner and Miquel (2001, 2005). The idea is to consider an n-period programme as a sequence of n one-period programmes and to control for selection at the beginning of each of the n periods. Selectivity between programme participants and the population of interest can then be controlled for at the beginning of the sequence (at $t = 0$), whereas endogeneity of programme durations is accounted for by selection correction in the $n - 1$ following periods. The effect of a programme of length n relative to a programme of length $n - 1$ can then be obtained by comparing the sequence 'participating up to $t = n$' with the sequence 'participating up to $t = n - 1$ and not participating in $t = n$'.[16]

Conditions for identification of the effects of JA and TR of different durations for the reference population of interest and their validity in this application are discussed in detail in Appendix D. The effects are estimated using an adapted version of the sequential matching estimator proposed by Lechner (2004, 2006). See Appendix D for all details on this estimator. For each of the two programmes, JA and TR, the estimation is performed separately in two subgroups of high and low search effectiveness and skills, respectively (sample divided at the median of the respective variable). For JA the effects of completing one half-month of JA compared with nonparticipation as well as two versus one, three versus two, four versus three and five versus four half-months of JA are estimated. For TR there are not enough observations to perform the

[16] An alternative way of estimating the effect of programmes of different durations would be to apply the methodology of Hirano and Imbens (2004) for continuous treatments (see e.g. Flores-Lagunes, Gonzalez, and Neumann, 2007). However, this method assumes that selection into programmes of different durations is fully determined at the beginning of the programme. That is, different programme durations are not the outcome of intermediate decisions after the programme started.

estimation half-monthly. Here, six months of TR is compared with nonparticipation as well as twelve months versus six months of TR.

It turns out that none of the programmes succeeds in raising search effectiveness or human capital independent of programme duration (see the technical appendix).[17] Since programme participation is costly, it is, therefore, obvious that JA and TR should not be used in the optimal policy. However, the simulation allows to identify regions of parameter values, for which it is optimal to use these programmes, thus providing reference values, e.g. for the minimum effectiveness such programmes must exhibit in order to be optimally used at all. Concretely, different values of the success rates of JA and TR, θ^{JA} and θ^{TR}, that match specific accumulation rates are simulated. Moreover, since the depreciation rate of search effectiveness cannot be calibrated from the data because the effects of JA are zero, different values of δ^p will be tested in the simulation.

Values for the cost parameters of JA and TR are calculated from total expenditures on the respective programme per year (excluding benefit payments), the number of participants and average programme durations from official statistics (BA, 2001-2005). This yields programme costs of $\kappa^{JA} = 190$ and $\kappa^{TR} = 300$ per half-month in EUR.

To parameterise the cost of a (currently unavailable) tight monitoring technology, the average gross salary of a caseworker per half-month (about 1200 EUR according to BA, 2001-2005) is divided by the number of unemployed a caseworker can reasonably take care of in a half-month. A conservative value would be 20 unemployed which yields a value of 60 EUR. Allowing for some administrative cost, the baseline value for κ^{JM} is set to 100 EUR. This value will be varied to check the sensitivity of the optimal use of JM and the other policies to κ^{JM}.

The parameters of the policy implemented in West Germany in the period under consideration, i.e. amount and duration of benefits as well as programme durations and fraction of unemployed subject to each policy in the period 2000-2002, are calculated directly from the data. Figure 8 in Appendix E displays the values of these parameters as a function of human capital.

[17] This is consistent with the findings of Wunsch and Lechner (2007), who analyse the effectiveness of different forms of JA and TR using the same database.

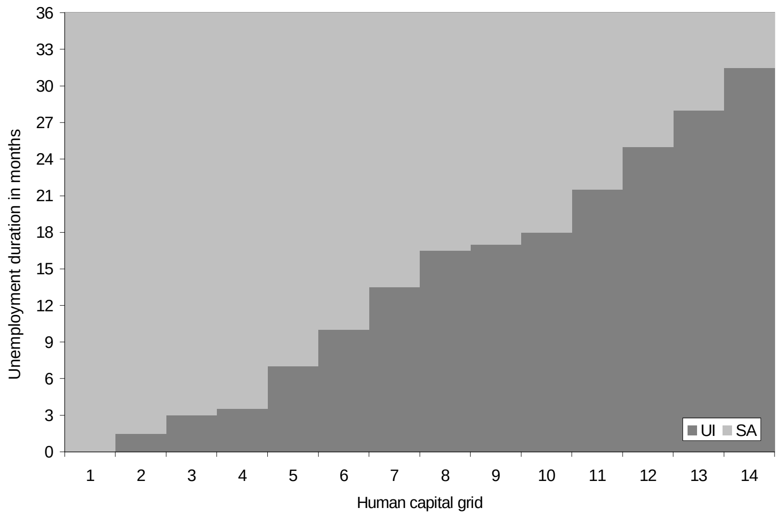
4.4 Results

4.4.1 Optimal policy and benefit scheme for West Germany

Figure 1 shows the sequence of policies that would have been optimal for West Germany in the period 2000-2002 in the absence of JM given that the same level of initial utility $U_0(h, p)$ is provided to a worker with characteristics (h, p) as under the actual policy. As already suggested by the ineffectiveness of JA and TR, these policies are not part of the optimal scheme.

If JM is unavailable, the duration for which UI is optimally used is completely determined by human capital because human capital depreciation is the dominant factor in this case. At the lowest level of h, the returns to search are too low relative to promised utility, so that it is not worthwhile to provide search incentives at all. However, the higher the initial level of h the higher the returns to search, and the longer it is optimal to incentivise the worker. Note that not providing search incentives for the lowest-skilled is a feature of many existing policies since, often, the level of benefits for this group of people is so high that their reservation wages lie above their expected wages.

Figure 1: Optimal policy for West Germany



Note: The human capital grids are geometrically spaced at rate 10% with the lowest grid absorbing all wages below the second lowest grid. Grip points of 14 and above all lie in the top 5%-quantile of the human capital distribution.

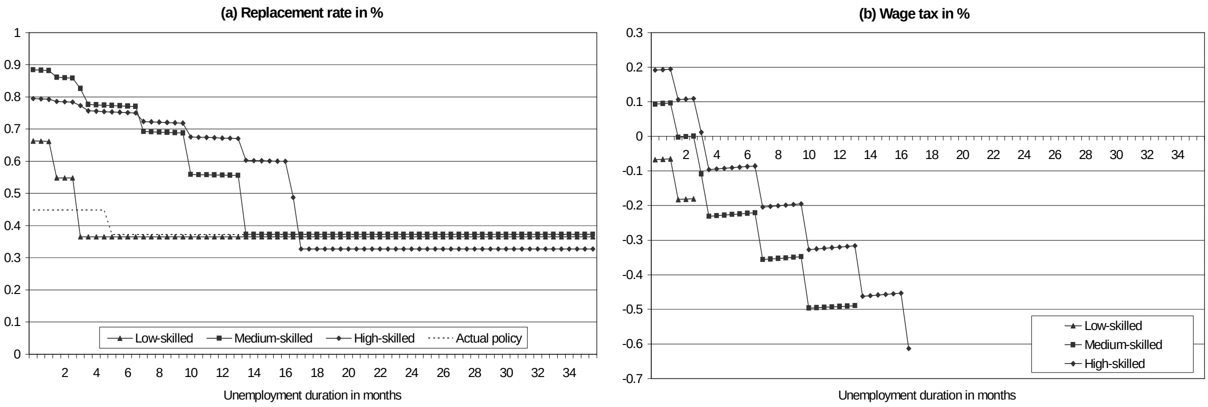
There are several ways how the result of different optimal durations of UI, i.e. UI claims, for jobseekers with different levels of human capital could be interpreted. Clearly, in contrast to the amount of benefits, it is probably very difficult to implement a policy that discriminates claim durations according to previous wages. Under the assumption that wages increase with tenure and age, one way to (at least partly) implement such a scheme would be to differentiate claims by previous time of contribution to the unemployment insurance and age. This is already common practice in many countries, including Germany.

Another way of looking at this result would be to use it to choose the optimal policy scheme for different populations of unemployed according to their median or mean characteristics. That means that countries where unemployment is a problem mainly of the very low-skilled can choose shorter durations of UI, while countries where higher-skilled people are affected as well should allow for longer UI claims. This would imply that the difference, e.g. between the short UI claims in the U.S. or Australia and the longer claims in Germany, is actually in line with the characteristics of the optimal scheme.

For West Germany, the optimal duration of UI at median human capital would be 13 months. Interestingly, this is very close to the maximum unemployment benefit claim of 12 months that has been introduced with the last large reform of unemployment insurance in Germany for the majority of unemployed (aged below 55).

Figure 2 displays the optimal sequence of benefits in terms of the replacement rate (Panel (a)) as well as the wage tax/subsidy upon reemployment (Panel (b)) for high, medium and low-skilled workers. As derived from the model, benefits decrease during UI in order to keep search effort high, and they remain constant during SA where positive effort is no longer required. The optimal level of social assistance ranges between 350-750 EUR with about 600 EUR at the median values of human capital and search effectiveness. Interestingly, the latter approximately equals the sum of the so-called unemployment benefits II (345 EUR) and the supplementary housing benefits that have been introduced recently in Germany as the baseline welfare payment that can be received after unemployment insurance payments have been exhausted.

Figure 2: Optimal benefits and wage tax/subsidy



Note: The skill levels correspond to the 5th, 10th and 15th 5%-quantile of the human capital distribution. The replacement rate is the last (gross) wage before entering unemployment divided by the benefit. The wage tax is (gross) wage minus consumption upon reemployment in % of the (gross) wage.

The wage tax decreases during UI because wages decline faster than benefits. For low values of human capital, a wage subsidy is paid from the beginning of the unemployment spell to keep the value of working sufficiently high to make search attractive. For high values of human capital, the wage tax turns into a subsidy only after some time when the value of working has deteriorated due to human capital depreciation.

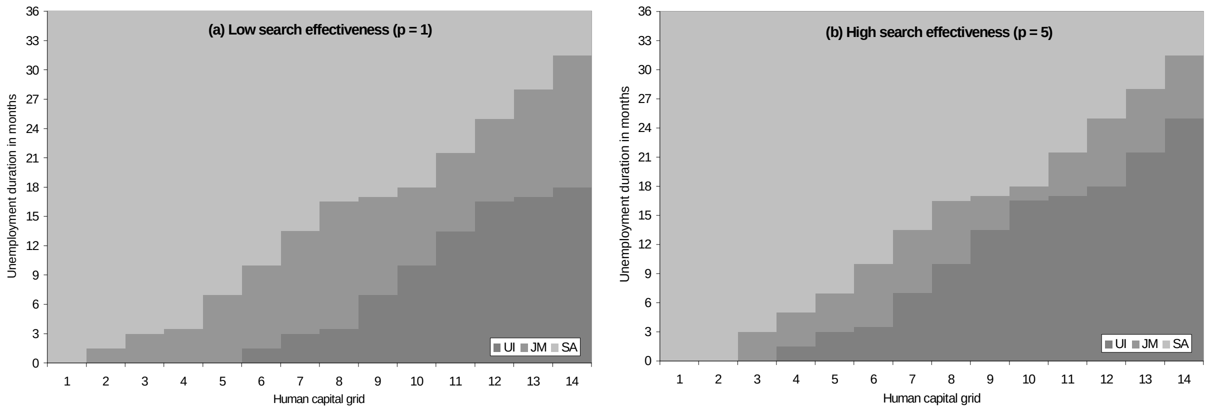
Panel (a) of Figure 2 also displays the effective replacement rates (mean benefits as a fraction of the last gross wage before entering unemployment) of unemployment benefits (UB) and assistance (UA) under the policy actually implemented in the period 2000-2002. It turns out that optimal benefits are considerably larger at the beginning of the unemployment spell. For low to medium levels of human capital, they decline to a level which is about the same as actual unemployment assistance (UA). However, for higher-skilled workers the long-run replacement rate is considerably lower.

The fact that optimal long-run benefits are often smaller than actual benefits and the ineffectiveness of JA and TR (i.e. the waste of their cost) lead to considerable government budget savings that can be realised by switching from the actual to the optimal policy. When providing the same initial utility to agents as under the actual policy, 32% of actual expenditures could be saved (20-40% by type of unemployed). For West Germany in the period 2000-2002 this would have implied savings in the expenditures on benefits, JA and TR of about 28 billion EUR.

4.4.2 Optimal use of job search monitoring

Figure 3 shows that if JM were available it would be used in the optimal scheme. The total duration for which it is optimal to incentivise agents remains constant because it is determined by human capital depreciation which is unaffected by JM. However, the relative use of JM and UI depends on the relation of the incentive costs during UI to the monitoring cost. The larger human capital and search effectiveness, the higher the exit rate to employment so that it is less costly to provide search incentives relative to incurring the fixed monitoring cost $\kappa^{JM} = 100$ EUR. Therefore, UI is used for longer periods while JM is used for shorter periods the higher h and p. Note that in the absence of p in the model it is not possible to fully determine the optimal timing and duration of JM, which provides an important role for p in the model even in the absence of job search assistance.

Figure 3: Optimal use of job search monitoring



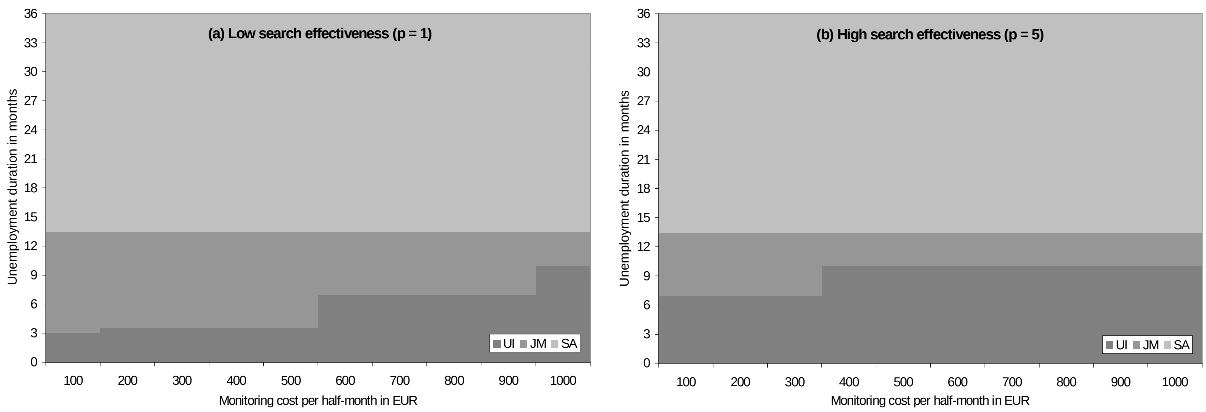
Note: Monitoring cost per half-month $\kappa^{JM} = 100$ EUR. The human capital grids are geometrically spaced at rate 10% with the lowest grid absorbing all wages below the second lowest grid. Grid points of 14 and above all lie in the top 5%-quantile of the human capital distribution. $p = 1, 5$ refers to the first and fifths 20%-quantile of the distribution of search effectiveness.

In case JM is used, benefits remain constant as predicted by the model. However, they are usually smaller than during UI because of the absence of the incentive costs. Instead, the monitoring cost has to be incurred. Overall, the budget savings from using JM as well are negligible. What happens is that about the same amount of money is spent in a different way.

It remains to be assessed how the use of the monitoring technology depends on the monitoring cost. In Figure 4, κ^{JM} is varied from 100-1000 EUR per half-month and the results are displayed for a worker with median skills and two levels of search effectiveness. Naturally, the use of JM is reduced with increasing cost. Yet, even at rather extreme values of κ^{JM} JM is still used. What changes is that JM is used later in the unemployment spell after increasing use of UI because the cost of providing search incentives decreases relative to the monitoring cost. As a result of increased use of UI, benefits decline to lower levels the larger κ^{JM}, which compensates for the higher fixed cost of monitoring.

To conclude, it would be optimal to use monitoring for almost all levels of human capital and even when monitoring costs are high. However, since the budget savings from using JM are negligible and the job finding rates are unchanged, there is no benefit from introducing it in West Germany. This conclusion is strengthened by the argument that monitoring can probably be only imperfect in reality and that its implementation would require hiring a large number of caseworkers.

Figure 4: Sensitivity of the use of UI and JM to the cost of JM (κ^{JM})



Note: Human capital is set to its median grid. Simulation for 3 years. $p = 1, 5$ refers to the first and fifths 20%-quantile of the distribution of search effectiveness.

4.4.3 Optimal use of job search assistance

The model allows to identify conditions under which JA and TR are used in the optimal policy scheme. For this purpose, the effects of the programmes are simulated as the probability that JA and TR rspectively increase search effectiveness and human capital by 10% within a half-month.[18] Note that this probability can be interpreted as the success rate of the programmes, θ^{JA} and θ^{TR}, when h and p change in 10%-steps.

In the first step, the choice between UI, JM, JA and SA is simulated. The success rate of JA for a 10% increase in p is varied between 10-100%, which corresponds to an accumulation rate of p, denoted by α^p, of 1-10% per half-month. Note that even if $\alpha^p = 10\%$ the impact on the exit rate to employment would range only between 2-10 %-points per half-month, and in the long run it would not exceed an overall gain of 20 %-points. Thus, the range of accumulation rates considered is reasonable.

The overall duration for which it is optimal to incentivise agents is, again, unchanged because it is determined by human capital depreciation which is unaffected by JA. As expected, JA is never used for agents with the highest level of search effectiveness ($p = 5$) because the returns to JA would be zero. It is also never used for the lowest-skilled or once h has depreciated to its lowest level ($h = 1$) because the value of working it too low to make search attractive in this case, even if the job finding rate could be increased considerably by use of JA.[19]

The simulation shows that the accumulation rate of p must exceed 3% per half-month for JA to be optimally used at all. This result is independent of the cost of JA because the incentive costs are the main determinant here. Once they are too high relative to the (expected) returns to JA, JA will not be used independent of its direct cost. If the accumulation rate of p is sufficiently large, then the following results are obtained. First, upon failure of JA it is always optimal to stay in the programme. Second, once p has depreciated after successful use of JA, it is always optimal to return to JA as long as $h > 1$. Moreover, the larger the depreciation rate of p, the earlier the worker should be re-assigned to JA. In both cases, the returns to JA do not fall, the direct and incentive costs are unchanged, and the effort compensation cost either falls with promised utility after failure of JA, or its increase after successful use of JA is only small and in most cases compensated by a fall during the policy that followed JA. Third, after use of JA, that policy is used which would be optimal in that period in the absence of JA given

[18] A 10% increase in h or p corresponds to an increase in the respective variable by one grid point in the simulation.

[19] In this and the following sections the values of p refer to the 20%-quantiles of the distribution of search effectiveness because the exit rate to employment only changes across these quantiles in the simulation. In contrast, the values of h refer to the grids used in the simulation.

characteristics (h, p). Since p increases after successful use of JA this means that the use of UI is effectively prolonged and that of JM is shorted.

When $\alpha^p \geq 8\%$, the returns to JA are so high and the incentive costs so low that it is always optimal to assign everyone with $p < 5$ and $h > 1$ immediately to JA when entering unemployment as long as initial promised utility is not too high. Since the latter increases with h, JA should be used for high-skilled workers only later in the unemployment spell after UI when U and h have decreased sufficiently. In particular, JA will not be used at the beginning of unemployment if h is high even if $\alpha^p = 10\%$ per half-month. Once assigned to JA it is optimal to stay in the programme until the highest level of p is reached when α^p is high.

Note that the lower α^p, the higher the incentive costs during JA and the lower the returns to JA. Moreover, after success of JA promised utility and, therefore, the benefit and effort compensation costs increase whereas they fall during UI and remain constant during JM. This is the reason why, for $\alpha^p < 8\%$ it is no longer optimal to always stay in JA until the maximum p is reached. Furthermore, for agents with $p = 4$ and $\alpha^p = 7\%$, or $p = 3$ and $\alpha^p = 5.5\%$, or $p = 2$ and $\alpha^p = 4\%$ JA should not be used until later in the unemployment spell after use of UI when promised utility has declined sufficiently and the incentive costs during UI have become too high relative to those during JA because of human capital depreciation. This is also the reason why it is sometimes optimal to return to JA even if p has not declined. When α^p is below the above values at the respective levels of p, JA will not be used at all because the incentive costs have become too high relative to the returns to JA.

For the case where the accumulation rate of p changes with p (or h), the same arguments apply as above because it makes no difference for the results whether the variation in θ^{JA} happens across individuals, within the unemployment spell or across otherwise equal economies.

The results on the optimal use of JA can be summarised as follows. The accumulation rate of p must exceed 3% in order for JA to be used at all. The higher search effectiveness, the higher the required minimum accumulation rate. When JA is part of the optimal scheme, it is usually optimal to use it at the beginning of the unemployment spell. If p depreciates after successful use of JA, jobseekers should be re-assigned to JA later in the unemployment spell. If the accumulation rate of p or the success rate of JA is either high or low, the optimal duration of JA is relatively short (0.5-2 months). For intermediate cases it can be longer (1-5 months).

With respect to optimal benefits, the predictions of the model apply. They should fall - as under UI - when JA fails and they should increase if it succeeds. In the simulation, the increases in benefits are small. They could be implemented as premia that are paid to jobseekers during or after participation in JA.

Now, what are the implications of the availability of an effective job-search-assistance technology for the government budget? If $\alpha^p > 3\%$ so that JA is part of the optimal policy, budget savings of 3-8% could be realised compared to the optimal policy without JA.

4.4.4 Optimal use of training

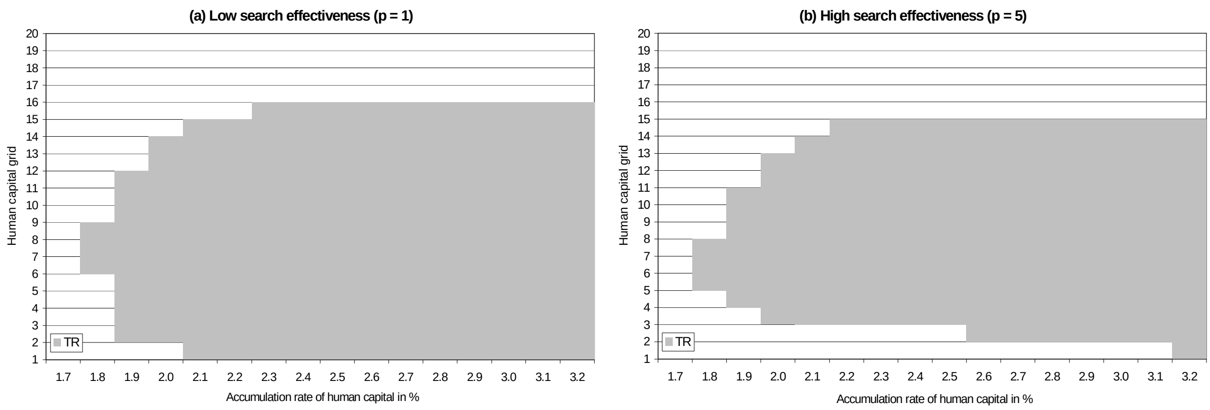
In contrast to JA, TR not only affects the exit rate to employment but also wages, which are the main determinant of the value of employment. The latter is very important because even if the exit rate is high, search incentives are low if the value of employment is low. Moreover, the effect of h on the job finding rate is relatively small, especially for larger values of h.

Since in the simulation increases in h directly translate into equivalent changes in wages, the range of reasonable values for the returns to TR are different from those for JA. According to Heckman, LaLonde, and Smith (1999), Martin and Grubb (2001) and Kluve and Schmidt (2002), the returns to training programmes in terms of wage increases usually do not exceed 10%, and in most cases such programmes have durations of more than a half-month, usually of 1-6 months but in some countries like Germany even of 12 months or more. Thus, accumulation rates of more than 3-4% per half-month seem very unlikely.

As h accumulates and depreciation of h is delayed during TR, the point from which SA is optimal is delayed as well, thus, increasing the period for which it is optimal to incentivise agents. Note that this justifies policies where participation prolongs benefit claims as implemented, for example, in several Scandinavian countries and Germany. Yet, under the optimal policy benefits should necessarily decrease during UI and after failure of TR, which is usually not the case for existing policies.

Figure 5 summarises the optimal use of TR in the absence of JA at different accumulation rates of human capital for the human capital grids used in the simulation. The grids are geometrically spaced at rate 10% with the lowest grid absorbing all levels of h below the second lowest grid. Because of the latter, the results for the lowest grid are not displayed, because an increase by one grid point corresponds to an increase in h by much more than 10% so that the return to TR would be much higher than for the other grids and the results would not be comparable.

Figure 5: Optimal use of training



Note: The human capital grids are geometrically spaced at rate 10% with the lowest grid absorbing all levels of h below the second lowest grid. $p = 1, 5$ refers to the first and fifths 20%-quantile of the distribution of search effectiveness.

It turns out that a wage accumulation rate of at least $\alpha^h = 1.8\%$ is necessary for TR to be used at all, which means that the effect of TR must be relatively large. Below that, the returns to TR are just too low given that both the incentive costs and direct cost during TR are high. As α^h increases the incentive costs during TR decrease while the returns to TR increase. Yet initially, TR will only be used for intermediate values of h. At high levels of h the incentive cost is much smaller during UI than during TR and zero during JM so that these policies are preferred if the returns to TR are not sufficiently high. In contrast, at low levels of h the fixed cost of TR is prohibitively high given the returns to TR relative to what would have to be paid in terms of the benefit and potential incentive and effort compensation cost during UI, JM and SA. Clearly, the highest-skilled will never be trained because the returns to TR are zero.

As α^h increases further, the range of h for which it is optimal to use TR increases. Note that the range increase faster the lower the fixed cost of TR, and also the lower p because the incentive costs during UI are higher so that the switch to TR occurs earlier. However, TR will not be used as long as promised utility U is too high. Since under the actual policy initial promised utility increases with h this implies that for workers who enter unemployment with high human capital, TR should not be used until later in the unemployment spell after UI when U and h have decreased sufficiently. Therefore, TR will not be used at the beginning of the unemployment spell if h is high even if $\alpha^h = 10\%$ per half-month.

Upon failure of TR, it is optimal to stay in the programme as long as h does not depreciate to a level at which it is no longer optimal to use TR. If TR is successful, agents should stay in the programme as long as promised utility U does not become prohibitively high. This implies that if h and α^h are either low or high, TR should be relatively short (0.5-4 months), while at intermediate values it can be rather long (more than 12 months possible) because of its strong impact on the value of employment which makes TR a dominant policy.

What happens if, in addition to training, there is an effective job-search-assistance technology available as well? For TR α^h is set to 2% while α^p is varied between 4-10% per half-month. In this case, the accumulation rate of p must exceed 6% for JA to be used at all because the effect of TR on the value of employment dominates the impact of JA on the job finding rate. Moreover, like TR, JA is never used at $h \geq 17$ because promised utility is too high. If $\alpha^p = 7\%$, JA is only used at low values of h where, in the absence of JA, it is not optimal to use TR ($h \leq 2$). As α^p increases, JA is used both at the lower and the upper end of $h = 1, ..., 16$ and for all values of this range when $\alpha^p = 10\%$. In the latter case and at the upper end of the human capital range JA is used at the beginning of the unemployment spell for 2-3 half-months, then followed by TR. As p depreciates after successful use of JA, it is optimal to re-assign the worker to JA whenever it falls to $p < 4$ if promised utility is relatively high and whenever it falls to $p < 5$ otherwise. In case of reassignment, the optimal duration of JA is usually 1-2 half-months. At the lower end of h when $\alpha^p < 10\%$ optimal durations are usually longer, ranging from 2-4 months. Of course, the larger α^h the more TR dominates JA because of its impact on the value of employment and the less JA is used. In particular, JA will no longer be used if $\alpha^h > 3.2\%$.

Unfortunately, the budget implications of TR cannot be assessed. Clearly, expenditures during unemployment are much higher because benefits increase with h and relatively high direct programme cost have to be incurred. However, since the model is simulated for the case where workers remain unemployed, the returns in terms of tax revenue and saved expenditures during unemployment once a jobseeker has found employment are not fully considered while the worst case maximum expenditures during unemployment are calculated.[20]

4.5 Sensitivity analysis

The sensitivity of the results with respect to key parameters of the model is tested. However, it has to be pointed out that the scope of the sensitivity analysis is limited by the numerical load of the simulation (approximation of the value functions), in particular when key parameters of the model are changed.

To simulate the case where there is uncertainty about the exit rate to employment or where it changes once the optimal policy is implemented, the exit rate is perturbated by 10% in both directions. It is found that a variation of the exit rate within this range does not change the results.

The sensitivity of the results with respect to the choice of the utility function of the worker is checked as well. Rather than log utility, $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ is chosen with intertemporal elasticity of substitution once above and once below one ($\sigma \in \{0.8, 1.2\}$).[21] Note that if σ is below (above) one, initial utility provided by the actual policy is higher (lower) than in the baseline case because the same level of consumption implies higher (lower) utility. Moreover, because of (13) benefits must decline more slowly (faster) upon failure of the activity when the incentive constraint binds under UI, JA or TR. However, it is found that the optimal sequence of policies remains unchanged. Because of the changed evolution of benefits the main changes are with respect to the implications for the government budget. For the baseline case without JM, JA and TR the budget savings are reduced (increased) by 5 (3) %-points.

[20] In each period where UI or JM are used, the returns are taken into account with weight one minus the job finding rate. Under JA only the job finding rate is affected so the full return is taken into account. However, TR also affects the value of employment and that is only taken into account with a weight considerably smaller than one.

[21] The utility functions are normalised to match $u(1) = 0$ as under log utility.

Finally, the disutility of effort, e, is allowed to vary between work and search. The simulation is repeated for the case where only half of the time spent working is spent on search or programme participation during unemployment. This implies a stronger preference for leisure during unemployment and, hence, a higher disutility of effort of 0.93 rather than 0.84 in the baseline case. Higher e implies higher incentive costs during UI, JA and TR but also makes work relatively more attractive. In the simulation these effects seem to offset each other because all results remain unchanged.

5 Conclusion

Labour market policies for the unemployed combine passive income support with active measures that aim at improving jobseekers' employment prospects. This paper extends the theoretical framework developed by Pavoni and Violante (2005a) for the optimal choice between different active and passive policies for the unemployed to a setting which allows for policies that affect the exit rate to employment but not productivity in the job. The most prominent example of such a policy are job search assistance programmes which aim at increasing search effectiveness and are one of the most widely used activation measures in OECD countries (OECD, 2007). The advantages of these programmes are that they are relatively short and, hence, inexpensive and that they have proven to be quite effective in many countries (e.g. Fay, 1996; Heckman, LaLonde, and Smith, 1999; Martin and Grubb, 2001; Kluve and Schmidt, 2002). Therefore, it is important to take these programmes into account when considering the optimal design of labour market policies.

The economy considered by Pavoni and Violante (2005a) is characterised by workers who differ with respect to human capital which depreciates during unemployment but can be raised by training. With respect to policy instruments the insurer can choose between unemployment benefits, social assistance, job search monitoring, training and wage taxes/subsidies. This setting is enriched by introducing search effectiveness as additional source of worker heterogeneity and allowing for it to be raised by a job-search-assistance technology. Thereby, the model allows to optimise over the most important instruments of active and passive labour market policies in all critical respects, namely the allocation among heterogeneous workers, the amount of benefits, and the duration and timing within the unemployment spell. Thus, it can be used to answer a wide range of interesting policy questions.

The theoretical results that can be derived from the enriched model confirm that the presence of the information asymmetry with respect to the worker's effort requires benefits to increase upon success of his activity and to fall upon failure in order to incentivise the worker, which underlines the generality of this result first derived by Shavell and Weiss (1979). It is shown that social assistance remains an optimal policy of last resort in the enriched setting. Furthermore, it is confirmed that it is never optimal to switch from monitored to unmonitored search. However, human capital dynamics are no longer necessary for policy transition to occur because of the dynamics implied by search effectiveness as additional state variable.

In order to derive more detailed insights on the optimal sequence of policies the model is calibrated to the West German economy in the period 2000-2002. Assessing the optimality of the West German policy in this period it turns out that about 32% or 28 billion EUR of government expenditure could have been saved by switching to the optimal scheme. It is shown that for low levels of human capital wage subsidies should be used which, because of human capital depreciation, implies that their use becomes more likely the longer unemployment duration. Moreover, when comparing the features of the optimal scheme with the recent reforms of the German system the optimal duration of an unemployment insurance claim and the optimal level of social assistance for a worker with median characteristics are strikingly close to what has been introduced recently in Germany.

The simulation is also used to assess whether a currently unavailable tight monitoring technology would be beneficial and worth introducing in West Germany. It is found that it would be optimal to use monitoring for almost all levels of human capital and even when monitoring costs are high. Moreover, the optimal timing and duration of monitoring within the unemployment spell depends on the worker's search effectiveness. Thus, there is an important role for search effectiveness in the model even in the absence of a job-search-assistance technology. However, the net gains from monitoring are too small to justify its introduction.

Another salient feature of the simulation is that by varying the parameters of the job-search-assistance and training technologies it is possible to obtain more general insights on their optimal use. It is found that existing ALMPs already share some but not all features of the optimal scheme. In line with existing ALMPs it is shown that participation prolongs the period for which unemployment insurance payments rather than social assistance are optimal. Moreover, for job search assistance (JA) it is usually optimal to use it at the beginning of unemployment as well as after depreciation of search effectiveness following the successful use of JA later in the unemployment spell. Furthermore, the optimal duration of JA is usually short. However, it is shown that the accumulation rate of search effectiveness must exceed 3% per half-month in order for JA to be used at all.

For training (TR) it turns out that a wage accumulation rate of at least 1.8% per half-month is necessary for TR to be used at all, which means that the effect of TR must be relatively large. Then, it is optimal to use TR for a range of intermediate values of human capital, which increases the larger the effect of TR. The optimal duration of TR can be long for intermediate accumulation rates while it should be relatively short otherwise. Because TR affects both the job finding rate and wages, it usually dominates JA. However, if JA is sufficiently effective, it is optimal to use TR only after a short period of JA at the beginning of the unemployment spell.

References

Abadie, A., and G. Imbens (2006): “Large Sample Properties of Matching Estimators for Average Treatment Effects,” *Econometrica*, 74, 235–267.

BA (2001-2005): “Arbeitsstatistik - Jahreszahlen 2000, 2001, 2002, 2003, 2004,” *Amtliche Nachrichten der Bundesagentur für Arbeit*, 49-53.

BMAS (2003): “Statistisches Taschenbuch 2003: Arbeits- und Sozialstatistik,” Bundesministerium für Arbeit und Soziales.

Chari, V., L. Christiano, and P. Kenhoe (1995): “Policy Analysis in Business Cycle Models,” in *Frontiers of Business Cycle Research*, ed. by T. Cooley, pp. 357–391, Princeton, New Jersey. Princeton University Press.

Dehejia, R. (2005): “Program Evaluation as a Decision Problem,” *Journal of Econometrics*, 125, 141–173.

Dehejia, R., and S. Wahba (2002): “Propensity-Score-Matching Methods for Nonexperimental Causal Studies,” *Review of Economics and Statistics*, 84, 151–161.

Fay, R. (1996): “Enhancing the Effectiveness of Active Labour Market Policies: Evidence from Program Evaluations in OECD Countries,” Labour Market and Social Policy Occasional Papers, 18, OECD.

Flores-Lagunes, A., A. Gonzalez, and T. C. Neumann (2007): “Estimating the Effects of Length of Exposure to a Training Program: The Case of Job Corps,” IZA Discussion Paper No. 2846.

Frölich, M. (2007): “Statistical Treatment Choice: An Application to Active Labour Market Programs,” *Journal of the American Statistical Association*, forthcoming.

Frölich, M., M. Lechner, and H. Steiger (2003): “Statistically Assisted Program Selection - International Experiences and Potential Benefits for Switzerland,” *Swiss Journal of Economics and Statistics*, 139(3), 311–331.

Gerfin, M., and M. Lechner (2002): “Microeconometric Evaluation of the Active Labour Market Policy in Switzerland,” *Economic Journal*, 112(482), 854–893.

Heckman, J. (1979): “Sample Selection Bias as a Specification Error,” *Econometrica*, 47(1), 153–162.

Heckman, J., R. LaLonde, and J. Smith (1999): “The Economics and Econometrics of Active Labor Market Programs,” in *Handbook of Labour Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3, pp. 1865–2097, Amsterdam. North-Holland.

Hirano, K., and G. Imbens (2004): “The Propensity Score with Continuous Treatments,” in *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, ed. by A. Gelman, and X.-L. Meng, pp. 73–84, West Sussex. John Wiley and Sons.

Hirano, K., and J. Porter (2006): “Asymptotics for Statistical Treatment Rules,” MPRA Paper No. 1173.

Hopenhayn, H., and J. Nicolini (1997): “Optimal Unemployment Insurance,” *Journal of Political Economy*, 105(2), 412–438.

——— (2005): “Optimal Unemployment Insurance and Employment History,” mimeo, Universidad Torcuato Di Tella.

Jespersen, S., J. Munch, and L. Skipper (2004): “Costs and Benefits of Danish Active Labour Market Programs,” IDanish Economic Council, Working Paper 2004:1.

Joffe, M., T. Ten Have, H. Feldman, and S. Kimmel (2004): “Model Selection, Confounder Control, and Marginal Structural Models,” *The American Statistician*, 58(4), 272–279.

Kluve, J., and C. Schmidt (2002): “Can Training and Employment Subsidies Combat European Unemployment?,” *Economic Policy*, 35, 411–448.

Lechner, M. (2004): “Sequential Matching Estimation of Dynamic Causal Models,” Discussion Paper 2004-06, Department of Economics, University of St. Gallen, revised 2005.

——— (2006): "Matching Estimation of Dynamic Treatment Models: Some Practical Issues," Discussion Paper 2006-03, Department of Economics, University of St. Gallen.

Lechner, M., and R. Miquel (2001): "A Potential Outcome Approach to Dynamic Program Evaluation - Part I: Identification," Discussion Paper 2001-07, Department of Economics, University of St. Gallen.

——— (2005): "Identification of the Effects of Dynamic Treatments by Sequential Conditional Idependence Assumptions," Discussion Paper 2005-17, Department of Economics, University of St. Gallen.

Lechner, M., R. Miquel, and C. Wunsch (2006): "Long-run Effects of Public Sector Sponsored Training in West Germany," Revised version of Discussion Paper 2004-19, Department of Economics, University of St. Gallen.

Lechner, M., and J. Smith (2007): "What Is the Value Added by Caseworkers?," *Labour Economics*, 14(2), 135–151.

Lechner, M., and S. Wiehler (2007): "Does the Timing of Active Labour Market Programmes Matter?," University of St. Gallen, mimeo.

Lechner, M., and C. Wunsch (2006): "Are Training Programs More Effective When Unemployment is High?," Discussion Paper 2006-23, Department of Economics, University of St. Gallen.

Machin, S., and A. Manning (1999): "The Causes and Consequences of Longterm Unemployment in Europe," in *Handbook of Labour Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3, pp. 3085–3139, Amsterdam. North-Holland.

Manski, C. (2000): "Identification Problems and Decisions under Ambiguity: Empirical Analysis of Treatment Response and Normative Analysis of Treatment Choice," *Journal of Econometrics*, 95, 415–442.

——— (2001): "Designing Programs for Heterogeneous Populations: The Value of Covariate Information," *American Economic Review*, 91(2), 103–106, Papers and Proceedings of the 113th Annual Meeting of the American Economic Association.

——— (2002): "Treatment Choice under Ambiguity Induced by Inferential Problems," *Journal of Statistical Planning and Inference*, 105, 67–82.

——— (2004): "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 72(4), 1221–1247.

——— (2005): "Minimax-Regret Treatment Choice with Missing Outcome Data," *Journal of Econometrics*, forthcoming.

Martin, J., and D. Grubb (2001): "What Works and for whom: A Review of OECD Countries' Experiences with Active Labour Market Policies," *Swedish Economic Policy Review*, 8, 9–56.

Mortensen, D. (1970): "Job Search, the Duration of Unemploymend and the Phillips Curve," *American Economic Review*, 60(5), 847–862.

OECD (2007): "OECD Employment Outlook 2007," Paris.

Pavoni, N. (2003): "Optimal Unemployment Insurance, with Human Capital Depreciation and Duration Dependence," University College London Discussion Paper No. 03/08.

Pavoni, N., and G. Violante (2005a): "Optimal Welfare-to-Work Programs," Federal Reserve Bank of Minneapolis Discussion Paper 143.

——— (2005b): "Optimal Welfare-to-Work Programs: Technical Appendix," mimeo.

——— (2006): "Optimal Welfare-to-Work Programs," *Review of Economic Studies*, forthcoming.

Phelan, C., and E. Stacchetti (2001): "Sequential Equilibria in a Ramsey Tax Model," *Econometrica*, 69, 1491–1518.

Phelan, C., and R. Townsend (1991): "Computing Multi-Period, Information-Constrained Optima," *Review of Economic Studies*, 58, 853–881.

Rubin, D. (1979): "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies," *Journal of the American Statistical Association*, 74(366), 318–328.

Shavell, S., and L. Weiss (1979): "The Optimal Payment of Unemployment Insurance Benefits over Time," *Journal of Political Economy*, 87(6), 1347–1362.

Sianesi, B. (2004): "An Evaluation of the Swedish System of Active Labour Market Programs in the 1990s," *Review of Economics and Statistics*, 86(1), 133–155.

——— (2007): "Differential effects of Swedish active labour market programs for unemployed adults during the 1990s," Revised version of IFS Working Paper W01/25.

Statistisches Bundesamt (2000-2002): "Einnahmen und Ausgaben privater Haushalte," *Wirtschaftsrechnungen*, Fachserie 15, Reihe 1.

van Ours, J. (2004): "The Locking-in Effect of Subsidized Jobs," *Journal of Comparative Economics*, 32(1), 37–52.

Wunsch, C. (2006): "Labour Market Policy in Germany: Institutions, Instruments and Reforms since Unification," Revised version of Discussion Paper 2005-06, Department of Economics, University of St.Gallen.

——— (2007): "Optimal Use of Labour Market Policies: Technical Appendix," University of St. Gallen, mimeo.

Wunsch, C., and M. Lechner (2007): "What Did All the Money Do? On the General Ineffectiveness of Recent West German Labour Market Programs," Discussion Paper 2007-19, Department of Economics, University of St. Gallen.

Zhao, R. (2000): "The Optimal Unemployment Insurance Contract: Why a Replacement Ratio?," mimeo, University of Chicago.

Appendix

A Proofs

Some of the proofs are only sketched. In these cases, the detailed proofs are relegated to the technical appendix (Wunsch, 2007).

Proposition 0: *(i) The functions W, $\boldsymbol{V}$ and V^i are concave in U (for all h and p). (ii) They are jointly continuous in (U, h, p) and monotonically increasing in h and p. (iii) Let $i^*(U, h, p)$ be the implemented policy and $c^*(U, h, p) > 0$ be the optimal payment at (U, h, p). Then both $\boldsymbol{V}$ and V^i are decreasing and continuously differentiable with respect to the first argument at (U, h, p) with*

$$\mathbf{V}_U(U, h, p) = V_U^{i^*(U,h,p)}(U, h, p) = -\frac{1}{u'(c^*(U, h, p))} \tag{15}$$

Proof of Proposition 0: See the technical appendix (Wunsch, 2007).

Proof of Proposition 1:

(i) Given the first-order conditions and the envelope condition for JM

$$-V_U^{JM}(U, h, p) = \frac{1}{u'(c_{JM})} \tag{16}$$

$$-\mathbf{V}_U(U^y, h^f, p^f) = \frac{1}{u'(c_{JM})} = \frac{1}{u'(c^y)}, \qquad y = f, s, \tag{17}$$

where $\mathbf{V}_U(U^s, h^f, p^f) = W_U(U^s, h^f, p^f)$, strict concavity of u implies that benefits and the net wage c^s are constant. For SA, the first-order conditions and the envelope condition yield

$$-V_U^{SA}(U, h, p) = \frac{1}{u'(c_{SA})} \tag{18}$$

$$-\mathbf{V}_U(U^f, h^f, p^f) = \frac{1}{u'(c_{SA})} = \frac{1}{u'(c^f)}. \tag{19}$$

Then, the strict concavity of u immediately implies the desired result. **Q.E.D.**

(ii) The relevant first-order conditions for UI yield

$$-V_U^{UI}(U, h, p) = \frac{1}{u'(c_{UI})} \tag{20}$$

$$-\mathbf{V}_U(U^f, h^f, p^f) = \frac{1}{u'(c_{UI})} - \mu^{UI} \frac{\pi(h, p)}{1 - \pi(h, p)} \tag{21}$$

$$-W_U(U^s, h^f, p^f) = \frac{1}{u'(c_{UI})} + \mu^{UI}, \tag{22}$$

where $\mu^{UI} \geq 0$ is the multiplier on the incentive compatibility constraint in (9). The result follows immediately from next period's envelope condition

$$-\mathbf{V}_U(U^f, h^f, p^f) = \frac{1}{u'(c^f)}, \tag{23}$$

where c^f is consumption if job search failed during UI, and the strict concavity of u.

Correspondingly, the relevant first-order conditions for programme $i \in \{JA, TR\}$ yield

$$-V_U^i(U,h,p) = \frac{1}{u'(c_i)} \quad (24)$$

$$-\mathbf{V}_U(U^f,h^f,p^f) = \frac{1}{u'(c_i)} - \mu^i \frac{\theta^i}{1-\theta^i} \quad (25)$$

$$-\mathbf{V}_U(U^s,h^i,p^i) = \frac{1}{u'(c_i)} + \mu^i, \quad (26)$$

where $\mu^i \geq 0$ is the multiplier on the respective incentive compatibility constraint and $(h^{JA}, p^{JA}) = (h^f, p^s)$, $(h^{TR}, p^{TR}) = (h^s, p^f)$. The desired result then follows immediately from next period's envelope condition

$$JA: \quad -\mathbf{V}_U(U^y,h^f,p^y) = \frac{1}{u'(c^y)}, \quad y = f, s \quad (27)$$

$$TR: \quad -\mathbf{V}_U(U^y,h^y,p^f) = \frac{1}{u'(c^y)}, \quad y = f, s, \quad (28)$$

and the strict concavity of u. **Q.E.D.**

(iii) The wage tax is defined as $\tau(h^f) = w(h^f) - c^s$. Since during JM $c_t = c_t^s = c_t^f$ and $h \geq h^f$ due to human capital depreciation the result is immediate. **Q.E.D.**

Proof of Proposition 2: (i) The proof follows the same line of reasoning as Pavoni and Violante (2005a,b) showing that in order for a policy to be strictly optimal at node (U_t, h_t, p_t), it cannot be that at such a node SA is implemented with positive probability and, in period $t+1$, a different policy is implemented after SA with positive probability. See the technical appendix (Wunsch, 2007) for all details of the proof.

(ii) Lemma A1 and A2 establish a ranking of the slopes of the value functions with respect to the three state variables (U, h, p) under JA, JM, UI and SA in the general case and show that this ranking does not depend on whether human capital dynamics are present or not. In particular, they show that V^{JM} declines more slowly with increasing U and declining p than V^{UI}. Lemma A3 then describes the dynamics of U under JM

Lemma A1: *Let $\nu(U,h,p)$ be the real number that solves $\mathbf{V}_U(U,h,p) = -q'\big((1-\beta)(U + \nu(U,h,p))\big)$, where $q \equiv u^{-1}$ is the inverse utility function and assume that $\nu(U,h,p)$ is non-increasing in U for every (h,p). Then value functions V^i, $i \in \{JM, UI\}$, with respect to U satisfy*

$$V_U^{UI}(U,h,p) \leq V_U^{JM}(U,h,p) \quad (29)$$

for all triples (U,h,p).

Proof of Lemma A1: See the technical appendix (Wunsch, 2007).

Lemma A2: *If $\mathbf{V}$ is submodular and both $\mathbf{V}$ and the V^i, $i \in \{SA, JM, UI\}$, are differentiable in h and p, then their slopes with respect to h and p satisfy, respectively,*

$$V_j^{UI}(U,h,p) \geq V_j^{JM}(U,h,p) \geq \hat{V}_j^{SA}(U) = 0 \quad (30)$$

for $j \in \{h,p\}$ and all triples (U,h,p).

Proof of Lemma A2: See the technical appendix (Wunsch, 2007).

Lemma A3: *Recall that $\mathbf{V}$ is concave and assume that it is submodular. (i) If the imple-*

mented policy at (U,h,p) is JM and either concavity or submodularity is strict, then $U^f(x) \geq U$ almost surely for all $x \in [0,1]$.

Proof of Lemma A3: From the first-order conditions for JM we have that

$$\mathbf{V}_U(U,h,p) = V_U^{JM}(U,h,p) = \mathbf{V}_U(U^f,h^f,p^f) = \mathbf{V}_U(U^f(x),h^f,p^f) \tag{31}$$

for (almost) all $x \in [0,1]$. Since $p^f \leq p$ and $h^f \leq h$, submodularity implies $\mathbf{V}_U(U,h^f,p^f) \geq \mathbf{V}_U(U,h,p)$. Then the desired result follows when either concavity or submodularity is strict at (U,h,p). **Q.E.D.**

Since $p^f \leq p$ and U weakly increases during JM, Lemma A1 and A2 establish that V^{UI} declines faster than V^{JM} so that a switch to UI will never occur after optimal use of JM. **Q.E.D.**

(iii) A necessary condition for JA to be used at all in period t is that search effectiveness increases in expectation, i.e. that $E(p_{t+1}) = \theta^{JA}p^s_{t+1} + (1-\theta^{JA})p^f_{t+1} > p_t$. Otherwise, it will never be worthwhile to pay the cost κ^{JA}. Now note that $E(p_{t+1}) = \theta^{JA}f(p_t) + (1-\delta^p)p_t$ so that the above condition can be rewritten as $f(p_t) > \frac{\delta^p}{\theta^{JA}}p_t$. Thus, this condition continues to hold after use of JA if $f'(p) \geq \frac{\delta^p}{\theta^{JA}}$ for all p. The right-hand side is constant while the left-hand side is strictly decreasing in p when f is strictly concave. This implies that upon failure of JA, $f'(p^f) \geq \frac{\delta^p}{\theta^{JA}}$ will always hold because $p^f \leq p$ and JA has been used at p. However, after continuous success of JA the condition will fail to hold at some point when f is strictly concave because $p^s > p$. **Q.E.D.**

(iv) The proof follows the same line of reasoning as the one of Proposition 2 (iii). A necessary condition for TR to be used at all in period t is that human capital increases in expectation, i.e. that $E(h_{t+1}) = \theta^{TR}h^s_{t+1} + (1-\theta^{TR})h^f_{t+1} > h_t$. Otherwise, it will never be worthwhile to pay the cost κ^{TR}. Now note that $E(h_{t+1}) = \theta^{TR}g(h_t) + (1-\delta^h)h_t$ so that the above condition can be rewritten as $g(h_t) > \frac{\delta^h}{\theta^{TR}}h_t$. Thus, this condition continues to hold after use of TR if $g'(h) \geq \frac{\delta^h}{\theta^{TR}}$ for all h. The right-hand side is constant while the left-hand side is strictly decreasing in h when g is strictly concave. This implies that upon failure of TR, $g'(h^f) \geq \frac{\delta^h}{\theta^{TR}}$ will always hold because $h^f \leq h$ and TR has been used at h. However, after continuous success of TR the condition will fail to hold at some point when g is strictly concave because $h^s > h$. **Q.E.D.**

Proof of Proposition 3: (i) If p remains constant during JM, it is sufficient to consider the dynamics with respect to U. The first-order and envelope conditions for JM yield $\mathbf{V}_U(U,p) = V_U^{JM}(U,p) = \mathbf{V}_U(U^f,p)$ implying that setting $U = U^f$ is optimal. Thus, after use of JM all state variables have the same values as one period before. Consequently, implementing JM every period is always optimal, and if $\mathbf{V}$ is strictly concave this optimal policy is unique.

(ii) From Proposition 3 (i) and 2 (i) we know that under the conditions stated in Proposition 3 (ii), whenever JM or SA have been optimal at (U,p) they are absorbing. Thus, JA will never be used thereafter. Now consider the case where UI is optimal at (U,p). If p remains constant during UI, it is sufficient to consider the dynamics with respect to U. Lemma A4 shows that the IC constraint is binding during UI in this case. Lemma A5 then establishes that whenever UI is optimal at (U,p), V^{JA} is more negatively sloped than V^{UI} if $\theta^{JA} \leq \pi(p)$ and, hence, JA will never be optimal in the next period.

Lemma A4: *At any (U,p) where UI is optimal and either $p = \hat{p}$ or $\delta^p = 0$, the incentive compatibility constraint binds with $\mu^{UI} > 0$ and we have $U > U^f$.*

Proof of Lemma A4: Concavity of $\mathbf{V}$ together with (20) and (21) imply that $U \geq U^f$. If the IC binds, this implies immediately that $U > U^f$. Now assume that $\mu^{UI} = 0$. The special form of W in the absence of human capital dynamics and (22) then imply that $U \geq (1-\beta)U^s + \beta U^f$.

However, if $U^f \geq U$, then $U^s > U$ from (13) and $e > 0$ and, hence, the IC constraint could never be satisfied. Thus, we must have $U > U^f$ whenever UI is optimal. Now note that the PK constraints of all potential future states are linear in U and that the corresponding payments can be written as $c = u^{-1}\big(u(c)\big)$. The PK constraints for all potential future states can be used to replace $u(c)$. Since p remains unchanged during UI if either $p = \hat{p}$ or $\delta^p = 0$, $u(c)$ and, hence, c are completely driven by the dynamics of U. It is then easy to show that $u^{-1}\big(u(c)\big)$ is strictly increasing and strictly convex in U for all potential future states (because u is strictly concave). Thus, because $U > U^f$ whenever UI is optimal, we have $c^{UI} > c^f$ and $\mu^{UI} > 0$ follows immediately from the strict concavity of u. **Q.E.D.**

Lemma A5: *For every (U, p) at which UI is optimal and $\theta^{JA} \leq \pi(p)$ we have that $V_U^{UI}(U, p) \geq V_U^{JA}(U, p)$.*

Proof of Lemma A5: From the first-order and envelope conditions for UI and JA we have, respectively,

$$V_U^{UI}(U,p) = -\frac{1}{u'(c_{UI})} \leq \mathbf{V}_U(U_{UI}^f, p^f) \tag{32}$$

$$V_U^{JA}(U,p) = \theta^{JA}\mathbf{V}_U(U_{JA}^s, p^s) + (1-\theta^{JA})\mathbf{V}_U(U_{JA}^f, p^f), \tag{33}$$

where c^{UI} and U_{UI}^f are optimal consumption and continuation utility under UI, and c^{JA}, U_{JA}^f and U_{JA}^s the respective values under JA. If $c^{UI} \leq c^{JA}$ the envelope condition and the concavity of u imply the desired result. So consider the case $c^{UI} > c^{JA}$. From Lemma A4 we know that the IC constraint binds when UI is optimal, i.e. $U = u(c^{UI}) + \beta U_{UI}^f$. Satisfying the IC constraint under JA, i.e. $U \geq u(c^{JA}) + \beta U_{JA}^f$, with the same U requires $U_{JA}^f > U_{UI}^f$ if $c^{UI} > c^{JA}$. Then the PK constraints imply that

$$U_{JA}^s - U_{UI}^s > \frac{e}{\beta\theta^{JA}} - \frac{e}{\beta\pi(p)} \geq 0 \tag{34}$$

if $\theta^{JA} \leq \pi(p)$. Then it follows from the concavity of $\mathbf{V}$ and the envelope conditions that

$$-\frac{1}{u'(c_{JA})} = \theta^{JA}\mathbf{V}_U(U_{JA}^s, p^s) + (1-\theta^{JA})\mathbf{V}_U(U_{JA}^f, p^f) \leq \mathbf{V}_U(U_{UI}^f, p^f) \leq -\frac{1}{u'(c_{UI})}. \tag{35}$$

Since u is strictly concave this implies $c^{UI} \leq c^{JA}$, a contradiction to the assumption that $c^{UI} > c^{JA}$. **Q.E.D.**

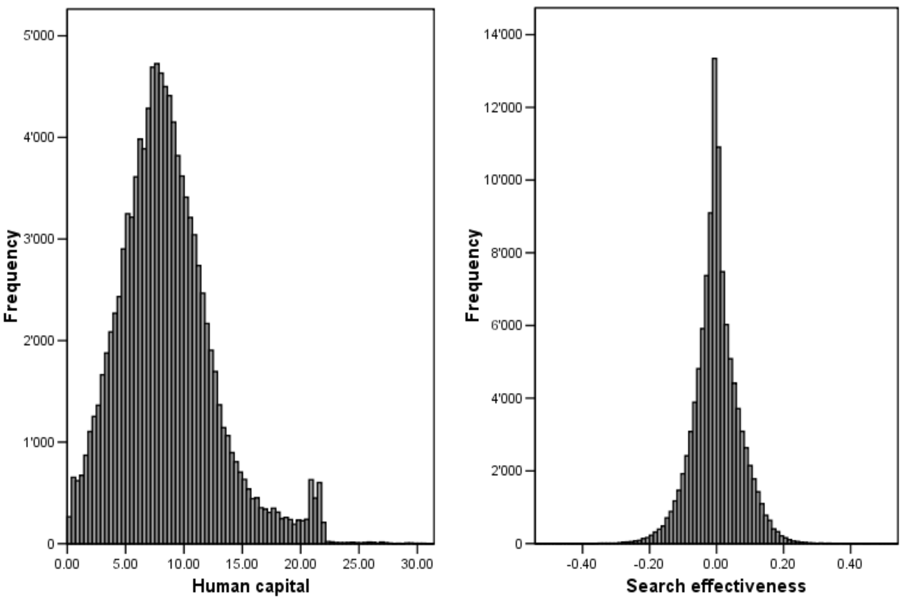
Since search effectiveness does not change during UI if either $p = \hat{p}$ or $\delta^p = 0$, while promised utility falls, Lemma A5 completes the proof of Proposition 3 (ii). **Q.E.D.**

(iii) Pavoni and Violante (2006) show that in the absence of JA and p as additional state variable UI will never be followed by SA or JM (Proposition 2 in Pavoni and Violante, 2006). From Proposition 3 (ii) we know that JA will never be used under the stated conditions and that p remains unchanged. Thus, the setting is identical to the one of Pavoni and Violante (2005b) with the conditioning on p kept implicit. Then their proof (Pavoni and Violante, 2005b) can be used to show the desired result. **Q.E.D.**

B Human capital and search effectiveness

Human capital is calibrated using the wage from the last employment spell before entering unemployment. Most individuals enter directly from employment (71%). For all other persons, the wage is depreciated according to the calibrated wage depreciation rate ($\delta^h = 0.985$) for the time between leaving employment and entering unemployment. Moreover, for parttime-workers the wage is scaled up to obtain an approximate fulltime-equivalent.

Figure 6: Distribution of human capital and search effectiveness in the sample



Search effectiveness, which is not directly observable in the data, is estimated by the contribution of variables related to search effectiveness to the probability of finding employment within 3 months after entering unemployment. Concretely, search effectiveness of individual j is estimated as

$$\begin{aligned} \hat{p}_j &= P(Y=1|X=x_j, Z=z_j) - P(Y=1|X=x_j) \\ &= P(Y=1|X=x_j, Z=z_j) - E_Z[P(Y=1|X=x_j, Z=z_j)] \end{aligned} \tag{36}$$

where $Y = 1$ if the worker succeeded in finding employment within 3 months after entering unemployment, and where Z is a vector of variables that are related to search effectiveness while X contains all other variables that affect the job finding probability.

A probit model is estimated with X mainly including variables that are related to reservation wages: gender, durtion and amount of remaining benefit claim, educational attainment, qualification, occupational status, age, marital status, past earnings, variables summarising past employment histories that are not contained in Z. Because of the high dimensionality of X the index $x'\beta$ is used to integrate out Z.

The vector Z includes measurements of labour market attachment (number and total duration of times out of labour force, time since last employment), duration of past unemployment spells (mean duration, duration of last unemployment spell), compliance with benefit conditions (imposition of sanctions, attendance of interviews at PES, cooperation with PES staff), number of placement propositions by the PES in the past, labour market conditions (local labour market conditions, seasonality dummies, industry and occupation dummies), health, foreigner status (control for potential language problems), presence of (young) children, lone parent status, participation in JA in the past, realised wage increases and decreases in the past (control for ability to find better employment).

C Estimation of the hazard rates and the wage depreciation and accumulation rates

Nonparametric estimates of the baseline exit rates to employment without participation in JA or TR as well as the wage depreciation and accumulation rates can be obtained from nonparticipants who are matched to the reference population of interest using the radius matching estimator proposed and applied by Lechner and Wunsch (2006). The matching is performed because interest is in the values of these parameters for the reference population which also includes participants and, therefore, differs systematically from actual nonparticipants in characteristics that also affect the exit rate to employment and wages. Moreover, to estimate the wage depreciation and accumulation rates, selective subgroups of nonparticipants are used.

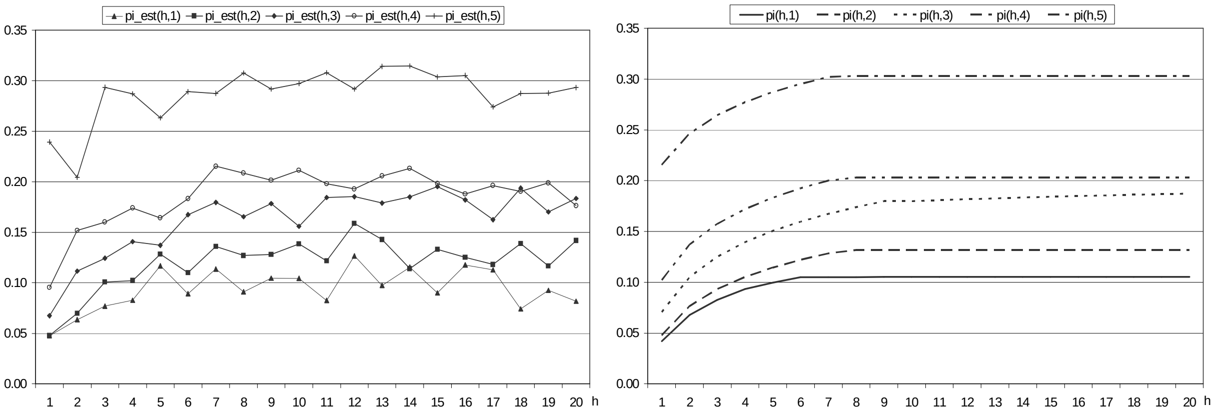
Lechner, Miquel, and Wunsch (2006) apply a matching procedure, that is based on the propensity score,[22] and has the following advantages. To allow for higher precision when many 'good' comparison observations are available, they incorporate the idea of calliper or radius matching (e.g. Dehejia and Wahba, 2002) into the standard algorithm used for example by Gerfin and Lechner (2002). Second, matching quality is increased by exploiting the fact that appropriate weighted regressions that use the sampling weights from matching have the so-called double robustness property. This property implies that the estimator remains consistent if either the matching step is based on a correctly specified selection model, or the regression model is correctly specified (e.g. Rubin, 1979; Joffe, Ten Have, Feldman, and Kimmel, 2004). Moreover, this procedure may reduce small sample bias as well as asymptotic bias of matching estimators (see Abadie and Imbens, 2006) and thus increase robustness of the estimator. For more information on this estimator and its performance see Lechner, Miquel, and Wunsch (2006).

The exit rates to employment are approximated by the probability to find employment within 2 months after entering unemployment. Before that, exit rates are extremely small due to labour market frictions. In order not to pick up these frictions, the 2-month probability is chosen rather than the immediate exit rate. They are estimated in 100 subgroups defined by 20 equidistant quantiles of the human capital distribution and 5 equidistant quantiles of the distribution of search effectiveness. The number of nonparticipants available in the subgroups varies between 155 and 1670 observations with the vast majority of cells having more than 700 observations. The results are displayed in the left panel of Figure 7. For the simulation a smooth function which is fitted to these estimates is used (see the right panel of Figure 7).

The wage depreciation rate is calculated from the weighted mean wage of nonparticipants who find employment within 3 month after entering unemployment and the weighted mean wage of nonparticipants who find employment only within 12-14 months. In turn, the wage accumulation rate is calculated from the weighted mean wage nonparticipants, who find employment, receive early (months 1-3) and later (months 12-14) in the employment spell.

[22] The propensity scores are estimated by binary probit models for each comparison of interest. The estimation equations have been subjected to extensive specification tests.

Figure 7: Estimated and calibrated exit rates to employment



Note: Left panel: probability to exit to employment within 2 months after entering unemployment. Right panel: fitted hazard functions using logarithmic functions.

D Identification and estimation of the effects of JA and TR of different durations

Here, the population for which the effects are to be estimated is an entry sample into unemployment that is defined independent of any programme participation later in the unemployment spell. Lechner and Miquel (2001, 2005) show that in this case, the effects of interest are identified under the so-called weak dynamic conditional independence assumption (W-DCIA). This assumption states, first, that conditional on confounding variables at $t = 0$, potential outcomes measured from time $\tau \geq 1$ onwards are independent of programme participation in period $t = 1$. This would account for any selectivity between programme participants and the population of interest. Second, conditional on participation status up to time t and confounding variables of all periods up to time t, potential outcomes are independent of participation in period $t + 1$. This would account for endogeneity of programme durations. The third part of the W-DCIA is a common-support requirement which demands overlap in the control variables between the populations involved in each of the selection steps.

To judge whether W-DCIA is plausible in this particular application, first, the confounding variables for programme participation relative to the reference population of interest have to be identified. This comparison is mainly driven by the difference between participants and nonparticipants in a particular programme. Second, those variables have to be detected which influence both outcomes and changes in treatment status, i.e. the decision between staying in the programme for another period and leaving.

Before discussing the determinants of selectivity, the outcome variables of interest have to be defined in order to identify the subset of confounding variables among these determinants. The choice of the outcome variables of interest is driven by the definition of search effectiveness and human capital in the calibration. For job search assistance the probability of finding employment within three months after completing a spell of JA is chosen. For training the impact on human capital is estimated by the effect of TR on average wages during employment. Besides the initial levels of search effectiveness and human capital, for both outcome variables the main driving factors are presumably gender, age, educational attainment, health, presence of (young) children, profession, industry, desired form of employment and local labour market conditions.

All these factors are captured by the data. Moreover, information on past employment histories and compliance with benefit conditions indirectly captures unobserved factors like ability and motivation.

Selection into programmes from the reference population is driven by programme eligibility, selection by caseworkers and self-selection by the unemployed. By construction of the sample, all unemployed are eligible because they receive unemployment insurance payments. Caseworkers select on the basis of an assessment of the employment prospects and the specific qualification needs of the unemployed. According to German legislation they, furthermore, have to take into account the chances of the unemployed for completing a specific programme successfully, and the situation in the local labour market. Thus, caseworker decisions are basically driven by the same factors as the outcome variables themselves. Similar arguments apply to self-selection by the unemployed because they also compare their employment prospects with and without a programme as well as the corresponding costs in terms of effort or potentially foregone benefits in case of refusal to participate. For training an additional and rather strong incentive to participate is the prolongation of total benefit claims. Since remaining benefit claims are observed in the data, this can be captured.

Decisions to leave or stay in the programme are driven by factors that change after entering a programme. The most important factors are probably the arrival of job offers, exhaustion of benefit claims and significant changes in health conditions. Other factors may be noncompliance with benefit conditions, changes in family status, moving to another place, or take up or loss of a minor employment,[23] and all of these as well as exhaustion of benefits are directly observed in the data. Changes in health conditions are observed if they are severe enough to affect unemployment insurance status. Arrival of job offers is not directly observed in the data, but the number of placement propositions by the PES per spell.[24] To approximate the arrival of job offers at or up to a specific point in time a Heckman (1979)-type selection model for the log number of placement propositions per day which accounts for zero propositions in the spell is estimated in the subsample of nonparticipants using a rich set of time-invariant (or deterministically changing) and time-varying variables measured before programme start as explanatory variables. The number of placement propositions at different points in time is then predicted for all individuals in the sample using the updated measurement of the time-varying covariates.

In summary, most of the potentially confounding factors are directly observed in the data. Moreover, those that are not directly observed either can be controlled for indirectly by information on past employment histories and compliance with benefit conditions, like e.g. ability or motivation, or can be approximated by use of observed variables, like arrival of job offers. Thus, the data are sufficiently rich to capture the main sources of selection bias at the different points in time before and during programme participation.

Lechner (2004, 2006) proposes a sequential nearest-neighbour matching estimator where the matching is based on propensity scores to estimate the effects of different sequences of programmes for a population defined within one of the sequences. Consider the case where interest is in estimating the effect of sequence $S^0 = (S_1^0, S_2^0)$ compared to $S^1 = (S_1^1, S_2^1)$ for the population defined by S_1^0. In the first step, the population defined by S_1^1 is matched to the population of interest S_1^0 based on the estimated propensity score of the corresponding selection equation. Second, the population defined by S_2^1 is matched to those observations in S_1^1 that served as

[23] To provide additional work incentives benefit recipients can earn additional labour income without losing their claim if they work less than 15 hours per week.

[24] A placement proposition is a job vacancy proposed to the jobseeker by the caseworker.

matches in the previous step based on the propensity scores from this and the first selection step. In a similar vein, the population defined by S_2^0 is matched to the one defined by S_1^0 based on the corresponding propensity score. To obtain an estimate of the effect of interest, the reweighed outcome of the population defined by S_2^0 is then subtracted from the reweighed outcome of the population defined by S_2^1.

Here, a modified version of this estimator is used. First, the effects are estimated for a population defined outside the sequences under consideration. This implies that an additional matching step has to be performed in each comparison: The populations defined by the first element of *each* sequence under consideration have to be matched to the population of interest for which the effects are to be estimated before matching within each of the sequences is performed. Second, radius matching as proposed by Lechner, Miquel, and Wunsch (2006) for static evaluation problems rather than nearest-neighbour matching is used to increase efficiency and potentially robustness given that the reference population of interest is potentially large compared to the populations defined by the last element of the sequences under consideration.[25]

Tables 2 and 3 display the estimated effects of JA and TR of different durations on, respectively, search effectiveness and human capital. For each programme, the effects are estimated separately for two subsamples of the reference population defined by persons with a value below/above the median of the respective variable of interest. The number of observation available for the estimation in each subgroup and programme type varies between 400 and 1200 observations (see Table 4). For each comparison, less than 1% of the reference population are deleted due a lack of common support.

Table 2: Estimated effects of JA of different durations

	Low search effectiveness	High search effectiveness
JA1-NP	-0.0028	-0.0468
JA2-JA1	-0.039	0.0181
JA3-JA2	-0.0259	-0.0332
JA4-JA3	0.0483	0.0268
JA5-JA4	0.0081	0.0496

Note: Difference in the probability to exit to employment within 3 months (%-points). NP: nonparticipation, JAn: completion of n half-months of JA. *Effect is significant on the 5% level.

Table 3: Estimated effects of TR of different durations

	TR1-NP		TR2-TR1	
	Wage^a	Zero wage^b	Wage^a	Zero wage^b
Low-skilled	-3.8294	0.0242	0.5960	-0.0182
High-skilled	-4.4626	0.0160	-2.5502	0.0409

Note: Difference in ^{a}the average daily gross wage if employed (EUR) and ^{b}the fraction of persons not employed (with zero average wage; %-points). NP: nonparticipation, TR1: completed up to 6 months of TR, TR2: completed more than 6 months of TR. *Effect is significant on the 5% level.

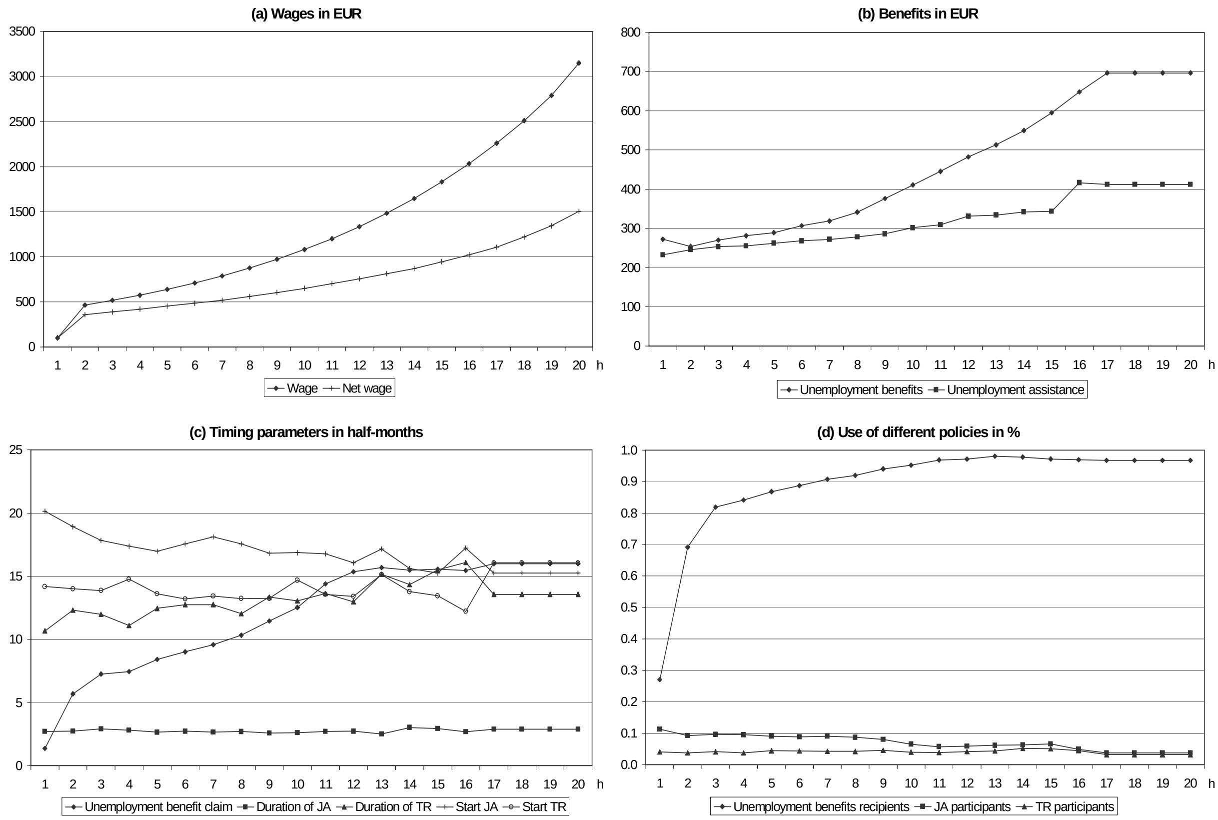
[25] If the comparison population is small compared to the population for which the effects are estimated, using more than one similar observation prevents that one particular observation gets too much weight.

Table 4: Numbers of observations available for the estimation

	Subsample	
	Low search effectiveness	High search effectiveness
1 half-month of JA	982	878
2 half-months of JA	668	620
3 half-months of JA	598	661
4 half-months of JA	504	509
More than 4 half-months of JA	373	386
	Low-skilled	High-skilled
Up to 6 months of TR	1171	1002
More than 6 months of TR	809	969

E Parameters of the actual policy in West Germany 2000-2002

Figure 8: Calibration of the actual policy parameters



Note: Averages calculated from the data for the reference population of interest. Time units are half-months. The net wage in panel (a) is calculated assuming German income tax class 1, tax rates as of 2002, public health insurance with rate 13.8%, no children and subject to statutory pension insurance as well as church tax. Start of JA and TR in panel (c): time to treatment within the unemployment spell.