Search and Work in Optimal Welfare Programs

Nicola Pavoni
European University Institute, Università Bocconi, and IFS

Ofer Setty
Tel Aviv University

Giovanni L. Violante
New York University, CEPR, and NBER

First Draft: June 2010 – This Draft: November 2010

Preliminary and incomplete: please, do not cite without permission

Abstract
Actual welfare programs often feature work requirements in exchange for benefits and job-search assistance. We study the optimal design of welfare programs for the unemployed when the principal/government can solicit either job-search effort or work effort from the agent or, alternatively, can relieve the agent from supplying effort and search on her behalf. Our main result is that when the generosity of the welfare program (i.e., its initial promised utility) is low, then the optimal program should be based on work activities. In contrast, when the initial generosity, or available budget, is high the optimal program should be based on search –and search assistance– activities.

JEL Classification: D82, H21, J24, J64, J65.

Keywords: Human Capital, Moral Hazard; Recursive Contracts, Welfare-to-Work.

1 Introduction

Government policies targeting the unemployed have the twofold objective of assisting these individuals while they are out of work and, at the same time, of leading them towards economic self-sufficiency through employment. Achieving both objectives is challenging because the provision of assistance interferes with individual incentives to find a suitable and lasting employment. In order to strike the right balance between assistance and incentives, governments use a wide range of policy instruments based on three mandatory activities for the individual: search, work and training. In this paper, we focus on the use of search-based and work-based activities.[1]

It is convenient to immediately define some concepts that we use throughout the paper. We define a policy as a government prescription of an activity (search, work, train or, simply, rest) to the individual with an associated cash transfer. A Welfare-to-Work (WTW) program is a government expenditure program that combines together several policy instruments. Clearly, every WTW program implicitly promises a certain level of ex-ante welfare to the unemployed agent. An *optimal* WTW program is a mix of policies that maximizes the expected discounted utility of the unemployed agent, subject to the government budget constraint.

In Pavoni and Violante (2007), we made a first step towards the development of a theoretical framework suitable to study WTW programs from a normative viewpoint. Our point of departure was the classic setup –originated largely from the seminal article of Shavell and Weiss (1979)– where the optimal unemployment insurance contract is studied in the presence of a repeated moral hazard problem: the risk-neutral principal (planner/government) cannot observe the risk-averse unemployed agent's effort (hidden action). Following the most recent contributions in the literature (Atkeson and Lucas, 1995; Wang and Williamson, 1996; Hopenhayn and Nicolini, 1997; Shimer and Werning, 2008; Pavoni, 2007; Hopenhayn and Nicolini, 2009; Pavoni 2009), we exploit the recursive representation of the planner's problem where the expected discounted utility promised by the contract to the unemployed agent becomes a state variable.

We enriched this standard environment by letting workers' wages and their job finding probabilities depend on *human capital* (skills) and allow human capital to depreciate along the unemployment spell. Human capital is our second key state variable in the recursive representation. The introduction of human capital depreciation in the problem permits a better representation of labor market data along two important dimensions. First, since wages depend on human capital, in our economy workers experience wage losses during unemployment, consistently with the findings of a vast set of empirical studies (for a survey, see Fallick, 1996). Second, since we let the job-finding probability depend on human capital, search effort becomes less effective as the unemployment spell progresses, inducing negative duration dependence in the unemployment hazard–a common feature of the data, as reported by Machin and Manning (1999) in their survey.[2]

[1] For an analysis of the use of training-based policies, see Pavoni and Violante (2005).

[2] In particular, several studies (e.g., Blank, 1989, for welfare recipients; Bover, Arellano and Bentolila, 2002, for UI

The innovation that allows us to use this enriched repeated moral hazard framework to study the optimal design of WTW programs is that we introduce additional "technologies" and associated worker "activities" (i.e. use of technologies and choice of effort level) besides search. In Pavoni and Violante (2007), we introduced a monitoring technology and investigated the extent to which search effort should be incentivized, monitored or stopped being enforced along the optimal program. Our general methodology has been, recently, followed in a number of other papers. Setty (2010) generalizes Pavoni and Violante (2007) by studying the optimal use of an imperfect monitoring technology of search effort during the unemployment spell. Wunsch (2009) analyzes the role of costly augmentation of job search skills. Pavoni and Violante (2005) and Spinnewijn (2010) study training programs that rebuild human capital.

Here, we continue this ongoing investigation by extending this framework to study the use of work activities in the optimal design of WTW programs. We view this extension as essential to this research agenda in light of the fact that under the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996 –which deeply reformed welfare– states require adults receiving benefits to be engaged in work-related activities after two years of receiving assistance.

There are a variety of work-site activities across U.S. states. Some of them simply represent an obligation for the individual to work in exchange for her welfare check. Others, instead, are meant to be a transition into private employment. For example, while the individual is mandated to work, the caseworker actively searches on her behalf for a suitable employment in a similar job. Or, the caseworker directly matches the unemployed to an employer with the expectation that she will be retained by that same employer. As we explain in some detail in the next section, this distinction is very clear to policy-makers. To distinguish the first type of work ("work in exchange for benefits") from the second ("stepping stone to private employment"), we label them, respectively, "Mandatory Work" and "Transitional Work".

Coherently with our dynamic mechanism design approach, to formally model these policies we enrich our framework with two new technologies. First, a *secondary production* technology that is less efficient than the (primary) one used in private employment but that, as the latter, requires effort to yield output. This feature captures that often work-based activities employ the welfare recipients on basic tasks with very low value added. Second, a costly *matching* technology that allows the agent to sample all her available job opportunities without search effort, upon payment of cost κ (e.g., to a third party who searches on behalf of the agent). This technology, therefore, frees up time from search to either work or rest.

We interpret the use of the secondary production technology alone as Mandatory Work activities, and the joint use of this production technology and matching as Transitional Work programs. Moreover, since the matching technology can be always used on its own, the model also includes a

benefits recipients) continue finding a rapidly declining hazard even after explicitly controlling for unobserved heterogeneity. Skill depreciation is also a central ingredient in a popular explanation of the comparative unemployment experience of the U.S. and Europe (e.g., Ljungqvist and Sargent, 1998).

Job Search Assistance policy where, at a cost, the worker can participate to job search without any effort. In addition to these three policy instruments, the framework yields very naturally Unemployment Insurance, corresponding to a state where the worker exerts search effort, and Social Assistance, corresponding to no effort requirements.

These five policy instruments arise as combination of choices of effort levels and use of technologies available in the economy. As in Pavoni and Violante (2007), we characterize the optimal WTW program which means studying: 1) in which region of the state space (the two dimensional space in promised utility and human capital) each policy dominates the others; 2) the optimal sequence of policies along the WTW program determined by the endogenous dynamics of promised utility and the exogenous human capital depreciation; 3) the optimal level and path of unemployment benefits, and wage subsidies upon employment, associated to each policy.

The main lesson we learn from our exercise is that there are two types of welfare programs that emerge as optimal, depending on the initial level of generosity of the program, a parameter of the economic environment determined, for example, by political economy or government budget constraints outside our model. A generous (or deep pocketed) principal would implement an optimal program based on *search* which follows the sequence Unemployment Insurance → Job-Search Assistance → Social Assistance. A less generous (or more budget constrained) principal would, instead, implement an optimal program based on *work* which follows the sequence Unemployment Insurance → Transitional Work → Mandatory Work. The reason is that, for low levels of promised utility, the effort compensation cost is smaller and it is convenient for the principal to require the agent to exert work effort and produce in exchange for welfare benefits. Another stark result we obtain is that, in order to be effective, Job Search Assistance must be combined with generous wage subsidies which compensate for the additional effort the agent will exert in the event she finds employment.

The second objective of the paper is to study quantitatively the features of the optimal WTW program for the typical welfare recipient in the U.S. economy and contrast them to the actual welfare system. We calibrate the parameters of our model to match some key labor market statistics. Next, we solve numerically for the optimal program and, by simulation, derive the optimal sequence of policies, their duration, the pattern of optimal benefits, taxes and subsidies. We then calculate the welfare gains for the worker (or, equivalently, the budget savings for the government) of shifting from the current scheme to the optimal scheme. For the model calibration, we exploit the National Evaluation of Welfare-to-Work Strategies (NEWWS). This is a large-scale longitudinal study, conducted by the U.S. Department of Health and Human Services, between 1991-1999. As part of the survey, 40,000 welfare recipients in seven distinct U.S. locations were randomly assigned to various treatment and control groups. These data allows us to set values for the key parameters of all our technologies.

Results from the quantitative analysis will be available in the next version of the paper.

The rest of the paper is organized as follows. Section 2 describes the different policies based on search and work activities that we aim at modelling. Section 3 formalizes the economic environment faced by the agent. Section 4 introduces the principal and describes the set of feasible contracts the principle can offer the agents. Here, we provide a mapping between the activities recommended by the principal to the agent and the actual policy instruments detailed in Section 2. Section 5 characterizes the optimal welfare-to-work programs, i.e., where the different policies emerge as optimal in the (U, h) space, the optimal sequences of policies and optimal consumption (i.e., unemployment benefits and wage taxes/subsidies). Section 6 calibrates the model based on the NEWWS data and studies quantitatively the optimal program, by comparing it with actual ones. Section 7 concludes.

2 Search-based and work-based policy instruments

Since the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996, welfare recipients are required to participate in search, training or work activities if they wish to qualify for TANF benefits. The legislation left states with ample freedom on how to organize these various activities into structured welfare-to-work programs. As a result, even abstracting from training and just focusing on search and work leaves an enormous variety of policy interventions and summarizing them is an arduous task. At the same time, distilling their key features is necessary for building a formal model and this is the route we take here.

Search-based activities: The starting point of any unemployment spell is *Unemployment Insurance*. The best way to describe unemployment insurance is that the worker receives benefits for a given period, during which she chooses her optimal level of search effort. In some locations, the worker search effort is monitored more or less severely. Since we have analyzed the optimal use of monitoring extensively (Pavoni and Violante, 2007; Setty, 2010), here we abstract from it. Often government agencies help the worker actively in her job search. Help takes two main forms: training and development of job-search skills, and assistance in locating vacancies. Wunsch (2010) studies the first type of intervention. Here, we study the second which we call *Job-Search Assistance.*

Work-based activities: The types of jobs performed by welfare recipients assigned to work activities always involve basic unskilled tasks such as food preparation and delivery in schools, janitorial and custodial tasks in low-income housing blocks, street and park cleaning, entry-level clerical tasks, etc. (Brock et al., 1993). However, the intent of the program changes substantially from location to location. According to Fagnoni (2000) –a comprehensive report to Congress on work-site activities in several U.S. locations– there is a "continuum" of work-based policies ranging from those which can be represented as "work in exchange for for benefits" to those which are heavily supplemented with job search assistance and therefore represent a "stepping stone to private employment." In the former class of programs, the emphasis is on the idea of personal responsibility: work must be a pre-condition to receive public assistance. Most individuals in these work-fare programs are placed in the public sector and non-profit organizations performing community service. We label this type of work-based activities *Mandatory Work*.

In the latter type of programs, instead, individuals are often carefully matched with a private employer and, while there is no contractual obligation on the part of the employer to hire the participant, there is mutual expectation that work-site employers will hire her if she performs well. Indeed the data on retention rate on transition of participants to private employment from these various types of work-based activities is consistent with this taxonomy. This second type of work-based activities is often called *Transitional Work* (see Kirby et al., 2002).

Additional instruments: When all types of active interventions fail (or are considered too expensive), and caseworkers cease requiring economic self-sufficiency from the welfare recipient, then she is exempted from any search or work requirement and simply receives a minimal amount of benefits, e.g. Food Stamps and Child Support. We call this income assistance policy of last resort *Social Assistance.*

Finally, many programs use heavily earnings subsidies in order to make employment more attractive to the unemployed. The Earned Income Tax Credit, introduced by the federal government in 1975, and greatly expanded since then, represents today the major wage subsidy program for low-income workers (see Moffitt, 2003, for a survey of the U.S. welfare system). Our model yields job subsidies and earnings taxes as part of the optimal payment structure promised by the principal to the agent.

3 Economic environment

Demographics and preferences: Individuals are infinitely lived. They have period utility over consumption c and effort a given by $u(c)-a$. Preferences are time-separable and the future is discounted at rate $\beta \in (0,1)$. We impose that $c \geq 0$, and that $u(\cdot)$ is strictly increasing, strictly concave and smooth, and u^{-1} has a convex first derivative. This last technical assumption will prove useful in our characterization and is satisfied by a wide range of utility functions, including the CRRA class with risk-aversion parameter greater than one half, and the entire CARA class (see Newman, 1995). Effort a is defined on a finite set, with the lowest effort normalized to zero without loss of generality.

Employment status: We denote the employment status of the agent by z. The agent can be either unemployed (z^u) or employed (z^w). Employment is an absorbing state.[3]

Human capital: At any point in time, agents are endowed with a stock of human capital (skills) $h \geq 0$. During unemployment human capital depreciates geometrically and deterministically at rate $\delta \in [0,1]$ and follows the law of motion:

$$h' = (1-\delta)\,h \tag{1}$$

Note that, given an initial level of human capital h_0 at the start of the unemployment spell, unemployment duration d of a worker with human capital h can be recovered as $d = \ln(h/h_0)/\ln(1-\delta)$.

[3] The optimal unemployment compensation contract with job separation and multiple unemployment spells is studied by Hopenhayn and Nicolini (2009). Their findings are relevant to our set up only in the sense that, while we assume an exogenous value for initial promised utility of the unemployed, with multiple spells this initial value would be endogenously determined by the employment history.

Production technologies: An agent who is working, i.e. employed on a production technology, needs to expend effort e to produce positive output. There are two types of production technologies in the economy which we call primary and secondary. An agent of type h employed on the primary production technology produces output $\omega(h)$. We let $\omega(\cdot)$ be a continuous and increasing function, with $\omega(h) \in [0, \omega_{\max}]$ and $\omega(0) = 0$. Note that, human capital depreciation induces wage depreciation –i.e., a deterioration of the agent's productivity in the primary sector– along the unemployment spell. A key feature of our model economy is that access to a work opportunity (i.e., a job) in the primary sector is frictional. Below we describe the friction in detail.

An agent working on the secondary production technology, upon applying effort e, produces an amount $\underline{\omega} \geq 0$ independent of h. Moreover, access to the secondary sector is frictionless, i.e. a secondary job is readily available. This dual sector structure is meant to represent a labor market where an agent who is willing to work can always find a job performing a simple task (e.g., janitor, fast-food cook, cashier, clerk at a supermarket, street sweeper, etc...), whereas finding a job vacancy that matches someone's occupational and industry skills, and hence paying proportionally to h, takes time.

Primary sector job search technology: We adopt the stock-flow approach of Coles and Smith (1998) with two distinct stages. In the first stage (matching), the agent locates a number of job opportunities, sends out her applications and she may be recontacted by an employer. The number of job opportunities is denoted by the discrete variable $\eta(h) \geq 0$, where $\eta(\cdot)$ is increasing in h. Exerting search effort means applying to all job opportunities and being re-contacted by each one with probability $\mu(e) = \mu > \mu(0) \equiv 0$. Hence, without sending applications ($a = 0$) the probability of being re-contacted is zero. Moreover, we assume that the matching stage yields at most one contact per period.[4] As a result the matching probability is given by

$$m(h, a) = 1 - (1 - \mu(a))^{\eta(h)}. \tag{2}$$

If the agent is re-contacted, a meeting (e.g., a job interview or a trial period), between employer and agent takes place. In this second stage, the firm and the agent meet and draw an idiosyncratic outcome: with probability $\lambda(r)$ the worker is retained by the firm, where r is a worker action. We let $r \in \{0, 1\}$ and $\lambda(1) = \lambda > \lambda(0) = 0$. The worker has control of the interview and can always, by choosing $r = 0$, make sure that it fails and that she does not receive a job offer.[5] We call this action r the "retention action". Putting both stages of the search technology together, the job finding probability is

$$\pi(h, a, r) = m(h, a)\, \lambda(r), \tag{3}$$

where it is useful to note that if $a = 0$ or $r = 0$, then $\pi = 0$; if $\eta \to \infty$, and $r = 1$, then $\pi = \lambda$. As a result, the job finding rate $\pi \in [0, \lambda]$. It is important to note that, as the unemployment

[4] This assumption is without loss of generality for two reasons. First, one can think of the period to be short (e.g., one day). Second, we can generalize the technology to allow for an arbitrary number N of contacts per period and, while the expression for π is more involved, the essence of the search process is unchanged.

[5] For example, the worker can appear "sloppy" and "uninterested" about the job at the interview, or pretend she is not competent in the required tasks.

spell progresses and h declines, so does the hazard rate since the set of job opportunities shrinks. The stock-flow approach suggests that this phenomenon is due to the fact that, initially, the unemployment agent draws quickly from the stock of vacancies. However, as the stock is exhausted, she can only draw from the new flow which is smaller. Finally, let y denote the outcome of the search activity during unemployment, with $y \in \{f, s\}$, where f denotes "failure" and s "success".

Matching technology: Upon payment of κ units of consumption, an agency can send out job applications on behalf of the unemployed and the unemployed saves on the search effort cost. An individual with human capital h who pays κ and uses the matching technology makes a contact with probability $m(h) = 1 - (1 - \mu)^{\eta(h)}$, and hence her job finding probability is still $\pi(h, e, r)$, exactly as if she exerted high search effort.

Financial and insurance markets: During unemployment the agent would like to purchase insurance (or self-insure through borrowing and saving) against the outcome y of her search/matching activity. We study the optimal contract when the worker has no access to storage, credit or insurance markets.[6]

4 Contractual relationship

We now introduce a risk-neutral planner/government (principal) who faces an intertemporal budget constraint with a real interest rate equal to $\beta^{-1} - 1$. At time $t = 0$, the planner offers the unemployed worker (agent) an insurance/credit contract that maximizes the expected discounted stream of net revenues (fiscal revenues minus expenditures) and guarantees the agent at least an expected discounted utility level U_0. The value of U_0 should be thought of as an exogenous parameter measuring the "generosity" of the welfare system (e.g., the outcome of voting or a political process).

Information structure: The use of the search and matching technologies and their outcome y is observable and contractible. Output during both primary and secondary work is fully observable and, since the technology is deterministic, work effort is contractible. However, search effort and the retention action are private information of the agent and under her control: these are the sources of moral hazard.

Contract: At every node, the contract specifies a consumption level for the worker, recommendations on the search or work effort level to exert, on the retention action, and on the use of available technologies: search, matching, or work on the secondary technology. During unemployment, the con-

[6] In Pavoni and Violante (2005), we show that when agents have *anonymous access* to credit markets, but face a no-borrowing constraint, the optimal contract outlined here can be implemented with a simple additional instrument: an interest tax on savings. When a tax on savings is ineffective (e.g., in the presence of hidden storage), Ábrahám and Pavoni (2008) prove that, when the source of moral hazard is hidden action (hence, different from the hidden information case of Cole and Kocherlakota, 2001), the constrained-efficient allocations improve upon self-insurance so, at least qualitatively, in our economy there is scope for government-sponsored programs. Quantitatively, the size of these welfare gains remains to be established.

sumption level corresponds to the unemployment compensation; during employment, the difference between the consumption level and the wage implies a wage subsidy/tax.

The period t components of the contract are contingent on all publicly observable histories up to t and the search-effort and retention-action recommendations must be incentive compatible. Moreover, at every t, we allow the planner to specify the contract contingent on the publicly observable realization $x_t \in [0, 1]$ of a uniform random variable X_t. This "randomization" may be used in the optimal contract to convexify the planner's problem and, thus, enhance welfare (Phelan and Townsend, 1991; Phelan and Stacchetti, 2001). A contract is a welfare-to-work (WTW) program.

4.1 Components of the contract as policy instruments of the WTW program

The combination of recommendations on the search/work effort level to expend, on the retention action, and on the use of technologies configures five possible options. Some combinations can be easily excluded: 1) simultaneously prescribe positive search effort and use of the matching technology, since the effort would be redundant, or to 2) simultaneously recommend zero retention action and pay the matching cost as this expenditure would be wasted, or to 3) simultaneously prescribe the zero retention action and use of search technology, since it could always recommend zero effort and save the agent the disutility of high effort. In particular, the contract always features $r = 1$.

As a result, the planner is left with five options, which we label "policies" of the WTW program, and we index with i. We denote as "Unemployment Insurance" ($i = UI$) the joint recommendation of search activity and positive effort. When zero effort is suggested together with the use of the matching technology, the policy will be labelled "Job Search Assistance" ($i = JA$). The zero effort recommendation without the use of any technology denotes "Social Assistance" ($i = SA$). A positive effort recommendation paired with the use of the secondary production technology denotes "Mandatory Work" ($i = MW$). Finally, since the costly matching technology does not require any effort, it can be used in conjunction with the secondary production function. We call this combination of work and search assistance "Transitional Work" ($i = TW$).

Table 1 summarizes these combinations. Finally, note that during primary employment, the difference between the wage $\omega(h)$ and the planner's transfer defines implicitly the employment tax (if positive) or subsidy (if negative).

Table 1: Mapping between effort level and technologies into policies

	Search	Matching	Search & Matching	Work	Work & Matching	None
High effort	UI	$\times$	$\times$	MW	TW	$\times$
Zero effort	$\times$	JA	$\times$	$\times$	$\times$	SA

4.2 Recursive formulation

Following Spear and Srivastava (1987) and Abreu, Pearce and Stacchetti (1990), we formulate and solve this problem recursively. The recursive formulation requires two state variables: human capital

h (or equivalently the unemployment duration d) and the continuation utility U promised by the contract.[7] The planner takes the initial conditions of this pair (U_0, h_0) as given.

Exploiting this recursive representation, consider an unemployed worker who enters the period with state (U, h). At the beginning of the period, the planner selects the optimal policy instrument $i(U, h)$ by solving

$$V(U,h) = \max_{i \in \{JA, MW, SA, TW, UI\}} V^i(U,h)$$

where the function V is the upper envelope of the values associated to the different policies which, in turn, we denote by V^i. In choosing a particular policy, implicitly, the planner also chooses an effort recommendation $a(U, h)$, the transfer $c(U, h)$ and the continuation utilities $U^y(U, h)$ conditional on the outcome y of the search or matching activity. We describe these additional choices in the next section.

As anticipated, the planner in general may decide to use randomizations through X. In this case, the value function for the planner solves

$$\begin{aligned} \mathbf{V}(U,h) &= \int_0^1 \max_{U(x) \in D} V(U(x), h)\, dx \\ s.t. &: \\ U &= \int_0^1 U(x)\, dx \end{aligned} \tag{4}$$

where the constraint says that the planner is committed to keep his promises: it must deliver to the agent continuation utility U in (ex-ante, with respect to the shock x) expected value terms.

4.3 Policies

We now describe in detail the planner problem during employment and for each of the five policy instruments available during the welfare-to-work program.

Primary employment (wage tax/subsidy): Consider an employed worker with state (U, h). Since employment is an absorbing state without informational asymmetries, the planner simply solves

$$\begin{aligned} W(U,h) &= \max_{c, U^s} \omega(h) - c + \beta W(U^s, h) \\ s.t. &: \\ U &= u(c) - e_w + \beta U^s \end{aligned} \tag{5}$$

where $e_w \geq 0$ is the work effort level on primary employment. The planner will provide full consumption smoothing for the agent, thus promised utility is constant over time, i.e., $U^s = U$. The promise-keeping constraint implies that in every period the optimal transfer c^e is constant and satisfies $c^e(U) = u^{-1}((1-\beta)U + e_w)$. Therefore, the magnitude

$$\tau(U,h) = \omega(h) - c^e(U) \tag{6}$$

[7] Employment status $z \in \{z_u, z_w\}$ is a trivial state since it equals z_u along the duration of the contract and then, upon finding a primary job, it switches to z_w forever since employment is an absorbing state.

is the implicit tax (or subsidy, if negative) the government imposes on employed workers. State-contingent taxes and subsidies are a key component of an optimal WTW plan.

By inspecting problem (5), it is easy to see that the value of employment has the following form:

$$W(U,h) = \frac{\omega(h)}{1-\beta} - \frac{u^{-1}\left((1-\beta)U + e_w\right)}{1-\beta} \tag{7}$$

and therefore W is a continuous function, increasing in h, and decreasing, concave and continuously differentiable in U.

Unemployment Insurance (UI): When the worker is enrolled in the unemployment insurance scheme, the problem of the planner is

$$\begin{aligned} V^{UI}(U,h) &= \max_{c,U^f,U^s} -c + \beta\left[\pi(h)W(U^s,h') + (1-\pi(h))\mathbf{V}(U^f,h')\right] \\ s.t. &: \\ U &= u(c) - e + \beta\left[\pi(h)U^s + (1-\pi(h))U^f\right], \\ U &\geq u(c) + \beta U^f, \\ U^s &\geq U^f \end{aligned} \tag{8}$$

where $e > 0$ is the effort level during search. Next period human capital h' is generated through the law of motion (1). The pair (U^s, U^f) are the lifetime utilities promised by the planner contingent on the outcomes (s or f) of search. Recall that the outcome of search is verifiable. To simplify notation, we have shortened $\pi(h, e, 1)$ as $\pi(h)$. The first constraint above describes the law of motion of the state variable U (promise-keeping constraint), the second constraint states that payments have to be incentive compatible to induce search, and the third states that payments have to be incentive compatible to induce the worker to the high retention action.

By combining the promise keeping constraint (PK) and the incentive compatibility constraint during search (IC-S), we can rewrite the latter as

$$U^s - U^f \geq \frac{e}{\beta\pi(h)}. \tag{IC-S}$$

Therefore, the additional incentive compatibility constraint on the retention action (IC-R) is implicitly satisfied and will never bind. Finally, the expressions for $\mathbf{V}$ and W are given by equations (4) and (5), respectively.

Job Search Assistance (JA): The problem of the planner that chooses to use the matching technology is

$$\begin{aligned} V^{JA}(U,h) &= \max_{c,U^f,U^s} -c - \kappa + \beta\left[\pi(h)W(U^s,h') + (1-\pi(h))\mathbf{V}(U^f,h')\right] \\ s.t. &: \\ U &= u(c) + \beta\left[\pi(h)U^s + (1-\pi(h))U^f\right], \\ U^s &\geq U^f. \end{aligned} \tag{9}$$

Notice the similarity between problem (JA) and problem (UI): the former is identical to (UI) except for the fact that there is no effort and no incentive-compatibility constraint in exchange for the additional per period cost κ. In this context, the matching cost κ can be interpreted as the salary of the government employee ("caseworker") who inspects available vacancies to find a suitable match for the agent, plus the additional administrative expenditures associated to this task.

Social Assistance (SA): In social assistance, the worker is "released" by the planner for the current period, in the sense that the planner does not demand high effort, but simply transfers some income to the worker. The problem of the planner is

$$\begin{aligned} V^{SA}(U,h) &= \max_{c,U^f} -c+\beta\mathbf{V}(U^f,h') \\ s.t. &: \\ U &= u(c)+\beta U^f. \end{aligned} \tag{10}$$

The expression for $\mathbf{V}$ is given by equation (4) and the constraint describes how the promised utility U can be delivered by a combination of current and future payments. It is natural to think of SA as a pure income-assistance program.

Mandatory Work (MW): When the planner assigns the worker to the secondary production technology the planner's problem is

$$\begin{aligned} V^{MW}(U,h) &= \max_{c,U^f} -c+\underline{\omega}+\beta\mathbf{V}(U^f,h') \\ s.t. &: \\ U &= u(c)-e+\beta U^f. \end{aligned} \tag{11}$$

The planner gives up search or matching in the labor market and the worker produces an amount $\underline{\omega}$ (which should be interpreted as the output net of the administrative costs) by paying a utility cost in terms of work effort e. Recall that work effort can be observed because output is deterministic. Thus, there is no incentive compatibility constraint during mandatory work. Under this policy, the agent works in exchange for benefits.

Transitional Work (TW): When the planner uses the matching technology and, in addition, assigns the worker to the secondary production technology, the planner's problem is

$$\begin{aligned} V^{TW}(U,h) &= \max_{c,U^f,U^s} -c-\kappa+\underline{\omega}+\beta\left[\pi(h)W(U^s,h)+(1-\pi(h))\mathbf{V}(U^f,h')\right] \\ s.t. &: \\ U &= u(c)-e+\beta\left[\pi(h)U^s+(1-\pi(h))U^f\right], \\ U^s &\geq U^f. \end{aligned} \tag{12}$$

The way to interpret this policy option, in light of our discussion of Section 2, is that while the agent is required to produce, the caseworker actively searches on her behalf for a suitable employment. In

Section 5.5, we discuss an alternative model for Transitional Work where the agent may be permanently hired by the same firm where she is temporarily producing $\underline{\omega}$.

In what follows, it is convenient to state some basic properties of these value functions. By applying fairly standard results in dynamic programming, the convexified upper envelope $\mathbf{V}$ inherits the same continuity, monotonicity and concavity properties of u, but two caveats are worth mentioning. First, monotonicity in U is guaranteed whenever at (U, h) the consumption level c associated to the optimal program is positive (e.g., whenever $u(0) = -\infty$). Second, the concavity of $\mathbf{V}$ in U is warranted thanks to the randomization in (4). Finally, the properties of $\mathbf{V}$ are inherited by the value functions of each single policy V^i. In particular, all the problems defining policies $i \in \{JM, MW, SA, TW, UI\}$ are also concave, and each V^i is continuously differentiable in U. See Pavoni and Violante (2007) for details.

5 The Optimal WTW Program

We are now ready to study the key characteristics of an optimal WTW program. We begin with a discussion of the economics behind the choice among alternative policies. Next we study in which regions of the state space (U, h) the various policies arise as optimal. Here, we heavily exploit the recursive formulation of the optimal contracting problem. By projecting the upper envelope $V(U, h)$ on the (U, h) state space, we obtain a graphical representation of which policy is optimally implemented at every (U, h) pair. The state space can be divided into different connected areas, each corresponding to a specific policy whose value dominates all the others. Finally, the state space can be thought of as a phase diagram, where U moves endogenously and h exogenously, which can be used to analyze the optimal sequence of policies along the unemployment spell.

5.1 Economic forces in the choice of policies

To understand the economic forces at work in the choice of policies, it is useful to compare, for a given pair (U, h) the costs and returns of each policy relative to social assistance. Social assistance is a useful benchmark because it has no returns for the government and, since effort is zero in SA, its cost to the government is simply that of delivering promised utility U by implementing full insurance, i.e. $c^{SA}(U) = u^{-1}((1-\beta)U)$.

Costs: All the policies that require effort to succeed (MW, TW, UI) entail an *effort compensation cost* for the planner. Since $u(c)$ is concave and disutility from effort is separable, as U increases the marginal utility of consumption falls whereas the marginal disutility of effort is fixed. Therefore, the higher is U, the higher is the transfer the planner has to pay to the agent to deliver the promised utility in order to compensate her for the fixed effort. To sum up, the effort compensation cost, a form of wealth effect due to the fact that leisure is a normal good in our model, increases with U.

The planner also faces a cost to satisfy the incentive compatibility constraint related to search

(IC-S) and related to retention (IC-R), respectively:

$$U^s - U^f \geq \frac{e}{\beta\pi(h)} \qquad (IC-S)$$

$$U^s \geq U^f \qquad (IC-R)$$

Because the agent has concave utility, she dislikes consumption (and, hence, promised utility) to be spread out across states. A planner facing an incentive compatibility constraint has to pay the agent a larger transfer, on average, to deliver a given level of promised utility. Both IC-S and IC-R are increasing in U, since u^{-1} has convex first derivative. The cost associated to IC-S is decreasing in h. As the unemployment spell progresses and the job-finding probability decreases, the employment outcome –that can only be achieved if the worker exerts the high job-search effort level– becomes less likely, and the planner needs to differentiate even more the future promised utilities across states to induce the agent to search. Recall that IC-S is present in UI and IC-R during both JA and TW.

The third cost component is the fixed cost of using the matching technology κ during JA and TW.

Returns: The return to the planner of using the secondary production technology (in MW and TW) is the production of output $\underline{\omega}$. The return to the planner of using the matching technology is that the planner saves on the effort compensation cost and on the costs associated to the incentive compatibility constraint IC-S, since job search can take place without unobservable effort.

The main return of using search or matching is that, with probability $\pi(h)$ a job in the primary sector is created. Recall from (7) that the net returns to employment in the primary sector for the planner are increasing in human capital h. When $e_w > 0$ there is also an effort compensation cost which makes the return to employment decreasing in promised utility U.

5.2 Optimal policies in the (U, h) space

Following Pavoni and Violante (2007), we characterize graphically the optimal policies in the (U, h) state space. When the upper envelope $V(U, h) = \max_i V^i(U, h)$ is projected onto the (U, h) space, as done in Figure 1, we obtain immediately the regions in the state space where each policy emerges as optimal. We start by interpreting Figure 1 as we move "horizontally" in the (U, h) space, i.e., we let h change for a given level of utility entitlement U. Next, we study the optimal policies as we move "vertically" through the diagram, i.e., we change U for a given level of human capital h.

Moving horizontally (along h): For high levels of U and high levels of h (top left region of Figure 1), the planner assigns the worker to UI because the human capital level that determines the success of search is relatively high. As human capital depreciates (still for this high level of promised utility) the job finding probability decreases and in order to save on the incentive cost associated to constraint IC-S, the planner shifts from UI to JA. Finally, as human capital further depreciates, the return to matching decreases because output in primary employment, a function of h, falls. The planner finds it optimal to save on the cost of matching and simply provides the agent with a constant transfer in SA. Social assistance tends therefore to emerge for low h and high U when the return to employment are too low and the effort compensation cost too high (top right region of Figure 1).

Consider now moving horizontally across the state space for lower levels of U. But for low enough levels of U, TW appears in the state space instead of JA. The effort compensation cost is low enough that, while the planner uses the matching technology, it simultaneously finds it optimal to require the agent to work as well. Similarly, moving to the right, MW appears instead of SA. The planner gives up the matching technology because the return is too small, since $\pi(h)$ and $\omega(h)$ are low, but for low levels of U it requires the agent to work in the secondary sector in exchange for benefits.

Moving vertically (along U): As U decreases, the effort compensation cost declines and the planner shifts from policies without effort (JA, SA) to policies requiring effort (UI, TW, MW). The shift from SA to JA is explained by two forces. First, as U falls, so does the cost of satisfying the IC-R constraint. Second, the effort compensation cost during primary employment (a possible outcome of JA only) is increasing in U. [8]

5.3 Some theoretical results

We now sketch a very preliminary and incomplete section with some theoretical results.

Proposition 1 ***(No human capital depreciation):*** *Assume $\omega(h) \equiv \bar{\omega} > 0$ and $\pi(h) \equiv \bar{\pi} > 0$ for all h. Moreover, assume $e_w = 0$. Then, every policy is absorbing: if policy i is chosen at the beginning of the program, choosing it thereafter is optimal. If, in addition,* $\mathbf{V}$ *is strictly concave, any optimal program must possess such absorbing characteristics.*

Proof (sketch). With $e_w = 0$ the incentive constraint IC-R is slack. Therefore, the only policy possibly involving transitions is UI, since in all other policies –from the necessary and sufficient first order conditions (recall that $\mathbf{V}$ is concave)– maintaining a constant level of promised utility U is optimal. Following the same line of proof as in Pavoni and Violante (2007) Proposition 2) one can show that the IC-S constraint is binding and $U^f < U$. As a consequence, both incentive and effort compensation costs decrease during a spell of UI. Since UI is the only policy with incentive costs and it has also effort requirements, it will continue to be optimal throughout the spell. **Q.E.D.**

Proposition 2 ***(Optimal policy sequence):*** *Assume $e_w = 0$ and allow $w(\cdot)$ and $\pi(\cdot)$ to increase with h. (i) SA are MW are absorbing. If* $\mathbf{V}$ *is submodular and twice differentiable: (ii) the only optimal transitions from TW are either into MW or SA; (iii) the only optimal transitions from JA is into SA.*

Proof (sketch). (i) The result can be shown by using the same line of proof as in Pavoni and Violante (2007), Proposition 1. The proof uses two facts. First, the value functions are increasing in

[8] The higher is h, the higher the level of promised utility U at which this switch takes place. The reason is that the return to primary employment, present in JA but not in SA, is increasing in h.

h. Second, the flow payoffs for SA and MW do not depend on h. Note that the absorbing property may only be weakly optimal.

(ii) We have to rule out that an optimal policy sequence might contemplate transitions from TW into UI or into JA. Let's first rule out the transition from TW into UI. The proof follows the argument of Proposition 5 in Pavoni and Violante (2007). Notice indeed that TW corresponds exactly to Job Search Monitoring, as modelled in Pavoni and Violante (2007), with a cost $\hat{\kappa} = \kappa - \underline{\omega}$. Since during TW the IC-R constraint is never binding, submodularity of $\mathbf{V}$ implies that, in the transition between TW and UI, utility must increase. But higher utility increases the cost associated to IC-S during UI.

Let's now rule out the transitions from TW to JA. One can show that –for each fixed (U, h) – the value function V^{TW} is steeper in U than V^{JA} hence IC-R would be binding during the transition from TW into JA. From the submodularity of $\mathbf{V}$, the first order conditions imply that U must decrease (see, Pavoni and Violante, 2007, Proposition 5). We hence get a contradiction: since the work effort compensation costs are increasing with U, either MW must still be preferred to SA next period or JA should have been preferred to TW in the current period.

(iii) We have to rule out the transition from JA into UI, TW and MW. To exclude the transition between JA and UI, we follow again the argument of Proposition 5 in Pavoni and Violante (2007). Notice that JA is similar to Job Search Monitoring, as modelled in Pavoni and Violante (2007) since the constraint IC-R is slack. Submodularity of $\mathbf{V}$ implies that U increases in the transition between JA and UI. But since during JA the worker does not supply search effort, it is easy to show that for larger utility levels –and higher effort compensation cost– one cannot prefer UI over JA. Similarly, the transition from JA into TW or MW can be excluded since the IC-R constraint is slack and submodularity of $\mathbf{V}$ implies that U increases. The result follows from the fact that the work effort (on the secondary production) compensation cost rises with U. **Q.E.D.**

Clearly, if the matching technology is too costly, both JA and TW disappear from the set of optimal policies and the only possible transitions are from UI into MW and from UI into SA.

Proposition 3 ***(Optimal payments):*** *(i) During UI and JA, benefits are decreasing. (ii) During SA and MW benefits are constant. (iii) In the transition from TW into MW benefits remain constant. (iv) If $e_w = 0$, benefits are constant in all policies but UI.*

Proof (sketch). Results (i)-(ii) descend easily from the fact that incentive constraints bind only during UI and JA. From the first order conditions, when incentive constraints are binding, benefits decrease, while when they are not binding benefits remain constant (see, for example, Pavoni and Violante, 2007, Proposition 7). (iii) We need to show that the IC-R constraint is not binding in the transition from TW into MW. This is true since MW is absorbing and since full insurance across employment states is incentive compatible because, in both states, the agent is required to supply effort forever after. (iv) If $e_w = 0$ all incentive compatibility constraints are slack but IC-S. The result comes easily from the first order and envelope conditions. **Q.E.D.**

Note that optimal wage subsidies tend to be high in the transitions from JA into employment and low in the transitions from TW into employment.

5.4 Two types of optimal welfare programs

The main lesson we learn from the characterization of the policy space and policy transitions is that there are two types of welfare programs that are most likely to emerge as optimal, depending on the initial level of generosity U_0. Figure 2 illustrates the dynamics of these two programs. A generous (or deep-pocketed) principal would implement an optimal program based on *search* which follows the sequence $UI \rightarrow JA \rightarrow SA$. A less generous (or more budget constrained) principal would, instead, implement an optimal program based on *work* which follows the sequence $UI \rightarrow TW \rightarrow MW$.

Figure 3 plots the optimal path of unemployment benefits and wage tax/subsidy in two examples of unemployment histories under the search-based program (top two panels) and the work-based program (bottom two panels). Unemployment benefits fall during UI because of the binding IC-S constraint. The top-left panel shows that satisfying the IC-R constraint during JA requires both declining benefits and a positive static wedge between promised consumption upon employment $c^e(U, h)$ and benefits $c^u(U, h)$ during unemployment. The reason is that the agent exerts no effort in JA whereas employment requires effort. Therefore, the agent will choose $r = 1$ only if the principal promises a level of consumption during employment high enough to compensate for the additional work effort.

Moving to the work-based program, the bottom-left panel shows that this static wedge is not present in TW since the agent supplies work effort also during TW. Recall that the incentive constraint on the retention action IC-R is not binding during TW since the program always features high effort from that point onward. As a result, the principal can fully insure the agent starting from her switch from UI into TW, as demonstrated by the constant consumption path.

Finally, note that, because of the higher average level of promised utility in the search-based program, wage subsidies are more generous and appear at earlier durations, compared to work-based welfare programs.

5.5 Extension: an alternative view of Transitional Work

As discussed in Section 2, there is a variety of different programs under the header of "Transitional Work". In some of these programs, the caseworker pairs the agent with a "suitable" firm for a trial period and the firm can then choose to retain the agent as its employee. To formalize this alternative form of Transitional Work, we introduce an improved version of the matching technology in our environment.

Suppose that by paying the cost $\kappa^m > \kappa$ a contact between a firm and the agent can be created with probability one. The additional cost may represent more accurate search or a side-payment to the candidate firm. Formally, this corresponds to setting $\mu = 1$ in equation (2) with an implied job finding rate of $\lambda(r)$. In other words, the only obstacle to employment on the primary sector for an agent who chooses the high retention action $r = 1$, while paired with the candidate firm, is the revelation of the match specific shock. The planner's problem becomes

$$\begin{aligned} V^{TW}(U,h) &= \max_{c,U^f,U^s} -c - \kappa^m + \underline{\omega} + \beta \left[\lambda W(U^s,h) + (1-\lambda)\mathbf{V}(U^f,h')\right] \\ s.t. &: \\ U &= u(c) - e + \beta \left[\lambda U^s + (1-\lambda) U^f\right], \\ U^s &\geq U^f. \end{aligned} \tag{13}$$

In the next draft of the paper, we'll provide a theoretical characterization for this richer environment as well.

6 Quantitative Analysis

6.1 Calibration

Data: The National Evaluation of Welfare-to-Work Strategies (NEWWS) is a longitudinal study that was administered by the US Department of Health and Human Services from 1991 to 1999. Its objective was to estimate the effectiveness of welfare-to-work programs, and specifically "what works best, and for whom?".

The study covered eleven mandatory Welfare-to-Work programs in seven distinct locations and included over 40,000 individuals over a five-year follow-up period. The study is based on random assignment of participants to experimental groups (subject to program requirements) and to control groups (without any requirement). The vast majority of program members was single mothers with two children and 11 years of education on average.[9]

The study was accompanied by several data sources, three of which are used in our analysis. The "full impact" sample collects five years of administrative records on demographic characteristics, earnings, and benefits for both treatment and control group members from all seven sites. Additional data on outcomes for adults and children were collected by interviewing a random sub-sample of about 5,000 members around two years after their date of random assignment and, in four of the seven sites, around their five-year anniversary. This survey includes ample data on the participants including the assignment of each participant to activities over the period, employment history, and receipt of non cash benefits. The third data source collects data on the costs of each activity drawn from state, county, and local fiscal records, supportive service payment records, administrative records, and case file participation records.

Parametrization: The unit of time is set to one month. It is useful to divide the parameters of the model into three groups. First, the preference parameters $\{u(\cdot), \beta, e\}$. Second, the labor

[9] Single fathers make up from three to eleven percent of the full impact sample, depending on site.

market parameters $\{\omega(h), \delta, \pi(h)\}$. Third, the parameters of the matching and secondary production technologies $\{\kappa, \underline{\omega}\}$.

We pick a value for the monthly discount factor of 0.9967 in order to match an interest rate of 4% on an annual basis, and use a logarithmic specification for the period utility over consumption. We assume that the disutility of search effort equals the disutility of work effort on both the primary and secondary sectors and, based on the evidence surveyed in Pavoni and Violante (2007), we set $e = 0.67$.

Without loss of generality, we use a linear monthly earnings function $\omega(h) = \omega h$ where human capital h is interpreted as efficiency units of labor in a competitive labor market. We normalize h so that one unit corresponds to monthly earnings of \$100. In the NEWWS data, initial monthly earnings upon entering the program are around \$1,000, and hence $h_0 = 10$. We choose an annual depreciation rate of 15 percent (see Pavoni and Violante, 2007, for details).

Finally, $\pi(h)$ is estimated from unemployment spell data in the NEWWS assuming a Weibull distribution $\alpha d^{\phi-1}$ where d is the duration of the unemployment spell. Our estimation yields parameter estimates of $\alpha = 0.351$, and $\phi = 0.780$. The estimated hazard displays therefore negative duration dependence.

The two technology parameters are estimated directly from the NEWWS data. The cost per worker of operating the matching technology for a month (κ) is based on "per month per worker" cost of the activity called "job clubs", which is essentially a form of job search assistance. This cost averages \$445 across the programs. The net gain to secondary production ($\underline{\omega}$) is obviously challenging to calculate as these activities have no clear market value. We follow Brock et al. (1993) that calculated the value of unpaid mandatory work by comparing the output of such tasks to the cost equivalent of regular workers. The result is \$849 per worker per month. However, these programs have also an administrative cost, which we estimate from the activity labelled, in NEWWS data, "work experience." This activity comprises of those programs that we described as "mandatory work" in Section 2. The estimated cost per worker per month is \$283 which subtracted from the return yields a net gain of mandatory work of \$566 per month per worker.

The algorithm used to solve numerically for the optimal contract is described in detail in the computational appendix.

6.2 Results

To be completed

7 Conclusions

The literature on the efficient provision of consumption insurance and search incentives to the unemployed in presence of private information has largely focused on the optimal path of benefits during the unemployment spell (Wang and Williamson, 1996; Hopenhayn and Nicolini, 1997; Shimer and

Werning, 2008; Pavoni, 2007; Hopenhayn and Nicolini, 2009; Pavoni 2009). Unemployment compensation during job search is a key pillar of the welfare state, but it is by no means the only instrument used by policy makers. In particular, many welfare programs directed to the unemployed do not elicit search effort ant all and, instead, require the agent to work in exchange for benefits.

In this paper, we have used the tools of recursive contract theory to study the optimal design of a welfare program that combines both search and work activities. The investigation is still in progress. The main result, at this stage, is that two types of welfare programs emerge as optimal, depending on the initial (exogenous) level of generosity of the program. A generous (or deep-pocketed) government would implement an optimal program based on *search*. A less generous (or more budget constrained) government would, instead, implement an optimal program based on *work*.

An issue which we plan to address in the next version of the paper is that of evaluating the effectiveness of policies. The standard empirical exercise in the evaluation literature is based on comparing labor market outcomes (e.g., employment and wages) of the treatment and control groups (see Heckman, LaLonde and Smith, 1999, for a survey). Our approach to the optimal design of welfare programs suggests that this is a narrow view of effectiveness of labor market policies: a policy may not be effective in isolation, but it can become so if combined with others which follow after a spell of the first policy. For example, consider the optimal sequence $UI \to TW$. It is easy to show that in our set up, the existence of this work-based policy where the unemployed is required to supply work effort, even when she does not find a job in the private sector, reduces the cost of providing search incentives during Unemployment Insurance since it lowers the value of remaining jobless for the agent.

References

[1] Ábrahám, Á. and N. Pavoni (2004a), "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending," *Review of Economic Dynamics*, 11(4), 781-803

[2] Abreu D., D. Pearce and E. Stacchetti (1990), "Towards a Theory of Discounted Repeated Games with Imperfect Monitoring,"*Econometrica*, 58(5), 1041-1064.

[3] Addison J., and P. Portugal (1989), "Job Displacement, Relative Wage Changes and Duration of Unemployment," *Journal of Labor Economics*, 7(3), 281-302.

[4] Atkeson, A. and R. Lucas (1995), "Efficiency and Equality in a Simple Model of Efficiency Unemployment Insurance," *Journal of Economic Theory*, 66, 64–88.

[5] Brock T., D. Butler and D. Long (1993), "Unpaid Work Experience for Welfare Recipients: Findings and Lessons from MDRC Research," *Mimeo MDRC*.

[6] Bover O., M. Arellano, and S. Bentolila (2002), "Unemployment Duration, Benefits Duration and the Business Cycle," *Economic Journal*, 112, 223-265.

[7] Cole H., and N. Kocherlakota (2001), "Efficient Allocations with Hidden Income and Hidden Storage," *Review of Economic Studies*, 68, 523-542.

[8] Coles, M. G. and E. Smith (1998), "Marketplaces and Matching," *International Economic Review*, 39(1), 239-54.

[9] Fagnoni (2010), "Work-Site Activities can Play an Important Role in TANF Programs," Report to Congressional Requesters, United States GAO.

[10] Flinn, C. (1986), "Econometric Analysis of CPS-Type Unemployment Data," *Journal of Human Resources*, 21(4), 457-484.

[11] Heckman J., R. LaLonde, and J. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs," in Handbook of Labor Economics, Vol. 3, Ashenfelter, A. and D. Card, eds., Amsterdam: Elsevier Science.

[12] Hopenhayn H., and J. P. Nicolini (1997), "Optimal Unemployment Insurance," *Journal of Political Economy,* 105(2), 412-438.

[13] Jacobson L. S., R. J. LaLonde, and D. G. Sullivan (1993), "Earning Losses of Displaced Workers," *American Economic Review*, 83(4), 695-709.

[14] Kirby G., H. Hill, L. Pavetti, J. Jacobson, M. Derr and P. Winston (2002), *Transitional Jobs: Stepping Stone to Unsubsidized Employment*, Mathematica Policy Research Inc.

[15] Ljungqvist L., and T.J. Sargent (1998), "The European Unemployment Dilemma," *Journal of Political Economy*, 106(3), 514-550.

[16] Machin S., and A. Manning (1999), "The Causes and Consequences of Long-Term Unemployment in Europe," in O. Ashenfelter and D. Card (Eds) *Handbook of Labor Economics*, Vol. 3C, 3085-3139.

[17] Meyer, B. D. (1990), "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58(4), 757-782.

[18] Moffitt, R. A. (2003), "The Temporary Assistance to Needy Families Program," in *Means tested Transfers programs in the United States*, R. A. Moffitt (Ed.), Chicago: Chicago University Press.

[19] National Evaluation of Welfare-to-Work Strategies (2001), "How Effective Are Different Welfare-to-Work Approaches? Five-Year Adult and Child Impacts for Eleven Programs," U.S. Department of Health and Human Services, and U.S. Department of Education.

[20] Neal D. (1995), "Industry-Specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 13(4), 653-677.

[21] Newman A. (1995), "Risk-Bearing, Entrepreneurship, and the Theory of Moral Hazard," mimeo, Boston University.

[22] Pavoni N. (2007), "On Optimal Unemployment Compensation," *Journal of Monetary Economics*, 54(6), 1612-1630.

[23] Pavoni N. (2009), "Optimal Unemployment Insurance, with Human Capital Depreciation and Duration Dependence," *International Economic Review,* 50(2), 323-362.

[24] Pavoni N., and G.L. Violante (2005), "Optimal Welfare-to-Work Programs," *Federal Reserve Bank of Minneapolis Discussion Paper 143.*

[25] Pavoni N., and G.L. Violante (2007), "Optimal Welfare-to-Work Programs," *Review of Economic Studies*, vol. 74(1), 283-318.

[26] Phelan C., and E. Stacchetti (2001); "Sequential Equilibria in a Ramsey Tax Model," *Econometrica*, 69, 1491-1518.

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

[28] Setty, O. (2010), "Optimal Unemployment Insurance with Monitoring," mimeo, NYU.

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

[30] Shimer R. and I. Werning (2008), "Liquidity and Insurance for the Unemployed" *American Economic Review*, 98(5), 1922-1942.

[31] Spear S. E., and S. Srivastava (1987), "On Repeated Moral Hazard with Discounting," *Review of Economic Studies,* 53, 599-617.

[32] Spinnewijn, J. (2010), "Training and Search During Unemployment," mimeo LSE.

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

[34] Wunsch C. (2010),"Optimal Use of Labor Market Policies: The Role of Job Search Assistance," *CESifo Working Paper Series.*

Appendix: Computational Method

This section describes the solution method for the planner's problems associated with each of the five policies: UI, JA, TW, MW and SA. We start from the UI problem. Recall that the IC-R constraint never binds. The two constraints (PK and IC-S) are used to express the decision variables U^s, c as functions of the third decision variable U^f as follows:

$$\begin{aligned} c &= u^{-1}\left(U - \beta U^f\right) \\ U^s &= U^f + \frac{e}{\beta\pi(h)}. \end{aligned}$$

Using these expressions we can transform the original problem with three decision variables and two constraints to the following problem with one decision variable, U^f, and no constraints:

$$\begin{aligned} V^{UI}(U,h) &= \max_{U^f} -u^{-1}\left(U - \beta U^f\right) \\ &\quad +\beta\left[\pi(h)\, W(U^f + \frac{e}{\beta\pi(h)}, h) + (1-\pi(h))\mathbf{V}(U^f, h')\right]. \end{aligned}$$

The first order condition (w.r.t U^f) of this problem is:

$$\left(u^{-1}\right)'\left(U - \beta U^f\right) + \pi(h)\, W'(U^f + \frac{e}{\beta\pi(h)}, h) + (1-\pi(h))\mathbf{V}'(U^f, h') = 0$$

The expression on the left-hand side is strictly monotone, continuous and it takes values of $(-\infty, +\infty)$ over U. Hence the UI problem has a unique solution.

Similarly for the other four problems (JA, TW, MW and SA) the first order conditions are, respectively:

$$\begin{aligned} \left(u^{-1}\right)'\left(U - \beta U^f\right) + \lambda W'(U^f, h) + (1-\lambda)\mathbf{V}'(U^f, h') &= 0 \\ \left(u^{-1}\right)'\left(U - \beta U^f + e\right) + \lambda W'(U^f, h) + (1-\lambda)\mathbf{V}'(U^f, h') &= 0 \\ \left(u^{-1}\right)'\left(U - \beta U^f + e\right) + \mathbf{V}'(U^f, h') &= 0 \\ \left(u^{-1}\right)'\left(U - \beta U^f\right) + \mathbf{V}'(U^f, h') &= 0 \end{aligned}$$

We approximate the slopes of the value functions $\mathbf{V}(\cdot)$ and $W(\cdot)$ by using linear interpolation on a grid over U, and we solve for the root of each condition. All the computations were done in Fortran.

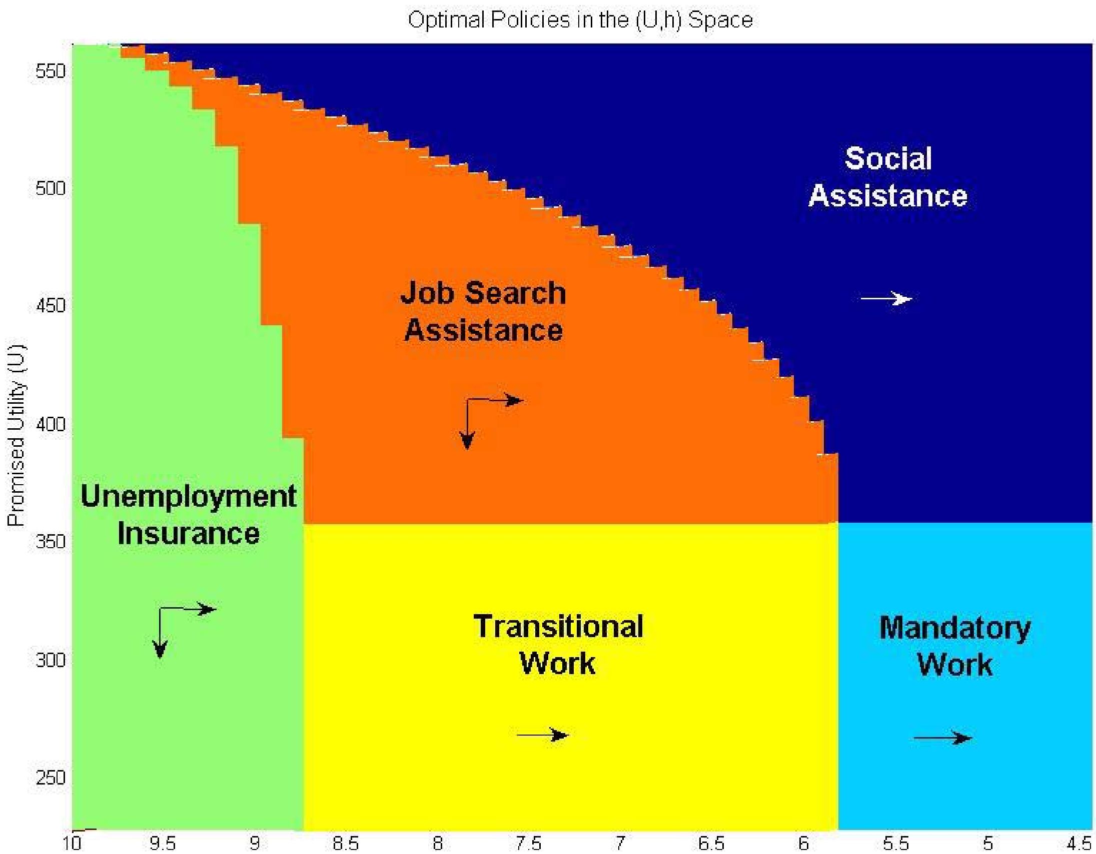

Figure 1: The policies of the optimal WTW program in the state space of human capital h and promised utility U. The arrows denote the dynamics of U and h in each policy region.

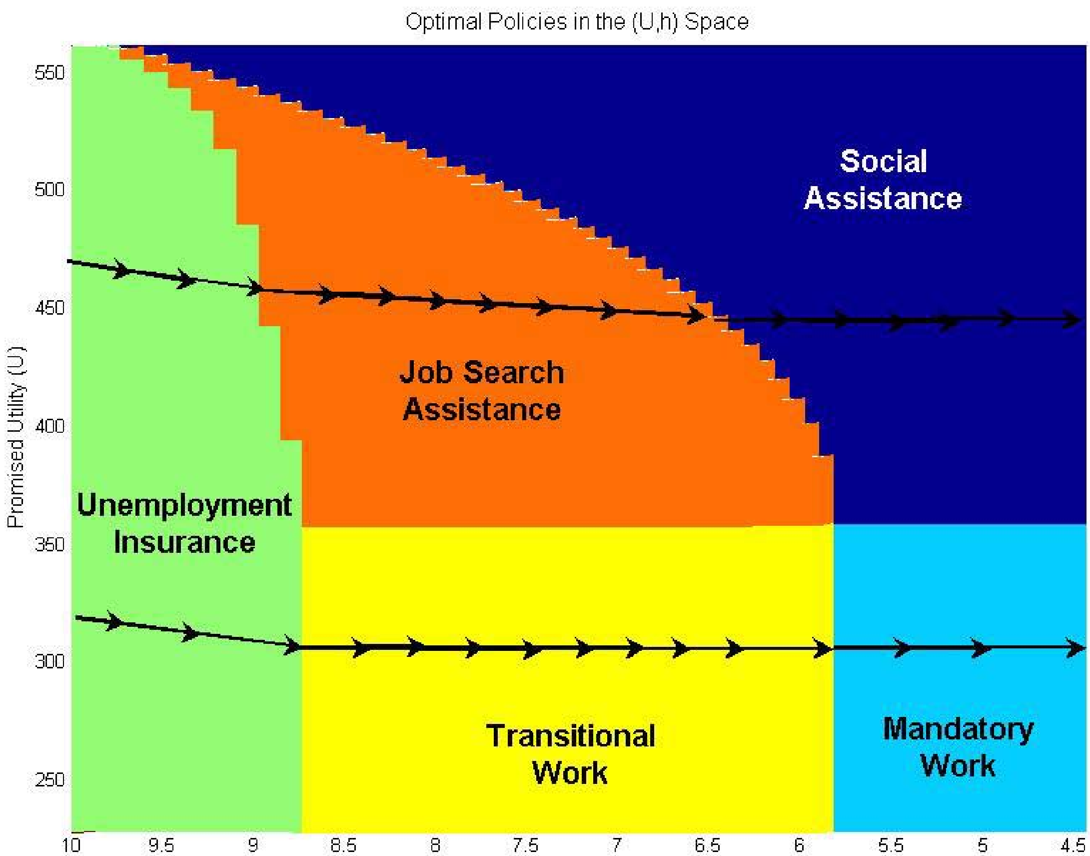

Figure 2: The optimal sequence of policies in the search-based program (top arrows) and the work-based program (bottom arrows).

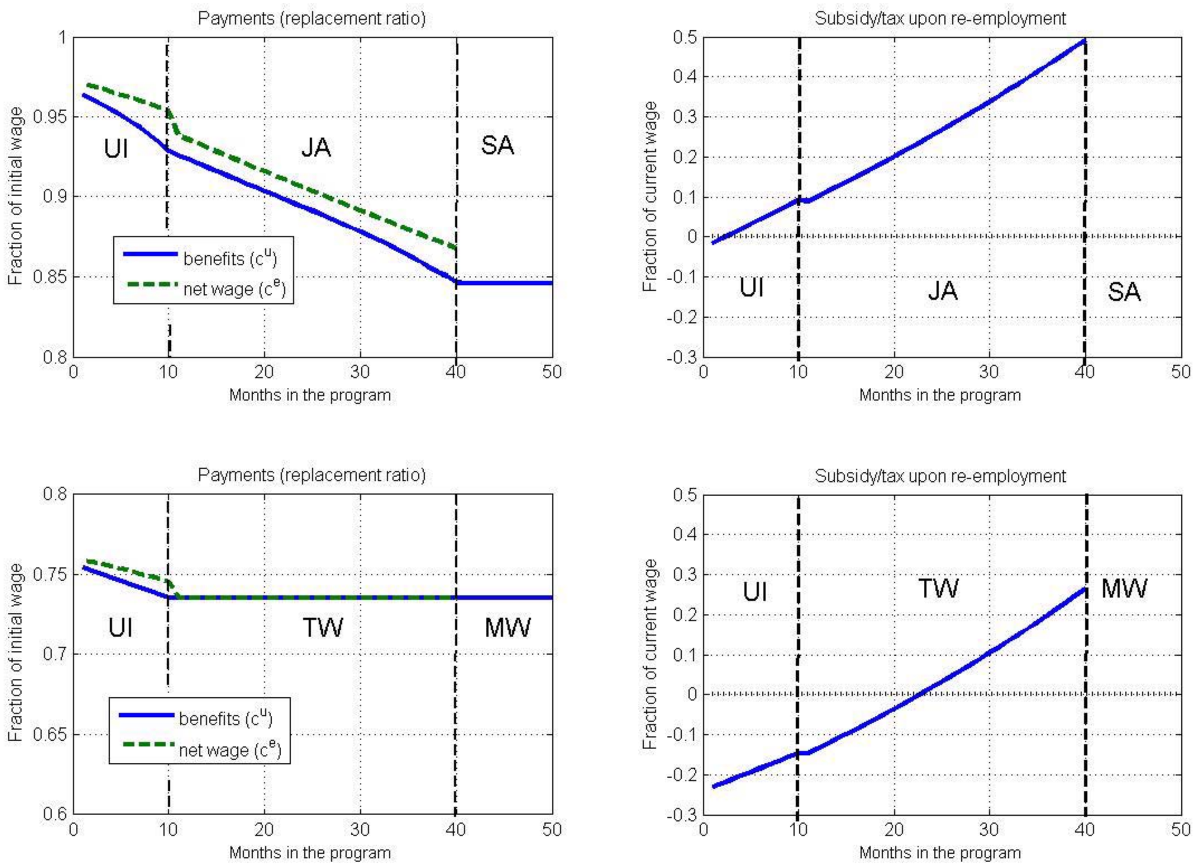

Figure 3: Dynamics of benefits and re-employment taxes/subsidies in two sample histories of unemployed workers in the search-based program (top panels) and work-based program (bottom panel).