

Optimal Unemployment Insurance with Worker Profiling*

[\[Click here for the latest version\]](#)

Sergio Cappellini[†]

April 25, 2022

ABSTRACT

In unemployment assistance programs, the government profiles recipients according to their traits with the twofold goal of facilitating their reemployment and eliminating overpayments. To this purpose, a profiling program establishes (i) which recipients to profile, (ii) when and (iii) how accurately, and (iv) the benefits to pay after it. This paper provides criteria to rank existing profiling programs, as well as an estimate of the welfare gains from the adoption of the optimal one. Two types of programs are possible at the optimum. The first type are programs that profile employability with full accuracy and envisage an active job search only for highly employable recipients. The second type are programs in which a fraction of poorly employable recipients are profiled as highly employable and persuaded to search at lower incentive costs. The reasons behind the second type of programs are (i) that expected returns on recipients' search are a concave function of the difference in the probabilities of success under search and no search, and (ii) that, absent incentive costs, also poorly employable recipients were better compensated for their search effort than left at rest. Profiling generates welfare gains also through fine-tuning of payments. Indeed, lowering the generosity of payments to those recipients who are mandated to search upon profiling reduces the cost of compensating them for the search effort.

JEL classification: D82, I38, J65, J68

Keywords: Bayesian Persuasion, Job-Search Assistance, Non-Contractible Effort, Social Assistance, Unemployment Insurance, Worker Profiling

*I would like to thank for helpful comments Nicola Pavoni, Nenad Kos, Antonella Trigari participants at the 19th Annual EEFS Conference, the 16th Annual Symposium on Economic Theory, Policy and Applications (ATINER), the 15th GRASS Workshop, the 2021 EALE Conference, and the 42nd ASSET Annual Meeting.

[†]Postdoctoral Fellow, University of Padua, e-mail: sergio.cappellini@unibocconi.it

1 Introduction

A renewed interest in optimal design of active labor-market policies (ALMPs) started in 2007 amid the financial crisis. Nowadays, following the outbreak of the Covid-19 pandemic, welfare support to the poor and the jobless is at the core of the political agenda of many governments worldwide. Nonetheless, the unprecedented increase in unemployment rates and the contemporaneous economic recession have led to a disproportion between public resources and the need for social security, which ultimately results in a push for optimizing public spending.¹ The trade-off between income support, incentive provision to job search and cost minimization for the public provider has led to policies tailored to recipients' characteristics. As a consequence, tracing a profile of any jobseeker who requests public financial support constitutes an aspect of first-order importance for the design of an effective welfare program. Profiling of welfare claimants is present in most OECD countries² and is usually employed as a tool to support and improve the design of existing ALMPs.

In the US, Worker Profiling and Reemployment Services (WPRS) and Reemployment and Eligibility Assessment (REA)³ are the two main Federal-funded programs dedicated to profiling of welfare claimants.. All workers who request access to public welfare support are asked to report their personal traits, such as education, past working experiences, family background, etc. This information allows for an *early assessment* of reemployment expectations, based on the statistical evidence provided by historical data on claimants' unemployment spells. In addition, both WPRS and REA may implement an *in-depth assessment* of the human capital of each claimant, in the form of one-on-one interviews and/or skill tests, to better tailor the assistance program to their needs.

The two programs generate savings for the provider through distinct channels. First, by improving upon the fit between workers and job-search methods. For instance, in WPRS "UI claimants who are identified through profiling methods as likely to exhaust benefits and who are in need of reemployment services to transition to new employment participate in reemployment services, such as job search assistance" (US Dept. of Labor).⁴ Second, by designing transfers based on recipients' needs during the unemployment spell. This holds especially for REA,⁵ that

¹Public unemployment spending in the US reached \$622 billions in 2021, accounting for 6.7% of the annual Federal budget (USASpending.gov, <https://www.usaspending.gov/explorer/agency>).

²Some examples are given by Worker Profiling and Reemployment Services and Reemployment and Eligibility Assessment programs (US), the Suivi Mensuel Personnalisé (France), 4-Phase Model (Germany) and Work Programme (UK).

³In 2015, REA has been replaced by the REemployment Services and Eligibility Assessment (RESEA) program, which provides greater access to reemployment services. I will nonetheless refer to the former version of the program, as it provides a clearer distinction between profiling and reemployment services which eases the exposition.

⁴<https://www.dol.gov/agencies/eta/american-job-centers/worker-profiling-reemployment-services>

⁵One of the several purposes of RESEA is to "[...] Strengthen UI program integrity" (US Dept. of Labor, <https://www.dol.gov/agencies/eta/american-job-centers/RESEA>). Hence, the two versions of the program have similar targets.

is devoted to “enhance the rapid reemployment of unemployed workers, identify existing and eliminate potential overpayments, and realize cost savings for UI trust funds” (Poe-Yamagata et al., 2011).⁶

Profiling is complementary to welfare policies, which instead deal with income support and provision of search incentives and assistance. US welfare assistance is funded partly by the Federal government and partly by single States, while the organization and design is mainly deferred to the latter. Profiling programs thus greatly differ along many dimensions, namely (i) who should be profiled, (ii) when and (iii) how accurately, (iv) whether profilees should be requested to search or rest in the meantime and/or upon it, based on the new information obtained, and (v) which payments should accompany it. All these dimensions must therefore be taken into account in the analysis of an optimal profiling program.

The first objective of this paper is to develop a framework suitable to study the main complementarities between profiling and welfare policies. Optimal welfare provision solves the problem of a risk-neutral public welfare provider (hereafter, ‘the government’), who needs to maximize the welfare of a risk-averse recipient (hereafter, ‘the worker’), subject to a budget constraint and to non-contractible job-search effort of the latter. Following [Shavell and Weiss \(1979\)](#) and [Hopenhayn and Nicolini \(1997\)](#), the design of a welfare program can be formalized as a dynamic principal-agent problem, where the state of the problem is composed by the current utility of the worker/agent, implicit in the structure of future payments, and the level of her expected reemployment skills. Keeping track of the state allows for a recursive formulation of the problem. Job search failures are themselves informative about hidden reemployment perspectives, and cause a revision of expectations. Like in [Pavoni et al. \(2013\)](#), policy instruments arise as the combination of (i) job-search recommendation to workers (‘Search’ or ‘Rest’), (ii) a transfer scheme, made of current consumption and continuation utilities, indexed to future employment status and profiling outcome, and (iii) technologies adopted. The technologies available to the planner are job search assistance and profiling, and can be implemented jointly. The optimal program arises as a sequence of policies over time.

The paper finds that reemployment expectations and the generosity of payments toward workers are crucial determinants of optimal programs. The cost of search incentives being decreasing in the expected human capital, and the one of search-effort compensation being increasing in the generosity of transfers, make the government save on these costs by delegating the job-search to workers when expected human capital (resp., program’s generosity) is high (resp., low). For this reason, the government lowers the payments to job seekers who are profiled as highly employable, in the attempt to further ease incentive provision. Therefore, a REA-like program contemplating this form of ‘punishments’ should optimally be adopted jointly with

⁶The report was commissioned by the Employment and Training Administration of the US Dept. of Labor.

direct job search. Furthermore, profiling possibly does not fully detect employability at the optimum. For low program's generosity and absent search incentives, indeed, the government would find it optimal to require also poorly employable workers to search. If the additional cost of search incentives is not too large, a fraction of poorly employable workers may be persuaded of being more employable than expected and requested to search afterwards. The paper also states sufficient conditions on workers' utility function that guarantee that the number of poorly employable workers requested to search declines in the level of generosity. On the one hand, indeed, search-effort compensation becomes more expensive and the gains from persuading poorly employable workers shrink accordingly. On the other hand, instead, the assumption on workers' utility function causes the decline of search-incentive costs in response to an increase of expectations to be more sizable when the program is more generous. These two forces are jointly conducive to a more accurate profiling when programs are more generous.

The second objective of this paper is to construct a benchmark to evaluate existing programs, as well as an estimate of the welfare gains of profiling (in the form of budget savings). To this aim, the paper solves for the optimal program with and without profiling, and uses the difference of returns in the two cases as a proxy of the value of profiling. The welfare gains for the State of Florida are estimated to be in the order of millions of dollars, a result which is comparable to the estimate contained in Poe-Yamagata et al. (2011). The analysis can be applied to all existing welfare programs, based on the generosity implicit in their transfers and the reemployment expectations of their recipients. While previous performance assessments of WPRS and REA focus on specific margins and targets (see Section 2), the advantage of this paper's approach, instead, lies in the absence of any arbitrary choice, neither about the margins to focus on, nor about the design of the program (sequence of policies and transfers, timing and accuracy of profiling, job-search methods, etc.).

The rest of the paper is organized as follows. Section 2 contains the literature review. Section 3 presents the economic environment. Section 4 describes the welfare policies. Section 5 solves for the optimal program when worker profiling is test-based. Section 6 solves for the optimal program when worker profiling is based on a statistical assessment. Section 7 conducts a quantitative analysis on REA program in the US. Section 8 extends the analysis to the case of private worker search. Section 9 concludes.

2 Literature Review

The main contribution of this paper is the development of a framework suitable to study worker profiling within a welfare program toward the jobless. The paper provides an analysis of the gains and losses of profiling, in conjunction with others labor-market policies, when workers'

human capital is not *ex-ante* observable.

Attempts have been made in the past to estimate returns of profiling programs. [Sullivan et al. \(2007\)](#) and [Poe-Yamagata et al. \(2011\)](#) are examples of such attempts. The first paper ranks WPRS programs in US States according to the occurrence of type-I error (i.e., the probability that a highly employable worker is profiled as lowly so). This paper finds that such a choice has a rationale. Indeed, optimal information design always signals low employability with full precision, but noisily detects high employability (i.e., positive probability of type-II error) whenever a share of poorly employable workers is persuaded to be highly so. In the second paper, authors conduct a field study on REA initiative in Florida, Idaho, Nevada and Illinois, and evaluate it over multiple dimensions, such as duration and total amount of unemployment benefits received, likelihood of reemployment and quarterly wage amounts received. In particular, the authors measure a positive impact of REA on public spending in three out of four States.⁷ Their estimates of cost savings have the same order of magnitude as this paper's estimates (millions of US Dollars, see [Section 7.4](#)).

The existence of an agency problem in the contractual relationship between the welfare provider and the recipients has long been acknowledged by the literature. The provider has the possibility to tackle it either by providing recipients with incentives ([Atkenson and Lucas, 1995](#); [Wang and Williamson, 1996](#); [Hopenhayn and Nicolini, 1997](#); [Chetty, 2008](#); [Shimer and Werning, 2008](#)), by monitoring them ([Pavoni and Violante, 2007](#); [Setty, 2019](#)), or else, by conducting the search on their behalf ([Pavoni et al., 2013](#); [2016](#)). In all cases, the job search produces an extra cost, which possibly outweighs the expected gain from re-employment. For this reason both active and passive policies coexist in a welfare program and only workers with better job opportunities are referred to the active ones. When human capital (i.e., job opportunities and on-the-job productivity) is allowed to depreciate during the unemployment spell, workers are reassigned to different policies. Likewise, in this paper any transition to a different policy follows the deterioration of expected reemployment perspectives. Yet, such a deterioration stems from an endogenous learning process which brings agents to revise their initial expectations. [Gonzalez and Shi \(2010\)](#) study unemployment-to-job transitions in a context where workers are heterogeneous in (unobservable) skills and get discouraged by long-lasting unemployment spells. Permanence in unemployment makes them more inclined to accept lower wage proposals. Therefore, the reemployment equilibrium wage is increasing in the perceived probability of being high-skilled. Similarly, in my framework the duration of unemployment spells has a discouragement effect on job-seekers. However, the need of larger search incentives for more discouraged workers produces a contrasting effect on net reemployment wages.

⁷The absence of any positive impact of REA in Illinois is attributed by the authors to the small number of eligible participants (3,122 in 2009).

Expectation revision can also be caused by profiling. Profiling can thus become the mean used by the government to persuade jobless recipients to search for new jobs by manipulating their expectations. The paper is related to the vast and growing literature on information design initiated by [Kamenica and Gentzkow \(2011\)](#), that deal with the design of an optimal signaling strategy about a payoff-relevant, yet unknown, state from a principal/sender to an agent/receiver. Strategic signalling may lead to partial information disclosure, with the aim of making the agent choose an action which is most favorable to the principal, yet not to the agent itself. The optimal signal has a geometric characterization, as its return coincides with the concave closure of the pre-signal payoff function. Previous works ([Bergemann and Morris, 2018](#); [Kolotilin, 2018](#); and [Galperti and Perego, 2018](#)) have highlighted the limitations of this method and proposed a new formulation of the problem that can be solved with linear programming.⁸ The first attempt to model profiling in the context of unemployment insurance has been made by [Cappellini \(2020\)](#). The paper outlines the trade-off in profiling between information detection and principal-agent incentive alignment and shows that it may lead to partial detection of human capital, aimed at persuading the worker to search at lower incentive costs. This paper builds on that result, by highlighting a second channel of gains from profiling that originates from reduction of effort-compensation costs on workers referred to direct search upon profiling. The peculiarity of this paper’s framework is the ‘hybrid’ nature of the problem, that requires the principal to deal with both information and incentive design. Consequently, the well-known concavification result à la Kamenica and Gentzkow only holds when utility promised to the agent is kept constant over profiling outcomes, and so the problem can be framed as one of pure information design (like in [Cappellini, 2020](#)). [Boleslavsky and Kim \(2021\)](#) extend the concavification result to a setting with three players (sender, agent and receiver) and incentive provision. The sender commits to a signal about a hidden state, with the twofold objective of convincing the agent to exert private effort and so affecting the state distribution with some probability, and persuading the receiver to take the sender’s preferred action. [Rodina \(2020\)](#) considers a similar setting where the agent effort is not private. Similarly this paper, both works find that incentive provision determines a change in the concavification result by Kamenica and Gentzkow. However, in my paper the agent/receiver’s effort affects the employment state, not the hidden one, and a combination of detected information and payments provides search incentives. [Bloedel and Segal \(2018\)](#), [Habibi \(2020\)](#) and [Zapechelnuyk \(2020\)](#) also study the tension between incentive and information provision in the Bayesian persuasion framework applied to situations of agent’s rational inattention, agent’s time-inconsistency and quality certification, respectively. However, to the best of my knowledge, no work has studied the relationship between information design

⁸In particular, Galperti and Perego have proved the existence of a dual of the original problem, where the ‘shadow price’ of the probability of each state is declining in a measure of agent’s persuasion.

and incentive provision in unemployment insurance so far.

3 Economic Environment

Players' Interaction. A risk-neutral government (principal, it) and a risk-averse worker (agent, she) are infinitely-lived. Time is discrete and their discount factor is $\beta \in (0, 1)$. The worker can be employed or not, and the government observes her employment state. In period 0, (i) the worker is unemployed, (ii) her human capital is unknown and both players hold common expectations about it, and (iii) the government offers her a contract contingent on any possible future employment state and new information about human capital. The contract is so designed to minimize the expected discounted value of net transfers to the worker, conditional on delivering to her some promised utility. For each history node, the contract specifies the technology/ies adopted by the government (assisted-search and/or profiling), the effort recommendation to the worker ('Search' or 'Rest') and transfers. Uncertainty about worker's employment state clears at the beginning of each period.

Human Capital and Job Search. Worker's human capital can be high ($h = H$), or low ($h = L$). Workers with high (resp., low) human capital are labelled as high- (resp., low-)skilled. If unemployed, the worker can either rest ($a = 0$) or search for a job ($a = 1$). In the first case, her job-finding probability is null. In the second case, the high-skilled worker finds a job with probability π_H , while the low-skilled one with probability $\pi_L \in (0, \pi_H)$. The job search is public, but non-contractible, and makes any worker incur effort cost e , with worker's utility over consumption c and effort a being separable and given by $v(c, a) = u(c) - e \cdot a$. The first-order derivative of u^{-1} , $1/u'$, is convex.

Market-sector production. Labor productivity is increasing in human capital ($\omega_H > \omega_L$). In the economy there is one market sector only, populated by identical atomistic firms competing à la Bertrand over job offers, and paying wages equal to labor productivity. Reemployment is an absorbing status, since the worker faces no risk of any future lay-off.

Expectations. Any worker who applies to welfare support undergoes an early assessment. The assessment attaches to the worker a probability μ of being high skilled ($\mu = Prob(h = H)$), which is henceforth referred to as *expectation*.⁹ In actual programs, welfare claimants report personal information (social background, past working experiences, education, etc.), according to which the welfare provider makes an initial evaluation of their human capital. The evaluation of any claimant is based on historical data that measure the reemployment frequency of claimants with same characteristics. Highly-educated and more experienced workers, for instance, are statistically more likely to exit unemployment than workers with less experience

⁹By the law of large numbers, such a probability is unbiased, meaning that the fraction of high-skilled workers among all workers with same expectation coincides with the expectation itself.

and/or lower educational attainment.

Assisted-search technology. The government can search on behalf of the worker at cost κ^{ja} . The cost includes the administrative expenses of the offices which are in charge of looking for vacancies, create a network with prospective employers and maintain contacts with them, circulate the worker's CV, etc.

Profiling technology. Profiling detects human capital with some accuracy, and returns a publicly observable outcome, at cost κ^{wp10} . Profiling can be thought of as a lottery that returns a binary outcome -'Pass' ($r = p$) or 'Fail' ($r = f$)- with predetermined probabilities, conditional on human capital realization h . The government can choose to profile with different levels of accuracy workers holding different expectations. This means that the lottery odds are indexed by expectation μ and program's generosity U

$$\{\sigma(p|h, \mu, U), \sigma(f|h, \mu, U)\}_{h \in \{H, L\}}$$

The timing of profiling is as follows. In the current period, the worker is profiled. In the next period, the new employment state realizes, before the profiling outcome is disclosed. The government is assumed not to be allowed to base worker's current and future consumption upon reemployment on the new information derived from profiling.¹¹

Policies. Any policy arises as the composition of (i) recommended search effort, (ii) consumption contract, and (iii) assisted-search and/or profiling technology (or neither of them). Combinations of search effort levels and technologies give rise to eight ($2 \times 2 \times 2$) possible policy instruments. However, when the assisted search technology is implemented, it would be redundant to prescribe positive search effort to the worker, which reduces to six the number of policies. If no technology is implemented, the government can decide whether to recommend positive search effort and pay incentives ('Unemployment Insurance', $i = UI$), or not ('Social Assistance', $i = SA$). If only the assisted search technology is implemented, it gives rise to 'Job-Search Assistance' ($i = JS$). Profiling without any search gives rise to 'Assistance and Profiling' ($i = AP$), whereas 'Insurance and Profiling' ($i = IP$) arises when the technology is adopted together with worker's search. Finally, 'Search-Assistance and Profiling' ($i = SP$) originates if both technologies are jointly adopted.

¹⁰The cost includes administrative expenses, as in the case of assisted search.

¹¹This assumption only binds when profiling is based on worker performance (Section 5).

No Profiling			
Recommendation	Assisted Search	Delegated Search	No Search
‘Search’	x	Unemployment Insurance (UI)	x
‘Rest’	Job-Search assistance (JS)	x	Social Assistance (SA)
Profiling			
Recommendation	Assisted Search	Delegated Search	No Search
‘Search’	x	Insurance & Profiling (IP)	x
‘Rest’	Search-assistance & Profiling (SP)	x	Assistance & Profiling (AP)

Table 1: Policy Instruments

At time $t = 0$, the planner offers the unemployed agent an insurance contract that minimizes transfers and guarantees her an expected discounted utility equal to U . The planner’s problem can be written recursively by keeping track of worker’s expected human capital and promised utility -henceforth, a proxy for program’s generosity- along the unemployment spell. The consumption contract of policy i consists of a menu of today’s consumption c^i and tomorrow’s continuation utilities $U_i^{s,r}$, contingent on reemployment ($s = w$) or not ($s = u$), and ‘Pass’ ($r = p$) or ‘Fail’ ($r = f$) outcome, if job search and/or profiling are conducted. Current expectation μ and program’s generosity U jointly determine the choice of the policy instrument. The government chooses the optimal policy $i(\mu, U)$ by solving

$$V(\mu, U) = \max_{i \in \{SA, JS, UI, AP, SP, IP\}} V^i(\mu, U) \quad (1)$$

The planner is allowed to randomize over worker’s utility U under the constraint that the promised utility must be delivered in expectation. To this end, the operator \mathbf{V} is defined as

$$\begin{aligned} \mathbf{V}(\mu, U) &= \max_{\{q(x), U(x)\}_{x \in [0,1]}} \int_0^1 q(x) V(\mu, U(x)) dx \\ \text{sub: } U &= \int_0^1 q(x) U(x) dx, \quad \int_0^1 q(x) dx = 1 \end{aligned} \quad (2)$$

4 Welfare Policies

I now introduce the problem of the welfare provider, who has to design the wage tax (or subsidy) upon reemployment (W) and the welfare policies (SA, JS and UI), defined as those instruments that do not resort to worker profiling.

Wage Tax/Subsidy (W). In case of successful job search, the worker’s productivity is revealed.

Therefore, the market-sector value when human capital is equal to $h \in \{H, L\}$ reads

$$W(h, U) = \max_{\tau, U^w} \tau + \beta W(h, U^w) = \max_{c^w, U^w} \omega_h - c^w + \beta W(h, U^w)$$

$$\text{sub: } U = u(c^w) + \beta U^w \quad (\text{PK})$$

Since reemployment is assumed to be an absorbing state (the separation rate between employees and firms is assumed null), the planner is sure to raise tax/pay subsidy also in the next period. The labor tax τ is the wedge between gross (ω_h) and net wage (c^w). The Promise-Keeping (hereafter, (PK)) constraint is the recursive expression of worker's utility. It guarantees that utility flow from current period $u(c^w)$ and continuation utility U^w are large enough to match current utility level U . The optimal contract prescribes constant continuation utility ($U^w = U$).¹² Hence from (PK) one can obtain the closed-form expression for consumption $c^w = u^{-1}((1-\beta)U)$. The expression for labor tax/subsidy thus is

$$W(h, U) = \frac{\omega_h - u^{-1}((1-\beta)U)}{1-\beta}$$

Social Assistance (SA). The planner's problem when neither job search, nor profiling is performed reads

$$V^{SA}(\mu, U) = \max_{c^{sa}, U^{sa}} -c^{sa} + \beta \mathbf{V}(\mu, U^{sa})$$

$$\text{sub: } U = u(c^{sa}) + \beta U^{sa} \quad (\text{PK})$$

The planner transfers c^{sa} and pledges continuation utility U^{sa} , without requiring the worker to exert any effort. SA is a passive measure, fully devoted to income support, and does not envisage any form of job search. Thus, there is no chance of reemployment for the worker, nor any chance for the provider of raising a labor tax in the incoming period. Differently from the definition of wage tax/subsidy, where reemployment is an absorbing state, the planner can freely select the best policy instrument in the next period. However, the following holds.

Proposition 1 (Absorbing SA). *Social Assistance is an absorbing policy and its continuation utility equals current utility ($U^{sa} = U$).*

Proof. See [Appendix A: Properties of SA, JS and UI](#). ■

The proof follows the same steps as in Pavoni et al. (2016). The result implies that, once the

¹²At the optimum, U^w solves

$$W_U(h, U) = -\frac{1}{u'(c^w)} = W_U(h, U^w) \implies U^w = U$$

worker enters SA, she is never reallocated to any other policy, neither she can exit unemployment, as no search is conducted. This result is admittedly quite extreme for policymakers, who may find it hard to politically defend a welfare program granting life-time financial support to people who will never have the chance of getting reemployed. Yet, the result is remarkable in that it establishes that any passive policy should be regarded as a policy of last resort, to target only workers with low expected human capital. Current consumption solves (PK) with $U^{sa} = U$.¹³ The value of SA is independent of μ and has a closed-form expression

$$V^{SA}(U) = -\frac{u^{-1}((1-\beta)U)}{1-\beta} \quad (3)$$

No revision of expectations occurs during SA, as no job search is conducted. When, instead, the search is unsuccessful, both the government and the worker downward revise initial expectation μ . The revised expectation that human capital is $h = H$ follows a Bayesian updating rule

$$\mu' \equiv \frac{\mu(1-\pi_H)}{\mu(1-\pi_H) + (1-\mu)(1-\pi_L)} \leq \mu \quad (4)$$

The revised probability is lower than initial one, with equality holding only if human capital is known already ($\mu \in \{0, 1\}$). The reason lies in the unbiasedness of μ , that is equal to the actual share of high-skilled workers among those who hold that expectation. Thus, a fraction $\pi(\mu) \equiv \mu\pi_H + (1-\mu)\pi_L$ of them manages to find a new employment, which implies that the high-skilled ones who remain unemployed in the next period are a fraction $\mu(1-\pi_H)/(1-\pi(\mu))$ of the initial group. Therefore, in case of failed search, a higher probability is attached to low human capital realization ($h = L$).¹⁴

Job-Search assistance (JS). When resorting to assisted search, the government looks for

¹³At the optimum, U^{sa} solves

$$V_U^{SA}(\mu, U) = -\frac{1}{u'(c^{sa})} = V^{SA}(\mu, U^{sa}) \implies U^u = U$$

¹⁴If the failed attempts to exit unemployment are t , then initial expectation μ is updated t times according to the formula

$$\mu^{(t)} = \mu^{(t-1)'} = \frac{\mu(1-\pi_H)^t}{\mu(1-\pi_H)^t + (1-\mu)(1-\pi_L)^t} \quad (5)$$

where the convention that $\mu^{(0)} = \mu$ is used. It is easy to see that:

- $\mu = 0$ and $\mu = 1$ are the only two expectations such that $\mu^{(t)} = \mu$. When players know human capital, no update ever occurs;
- $\lim_{t \rightarrow \infty} \mu^{(t)} = 0$, if $\mu^{(0)} < 1$.

employment on worker's behalf at cost κ^{ja} . The value of JS reads

$$V^{JS}(\mu, U) = \max_{c^{js}, U_H^w, U_L^w, U^u} -c^{js} - \kappa^{ja} + \beta [\mu\pi_H W(H, U_H^w) + (1 - \mu)\pi_L W(L, U_L^w) + (1 - \pi(\mu))\mathbf{V}(\mu', U^u)]$$

$$\text{sub: } U = u(c^{js}) + \beta [\mu\pi_H U_H^w + (1 - \mu)\pi_L U_L^w + (1 - \pi(\mu))U^u] \quad (\text{PK})$$

Two are the sources of risk related to the job search. The first risk is related to its outcome (success or failure). The second one, instead, is related to human capital realization, conditional on finding a new job for the worker. While the government finds it optimal to insure the worker against the latter, due to her risk aversion ($U_H^w = U_L^w$), the same holds for the former only if no search incentive is to be paid ahead, i.e. if the worker will in no case be referred to UI during the spell.¹⁵ In either case, then the optimal contract solves

$$-\frac{1}{u'(c^{js})} = W_U(\mu, U^w) = \mathbf{V}_U(\mu', U^u)$$

with

$$W(\mu, U) \equiv \frac{\mu\pi_H}{\pi(\mu)} W(H, U) + \frac{(1 - \mu)\pi_L}{\pi(\mu)} W(L, U)$$

being the expected wage tax/subsidy, conditional on reemployment.

Unemployment Insurance (UI). The planner may delegate the job search to the agent and provide her with incentives to conduct it. Incentive provision originates from the fact that worker's effort is non-contractible, and boils down to adding an Incentive Compatibility constraint (hereafter, (IC)) to the planner's problem.

$$U \geq u(c^{ui}) + \beta U^u \quad (\text{IC})$$

The (IC) constraint guarantees incentive compatibility of the contract against agent's deviation from recommended search effort. Promise Keeping in UI takes into account the effort cost e exerted by the job-seeker agent

$$U = u(c^{ui}) - e + \beta [\pi(\mu)U^w + (1 - \pi(\mu))U^u] \quad (\text{PK})$$

The problem of the planner reads

$$V^{UI}(\mu, U) = \max_{c^{ui}, U^w, U^u} -c^{ui} + \beta [\pi(\mu)W(\mu, U^w) + (1 - \pi(\mu))\mathbf{V}(\mu', U^u)]$$

$$\text{sub: } (\text{PK}) - (\text{IC})$$

¹⁵The proof is reported in [Appendix A: Properties of SA, JS and UI](#).

(IC) and (PK) constraints imply the following condition on the difference in continuation utilities between successful (U^w) and failed (U^u) search

$$U^w - U^u \geq \frac{e}{\beta\pi(\mu)} \quad (6)$$

The condition in (6) is binding at the optimum and accounts for the planner's cost of incentive provision. Incentive cost can be defined by the difference in costs between the cases of contractible and non-contractible effort. Incentive costs are increasing in the cost of effort and decreasing in the level of patience (β) and confidence (μ). Intuitively, it is less expensive to convince the agent to search when she expects larger return on search and weighs more the prospective reward ensuing from it. Condition (6) shows that search incentives have a convex hyperbolic shape in the space of expectations. Concavity of V^{UI} in μ follows from convexity of incentive costs and the linearity of returns. Lemma 1 proves these two facts.

Lemma 1 (Slopes of the value functions with respect to μ and U). *Every policy return V^i is concave increasing in expectations, and concave decreasing in promised utility. V^{UI} is supermodular (i.e., $V_{\mu U}^{UI}(\mu, U) \geq 0$). And so is V^{JS} , whenever $U_{JS}^w \geq U_{JS}^u$.*

Proof. See [Appendix A: Properties of SA, JS and UI](#). ■

Incentive costs depending negatively on expectations through the utility dispersion generates a comparative advantage of UI for high-end expectations. On the contrary, convexity of $1/u'$ is a sufficient condition for the costs of incentive provision and effort compensation to be increasing in U , and so for UI to suffer a comparative disadvantage for high-end generosities. For this second reason, more generous programs are mainly focused on assistance, in the form of income support and assisted search, and less on search incentives. The following proposition establishes how policies are located over the (μ, U) space.

Proposition 2 (Welfare Policies in the (μ, U) Space). *Assume \mathbf{V} is 'locally' supermodular, that is, for every (μ, U) , there exist $\epsilon_\mu, \epsilon_U > 0$ such that*

$$\mathbf{V}_U(\tilde{\mu}, U) \geq \mathbf{V}_U(\mu, U) \quad \text{and} \quad \mathbf{V}_\mu(\mu, \tilde{U}) \geq \mathbf{V}_\mu(\mu, U)$$

for every $\tilde{\mu} \in (\mu, \mu + \epsilon_\mu)$ and $\tilde{U} \in (U, U + \epsilon_U)$. Then,

$$V_U^{UI}(\mu, U) \leq V_U^{JS}(\mu, U) \leq V_U^{SA}(U) = W_U(\mu, U) < 0 \quad (7)$$

and

$$0 = V_\mu^{SA}(U) \leq V_\mu^{JS}(\mu, U) \leq V_\mu^{UI}(\mu, U) \quad (8)$$

where the last inequality holds (at least) at the crossing point. Lastly, if the program does not envisage transition from JS to UI, then $V_{\mu\mu}^{JS}(\mu, U) = 0$ and $V_U^{JS}(\mu, U) = V_U^{SA}(U)$.

Proof. See [Appendix A: Properties of SA, JS and UI](#). ■

The assumption about local supermodularity of \mathbf{V} descends from supermodularity of each V^i and requires the marginal gain of μ to be increasing in the level of generosity (i.e., $\mathbf{V}_{\mu U}(\mu, U) \geq 0$). However, the assumption only holds locally. Indeed, the shape of \mathbf{V} is determined by two contrasting forces, within-policy supermodularity and between-policy submodularity. While the former dominates locally, the latter has a global impact on \mathbf{V} .

A rise in $\pi(\mu)$ increases the job-search return, and lowers the cost of incentive provision and effort compensation connected to direct search. On the contrary, a rise in U produces an increase in the contract costs of direct search. Therefore, fixing generosity and spanning the space of expectations, SA is replaced by active labor-market policies at larger expectations. And UI dominates JS for high-end expectations and low-end generosityes.

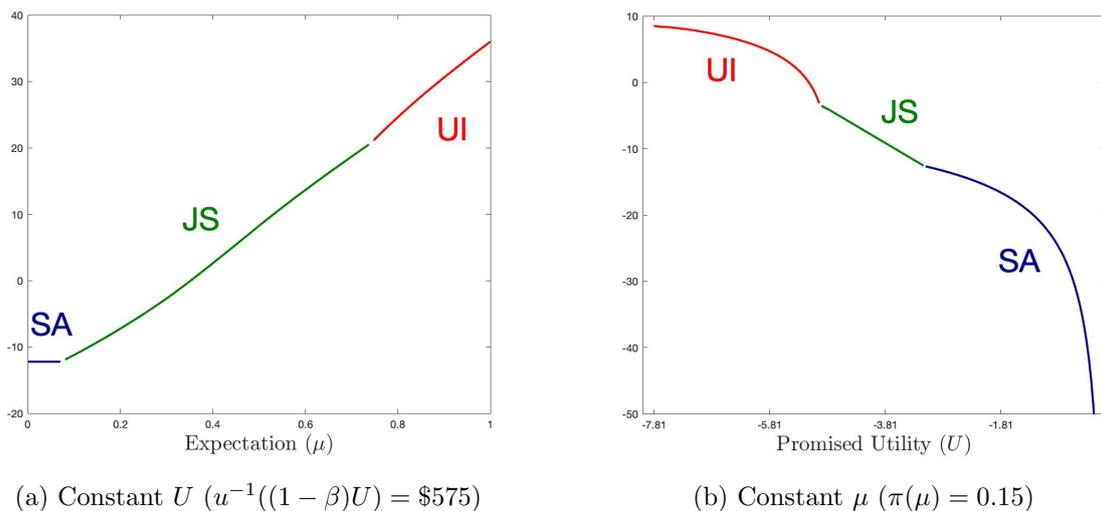


Figure 1: Value of welfare policies.¹⁶

The upper envelope V thus displays a tendency toward between-policy convexity in the space of μ (see Fig. 1a). Such shape of V has deep implications on the choice of optimal profiling. Similarly, Fig. 1b plots V over U for constant μ . UI is optimal for low generosityes, while SA is first-best for high generosityes, when no search is worth conducting because of the expensive contract costs. The slope of JS lies in between the one of UI and SA and depends on the downstream program in case the job search is unsuccessful. The following proposition clarifies the point.

¹⁶The parameter values and functional forms used in this Section 4, 5, and 6 are: CRRA worker's utility with $\sigma = 3$, $\beta = 0.9$, $e = 0.7$, $\kappa^{ja} = 8$, $\kappa^{wp} = 0.5$, $\omega_H = 10$, $\omega_L = 3$, $\pi_H = 0.27$, $\pi_L = 0.14$. All monetary values are divided by 100.

Proposition 3 (Optimal Dynamics of Promised Utility and Benefits). *Continuation utility upon failed search is*

- decreasing when UI is part of the policy sequence ahead;
- constant, otherwise.

Unemployment benefits are constant in SA and JS, and decreasing in UI.

Proof. See [Appendix A: Properties of SA, JS and UI](#). ■

Since contract costs of direct search increase in the level of worker's utility, search incentives are a mix of reward upon reemployment and punishment upon job-search failure. Hence, promised utility decreases during UI. Furthermore, whenever direct search ensues assisted search, worker's utility starts decreasing while the worker is still enrolled in JS. This finding sheds light on the possible policy patterns that can arise as a function of worker's initial expectation and program's generosity. Fig. 2 shows two instances of optimal policy sequences, for same initial expectation ($\mu_0 = 0.9$) and different generosity. When generosity is higher ($c(\bar{U}_0) = 290$ in consumption-equivalent terms),¹⁷ the worker never enters UI and only moves from JS to SA. Thus, she is required to exert no effort at any time and granted both consumption and utility insurance along the spell. So, she eventually exits unemployment or enters SA with the same utility level and consumption as she entered the program. When, instead, generosity is lower ($c(\underline{U}_0) = 190$), the worker is relocated from JS to UI at some point of the spell. For this reason, her utility decreases over time until she finds a job and exits unemployment.

One may ask whether a worker in UI can be referred to JS. Two contrasting forces have an impact on the policy transition over the UI-JS frontier. First, the optimal contracts of UI and JS prescribe a decline in promised utility for values of the (μ, U) space that are close enough to the frontier. This produces a drop in the contract costs of direct search, and makes UI more appealing with respect to JS *ceteris paribus*. Second, the downward revision of expectations causes an increase in the incentive cost of UI, which makes JS more appealing. The following proposition establishes a relationship between the concavity of V and the slope of the UI-JS frontier in the μ space that avoid the possibility of any transition from UI to JS.

Proposition 4 (Policy Transitions). *Let $\hat{U}(\mu)$ be the promised utility level that makes the government indifferent between administering UI or JS to a worker with expectation μ , i.e.*

$$V^{UI}(\mu, \hat{U}(\mu)) \equiv V^{JS}(\mu, \hat{U}(\mu))$$

¹⁷The expression of consumption equivalent of U is

$$c(U) = u^{-1}((1 - \beta)U)$$

In addition, define $\eta^i(\mu, U)$ as the real non-negative number such that

$$V_U^i(\mu, U) \equiv -g'((1 - \beta)U + \eta^i(\mu, U)), \quad i = SA, UI, JS$$

where $g \equiv u^{-1}$ is the inverse utility function. Assume that $\eta^i(\mu, U)$ is non-increasing in U for every μ and that

$$\beta[\hat{U}(\mu) - \hat{U}(\mu')] \leq \eta^{UI}(\mu, \hat{U}(\mu)) \quad (9)$$

Then, any worker in UI either remains in UI, enters SA or exits unemployment. In particular, the optimal program never switches from UI to JS.

Proof. See [Appendix A: Properties of SA, JS and UI](#). ■

The sufficient condition (9) establishes an upper bound on the slope of the UI-JS frontier in the μ space. Such slope is determined by the change in the difference of contract costs between UI and JS in response to an increase of μ and U . In particular, (i) incentive costs in UI fall in response to any increase of μ and (ii) both incentive and effort-compensation costs in UI rise in response to any increase in U . An upper bound on the slope of the frontier thus requires that the first determinant does not have a major effect on the difference of contract costs, so that every variation of μ only requires a relatively small variation of U (of the same sign) to reestablish the parity between the value of UI and JS. Condition (9) guarantees that the former effect prevails over the latter one.

5 Test-Based Profiling

Profiling publicly discloses worker's human capital, up to a level of accuracy chosen by the government. For this purpose, profiling programs implement different strategies to infer the level of human capital of the worker. Some assign the profiled claimant a given task and assess human capital based on her performance. An example of this type of profiling is adopted by Missouri, where claimants selected for REA are required to fill in a self-eligibility assessment ([Annex I](#)). In cases like this, profiling takes the form of a test with two mutually exclusive outcomes, 'Pass' ($r = p$) and 'Fail' ($r = f$). Conditional on the test outcome, the worker is referred to a different policy instrument. The test accuracy is determined by the minimum requirement to pass it. Assume that each profiled worker is assigned a score as a proxy for their reemployment chances, and is returned a 'Pass' whenever the score is above a given threshold, established by the government. A low threshold allows for a fraction of low-skilled workers to obtain the 'Pass' together with high-skilled ones, whereas a high threshold pools low-skilled and

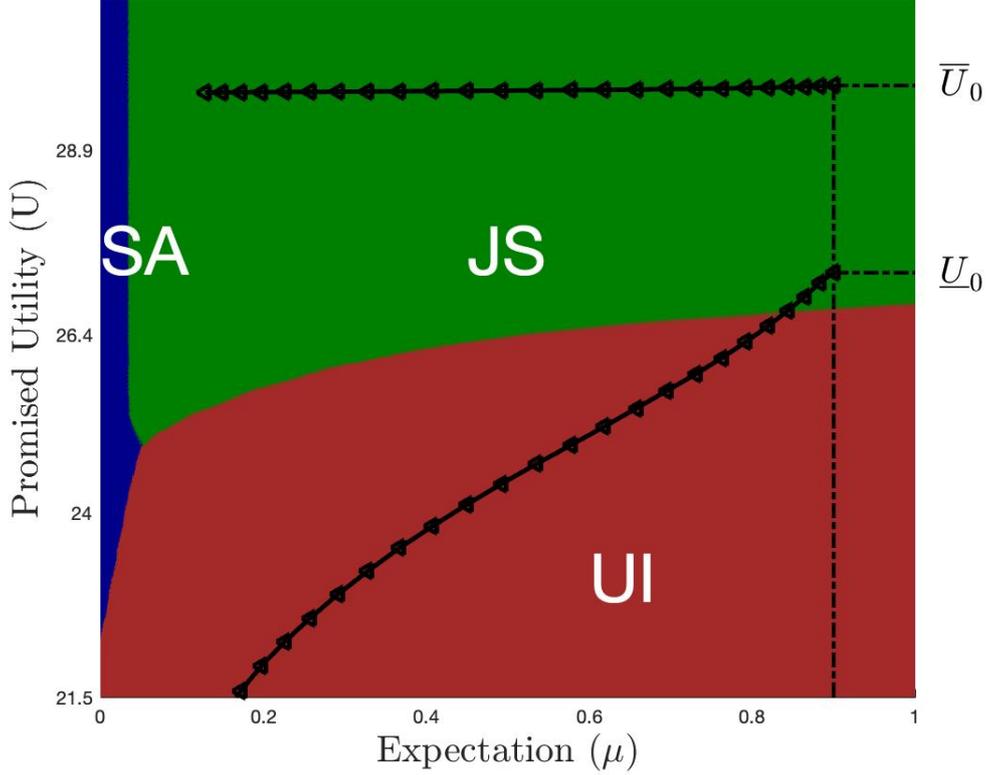


Figure 2: Optimal Policy Transitions in the Space of Expectation and Generosity.

a fraction of high-skilled workers into the ‘Fail’ outcome. The probability schedule brings both the worker (i.e., the profilee) and the government (i.e., the profiler) to revise prior expectation μ upon observing the public outcome r , following a Bayesian updating rule

$$\mu_r \equiv \frac{\mu\sigma(r|H, \mu)}{\mu\sigma(r|H, \mu) + (1 - \mu)\sigma(r|L, \mu)}$$

A necessary and sufficient condition for profiling to induce a change in expectations is to make the probability of each outcome dependent on the underlying human capital realization (e.g., $\sigma(r|H, \mu) \neq \sigma(r|L, \mu)$). In addition, profiling does not create any type of bias in the aggregate, since expectations are correct on average. Which boils down to require that the revised expectations are equal in mean to the prior (so called Martingale Property, (MP) henceforth).

$$q\mu_p + (1 - q)\mu_f = \mu, \quad \mu_f, \mu_p \in [0, 1], \quad \mu_f \leq \mu \leq \mu_p^{18} \quad (\text{MP})$$

where q is the probability that a ‘Pass’ is returned to any worker whose initial expectation is equal to μ . (MP) can be interpreted as a restriction to guarantee the credibility of profiling. Indeed, considering all workers who share the same initial expectation μ , some must be induced to

¹⁸Without loss of generality, the posterior upon ‘Fail’ (μ_f) is set to be lower than the posterior upon ‘Pass’ (μ_p).

revise their expectation down to μ_f upon ‘Fail’, for someone else to revise their expectation up to μ_p upon ‘Pass’. Given the nature of this profiling methodology, any profiled worker could pretend to be low-skilled by intentionally failing the test. It is thus necessary that the continuation utility upon ‘Pass’ is non smaller than the continuation utility upon ‘Fail’. If otherwise, indeed, the worker would retain the same expectation she was holding before undergoing profiling, since an intentional ‘Fail’ is not informative about human capital. Such a requirement, labelled No Discrimination constraint (ND, hereafter), imposes a restriction on the contract offered under either profiling outcome. In principle, the worker’s utility for intentionally failing the test has a complex expression, as no learning occurs for her in this case and transfers must be evaluated with the initial expectation μ . The problem arises whenever expectations determine the generosity of payments upon ‘Fail’. For instance, if the worker is requested to search right after, the consumption dispersion present in the UI contract is weighed differently as

$$U_f^u = u(c^{ui}) - e + \beta[\pi(\mu_f)U_{ui}^w + (1 - \pi(\mu_f))U_{ui}^u(\mu_f')] < u(c^{ui}) - e + \beta[\pi(\mu)U_{ui}^w + (1 - \pi(\mu))U_{ui}^u(\mu')]$$

However, the following result applies.

Proposition 5 (No Type-I Error). *Define type-I error as the occurrence where a high-skilled worker is profiled as low skilled.¹⁹ Therefore, at the optimum the probability of type-I error is null ($\mu_f = 0$), that is, only low-skilled workers are given a ‘Fail’ and referred to Social Assistance henceforth.*

Proof. See [Appendix B: Properties of AP, SP and IP](#). ■

Any disclosure of new information via profiling is equivalent to a randomization over the space of expectations with mean equal to the prior μ . Therefore, the government finds it convenient to implement profiling whenever it can exploit the between-policy convexity generated by the different policy slopes as in (8). The linearity of V^{SA} in μ causes the optimal profiling strategy to refer only low-skilled workers to SA ($\mu_f = 0$). Intuitively, if the threshold for ‘Pass’ eligibility is high, lowering it reduces the chances that high-skilled workers receive the ‘Fail’ outcome, while contemporaneously increasing the frequency of ‘Pass’ (q) and the ability of ‘Fail’ to signal low human capital. Lowering μ_f produces a constant marginal loss (by linearity of V^{SA} in expectations), as well as a gain from referring more workers to direct search (higher q). And since the marginal gain produced by the rise of q is equal to the difference between the program’s return in μ_p and μ_f , hence larger than the slope of V^{SA} by between-policy convexity of \mathbf{V} , the gain of a smaller μ_f always outweighs the loss. Hence, a null type-I error generates a better

¹⁹Profiling can be framed as a binary hypothesis test. To this purpose, the null hypothesis is that the worker is high skilled. As a consequence, the type-I and type-II errors coincide with the occurrence of ‘false positive’ ($\sigma(f|H, \mu)$) and ‘false negative’ ($\sigma(p|L, \mu)$), respectively.

principal-agent incentive alignment *ceteris paribus*.

As a consequence of Prop. 5, since SA is an absorbing policy and no job search is conducted ever after, the generosity of payments does not depend on the expected job-finding probability of reemployment, nor on current and future expectations thereof. Therefore, (ND) constraint admits a closed-form expression.

Assistance-and-Profiling (AP). AP does not envisage any job search. Thus, the planner's problem reads

$$\begin{aligned}
V^{AP}(\mu, U) &= \max_{c^{ap}, (U_r^u)_{r=\{p,f\}}, \mu_p} -c^{ap} - \kappa^{wp} + \beta[q\mathbf{V}(\mu_p, U_p^u) + (1-q)\mathbf{V}(0, U_f^u)] \\
\text{sub: } U &= u(c^{ap}) + \beta[qU_p^u + (1-q)U_f^u] \quad (\text{PK}), \quad (\text{MP}) \\
U &\geq u(c^{ap}) + \beta U_f^u \quad (\text{ND})
\end{aligned}$$

Using (PK), one can rewrite (ND) as $U_p^u \geq U_f^u$. Absent (ND), the contract would equate the slopes of \mathbf{V} across the different outcomes (i.e., $\mathbf{V}_U(\mu_p, U_p^u) = \mathbf{V}_U(0, U_f^u)$). By (7), this would boil down to set

$$U_f^u \geq U_p^u \tag{10}$$

Intuitively, if the government expects to incur an extra cost for incentive provision later during the unemployment spell, then it would save on this cost by decreasing the promised utility of the worker who passes the test and may at some point be referred to UI. Condition (10) and (ND) constraint determine a contract that insures the agent's utility against the risk of profiling outcome ($U_f^u = U_p^u = U^u$).

Passing to the choice of μ_p , one may be tempted to guess that the 'Pass' outcome only targets high-skilled workers at the optimum (e.g. $\mu_p = 1$). However, it may be the case that the probability of type-II error is positive at the optimum. To clarify the point, I will make use of Fig. 1a and assume that, if incentive costs for having the agent search were null, also low-skilled workers would be required to actively search for a job in first-best.²⁰ In other words, the planner is willing to bear an extra cost for compensating low-skilled workers for the search effort (effort compensation), but not for the risk entailed by the search outcome (incentive costs). If so, the government might want to persuade a fraction of low-skilled workers to be high skilled and have them search at a rather low incentive cost, by pooling them in the same ('Pass') outcome with high-skilled ones. This would lead to the posterior μ_p to be smaller than 1, as now 'Pass' does not entirely disclose high human capital. What profiling is effectively doing is to make the

²⁰For instance, one may think of situations in which the limited liability constraint is no longer in place, as the government can infinitely punish the worker for shirking. In this case, (IC) would not hold any longer.

high-skilled workers in the pool a bit more pessimistic compared to the perfect signal ($\mu_p = 1$), while proportionally increasing the optimism of the low-skilled ones. Due to the convex shape of the incentive costs, such pooling might actually be optimal: the increase in the incentive costs paid to high-skill workers is more than compensated by the additional return of moving the low-skilled ones from passive to active search policy. The internal solution satisfies

$$\mathbf{V}_\mu(\mu_p, U^u) = \frac{\mathbf{V}(\mu_p, U^u) - \mathbf{V}(0, U^u)}{\mu_p} \quad (11)$$

Eq. 11 shows that μ_p does not depend on worker's initial expectations, which means that all profiled workers hold the same revised expectation after receiving a 'Pass'. The downside is that the value of information for the government is negative when workers' initial expectation is larger than μ_p , even with null administrative cost of profiling, as the low-skilled ones among them are overconfident about their reemployment chances and thus mistaken in a direction that is favorable to the government. Hence, disclosing any information about their actual human capital causes the planner a net loss.²¹

Search-Assistance-and-Profiling (SP). Whenever the planner jointly adopts assisted search and profiling technologies, its problem reads

$$V^{SP}(\mu, U) = \max_{c^{sp}, U^w, U_p^u, U_f^u, \mu_p} -c^{sp} - \kappa^{wp} - \kappa^{ja} + \beta[\pi(\mu)W(\mu, U^w) + q(1 - \pi(\mu_p))\mathbf{V}(\mu'_p, U_p^u) + (1 - q)(1 - \pi_L)\mathbf{V}(0, U_f^u)]$$

$$\text{sub: } U = u(c^{sp}) + \beta[\pi(\mu)U^w + q(1 - \pi(\mu_p))U_p^u + (1 - q)(1 - \pi_L)U_f^u] \quad (\text{PK})$$

$$U \geq u(c^{sp}) + \beta[\pi(\mu)U^w + (1 - \pi(\mu))U_f^u] \quad (\text{ND}), \quad (\text{MP})$$

As for the case of AP, the planner insures the worker against the risks related to profiling, by committing to a constant continuation utility (i.e., $U_p^u = U_f^u$). About the informativeness of the profiling strategy, the posterior expectation μ_p induced by 'Pass' outcome, is either 1 or solves

$$\mathbf{V}_\mu(\mu'_p, U^u) = \frac{\mathbf{V}(\mu'_p, U^u) - \mathbf{V}(0, U^u)}{\mu'_p} \quad (12)$$

Indeed, if in AP the randomization over expectations concavifies the upper envelope \mathbf{V} , now instead profiling informs about human capital only conditional on job-search failure. Therefore, optimal profiling in SP delivers the concave closure of $(1 - \pi(\mu))\mathbf{V}(\mu', U^u)$.

Insurance-and-Profiling (IP). When profiling is implemented together with direct search,

²¹The necessary and sufficient condition for a positive type-II error is that the planner prefers not to disclose information about human capital to any low-skilled worker who expects to be high skilled with (almost) 100% probability.

the planner's problem reads

$$\begin{aligned}
V^{IP}(\mu, U) = & \max_{c^{ip}, U^w, U_p^u, U_f^u, \mu_p} -c^{ip} - \kappa^{wp} + \\
& + \beta [\pi(\mu)W(\mu, U^w) + q(1 - \pi(\mu_p))\mathbf{V}(\mu_p', U_p^u) + (1 - q)(1 - \pi_L)\mathbf{V}(0, U_f^u)] \\
\text{sub: } & U = u(c^{ip}) - e + \beta [\pi(\mu)U^w + q(1 - \pi(\mu_p))U_p^u + (1 - q)(1 - \pi_L)U_f^u] \quad (\text{PK}) \\
& U \geq u(c^{ip}) + \beta [qU_p^u + (1 - q)U_f^u] \quad (\text{IC}), \quad U_p^u \geq U_f^u \quad (\text{ND}), \quad (\text{MP})
\end{aligned}$$

Like in SP, profiling selects a posterior upon 'Pass' which equals 1 or solves (12).

5.1 Optimal Profiling Policies in the (μ, U) Space

I am now ready to locate the optimal profiling policies in each region of the (μ, U) state space. First, no profiling policy is optimal for very high or very low expectations, as the administrative expenses (κ^{wp}) outweigh the prospective gains from relocating workers across different policy instruments. Indeed, under the assumption that expectations are correct on average, workers with very high or very low expectations are likely to be assigned to a suitable policy instrument already, and prospective benefits from reallocation are quite small. Second, there exists a correspondence between profiling policies and their welfare counterparts. Indeed, each profiling policy dominates the other two in a region of the space of expectations where the optimal welfare policy is the one that adopts the same job-search method.

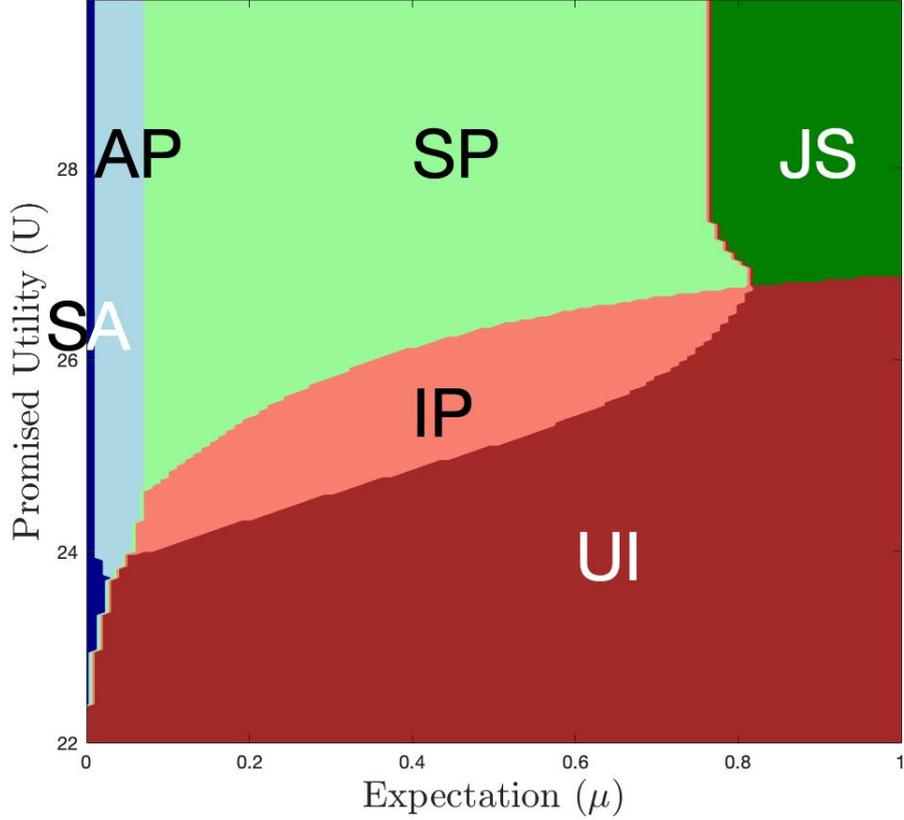


Figure 3: Optimal Policies in the Space of Expectation and Generosity.

Fig. 3 reports the optimal policies in the (μ, U) space. The choice of the optimal profiling policy mirrors the complementarity of search effort with expectations and its substitutability with promised utility. In particular, AP does not implement search and is thus optimal for lower-intermediate expectations. SP and IP, on the contrary, contemplate search, and are thus optimal for upper-intermediate expectations. As generosity rises (moving vertically from bottom to top of Fig. 3), higher costs of search-effort compensation and incentive provision (due to convex $1/u'$ ²²) lower the return of direct search (UI and IP), which is replaced by assisted search (JS and SP) for high-end expectations, and by no search (SA and AP) for low-end expectations.

An important aspect of a profiling policy is the optimal probability of type-II error. Indeed, as outlined in Eq. 11 and 12, the value of information to the planner is possibly negative when the within-policy concavity of \mathbf{V} in μ prevails over its between-policy convexity. If so, the optimal profiling strategy does not completely disclose high human capital ($\mu_p < 1$) and refers also a fraction of low-skilled workers to active policies, after pairing them with the ‘Pass’ outcome. Thus, any worker who receives a ‘Pass’ and is referred to an active policy can downward revise her expectation and reenter into IP, SP or AP at any later stage (unless she exits unemployment in the meantime). On the contrary, if profiling is fully accurate, workers do not revise expectation

²²See Lemma 1.

henceforth and never undergo profiling at any successive stage. The following result establishes a monotone relationship between profiling accuracy and program generosity.

Proposition 6 (Optimal Type-II Error). *Fix μ . Then, the probability of type-II error is decreasing in the level of generosity ($\partial\mu_p/\partial U \geq 0$). In particular, such probability is null under Search-Assistance and Profiling (SP), whenever no direct search is implemented in the downstream policy sequence after ‘Pass’.*

Proof. See [Appendix B: Properties of AP, SP and IP](#). ■

The intuition of the result hinges on the sensitivity of the slope of \mathbf{V} in μ to changes of U . As anticipated in Prop. 2, \mathbf{V} is within-policy supermodular and between-policy submodular. Looking at the determinants of μ_p , within-policy supermodularity produces an increase in the left-hand side of Eq. 11 and 12 in response to a rise in U , which is accompanied by a drop in the right-hand side due to between-policy submodularity. To reestablish equality between the two sides of the equation, μ_p must increase.

For high-end generosity, the optimal program never resorts to direct search, making the job-search return linear in μ , and worker’s utility does not fall over time (see Prop. 2 and 3). In such a case, human capital is fully revealed ($\mu_f = 0$, $\mu_p = 1$) and worker’s utility is kept constant over time ($U_{SP}^u = U$). The gains from profiling exclusively originate from the efficient match between workers and policy instruments, as high- (resp., low-)skilled workers are referred to Reemployment Services (resp., passive welfare support). This type of profiling is reminiscent of the WPRS program, where information on human capital is used as a criterion to allocate services. On the contrary, for low-end generosity, the government may use information design strategically to align incentives between itself and workers. A larger type-II error favors incentive alignment with the fraction of low-skilled workers who are classified as high skilled and referred to direct search, but at the same time causes incentive misalignment with high-skilled workers, who are made more pessimistic compared to the fully informative signal (no type-II error). Therefore, information on human capital is used also to fine-tune incentives and transfers, as it is the case in the REA program. Savings originating from this second form of profiling are even more relevant when profiling is not test-based and no opportunistic behavior is possible for profilees. Indeed, the possibly of referring workers to ‘more generous’ or ‘less generous’ programs based on their profiles is exploited to ease the cost of incentive provision on those who are classified as high skilled (see Section 6).

6 Statistical Profiling

The presence of (ND) constraint in test-based profiling prevents ‘punishment’ of workers who are referred to direct search upon it. However, this restriction does not apply whenever human capital is assessed via statistical methods. In particular, some programs profile welfare claimants by conducting a background check and gathering observable data, and estimate a probability of reemployment in accordance with the information collected, based on a large number of past observations. Given that the outcome of statistical profiling does not rely on worker’s commitment to it, (ND) constraint is not present in this new context. Consequently, the planner exploits the additional flexibility originating from the removal of ‘no-punishment’ restrictions as a leverage for incentive provision, with the target of reducing expected future transfers to recipients.

Profiling and reduction of transfers over time are two complementary instruments that open the way for sizable efficiency gains in the design of the optimal assistance program. Indeed, the planner now finds it optimal to index future transfers to the information detected by profiling. Therefore, the contract of any profiling policy is not only consisting of the lottery odds of each outcome, but also of the schedule of continuation utilities depending on it.

Proposition 7 (Optimal Statistical Profiling). *When profiling refers workers to Job-Search assistance (for high-end generosities), it admits no type-II error. When, instead, profiling refers workers to UI (for low-end generosities), the ‘Pass’ posterior is either 1 (i.e., no type-II error) or solves*

$$\mathbf{V}_\mu(\hat{\mu}, U_p^u) = \frac{\mathbf{V}(\hat{\mu}, U_p^u) - \mathbf{V}(0, U_f^u) + \mathbf{V}_U(0, U_f^u)(U_f^u - U_p^u)}{\hat{\mu}} \quad (13)$$

with $\hat{\mu} = \mu_p$ when no search is conducted (i.e., in AP) and $\hat{\mu} = \mu_p'$ when it is (i.e., in SP and IP). As in case of test-based profiling, the frequency of type-II error is decreasing in generosity (i.e., $\partial\hat{\mu}/\partial U > 0$).²³

Proof. See [Appendix C: Statistical Profiling](#). ■

The government sets different continuation utilities according to the profiling outcome. As shown in Prop. 2, the cost of incentive provision and effort compensation is increasing in generosity, which makes the marginal cost of U larger in UI than in SA. Hence, the government finds it

²³For $i = AP, SP$, statistical profiling at the optimum delivers the concave closure of \mathbf{V} over the two dimensions. Indeed, condition (13) can be rewritten

$$[\mathbf{V}_\mu(\mu_p, U_p^u) \quad \mathbf{V}_U(\mu_p, U_p^u)] \begin{bmatrix} \mu_p \\ \frac{U_p^u - U_f^u}{\|\mathbf{h}\|} \end{bmatrix}^T = \frac{\mathbf{V}(\mu_p, U_p^u) - \mathbf{V}(0, U_f^u)}{\|\mathbf{h}\|}$$

with $\|\mathbf{h}\| = \sqrt{\mu_p^2 + (U_p^u - U_f^u)^2}$.

optimal to lower the net discounted value of future payments upon ‘Pass’ (i.e., $U_p^u < U_f^u$). The result matches a characteristic of the actual REA program, where any worker who is found high-skilled is referred to minimum welfare support in the form of SNAP transfers up until reemployment. The criterion at the base of this rule is that any high-skilled worker does not need more generous transfers as she is likely to find reemployment soon.

The possibility to randomize over continuation utilities modifies the informativeness of the ‘Pass’ outcome. Eq. 13 strikes a new balance between incentive cost reduction of UI contract, the likelihood of being referred to it, and the new channel arising from the relaxation of the Incentive-Compatibility constraint.²⁴ Increasing informativeness, indeed, also increases the possibility of a ‘Fail’, conditional on which the planner pledges a larger utility. Hence, expected continuation utility for the agent is larger if the ‘Pass’ outcome is made more informative (and less likely) *ceteris paribus*, which allows the planner to further lower promised payments in order to restore contract efficiency (i.e., a binding (PK) constraint).

Proposition 8 (Statistical Profiling Contracts).

- *If the policy sequence after ‘Pass’ includes UI, utility upon ‘Pass’ is lower than current utility. Otherwise, it remains constant.*
- *Unemployment benefits fall over time in IP, and remain constant otherwise. In particular, in IP benefits fall to a larger extent once workers receive a ‘Pass’.*
- *In IP (resp., SP), the net wage upon reemployment is larger than (resp., equal to) current unemployment benefits.*

Proof. See [Appendix C: Statistical Profiling](#). ■

Fig. 4 plots the patterns of policies, expectation, utility and unemployment benefits for a worker who enters the program with initial expectation of $\mu_0 = 0.5$ and the promised utility whose consumption equivalent is $c(U_0) = \$113$. The worker is referred to direct search and profiled right upon entering the program. If she is found low-skilled, she is referred to SA with constant transfers. If, instead, she is found high-skilled, she keeps searching with transfers declining over time. Profiling does not fully disclose human capital upon ‘Pass’ (positive type-II error), and workers profiled as high skilled revise expectations along the spell.

²⁴The first two forces were already at play in the problem with (ND) constraint (see Section 4).

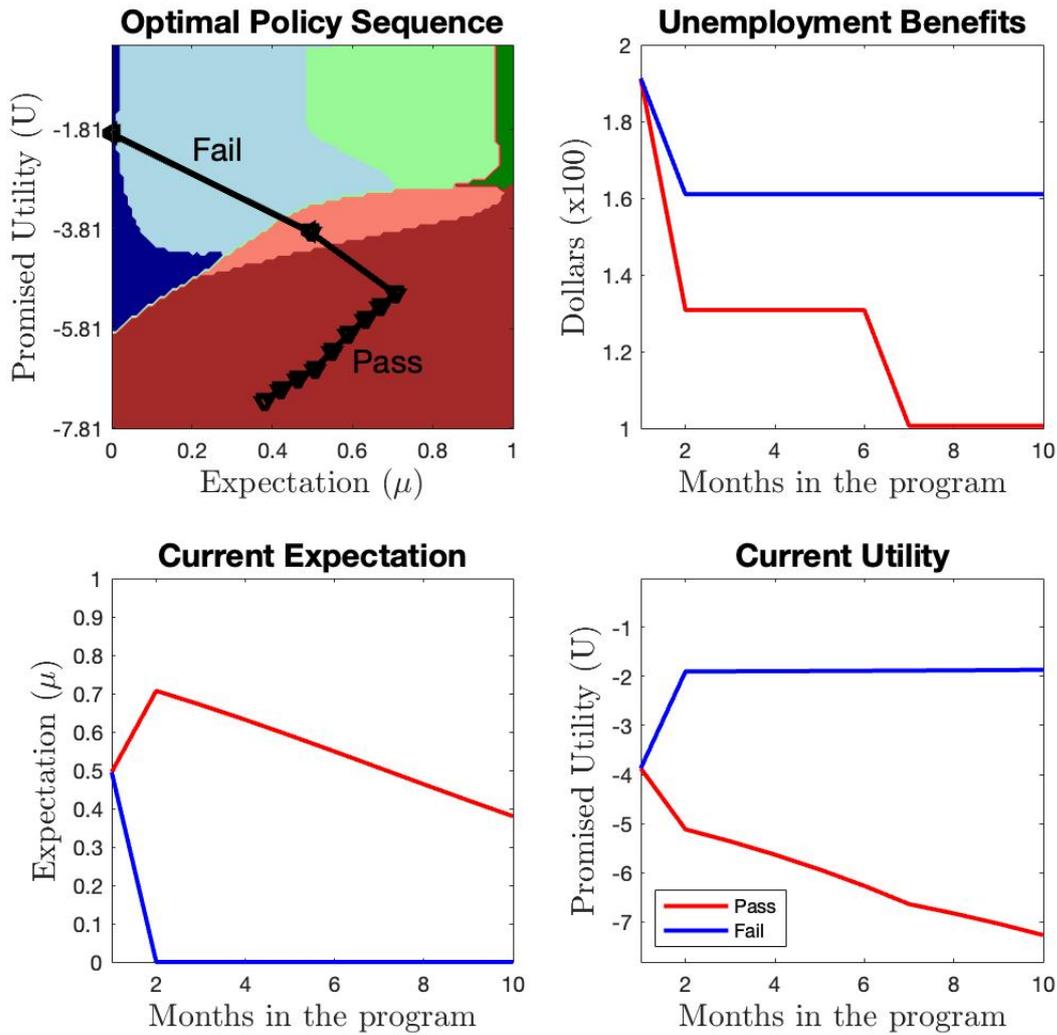


Figure 4: Consumption pattern upon profiling in a program with decreasing utility and $\mu_0 = 0.5$ and $c(U_0) = 100 \times u^{-1}((1 - \beta)U_0) = \113 .²⁵

7 Quantitative Analysis

The main objectives of this paper are to present optimal welfare programs with profiling, and to estimate welfare gains from their implementation. To this aim, this section estimates the return of optimal profiling in the State of Florida, by comparing it to an alternative scenario where only welfare policies SA, JS and UI are implemented. The estimated welfare gains of profiling equal the difference in returns between the two cases.

7.1 Program Description

Many welfare programs worldwide combine UI benefits, profiling and job-search assistance, in the attempt to improve compliance to program requirements and the effectiveness of job search

²⁵Monetary values are scaled down by a factor of 100.

(see Section 1). In US, for example, two are the operating programs that profile workers: the Worker Profiling and Reemployment Services (WPRS) and the Reemployment and Eligibility Assessment (REA). WPRS is a federally-mandated program that supplies job-search assistance to welfare claimants who face a high risk of benefit exhaustion prior to reemployment. REA is, instead, a voluntary program each State can opt in, whose goal is to reduce fraud and fund misallocation by excluding from UI benefits those recipients who either do not conduct any search activity, or do not need any form of welfare support (because they are highly re-employable). In other words, REA and WPRS differ in the use of information they collect with profiling, as with REA efficiency gains realizes via reduction of transfers, whereas with WPRS by implementing the job-search with proper methodology. To meet their target, both programs conduct an in-depth assessment of individual skills, based on which workers receive job-counseling, learn how to develop a resume and/or are directly referred to employers (see [Manoli et al., 2018](#)). Moreover, neither program allows workers enrolled in a training program to access any of these services.

Poe-Yamagata et al. (2011) conduct a cost-benefit analysis of the REA programs in Florida, Idaho, Illinois and Nevada, which assisted a total of 134,550 claimants in 2009. Of all claimants, 58% were men, 66% were white and 13% black. The report distinguishes between high and low skilled workers. The weighted mean share of high skilled participants is 48%.²⁶ Each State adopts a different selection process to refer claimants to either program. In particular, Florida randomly assigns eligible claimants to REA, PREP²⁷ or a control group, where no reemployment or profiling service is supplied ([Annex II](#)). While PREP participants are referred to group orientation where they receive information and are supplied job-search services, REA participants are required to undergo an in-person eligibility review at a public employment office.

7.2 Parameterization

Turning to to the choice of parameters (see [Table 2](#)), the parameters to be chosen are: the functional form of period utility ($u(\cdot)$), the discount factor (β), the effort cost of searching (e), the on-the-job productivity (i.e., the gross wage) and reemployment hazard rates of high- and low-skilled workers ($\{\omega_h, \pi_h\}_{h \in \{H, L\}}$), and the cost of administering profiling (κ^{wp}) and assisted search (κ^{ja}). The unit of time is set to one month.

I use a logarithmic specification of utility and set the monthly discount factor equal to $\beta = 0.9$. Based on [Pavoni et al. \(2013\)](#), the working effort cost is 49% of the consumption equivalent for men and 62% for women, corresponding respectively to $\bar{e}^m = 0.67$ and $\bar{e}^w = 0.97$ given the

²⁶The relative weight assigned to each State depends on the number of participants it assisted. In 2009, Florida, Idaho, Illinois and Nevada supplied UI to 80,531, 18,156, 3,112 and 32,751 jobless workers, respectively (Poe-Yamagata et al., 2011). The report does not make a distinction between high- and low-skilled workers in Illinois. However, this is not a source of major concern, given the small number of welfare recipients in the State.

²⁷In Florida, the WPRS program is called Priority RE-employment Planning (PREP).

Parameter	Symbol	Value	Source
Preferences			
Discount Factor	β	0.9	various sources
Search Effort Cost	e	0.27	
Labor Market			
Job Search Hazard	$\{\pi_H, \pi_L\}$	{0.27, 0.14}	basic monthly CPS, y. 2019 Poe-Yamagata et al. (2011) EIC, FICA
Net Wage	$\{c_H^w, c_L^w\}$	{\$2,498, \$1,128}	
Wage Tax	$\{\tau_H, \tau_L\}$	{\$178, -\$224}	
Worker Profiling			
Administrative Cost	κ^{wp}	\$50	Poe-Yamagata et al. (2011)
Assisted Search			
Administrative Cost	κ^{ja}	\$430	Balducchi and O’Leary (2018)
REA programs (FL, ID, IL, NV)			
Generosity (consumption equivalent)	c^i , $i = FL, ID, IL, NV$	[\$1,519,\$1,697]	Nicholson and Needels (2011)

Table 2: Choice of Parameters Value

logarithmic specification²⁸. And given that the percentage of male participants within the four programs is 58%, the working effort cost of the average participant amounts to $\bar{e} = 0.58\bar{e}^m + 0.42\bar{e}^w = 0.8$. Krueger and Muller (2010) conduct an analysis on the cost of search effort based on the American Time Use Survey (ATUS) and find that jobseekers spend on average 160 minutes every day looking for a job. Following Pavoni et al. (2013), I target the search effort to 1/3 (160/480) of the working effort, hence $e = 0.8/3 = 0.27$. The value is consistent with Pavoni et al. (2013), who estimate a cost of effort of $e = 0.22$.

Poe-Yamagata et al. (2011) reports data about net wages earned in the last 10 quarters prior to the start of UI claim. Quarterly wages in all States display a hump-shaped pattern, which increases until it reaches a peak about three quarters before displacement and steadily declines later on. The decline is consistent with the phenomenon known as *Ashenfelter’s dip*, suggesting that wages fall in the pre-layoff period (Ashenfelter, 1978). Preventing this effect from distorting estimates requires to exclude the last three quarters of pre-layoff wage. However, the paper does not consider human capital depreciation along the unemployed spell, which is instead well documented by the empirical literature (Keane and Wolpin, 1997; Neal, 1995) and requires to lower the prospective reemployment wage, in accordance with the duration of unemployment spell. As the two effects tend to offset each other, I simply consider the wage earned in the last quarter. As a consequence, the monthly net wages of Florida, Idaho and Nevada are \$1,833, \$1,367 and \$1,900, respectively.²⁹ The report, however, does not distinguish between wages of high- and low-skilled workers. Thus, I exploit the cross-sectional variation in wages and the share of high-skilled participants across States, reported by Poe-Yamagata et al. (2011). Given that there are two unknowns and three States, I compute $\{c_H^w, c_L^w\}$ as the pair that minimizes the

²⁸Logarithm allows for separation of consumption utility from working disutility in a natural way, according to the formula

$$\log((1 - \xi)c) = \log(c) + \log(1 - \xi) = \log(c) - \bar{e}$$

with $\xi \in \{0.49, 0.62\}$ being the consumption equivalent of working disutility.

²⁹Poe-Yamagata et al. (2011) does not report the percentage of high-skilled recipients in Illinois, which makes their data on wages useless for the estimation of $\{c_H^w, c_L^w\}$.

loss function

$$\Lambda(\hat{c}_H^w, \hat{c}_L^w) = \sum_{i=1}^3 \varphi_i (\theta_i \hat{c}_H^w + (1 - \theta_i) \hat{c}_L^w - c_i)^2, \quad i = \{FL, ID, NV\}$$

with φ_i being the fraction of welfare recipients in country i . The computation delivers monthly wages equal to $c_H^w = \$2,498$ and $c_L^w = \$1,128$. In order to compute their gross counterpart, I reverse engineer the gross labor income by computing the tax and deductibles from net amounts. In the US, employees are subject to the Federal Insurance Contribution Act (FICA) tax, which is comprehensive of Social Security and Medicare tax. FICA tax is a net payroll tax which is levied half on employers and half on employees, and amounted to 15.3% of Adjusted Gross Income (AGI) in 2009. Moreover, taxpayers are entitled to an Earned Income Credit (EIC), if their AGI is lower than a certain amount, depending on their marital status and number of children. Since no data on the marital status or the number of children of recipients is available, I assume that the representative recipient is married and has two children. Under 2009 FICA and EIC tax schemes, fiscal neutrality for a married couple with two children is achieved at a gross annual income of \$26,250, with the couple paying a tax (resp., receiving a subsidy) for an income above (resp., below) that threshold. Therefore, low-skilled recipients, whose net annual income is \$13,536, receive a tax credit under EIC, making their gross income lower than the net one, and precisely equal to \$10,844. High-skilled recipients, instead, have a gross income of \$32,112 and a net one of \$29,976.³⁰ Therefore, monthly gross wages are equal to $\omega_H = \$2,676$ and $\omega_L = \$904$.

I estimate the hazard rates $\{\pi_H, \pi_L\}$, using data from the basic monthly Current Population Survey (CPS). Following the method-of-moments estimation, the probability of reemployment after t periods is computed as the fraction of workers who exit unemployment in t . Reemployment probabilities are chosen as the ones that minimize the distance between the probabilities of reemployment so computed and the expected hazard rates, with weights given by the fraction of high- and low-skilled workers in the sample (for a more detailed description, see [Appendix E: Estimation of hazard rates](#)). The estimated monthly hazard rates are $\pi_H = 0.27$ and $\pi_L = 0.14$. The assumption that the worker can exit unemployment only upon search is quite extreme. I therefore assume the rate of reemployment in case no search is conducted either by the worker or by the government to be equal to half the after-search rate of low-skilled workers, i.e. $\hat{\pi} = 0.07$. A positive hazard rate in case of no search has a positive impact on the return of passive labor-market policies, like SA and AP, and a negative impact on the return of effort-incentivizing ones, like UI and IP, due to the increase in incentive costs. The value of off-the-equilibrium no-search

³⁰The net annual income of high- and low-skilled workers is $\$2,498 \times 12 = \$29,976$ and $\$1,128 \times 12 = \$13,536$, respectively. Low-skilled workers pay \$1,659 under FICA, i.e. the 15.3% of their gross income, but receive \$4,350 under EIC, hence receiving an annual subsidy of \$2,691. High-skilled workers, instead, pay a FICA tax of $15.3\% \times \$32,112 = \$4,913$, and are given a tax rebate of \$2,774, that account for an annual tax of \$2,139.

option, indeed, is higher with respect to the case with $\hat{\pi} = 0$, and the incentive constraint now requires satisfying a more stringent condition

$$U^w - U^u \geq \frac{e}{\beta(\pi(\mu) - \hat{\pi})}$$

Passing to the choice of κ^{wp} , the estimates of average per-capita cost of REA in 2009 contained in the report range from \$12 (Idaho) to \$134 (Illinois) and include cost of personnel and operative costs of centers supplying REA services (e.g., State Workforce Agencies and One-Stop Career Centers). I, therefore, set the administrative cost of profiling equal to the weighted average of REA per-capita cost among the four State programs, that is, $\kappa^{wp} = \$50$. Instead, the cost of assisting any worker in the job search is based on Balducchi and O’Leary (2018), who estimate $\kappa^{ja} = \$430$. Such a figure is consistent with other estimates ($\kappa^{ja} = \$500$ in Pavoni et al. (2016)), as well as cost estimates of programs that perform different activities (for instance, search monitoring), but feature a similar set of operations (regular meetings with personnel at One-Stop Career Centers, phone calls to employers, etc.). For instance, Pavoni and Violante (2007) estimate a monthly cost of monitoring of \$478 per claimant.

The generosity of any program depends both on the amount of flow endowments and the duration. Poe-Yamagata et al. (2011) collects data about the average maximum and weekly benefit in each State, as well as the distribution of benefit duration among participants. The weekly benefit amount ranges from \$234 in Florida to \$299 in Nevada, suggesting a substantial variability in the generosity of State programs. In the US, unemployment benefits are paid under four distinct schemes, which are activated in succession, depending on the current labor market situation of each State. Unemployment Insurance (UI) benefits last up to 26 weeks in all States. Workers who are still unemployed at the end of the 26th week, are entitled to additional 53 weeks under the Emergency Unemployment Compensation (EUC) scheme. Moreover, States pay additional benefits up to 20 more weeks under the Extended Benefits (EB) scheme, if their unemployment rate exceeds 8.5%, which was the case for all four States in 2009. Exhaustees of UI, EUC and EB are finally referred to the Supplemental Nutrition Assistance Program (SNAP), which replaced the Food Stamps Program in 2008. This constitutes a typical purely income-support measure of last resort, consisting of a constant allowance for the purchase of food, with no eligibility assessment or time limit. Transfers decline over time, as claimants move from one program to another. WPRS and REA never profile workers after they have exhausted UI, EUC and EB, as no assisted search or further transfer reduction is possible once the worker enters SNAP. I assume that workers who are entitled to 26 weeks of regular UI benefits are assisted under EUC and EB programs for the whole prospective duration of the programs, i.e. 73 weeks, and that exhaustees who are still unemployed at the end of UI+EUC+EB receive an endowment

from the Supplemental Nutrition Assistance Program (SNAP). The average total payment is \$7,930 under EUC and \$3,844 under EB (Nicholson and Needels, 2011), hence constituting a monthly endowment of $c^{EUC/EB} = \$645$, while a family of four people is receiving a \$501 monthly benefit from SNAP.³¹ The program's generosity for each of the four States is computed backward from the moment the welfare recipient enters into SNAP or finds reemployment, up until the first month of regular UI benefits. Worker's utility of reemployment in case she is high-(resp., low-)skilled amounts to³²

$$U_H^w = \frac{u(c_H^w)}{1-\beta} = \frac{\log(24.98)}{1-0.9} = 32.18 \quad U_L^w = \frac{u(c_L^w)}{1-\beta} = \frac{\log(11.28)}{1-0.9} = 24.23$$

while in SNAP with no search it is equal to

$$U_{e=0}^{SNAP}(\mu) = \frac{u(c^{SNAP}) + \beta\hat{\pi}[\mu U_H^w + (1-\mu)U_L^w]}{1-\beta(1-\hat{\pi})} = 22.32\mu + 19.25(1-\mu)$$

Condition $U_L^w > U_{e=0}^{SNAP}(0) + \frac{e}{\beta(\pi_L - \hat{\pi})}$ implies that the worker always finds it convenient to search also under SNAP. Hence, the expected generosity of SNAP can be rewritten as function of entry expectation μ

$$U^{SNAP}(\mu) = \mu \frac{u(c_{SNAP}) - e + \beta\pi_H U_H^w}{1-\beta(1-\pi_H)} + (1-\mu) \frac{u(c_{SNAP}) - e + \beta\pi_L U_L^w}{1-\beta(1-\pi_L)}$$

If the worker is entitled to regular UI, EUC and EB, then her assistance program lasts for $26+53+20=99$ weeks, that is, around 25 months. Starting from month 25, the following recursion is implemented

$$U_{i,j,t}^{WEL} = u(c_t^j) - e + \beta[\mu_i^{t-1}\pi_H U_H^w + (1-\mu_i^{t-1})\pi_L U_L^w + (1-\pi(\mu_i^{t-1}))U_{i,j,t+1}^{WEL}], \quad 1 \leq t \leq 25,$$

$$j = \{FL, ID, IL, NV\}, \quad i = \{< HS, HS, < CD, CD, GD\}$$

with $U_{i,j,26}^{WEL} = U^{SNAP}(\mu_i^{26})$, j indexing States and i indexing the educational attainment. The initial probability of being high-skilled, μ_i^0 , equals the share of high-skilled individuals with same educational attainment, ϑ_i . $U_{i,j,1}^{WEL}$ represents the entry promised utility of the regular assistance program for people who are never profiled over the spell. If instead they do undergo profiling after t periods of benefits, their promised utility reads

$$U_{i,j,t}^{REA} = \mu_i^t \frac{u(c_{SNAP}) - e + \beta\pi_H U_H^w}{1-\beta(1-\pi_H)} + (1-\mu_i^t)U_{i,j,\infty}$$

The generosity of the program is the weighted average between the promised utilities in the two

³¹See SNAP Data Tables at: <https://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap>.

³²All monetary amounts are normalized so that 1 consumption unit corresponds to \$100.

alternative scenarios. According to Poe-Yamagata et al. (2011), welfare claimants are randomly assigned to REA after some time spent in the program, that ranges from 2-3 (Nevada) up to 4-6 weeks (Florida), and are scheduled for REA 2 weeks after the assignment. Hence, the expression for the program's generosity reads

$$U_{i,j,\bar{t}_j} = u(c_{\bar{t}_j}^j) - e + \beta[\mu_i^{\bar{t}_j-1}\pi_H U_H^w + (1 - \mu_i^{\bar{t}_j-1})\pi_L U_L^w + (1 - \pi(\mu_i^{\bar{t}_j-1}))(\nu_j U_{i,j,\bar{t}_j+1}^{REA} + (1 - \nu_j)U_{i,j,\bar{t}_j+1}^{WEL})]$$

$$U_{i,j,t} = u(c_t^j) - e + \beta[\mu_i^{t-1}\pi_H U_H^w + (1 - \mu_i^{t-1})\pi_L U_L^w + (1 - \pi(\mu_i^{t-1}))U_{i,j,t+1}], \quad 1 \leq t \leq \bar{t}_j = 1/2$$

where \bar{t}_j is the actual timing of REA implementation and ν_j is the fraction of REA-eligible claimants who actually benefit from the program in State j .³³ The generosity levels of each program and educational attainment, expressed in consumption-equivalent terms, are reported in Table 3. Unsurprisingly, the generosity of the program is increasing in the level of educational attainment, due to higher initial expectations and $U_H^w > U_L^w$. Among the four States, Nevada (resp., Florida) is the most (resp., least) generous one for all levels of education.

States	Less Than HS	HS Diploma	Some College	College	Graduate
Florida	\$ 1,288	\$ 1,447	\$ 1,485	\$ 1,623	\$ 1,675
Idaho	\$ 1,458	\$ 1,628	\$ 1,668	\$ 1,816	\$ 1,872
Illinois	\$ 1,418	\$ 1,562	\$ 1,596	\$ 1,718	\$ 1,763
Nevada	\$ 1,476	\$ 1,647	\$ 1,687	\$ 1,835	\$ 1,890

Table 3: Program generosity for any State and educational level (consumption equivalent).³⁴

7.3 Optimal REA Program

Workers are assessed via in-person interviews with REA staff. When issuing the call for the interview, States target those who are predicted to be likely to exhaust their UI benefit. The assessment is based on interviews, that last 45 min on average. After the interview, workers profiled as high skilled suffer a reduction of unemployment benefits. Therefore, the profiling methodology adopted in the REA program is statistical, as it allows 'punishment' of high-skilled recipients.

Fig. 5 reports the optimal policies in the state-space of programs' generosity and expectation, and locates Florida's welfare claimants on it according to their initial (μ_0, U_0) .³⁵ Recipients with

³³Poe-Yamagata et al. (2011) report that in 2009 ν_j ranged from 4% in Idaho to 99% in Illinois (see Exhibit 5, p. 16).

³⁴The expression of consumption equivalent of U is

$$c(U) = \exp((1 - \beta)U)$$

³⁵The initial generosity of welfare programs in the other three States is quite similar.

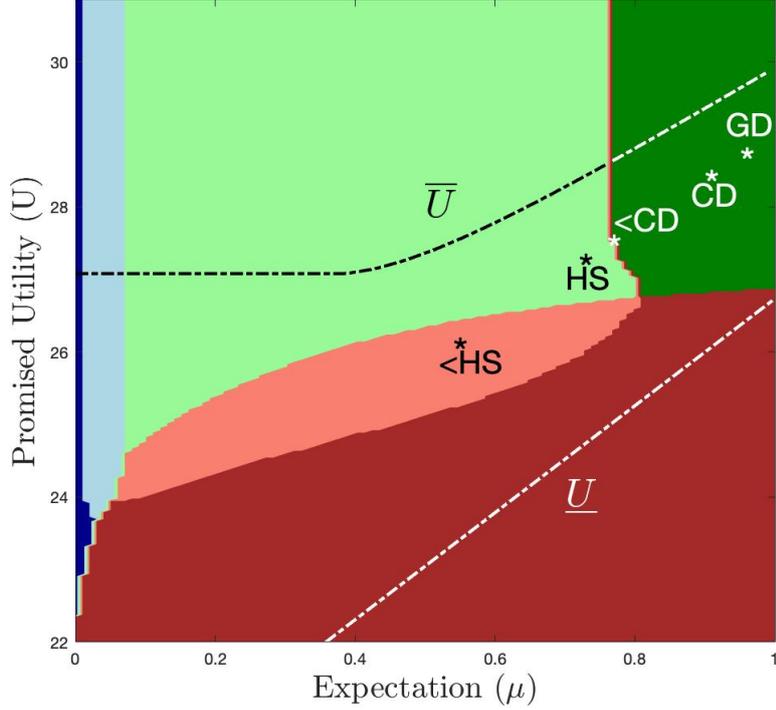


Figure 5: Optimal Policies for Florida's REA Recipients.

Note: <math><HS</math>=Less Than High School, HS=High School Diploma, <math><CD</math>=Some College, CD=College Degree, GD=Graduate Degree

less than a high-school degree should better search and at the same time undergo profiling right upon entry into the program. Those with a higher educational attainment should be initially assisted in the search and profiled as they enter the program (high-school graduates) or after some failed attempts by the government to find them a job (all those who have attended college, even with no final diploma). Contrary to the common idea, the ones who are assisted in the search are those recipients whose search would have a larger expected return. The reason is that a larger $\pi(\mu_0)$ is correlated with larger consumption upon reemployment and thus with higher implicit utility U_0 , which in turn makes graduates' effort too expensive to compensate for the government. The figure displays also upper and lower bounds for the initial location of claimants in the state space. The two lines, indeed, represent the initial generosity of programs that pay perpetual monthly instalments of \$1,500 and \$500 to claimants with initial expectation displayed on the x axis.³⁶ The upper bound becomes flat for low-end expectations, as claimants refuse to search when their reemployment perspectives are low enough.

³⁶The value of $\bar{U}(\mu)$ is defined as

$$\bar{U}(\mu) = \max \left\{ \frac{u(\bar{c})}{1-\beta}, u(\bar{c}) - e + \beta [\mu\pi_H U_H^w + (1-\mu)\pi_L U_L^w + (1-\pi(\mu))\bar{U}(\mu')] \right\}$$

with $\bar{c} = \$1,500$. An analogous definition holds for \underline{U} with $\underline{c} = \$500$.

7.4 Welfare Gains

A relevant aspect for policymakers is the savings that can be realized from adopting an optimal profiling program. In order to estimate such value, I compare two distinct programs, one with only welfare policies SA, UI and JS (\mathcal{W}) and the other featuring all six policies (\mathcal{P}). Fig. 6 reports the optimal patterns of promised utility, unemployment benefits and wage taxes/subsidies in the both \mathcal{W} and \mathcal{P} programs for Florida’s jobseekers with a high school degree (i.e., the most numerous group, accounting for 54% of all recipients in Florida in 2009), whose initial expectation and promised utility (in consumption-equivalent terms) are $\mu_0 = 0.72$ and $c(U_0) = \$1,447$, respectively. As shown in Fig. 5, this group of jobseekers are profiled under SP right after entering \mathcal{P} and referred to UI and SA, while in \mathcal{W} they never exit UI. The pattern of promised utility and unemployment benefits is thus falling over time for all workers in \mathcal{W} and only for high-skilled ones in \mathcal{P} (Fig. 6). However, the larger incentive costs in the latter case lead to a steadier decline in both benefits and utility. The reemployment tax displays a monotone increasing pattern in \mathcal{P} upon profiling (in the weak sense for low-skilled workers) due to the declining promised utility. In \mathcal{W} , instead, this component is contrasted by declining expected productivity (and gross wage thereof) of workers and increasing incentive costs, as expectations are revised downward. For this reason, the pattern of expected wage tax upon reemployment is non-monotonic over time.

Table 4 reports the per-capita welfare gain of profiling for each educational group, computed as a difference between the return of the welfare program with ($V_{\mathcal{P}}$) and without profiling ($V_{\mathcal{W}}$). The estimates increase in the educational level of the recipient and range from \$1,026 (less than high-school diploma) to \$1,305 (graduate). These figures are up to three times larger than the estimate of net per-capita savings in Poe-Yamagata et al. (2011), that amounts to \$395.³⁷ This is evidence that the current program design has large room for improvement. Given the relative size of each educational level and the number of REA-eligible workers in 2009 (459,021),³⁸ the annual aggregate welfare gain of Florida would have been \$540.4 millions. The expected total number of claimants to be profiled is 288,180, amounting to 62.8% of all welfare recipients, for an overall (discounted) cost of \$136,660. Therefore, I can conclude that every dollar spent on REA in Florida generates positive net savings equal to \$3,954.

	Less Than HS	HS Diploma	Some College	College	Graduate
Perc. in Program	0.13	0.54	0.17	0.12	0.04
$\pi(\mu_0)$	0.21	0.23	0.24	0.26	0.27
$c(U_0)$	\$ 1,288	\$1,447	\$1,485	\$1,623	\$1,675
Per-cap. Welfare Gain	\$1,026	\$1,128	\$1,140	\$1,250	\$1,305

Table 4: Welfare Gains of Profiling per Education Group in Florida (y. 2009)

³⁷The reader should recall that Poe-Yamagata et al. (2011) estimate the return of the current program, not of the optimal one.

³⁸Poe-Yamagata et al. (2011) report a number of 64,263 REA recipients in Florida in 2009, accounting for 14% of overall REA-eligible recipients (Exhibit 5, p. 16).

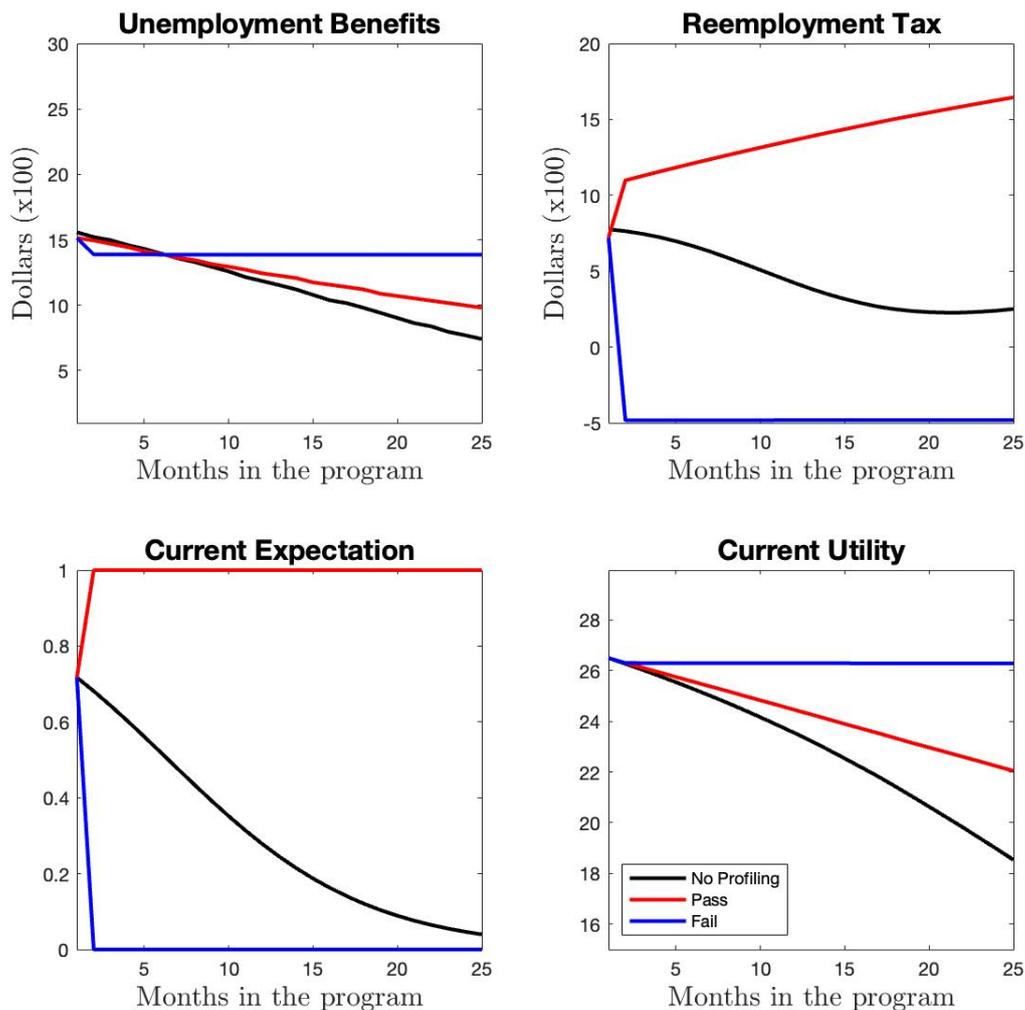


Figure 6: Optimal REA program of Florida for recipients with a high-school diploma over a 25-month horizon (UI+EUC+EB). Initial expectation and generosity are $\mu_0 = 0.72$ and $c(U_0) = \$1,447$, respectively.

7.5 Robustness Checks

A relevant dimension in which workers display a large heterogeneity is the search-effort cost. Various studies estimate different costs between men and women (Attanasio et al., 2008; Eckstein and Wolpin, 1989). In addition, they document the existence of a work-effort cost, which is not accounted for in the baseline model of this paper. I therefore conduct a robustness check by allowing for the search-effort cost to vary by $\pm 10\%$ with respect to the baseline value. Second, I assume that the reemployed worker incurs a working cost equal to the search-effort one. Optimal policies in the (μ, U) space for $e = 0.24$, $e = 0.3$ and $e^w = e = 0.27$ are reported in Fig. 7. When the search-effort cost is lower, then search-incentivizing policies (UI and IP) expand their areas at the expense of search-assisting ones (JS and SP). The opposite occurs when the search-effort cost is larger and all groups of recipients are offered search assistance upon entry. Positive

working-effort cost $e^w = 0.27$ produces a comparative disadvantage for active labor-market policies, as reemployment (and effort compensation upon it) is less likely in SA. Therefore, the gain from reallocating workers across SA and JS is larger the higher is the working-effort cost, via relaxation of the Promise-Keeping constraint. For this reason, SP replaces JS for high-end generosity and expectations (see Fig. 7c).

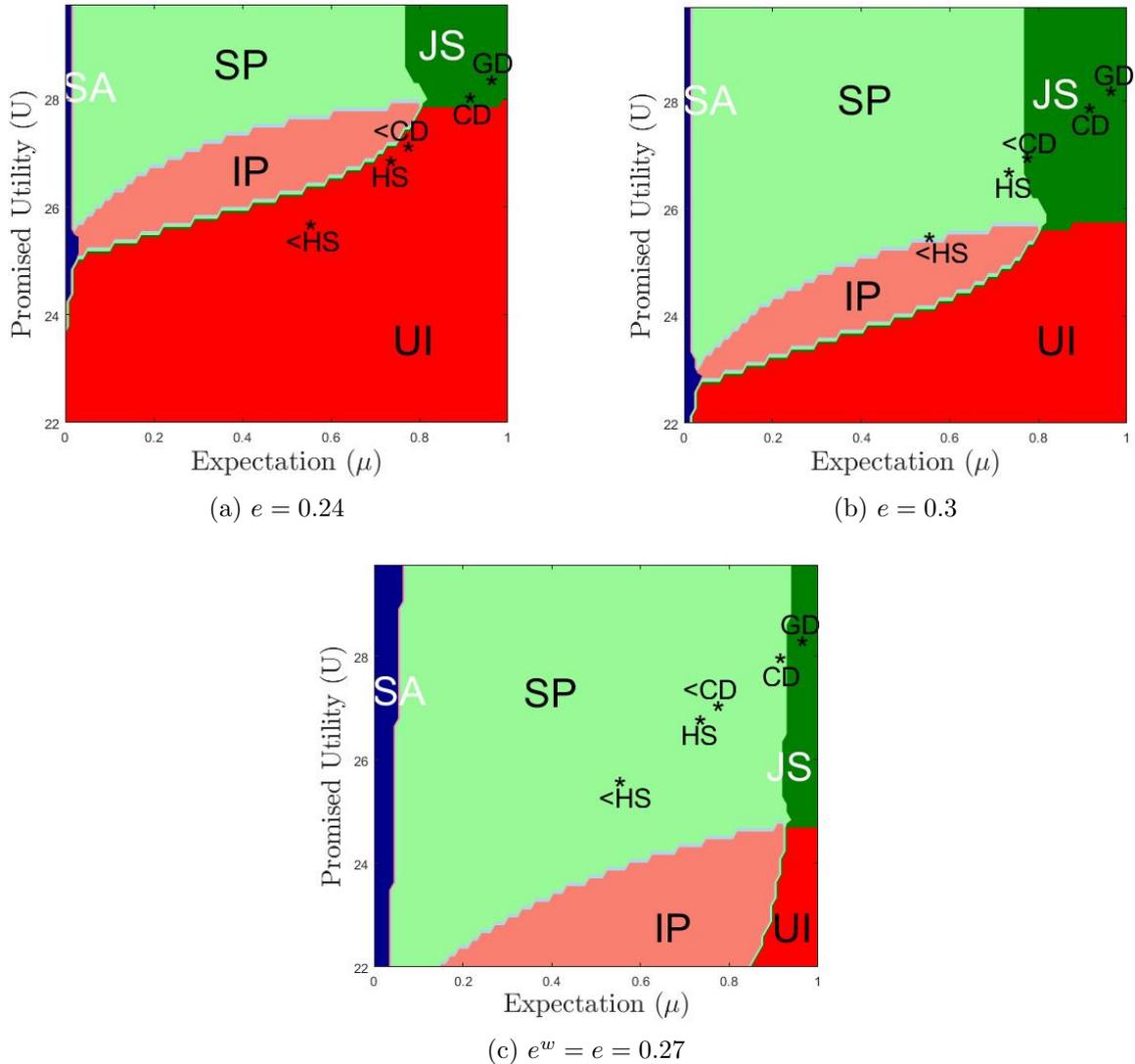


Figure 7: Optimal Policies over the (μ, U) Space for Different Effort Costs

A second parameter displaying great variability is the cost of assisted search and profiling technologies. Poe-Yamagata et al. (2011) estimate that in Nevada, where reemployment and profiling services are adopted together, κ^{ja} equals \$148. Such a figure is consistent with past estimates of administrative costs of assisted search. For example, Pavoni et al. (2013) compute an average cost of \$150 per person. Fig. 8a and 8b report the state space of policies for $\kappa^{ja} = \$387$ (-10%) and $\kappa^{ja} = \$473$ (+10%). When assisting workers is less costly, JS and SP are optimal also for lower generosity, and the opposite occurs when job-search assistance is more costly and incentivizing workers is more convenient also for higher effort-cost compensation.

The cost of profiling varies according to the design of the REA program. Given that all possible levels of accuracy are allowed, I select the largest cost among all four States, i.e. $\kappa^{wp} = \$134$ (Illinois). As a consequence, the areas of SP and IP shrink in favor of the respective welfare policies, JS and UI (see Fig. 8c).

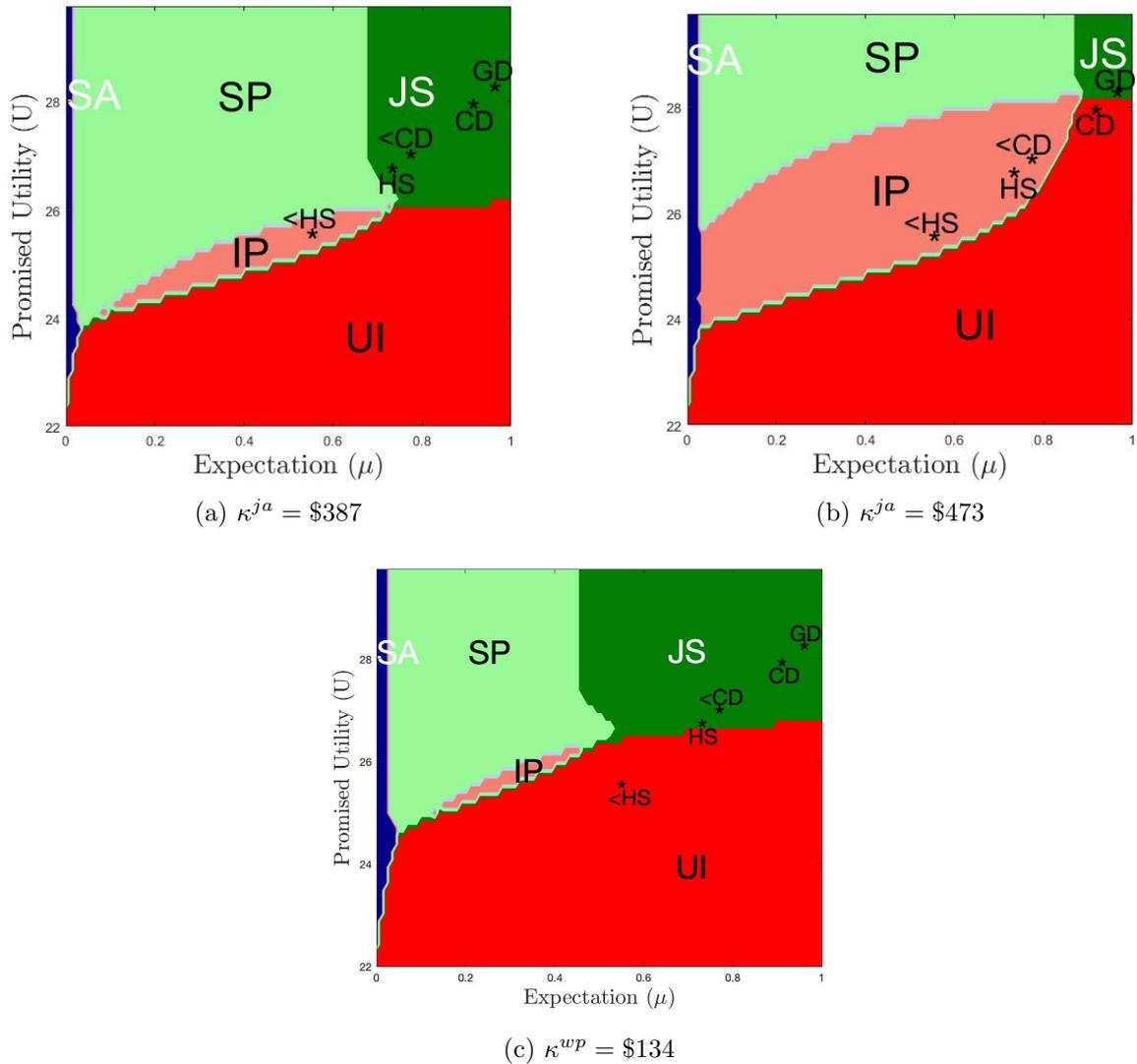


Figure 8: Optimal Policy Space for Different Costs of Assisted Search and Worker Profiling

8 Private Search and Moral Hazard

The government may be unable to observe worker's actions. In particular, worker's search may be private and thus unobservable for it. This may create a misalignment of expectations between the two parties, whenever the agent shirks effort and thus derives no new information about her human capital, while the principal assumes permanence in unemployment to be the result of a failed search and revises its expectations accordingly. The next period's contract provides larger incentives, consistently with the failed-search hypothesis. Therefore, as the worker retains a more optimistic expectation than the on-equilibrium one, she also expects larger transfers.

As a consequence, with private search the worker derives an additional advantage from shirking effort, other than sparing its cost. To contrast it, the planner promises larger transfers upfront in case of re-employment. Such high-powered incentives are increasing in the prospective duration of private search, following the rise in the number of possible deviations from recommended effort by the agent, and are driven to zero if the next-period contract does not request private search. Indeed, the agent can benefit from deviation only if a dispersion in promised transfers exists in the two alternative scenarios of reemployment and unemployment.

In consequence to the possibility of a covert misalignment of expectations, the planner incurs payment of *learning rents* to induce agent's search. The value of UI and IP thus incorporates such rents. Therefore, the next step consists of designing a contract which is robust to any possible deviation from $t = 0$ (today) to $t = T$ periods ahead. The worker could deviate in the first, second, ... T -th period after being assigned to UI. But she may also decide to deviate multiple -possibly not successive- times, before reverting to search, or even shirk forever after. For this reason, the design of a robust contract is in principle a complicated task. However, the following result enables to write the (IC) constraint in a compact way, thus making the problem well defined.

Proposition 9. *Any contract incentivizing search effort for T periods is robust against any possible deviation, whenever it is robust against one-shot deviations from the sequence of efforts.³⁹ Therefore, the (IC) constraint when implementing UI for T periods reads*

$$U \geq u(c^{ui}) + \beta[U^u + \Lambda(T, \mu)] \quad (14)$$

with $\Lambda(T, \mu)$ defined by the recursion

$$\Lambda(1, \mu^{T-1}) = 0 \quad (15)$$

$$\Lambda(T - j, \mu^j) = \left(\frac{\pi(\mu^{(j)})}{\pi(\mu^{(j+1)})} - 1 \right) e + \beta \pi(\mu^{(j)}) \left(\frac{1}{\pi(\mu^{(j+1)})} - 1 \right) \Lambda(T - j - 1, \mu^{(j+1)}), \quad 0 \leq j \leq T - 2$$

which is (i) independent of U^u , (ii) null in $\mu \in \{0, 1\}$ and/or $T = 1$, and (iii) increasing in T .

Proof. See [Appendix D: Moral Hazard](#). ■

³⁹This result is reminiscent of the way Euler equations are derived. Indeed, Euler equations are conditions imposed on the path of controls (consumption, investment, etc.), which guarantee that the decision maker is never willing to select any different path lying in the feasibility set and differing from the optimal one in one period only. The same property holds for Nash equilibrium strategies in repeated games, which are so if robust to deviations at any single node of the game tree.

The problem of the planner when private search ends in T periods reads

$$V_T^{UI}(\mu, U) = \max_{c^{ui}, U^w, U^u} -c^{ui} + \beta[\pi(\mu)W(\mu, U^w) + (1 - \pi(\mu))V_{T-1}^{UI}(\mu', U^u)]$$

$$\text{sub: } V_0^{UI}(\mu, U) := \max_{i \in \{SA, JS, SP, AP\}} V^i(\mu, U), \quad (\text{PK}) + (14)$$

The only difference with the case of observable effort (other than the (IC) constraint) is constituted by the restriction on the basket of policies to choose among in the next period, in order to be consistent with the current provision of learning rents. The dispersion in utilities between re-employment and unemployment now reads

$$U^w - U^u = \frac{e + \beta\Lambda(T, \mu)}{\beta\pi(\mu)}$$

The gap between U^w and U^u , which proxies the cost of incentives, is increasing in Λ . If UI is implemented for $t < T$ periods, then learning rents are lower and the planner incurs a lower contract cost.

Passing to IP, any worker can be profiled at any stage of the unemployment spell, possibly multiple times. If IP is designed to fully reveal the underlying state (zero type-I and type-II errors), the worker is certain to be high-skilled after receiving a ‘Pass’ and low-skilled otherwise. Hence, she does not revise her expectation henceforth, even if she fails to exit unemployment in the next stages of the welfare program. In other words, the worker whose human capital is detected with full precision does not undergo profiling a second time. In case profiling is not designed as a perfect signal, instead, the worker who passes it and is referred to any active policy can downward revise her expectation and reenter into IP at a later stage (unless she escapes unemployment in the meantime).

Proposition 10. *When worker’s search is private, accuracy of profiling under IP is determined by the need to reduce learning rents, as the ‘Pass’ posterior is either 1 or solves*

$$V_\mu(\mu'_p, U_p^u) = \frac{V(\mu'_p, U_p^u) - V^{SA}(U_f^u) + V_U^{SA}(U_f^u)(U_f^u - U_p^u)}{\mu'_p} + \quad (16)$$

$$+ [V_U^{SA}(U_f^u) - W_U(\mu, U^w)] \frac{\mu_p \Lambda_\mu(t, \mu_p) - \Lambda(t, \mu_p)}{\mu'_p}$$

Proof. See [Appendix D: Moral Hazard](#). ■

The result sheds light on the complementarity between profiling and private search. Indeed, in the case of AP and SP, the posterior upon ‘Pass’ is so chosen to equate the marginal gain from persuasion of low-skilled workers with the marginal cost of reducing profiling informativeness and incurring larger incentive costs. When the job search is private, instead, the reduction of

learning rents is an additional determinant of μ_p .

9 Conclusions

This paper provides an estimate of the welfare gains that can be obtained in programs of unemployment assistance via profiling of recipients. The rationale for embedding profiling into a welfare program stems from the difficulty of inferring recipients' skills and employability. At the optimum, active labor-market policies and workers' expectations about personal skills are complementary. Workers who are likely to be low-skilled are thus provided income support only, while those who have moderate or high expectations of being high-skilled are supplied with job-search assistance or search incentives in the form of lower wage taxes or higher wage subsidies. Looking at program generosity, instead, search-incentivizing policies are adopted for low-end generosities, while search-assisting ones are adopted for high-end generosities, due to increasing costs of effort compensation. This causes the dynamic of worker's utility which is implicit in the stream of payments to be decreasing along the spell whenever incentives to worker's search are provided.

Worker profiling enables the government to align incentives with workers. The implementation of worker profiling divides into gains and losses. First, profiled workers are referred to the least costly job-search method, rather than pooled into the same policy and contract, irrespective of their actual skills. In case the government would like to have all of them looking for reemployment at zero incentive costs, incentive alignment may entail a noisy profiling strategy that classifies as high skilled also a fraction of low-skilled workers and persuade them to search at low incentive costs in the sense of Kamenica and Gentzkow (2011). The rationale at the base of noisy profiling (i.e., positive type-II error) is that low-skilled workers are sufficiently employable and the generosity of the program sufficiently low that, absent incentive costs, the government would compensate their search effort rather than assign them to a passive labor-market policy. Therefore, some of them might better become overconfident about their reemployment chances and be given search incentives, even at the cost of higher incentive costs to high-skilled workers. While noisy profiling is purely driven by a cost-efficiency argument, the government may find it convenient also for other reasons. First, a fully accurate profiling could be discriminatory against ethnic minorities, which are likely to be less employable and, for this reason, referred to passive policies. Thus, the government would rather prefer to classify (at least) a fraction of recipients who are part of a minority group, as likely to be reemployed and provide them with search incentives. Second, noisy profiling may alleviate the unemployment-related stigma effect on recipients who have been jobless for longer periods,⁴⁰ by signalling higher job-finding

⁴⁰A stigma effect on the long-term unemployed has been documented to cause negative duration dependence in the escape rate from unemployment (see Heckman and Borjas, 1980, and Krug et al., 2019).

chances to prospective employers. These and similar beneficial effects provide further reasons for devising a strategic level of ‘cheating’ in worker profiling.

Profiling is also used to fine-tune the generosity of downstream transfers, whenever it follows a statistical criterion. In particular, an optimal program promises lower transfers to recipients who are profiled as high skilled and required to search for job, due to agency costs increasing in the level of promised utility. This result, which features actual REA programs, should be accompanied with a decreasing pattern of unemployment benefits, as opposed to the constant subsidy under SNAP. On the contrary, workers classified as low skilled should be granted a constant unemployment benefit. If the contract costs of direct search increase in the level of utility promised to recipients, this fine-tuning argument is outweighed by the consumption-insurance one in highly generous programs, where WPRS-like programs replace REA-like ones.

Some questions remain unanswered. First, the paper abstracts from the trade-off between profiling administrative cost and accuracy. Nonetheless, actual per-capita costs of profiling depend on the accuracy with which skills are detected. A more accurate detection, indeed, leads to a more expensive profiling process (e.g., longer in-person interviews, more elaborate tasks to perform, etc.). In addition, any actual profiling program, as well as any sort of tests aimed at detecting a hidden characteristic, contains a given amount of unavoidable noise that impedes an exact detection of skills. Assuming (i) the cost of profiling to vary in accordance to the change induced on the initial expectation (i.e., known as *entropy cost* in the information design literature),⁴¹ and/or (ii) the accuracy of information detection to be upper bounded, might lead to different estimates of the value of worker profiling. In this sense, the welfare gain computed in Section 7 can be read as an upper bound on the return from adopting optimal profiling. A second aspect on which further inspection is required is the absence of any information asymmetry. Both parties are assumed to share the same initial expectation, as claimants truthfully report their personal data to the provider at the beginning of the program. However, if claimants could conceal their personal information, those with high expectations in reemployment would anticipate being requested to search and choose to misreport it, so as to benefit from larger incentives.⁴² If this is the case, the government would need to make search-incentivizing payments robust to information misreporting and this would further exacerbate the problem of incentive provision. In this sense, as noticed in the case of private worker search (Section 8), profiling might constitute a way for the government to curb the information rents that originate from the agency problem. Third, many existing welfare programs feature hybrid policies devoted to worker profiling and training. Training under any such policy can improve worker’s human capital. Therefore, the

⁴¹For a description of costs of information detection that allow the concavification result to survive, see [Kamenica and Gentzkow \(2014\)](#).

⁴²In no other case they would find convenient to lie, as all contracts other than incentive-providing ones are independent of expectations.

government needs to design a profiling program to provide workers with incentives to train, as well as to detect hidden information about their employability. Boleslavsky and Kim (2021) constitute a useful starting point to understand how government's incentives to train workers and to profile them interplay.

References

- (1) Ashenfelter, O., 1978, “Estimating the Effect of Training Programs on Earnings”, *Review of Economics and Statistics*, 60, 47-50;
- (2) Atkinson, A., Lucas R. E. J., 1995, “Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance”, *Journal of Economic Theory*, 66(1), 64-88;
- (3) Attanasio, O., Low H., Sánchez-Marcos V., 2008, “Explaining Changes in Female Labor Supply in a Life-Cycle Model”, *American Economic Review*, 98(4), 1517-1552;
- (4) Bergemann, D., Morris S., 2018, “Information Design: A Unified Perspective”, *Journal of Economic Literature*, 57(1), 44-95;
- (5) Black, D. A., Galdo J., Smith J. A., 2007. “Evaluating the Worker Profiling and Reemployment Services System Using a Regression Discontinuity Approach”, *American Economic Review*, 97(2), 104-107;
- (6) Boleslavsky, R., Kim K., 2021, “Bayesian Persuasion and Moral Hazard”, *mimeo*;
- (7) Cappellini, S., 2020, “Worker Profiling and Reemployment Services”, *mimeo*;
- (8) Card, D., Kluve J., Weber A., 2018, “What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations”, *Journal of the European Economic Association*, 16(3), 894-931;
- (9) Chetty, R., 2008, “Moral Hazard versus Liquidity and Optimal Unemployment Insurance”, *Journal of Political Economy*, 116(2), 173-234;
- (10) Eckstein, Z., Wolpin K. I., 1989, “Dynamic Labour Force Participation of Married Women and Endogenous Work Experience”, *Review of Economic Studies*, 56 (3), 375–390;
- (11) Fallick, B. C., 1996, “A Review of the Recent Empirical Literature on Displaced Workers”, *Industrial and Labor Relations Review*, 50:1, 5-16;
- (12) Fuller, D. L., 2014, “Adverse Selection and Moral Hazard: Quantitative Implications for Unemployment Insurance”, *Journal of Monetary Economics*, 62, 108-122;
- (13) Galperti, S., Perego J., 2018, “A Dual Perspective on Information Design”, Available at SSRN: <https://ssrn.com/abstract=3297406>;
- (14) Gonzalez, F. M., Shi S., 2010, “An Equilibrium Theory of Learning, Search, and Wages”, *Econometrica*, 78(2), 509-537;

- (15) Habibi, A., 2020, "Motivation and Information Design", *Journal of Economic Behavior and Organization*, 169, 1-18;
- (16) Hagedorn, M., Kaul A., Mennel T., 2010, "An Adverse Selection Model of Optimal Unemployment Insurance", *Journal of Economic Dynamics and Control*, 34(3), 490-502;
- (17) Heckman, J. J., Borjas, G., 1980, "Does Unemployment Cause Future Unemployment? Definitions, Questions, and Answers from a Continuous Time Model of Heterogeneity and State Dependence", *Economica*, 47, 247-283;
- (18) Hopenhayn, H. A., Nicolini J. P., 1997, "Optimal Unemployment Insurance", *Journal of Political Economy*, 105(2), 412-438;
- (19) Kamenica, E., Gentzkow M., 2011, "Bayesian persuasion", *American Economic Review*, 101(6), 2590-2615;
- (20) Kamenica, E., Gentzkow M., 2014, "Costly Persuasion", *American Economic Review: Papers and Proceedings*, 104(5), 457-462;
- (21) Keane, M. P., Wolpin K. I., 1997, "The Career Decisions of Young Men", *Journal of Political Economy*, 105 (3), 473-522;
- (22) Kirby, G., Hill H., Pavetti L., Jacobson J., Derr M., Winston P., 2002, "Transitional Jobs: Stepping Stone to Unsubsidized Employment", *Mathematica Policy Research Inc.*;
- (23) Kluve, J., Schmidt C.M., 2002, "Can Training and Employment Subsidies Combat European Unemployment?", *Economic Policy*, 35, 409-448;
- (24) Kolotilin, A., 2018, "Optimal Information Disclosure: a Linear Programming Approach", *Theoretical Economics*, 13, 607-635;
- (25) Krueger, A. B., Mueller A. I., 2010, "Job Search and Unemployment Insurance: New Evidence from Time Use Data", *Journal of Public Economics*, 94(3-4), 298-307;
- (26) Krug, G., Drasch K., Jungbauer-Gans M., 2019, "The social stigma of unemployment: consequences of stigma consciousness on job search attitudes, behaviour and success", *Journal of Labour Market Research*, 53(11);
- (27) Machin, S., Manning A., 1999, "The Causes and Consequences of Long-Term Unemployment in Europe", in O. Ashenfelter and D. Card (Eds), *Handbook in Labor Economics*, Vol. 3C, 3085-3139;
- (28) Manoli, D., Michaelides M., Patel A., 2018, "Long-Term Effects of Job-Search Assistance: Experimental Evidence Using Administrative Tax Data", *mimeo*;

- (29) Meyer, B. D., 1990, “Unemployment Insurance and Unemployment Spells”, *Econometrica*, 58(4), 757-782;
- (30) Michaelides, M., Mueser P., 2020, “The Labor Market Effects Of US Reemployment Policy: Lessons From An Analysis Of Four Programs During The Great Recession”, *Journal of Labor Economics*, 38(4), 1099-1140;
- (31) Neal, D., 1995, “Industry-Specific Human Capital: Evidence from Displaced Workers”, *Journal of Labor Economics*, 13 (4), 653–677;
- (32) Nicholson, W., Needels K., 2011, “The EUC08 Program in Theoretical and Historical Perspective”, *Mathematica Policy Research*, n. 06863.320;
- (33) Pavoni, N., Setty O., Violante G. L., 2013, “Search and Work in Optimal Welfare Programs”, NBER Working Paper No. 18666;
- (34) Pavoni, N., Setty O., Violante G. L., 2016, “The Design of ‘Soft’ Welfare-to-Work Programs”, *Review of Economic Dynamics*, 20, 160-180;
- (35) Pavoni, N., Violante G. L., 2007, “Optimal Welfare-to-Work Programs”, *Review of Economic Studies*, 74(1), 283-318;
- (36) Poe-Yamagata, E., Benus J., Bill N., Carrington H., Michaelides M., Shen T., 2011, “Impact of the Reemployment and Eligibility Assessment (REA) Initiative”, IMPAQ International, Task Order No. DOLF091A21507;
- (37) Rodina, D., 2020, “Information Design and Career Concerns”, CRC TR 224 Discussion Paper Series, University of Bonn and University of Mannheim, Germany;
- (38) Setty, O., 2019, “Optimal Unemployment Insurance with Monitoring”, *Quantitative Economics*, 10, 693-733;
- (39) Shavell, S., Weiss L. (1979), “The Optimal Payment of Unemployment Insurance Benefits over Time”, *Journal of Political Economy*, 87(6), 1347-1362;
- (40) Shimer, R., Werning I., 2008, “Liquidity and Insurance for the Unemployed”, *American Economic Review*, 98(5), 1922-1942;
- (41) Sullivan, W. F. Jr, Coffey L., Kolovich L., McGlew C. W., Sanford D., Sullivan R., 2007, “Worker Profiling and Reemployment Services. Evaluation of State Worker Profiling Models: Final Report”, U.S. Department of Labor, Employment and Training Administration: Office of Workforce Security;

- (42) Wang, C., Williamson S., 1996, “Unemployment Insurance with Moral Hazard in a Dynamic Economy”, Carnegie-Rochester Conference Series on Public Policy, 44, 1-41;
- (43) Zapechelnuk, A., 2020, “Optimal Quality Certification”, AER: Insights, 2(2), 161–176.

APPENDIX

Appendix A: Properties of SA, JS and UI

Proof of Prop. 1

Proof. Envelope Theorem and first-order conditions imply

$$V_U^{SA}(\mu, U) = -\frac{1}{u'(c_{sa})} = \mathbf{V}_U(\mu, U^u)$$

Now, given that SA is optimal in (μ, U) , then $\mathbf{V}_U(\mu, U) = V_U^{SA}(\mu, U) = \mathbf{V}_U(\mu, U^u)$, and concavity of \mathbf{V} in U implies that $U^u = U$. Therefore, the state space (μ, U) is equal in the next period, proving that SA is optimal forever after. ■

Proof of Lemma 1

Proof. The problem of policy $i \in \{SA, JS, UI\}$ reads

$$\begin{aligned} V^i(\mu, U) &= \max_{(z, U^w, U^u) \in \Gamma(\mu, U)} -g(z) - \kappa^i + \beta [p^i(\mu)W(\mu, U^w) + (1 - p^i(\mu))\mathbf{V}(\mu^i, U^u)] \\ \text{sub: } \Gamma^i(\mu, U) &= \left\{ (z, U^w, U^u) : U = z - e^i + \beta [p^i(\mu)U^w + (1 - p^i(\mu))U^u], U \geq z + \beta U^u \right\} \end{aligned}$$

with $p^{SA}(\mu) = 0$, $p^{JS}(\mu) = p^{UI}(\mu) = \pi(\mu)$ and

$$(e^i, \kappa^i) = \begin{cases} (0, 0) & \text{if } i = SA \\ (0, \kappa^{ja}) & \text{if } i = JS \\ (e, 0) & \text{if } i = UI \end{cases}$$

The following holds.

Lemma 2. V^i is decreasing in U and increasing in μ . Moreover, if \mathbf{V} is concave in either argument, then so is V^i .

Proof. To prove concavity of V^i in U/μ , it suffices to show that:

- the objective function is concave in the choice variables and U/μ ;
- the graph of the feasibility set is convex.

Simply notice that $g = u^{-1}$ is convex, and that W and V are concave in U^w and U^u , respectively. Moreover, while $\pi(\mu)W(\mu, U)$ is linear in μ , $(1 - \pi(\mu))\mathbf{V}(\mu', U)$ is concave in μ if \mathbf{V} is concave

in the first argument, as

$$-2(\pi_H - \pi_L) \frac{\partial \mu'}{\partial \mu} V_\mu(\mu', U) + (1 - \pi(\mu)) \left[\frac{\partial^2 \mu'}{\partial \mu^2} V_\mu(\mu', U) + \frac{\partial \mu'}{\partial \mu} V_{\mu\mu}(\mu', U) \right] < 0$$

as $(1 - \pi(\mu)) \frac{\partial^2 \mu'}{\partial \mu^2} = 2(\pi_H - \pi_L) \frac{\partial \mu'}{\partial \mu}$.

Furthermore, PK constraint is linear in U , z , U^w and U^u , and so is IC constraint, since U^i is linear in U . This means that the graph of Γ_μ^i (i.e., for constant μ) defined as

$$Gr\Gamma_\mu^i = \{(z, U^w, U^u, U) : U = z - e^i + \beta[p^i(\mu)U^w + (1 - p^i(\mu))U^u], U \geq z + \beta U^u\}$$

is convex. Same applies to the graph of Γ_U^i (i.e., for constant U), since PK and IC are linear in μ .

To prove (negative) monotonicity in U , one needs to show:

- (negative) monotonicity of the objective function in U ;
- (negative) monotonicity of the feasibility set Γ_μ^i in U , i.e.

$$U < \tilde{U} \implies \Gamma^i(\mu, \tilde{U}) \subseteq \Gamma^i(\mu, U)$$

The objective function does not directly depend on U , while monotonicity can be shown by rewriting the IC constraint as

$$U^w - U^u \geq \frac{e^i}{\beta p^i(\mu)}$$

which does not depend on U . Therefore, the PK constraint is tightened by an increase of U , which thus leads to a shrinkage of Γ_μ^i .

Proving (positive) monotonicity of V^i in μ is analogous. Indeed, it follows from:

- (positive) monotonicity of the objective function in μ ;
- (positive) monotonicity of the feasibility set Γ_U^i in μ , i.e.

$$\mu < \tilde{\mu} \implies \Gamma^i(\mu, U) \subseteq \Gamma^i(\tilde{\mu}, U)$$

The objective function is always monotone in μ , as so are W and V in their first argument, μ' is an increasing function of μ and $W(\mu, U^w) \geq V(\mu^i, U^u)$. Monotonicity of $\Gamma^i(\cdot, U)$, instead, holds as an increase of μ leads to a relaxation of (IC)⁴³. Indeed, (IC) is more slack since

$$\mu < \tilde{\mu} \implies U^w - U^u \geq \frac{e^i}{\beta p^i(\mu)} \geq \frac{e^i}{\beta p^i(\tilde{\mu})}$$

⁴³(PK) is always relaxed by an increase in μ (recall that $U^w \geq U^u$ at the optimum).

■

V^i and \mathbf{V} concave in U

The proof of concavity of V^i in U follows a recursive argument. Guessing concavity of \mathbf{V} in U , it holds that V^i is concave by Lemma 2 and that, from definition (2) of \mathbf{V} ,

$$\mathbf{V}_{UU}(\mu, U) = V_{UU}^i(\mu, U)$$

when no randomization over U is conducted. Otherwise, $\mathbf{V}_{UU}(\mu, U) = 0$.

V^i and \mathbf{V} concave in μ

The proof of concavity in μ follows the same steps, as assuming concavity of \mathbf{V} in μ leads to concavity of V^i . However, the second derivative of \mathbf{V} in μ reads

$$\mathbf{V}_{\mu\mu}(\mu, U) = qV_{\mu\mu}^i(\mu, \bar{U}) + (1 - q)V_{\mu\mu}^j(\mu, \underline{U}) < 0, \quad \text{with: } q = \frac{U - \underline{U}}{\bar{U} - \underline{U}}$$

V^{UI} and V^{JS} supermodular

The derivative of V^{UI} and V^{JS} wrt U reads

$$\begin{aligned} V_U^{UI}(\mu, U) &= -\frac{1}{u(c_{UI})} = \pi(\mu)W_U(\mu, U_{UI}^w) + (1 - \pi(\mu))\mathbf{V}_U(\mu', U_{UI}^u) \\ V_U^{JS}(\mu, U) &= -\frac{1}{u(c_{JS})} = W_U(\mu, U_{JS}^w) = \mathbf{V}_U(\mu', U_{JS}^u) \end{aligned}$$

Thus

$$\begin{aligned} V_{\mu U}^{UI}(\mu, U) &= (\pi_H - \pi_L)(W_U(\mu, U_{UI}^w) - \mathbf{V}_U(\mu', U_{UI}^u)) + \pi(\mu)W_{UU}(\mu, U_{UI}^w)\frac{\partial U_{UI}^w}{\partial \mu} + \\ &+ (1 - \pi(\mu))\mathbf{V}_{UU}(\mu', U_{UI}^u)\frac{\partial U_{UI}^u}{\partial \mu} + (1 - \pi(\mu))\frac{\partial \mu'}{\partial \mu}\mathbf{V}_{\mu U}(\mu', U_{UI}^u) \\ &= (\pi_H - \pi_L)(W_U(\mu, U_{UI}^w) - \mathbf{V}_U(\mu', U_{UI}^u) + W_{UU}(\mu, U_{UI}^w)(U_{UI}^u - U_{UI}^w)) + \\ &+ \frac{\partial U_{UI}^u}{\partial \mu}(\pi(\mu)W_{UU}(\mu, U_{UI}^w) + (1 - \pi(\mu))\mathbf{V}_{UU}(\mu', U_{UI}^u)) + (1 - \pi(\mu))\frac{\partial \mu'}{\partial \mu}\mathbf{V}_{\mu U}(\mu', U_{UI}^u) \end{aligned}$$

Convexity of $1/u'$ implies concavity of W_U , which boils down to

$$W_U(\mu, U_{UI}^w) + W_{UU}(\mu, U_{UI}^w)(U_{UI}^u - U_{UI}^w) > W_U(\mu, U_{UI}^u) = W_U(\mu', U_{UI}^u) \geq \mathbf{V}_U(\mu', U_{UI}^u)$$

Assume *per contra* that $V_{\mu U}^{UI}(\mu, U) \leq 0$. Then, it must be that $\partial U_{UI}^u / \partial \mu > 0$, which in turn implies that $\partial c_{UI} / \partial \mu < 0$, as $u(c_{UI}) = U - \beta U_{UI}^u$. But this leads to a contradiction as

$$V_{\mu U}^{UI}(\mu, U) = -\frac{\partial}{\partial c_{UI}}\left(\frac{1}{u'(c_{UI})}\right)\frac{\partial c_{UI}}{\partial \mu} > 0$$

Passing to JS,

$$V_{\mu U}^{JS}(\mu, U) = W_{UU}(\mu, U_{JS}^w) \frac{\partial U_{JS}^w}{\partial \mu} = \mathbf{V}_{\mu U}(\mu', U_{JS}^u) \frac{\partial \mu'}{\partial \mu} + \mathbf{V}_{UU}(\mu', U_{JS}^u) \frac{\partial U_{JS}^u}{\partial \mu}$$

Per contra, assume that $V_{\mu U}^{JS}(\mu, U) < 0$. Then, it must be that $\partial U_{JS}^s / \partial \mu > 0$, $s \in \{w, u\}$.

However, the PK-JS constraint reads

$$U = (1 - \beta + \pi(\mu))U_{JS}^w + \beta(1 - \pi(\mu))U_{JS}^u$$

And so

$$\frac{\partial U_{JS}^w}{\partial \mu} = -\frac{\beta}{1 - \beta + \beta\pi(\mu)} \left[(\pi_H - \pi_L)(U_{JS}^w - U_{JS}^u) + (1 - \pi(\mu)) \frac{\partial U_{JS}^u}{\partial \mu} \right] < 0$$

where the inequality follows from assumption $U_{JS}^w \geq U_{JS}^u$. Hence, I have reached a contradiction.

$V_{\mu\mu}^{JS}(\mu, U) = 0$ and $V_U^{JS}(\mu, U) = V_U^{SA}(U)$ with no JS \rightarrow UI transition

The period before entering SA, the FOC condition is

$$V_U^{JS}(\mu, U) = -\frac{1}{u'(c_{JS})} = -\frac{1}{u'(c_{SA})} = V_U^{SA}(U_{JS}^u) \implies U_{JS}^u = (1 - \beta)u(c_{JS}) = U_{JS}^w = U$$

The value of JS wrt μ after imposing $U_{JS}^u = U_{JS}^w = U$ reads

$$V^{JS}(\mu, U) = -u^{-1}((1 - \beta)U) - \kappa^{ja} + \beta[\pi(\mu)W(\mu, U) + (1 - \pi(\mu))\mathbf{V}(\mu', U)]$$

Therefore, if \mathbf{V} is linear in the first argument, so is V^{JS} , given linearity of $\pi(\mu)W(\mu, U)$ in μ .

But then the proof follows from a recursive argument and linearity of V^{SA} in μ . ■

Proof of Prop. 2

Proof. The derivative of the value of each policy i with respect to U is

$$V_U^i(\mu, U) = \mathbf{V}_U(\mu, U) = -\frac{1}{u'(c_i)} \tag{17}$$

which can be obtained by applying the envelope theorem to the problem of each policy.

Third, first-order conditions in UI and JS impose

$$\mathbf{V}_U(\mu', U_{UI}^u) - V_U^{UI}(\mu, U) = \mathbf{V}_U(\mu', U_{UI}^u) + \lambda^{UI} - \chi^{UI} = \frac{\pi(\mu)}{1 - \pi(\mu)} \chi^{UI} > 0 \quad (18)$$

$$V_U^{UI}(\mu, U_{UI}) = -(\lambda^{UI} - \chi^{UI}) = W_U(\mu, U_{UI}^w) + \chi^{UI} > W_U(\mu, U_{UI}^w) \quad (19)$$

$$V_U^{JS}(\mu, U) = -\lambda^{JS} = W_U(\mu, U_{JS}^w) = \mathbf{V}_U(\mu', U_{JS}^u) \quad (20)$$

where λ_i (resp., χ_i) is the Lagrange multiplier associated to (PK) (resp., (IC)) constraint. Hence,

$$\mathbf{V}_U(\mu', U_{UI}^u) > V_U^{UI}(\mu, U) > \mathbf{V}_U(\mu', U) \implies U_{UI}^u < U \quad (21)$$

where the first inequality holds by FOC and the second by supermodularity of \mathbf{V} . Similarly, in JS consumption is constant over time and employment states by (20), which implies that

$$u(c_{JS}) = u(c_{JS}^w) = (1 - \beta)U_{JS}^w \implies U = (1 - \beta + \beta\pi(\mu))U_{JS}^w + \beta(1 - \pi(\mu))U_{JS}^u$$

and

$$\mathbf{V}_U(\mu, U) = V_U^{JS}(\mu, U) = \mathbf{V}_U(\mu', U_{JS}^u) \leq \mathbf{V}_U(\mu, U_{JS}^u) \quad (22)$$

where the inequality holds since \mathbf{V} is supermodular, the first equality as JS is optimal in (μ, U) and the second equality from FOC (20). Thus, by concavity of \mathbf{V} in U , it holds that $U_{JS}^u \leq U \leq U_{JS}^w$. Supermodularity of JS follows from Lemma 1.

Optimal Policies in the U Space

The proof of the first part of the statement consists of showing that at the crossing point

$$V_U^{UI}(\mu, U) \leq V_U^{JS}(\mu, U) \leq V_U^{SA}(U) = W_U(\mu, U) \quad (23)$$

First, the closed-form expressions of W and V^{SA} deliver

$$W_U(\mu, U) = -\frac{1}{u'(u^{-1}((1 - \beta)U))} = V_U^{SA}(U)$$

Then, by (20),

$$V_U^{JS}(\mu, U) = W_U(\mu, U_{JS}^w) \leq W_U(\mu, U) = V_U^{SA}(U)$$

where the inequality follows from $U \leq U_{JS}^w$.

By (21), we know that $U_{UI}^u < U$. Hence,

$$u(c_{UI}) = U - \beta U_{UI}^u > (1 - \beta)U \implies V_U^{UI}(\mu, U) < -\frac{1}{u'(u^{-1}((1 - \beta)U))} = V_U^{SA}(U)$$

If JS refers to SA, then

$$u(c_{JS}) = u(c_{JS}^u) = u(c_{SA}) = (1 - \beta)U_{JS}^u$$

So $U_{JS}^w = U_{JS}^u = U$ and

$$V_U^{JS}(\mu, U) = -\frac{1}{u'(u^{-1}((1 - \beta)U))} = V_U^{SA}(U)$$

So far, I have shown that if JS is followed by SA, then

$$V_U^{UI}(\mu, U) \leq -\frac{1}{u'(u^{-1}((1 - \beta)U))} = V_U^{SA}(U) = V_U^{JS}(\mu, U)$$

What is left to show is that $V_U^{UI}(\mu, U) \leq V_U^{JS}(\mu, U)$, even if JS does not refer to SA in (μ, U) .

Per contra, assume that $V_U^{UI}(\mu, U) > V_U^{JS}(\mu, U)$. First, it must be that the inequality

$$-\frac{1}{u'(u^{-1}((1 - \beta)U))} \geq V_U^{UI}(\mu, U) > V_U^{JS}(\mu, U)$$

holds and implies that $(1 - \beta)U \leq u(c_{UI}) < u(c_{JS})$. And from (22), it must be that $U_{JS}^u < U < \frac{u(c_{JS})}{1 - \beta}$.

Now, by FOCs (18) and (20), it holds that

$$\mathbf{V}_U(\mu', U_{UI}^u) > V_U^{UI}(\mu, U) > V_U^{JS}(\mu, U) = \mathbf{V}_U(\mu', U_{JS}^u)$$

By concavity of \mathbf{V} in U , it must hold that $U_{UI}^u < U_{JS}^u$. But this is impossible as

$$u(c_{UI}) + \beta U_{UI}^u = U = u(c_{JS}) + \beta U_{JS}^u + \beta\pi(\mu) \left[\frac{u(c_{JS})}{1 - \beta} - U_{JS}^u \right] > u(c_{UI}) + \beta U_{UI}^u$$

where the inequality follows from $c_{JS} > c_{UI}$ and $\frac{u(c_{JS})}{1 - \beta} > U_{JS}^u$.

Therefore, it has been shown that $V_U^{UI}(\mu, U) \leq V_U^{JS}(\mu, U)$. Hence, V^{UI} dominates V^{JS} for low-end generosity levels and crosses it from above.

Optimal Policies in the μ Space

Passing to the second part of the statement, it is enough to prove that at the crossing point

$$0 = V_\mu^{SA}(U) < V_\mu^{JS}(\mu, U) < V_\mu^{UI}(\mu, U)$$

The derivatives of V^{JS} and V^{UI} wrt to μ

$$\begin{aligned} V_{\mu}^{JS}(\mu, U) = & \beta(\pi_H - \pi_L)[W(\mu, U_{JS}^w) - \mathbf{V}(\mu', U_{JS}^u) - \lambda^{JS}(U_{JS}^u - U_{JS}^w)] + \\ & + \beta \left[\pi(\mu)W_{\mu}(\mu, U_{JS}^w) + (1 - \pi(\mu)) \frac{\partial \mu'}{\partial \mu} \mathbf{V}_{\mu}(\mu', U_{JS}^u) \right] \end{aligned} \quad (24)$$

$$\begin{aligned} V_{\mu}^{UI}(\mu, U) = & \beta(\pi_H - \pi_L)[W(\mu, U_{UI}^w) - \mathbf{V}(\mu', U_{UI}^u) - \lambda^{UI}(U_{UI}^u - U_{UI}^w)] + \\ & + \beta \left[\pi(\mu)W_{\mu}(\mu, U_{UI}^w) + (1 - \pi(\mu)) \frac{\partial \mu'}{\partial \mu} \mathbf{V}_{\mu}(\mu', U_{UI}^u) \right] \end{aligned} \quad (25)$$

Consider JS being implemented in the current period. Using FOC $-\lambda^{JS} = W_U(\mu, U_{JS}^w)$, it holds that

$$W(\mu, U_{JS}^w) + W_U(\mu, U_{JS}^w)(U_{JS}^u - U_{JS}^w) - \mathbf{V}(\mu', U_{JS}^u) > W(\mu, U_{JS}^u) - \mathbf{V}(\mu', U_{JS}^u) > 0 = V_{\mu}^{SA}(U)$$

where the first inequality follows from concavity of W in U and the second is a necessary condition for optimality of JS.

Consider a program that implements UI with the additional constraint that $U_{UI}^u \geq U_{JS}^u$, and label its value \hat{V}^{UI} . Moreover,

$$W_U(\mu, U_{JS}^w) = V_U^{JS}(\mu, U) \geq V_U^{UI}(\mu, U) > W_U(\mu, U_{UI}^w) \implies U_{JS}^w < U_{UI}^w$$

where the first inequality follows from the statement proved above and the second one from FOC of UI.⁴⁴ Hence, derivatives (24) and (25) can be rewritten

$$\begin{aligned} V_{\mu}^{JS}(\mu, U) = & \beta(\pi_H - \pi_L)[W(\mu, U_{JS}^w) - \mathbf{V}(\mu', U_{JS}^u) + W_U(\mu, U_{JS}^w)(U_{UI}^u - U_{JS}^w) + \mathbf{V}_U(\mu', U_{JS}^u)(U_{JS}^u - U_{UI}^u)] + \\ & + \beta \left[\pi(\mu)W_{\mu}(\mu, U_{JS}^w) + (1 - \pi(\mu)) \frac{\partial \mu'}{\partial \mu} \mathbf{V}_{\mu}(\mu', U_{JS}^u) \right] \end{aligned}$$

$$\begin{aligned} \hat{V}_{\mu}^{UI}(\mu, U) = & \beta(\pi_H - \pi_L)[W(\mu, U_{UI}^w) - \mathbf{V}(\mu', U_{UI}^u) + W_U(\mu, U_{UI}^w)(U_{JS}^w - U_{UI}^w) + W_U(\mu, U_{UI}^w)(U_{UI}^u - U_{JS}^w)] + \\ & + \beta \left[\pi(\mu)W_{\mu}(\mu, U_{UI}^w) + (1 - \pi(\mu)) \frac{\partial \mu'}{\partial \mu} \mathbf{V}_{\mu}(\mu', U_{UI}^u) \right] \end{aligned}$$

In order to prove the result, it is enough to show that

$$\begin{aligned} & W(\mu, U_{JS}^w) + W_U(\mu, U_{JS}^w)(U_{UI}^u - U_{JS}^w) - \mathbf{V}(\mu', U_{JS}^u) + \mathbf{V}_U(\mu', U_{JS}^u)(U_{JS}^u - U_{UI}^u) < \\ & < W(\mu, U_{UI}^w) + W_U(\mu, U_{UI}^w)(U_{JS}^w - U_{UI}^w) + W_U(\mu, U_{UI}^w)(U_{UI}^u - U_{JS}^w) - \mathbf{V}(\mu', U_{UI}^u) \end{aligned}$$

and

$$\pi(\mu)W_{\mu}(\mu, U_{JS}^w) + (1 - \pi(\mu)) \frac{\partial \mu'}{\partial \mu} \mathbf{V}_{\mu}(\mu', U_{JS}^u) < \pi(\mu)W_{\mu}(\mu, U_{UI}^w) + (1 - \pi(\mu)) \frac{\partial \mu'}{\partial \mu} \mathbf{V}_{\mu}(\mu', U_{UI}^u)$$

⁴⁴The additional constraint preserves the FOC $V^{UI}(\mu, U) > W_U(\mu, U_{UI}^w)$.

The first inequality holds since:

- $W(\mu, U_{JS}^w) < W(\mu, U_{UI}^w) + W_U(\mu, U_{UI}^w)(U_{JS}^w - U_{UI}^w)$, by concavity of W in U ;
- $W_U(\mu, U_{JS}^w)(U_{UI}^u - U_{JS}^w) < W_U(\mu, U_{UI}^w)(U_{UI}^u - U_{JS}^w)$, as $U_{UI}^u < U \leq U_{JS}^w < U_{UI}^w$;
- $\mathbf{V}(\mu', U_{UI}^u) \leq \mathbf{V}(\mu', U_{JS}^u) + \mathbf{V}_U(\mu', U_{JS}^u)(U_{UI}^u - U_{JS}^u)$, by concavity of \mathbf{V} in U .

The second inequality holds since $W_{\mu U} = 0$ and $\mathbf{V}_\mu(\mu', U_{JS}^u) \leq \mathbf{V}_\mu(\mu', U_{UI}^u)$, by assumption $U_{JS}^u \leq U_{UI}^u$ and supermodularity of \mathbf{V} . Therefore, it has been shown that \hat{V}^{UI} crosses V^{JS} from below in the μ space, and so does V^{UI} , which implies that UI dominates JS for high expectations. ■

Proof of Prop. 3

Unemployment Benefits

Thus, unemployment benefits fall over time during UI and stay constant in JS, as

$$\begin{aligned} \mathbf{V}_U(\mu', U_{UI}^u) > V_U^{UI}(\mu, U) &\implies c_{UI}^u < c_{UI} \\ V_U^{UI}(\mu, U_{UI}) > W_U(\mu, U_{UI}^w) &\implies c_{UI} < c_{UI}^w \\ V_U^{JS}(\mu, U) = W_U(\mu, U_{JS}^w) = \mathbf{V}_U(\mu', U_{JS}^u) &\implies c_{JS} = c_{JS}^w = c_{JS}^u \end{aligned}$$

where the implications follow from (17).

Continuation Utility

If JS never refers to UI, then one can start computing backward from the point in time where JS refers to SA. Hence, $V_U^{JS}(\mu, U) = V_U^{SA}(U)$. Therefore $U_{JS}^w = U_{JS}^u = U$ and $\mathbf{V}_{\mu U}(\mu, U) = V_{\mu U}^{JS}(\mu, U) = 0$. The last period before the worker enters SA, the contract satisfies

$$V_U^{JS}(\mu, U) = \mathbf{V}_U(\mu', U_{JS}^u) = \mathbf{V}_U(\mu, U_{JS}^u) \implies U_{JS}^u = U$$

The result is shown by induction argument.

Proof of Prop. 4

Assume that UI is the optimal policy in (μ, U) . From the first-order condition on UI, it holds that

$$-g'((1 - \beta)U + \eta^{UI}(\mu, U)) = V_U^{UI}(\mu, U) = -g'(U - \beta U_{UI}^u)$$

From which it follows that

$$U - \frac{\eta^{UI}(\mu, U)}{\beta} = U_{UI}^u$$

Since by assumption $\eta_{UI}^{UI}(\mu, U) \leq 0$, the left-hand side is increasing in U . Thus,

$$U_{UI}^u = U - \frac{\eta^{UI}(\mu, U)}{\beta} \leq \hat{U}(\mu) - \frac{\eta^{UI}(\mu, \hat{U}(\mu))}{\beta} \leq \hat{U}(\mu')$$

The condition guarantees that anytime UI is adopted in (μ, U) , then the next-period state in case of job search failure becomes (μ', U_{UI}^u) where it is still optimal to implement UI (or switch to SA) as $U_{UI}^u \leq \hat{U}(\mu')$. To conclude, the program never switches from UI to JS.

Appendix B: Properties of AP, SP and IP

Proof of Prop. 5

Proof. The problem looks for $i \in \{AP, SP, IP\}$ reads

$$\begin{aligned} V^i(\mu, U) &= \max_{(c, U^w, U_p^u, U_f^u)} -c - \kappa^i + \beta[p^i(\mu)W(\mu, U^w) + q(1 - p^i(\mu_p))\mathbf{V}(\mu_p^i, U_p^u) + (1 - q)(1 - p^i(\mu_f))\mathbf{V}(\mu_f^i, U_f^u)] \\ \text{sub: } U &= u(c) - e^i + \beta[p^i(\mu)U^w + q(1 - p^i(\mu_p))U_p^u + (1 - q)(1 - p^i(\mu_f))U_f^u] \quad (\lambda) \\ U &\geq u(c) + \beta[qU_p^u + (1 - q)U_f^u] \quad (\chi), \quad U_p^u \geq U_f^u \quad (\xi), \quad q = \frac{\mu - \mu_f}{\mu_p - \mu_f} \end{aligned}$$

with $p^{AP}(\mu) = 0$, $p^{SP}(\mu) = p^{IP}(\mu) = \pi(\mu)$.

The first-order derivative with respect to μ_f reads

$$\begin{aligned} & -\beta p^{i'}(\mu_f)(1 - q)[\mathbf{V}(\mu_f^i, U_f^u) + \lambda U_f^u] + \beta(1 - q)(1 - p^i(\mu_f))\mathbf{V}_\mu(\mu_f^i, U_f^u) \frac{\partial \mu_f^i}{\partial \mu_f} + \\ & + \frac{\partial q}{\partial \mu_f} \beta [(1 - p^i(\mu_p))(\mathbf{V}(\mu_p^i, U_p^u) + \lambda U_p^u) - (1 - p^i(\mu_f))(\mathbf{V}(\mu_f^i, U_f^u) + \lambda U_f^u) - \chi(U_p^u - U_f^u)] \\ & = -A \left\{ p^{i'}(\mu_f) [\mathbf{V}(\mu_f^i, U_f^u) + \lambda U_f^u] - (1 - p^i(\mu_f)) \mathbf{V}_\mu(\mu_f^i, U_f^u) \frac{\partial \mu_f^i}{\partial \mu_f} + \right. \\ & \quad \left. + \frac{1}{\mu_p - \mu_f} [(1 - p^i(\mu_p))(\mathbf{V}(\mu_p^i, U_p^u) + \lambda U_p^u) - (1 - p^i(\mu_f))(\mathbf{V}(\mu_f^i, U_f^u) + \lambda U_f^u) - \chi(U_p^u - U_f^u)] \right\} \\ & = -A \left\{ - (1 - p^i(\mu_f)) \mathbf{V}_\mu(\mu_f^i, U_f^u) \frac{\partial \mu_f^i}{\partial \mu_f} + \frac{1 - p^i(\mu_p)}{\mu_p - \mu_f} [\mathbf{V}(\mu_p^i, U_p^u) - \mathbf{V}(\mu_f^i, U_f^u) + \lambda(U_p^u - U_f^u)] - \frac{\chi(U_p^u - U_f^u)}{\mu_p - \mu_f} \right\} \\ & < -A \left\{ (1 - p^i(\mu_f)) \frac{\partial \mu_f^i}{\partial \mu_f} (\mathbf{V}_\mu(\mu_p^i, U_p^u) - \mathbf{V}_\mu(\mu_f^i, U_f^u)) + \frac{1 - p^i(\mu_p)}{\mu_p - \mu_f} [\mathbf{V}(\mu_p^i, U_p^u) - \mathbf{V}(\mu_p^i, U_f^u) + \lambda(U_p^u - U_f^u)] - \right. \\ & \quad \left. - \frac{\chi(U_p^u - U_f^u)}{\mu_p - \mu_f} \right\} \\ & < -A \left\{ (1 - p^i(\mu_f)) \frac{\partial \mu_f^i}{\partial \mu_f} (\mathbf{V}_\mu(\mu_p^i, U_p^u) - \mathbf{V}_\mu(\mu_f^i, U_f^u)) + \frac{U_p^u - U_f^u}{\mu_p - \mu_f} [(1 - p^i(\mu_p))(\mathbf{V}(\mu_p^i, U_p^u) + \lambda) - \chi] \right\} \\ & = -A(1 - p^i(\mu_f)) \frac{\partial \mu_f^i}{\partial \mu_f} (\mathbf{V}_\mu(\mu_p^i, U_p^u) - \mathbf{V}_\mu(\mu_f^i, U_f^u)) < 0 \end{aligned}$$

with $A = \beta(1 - q)$.

The first inequality follows from concavity of \mathbf{V} in the first argument

$$\mathbf{V}(\mu_f^i, U_f^u) < \mathbf{V}(\mu_p^i, U_f^u) + (\mu_f^i - \mu_p^i) \mathbf{V}_\mu(\mu_p^i, U_f^u) = \mathbf{V}(\mu_p^i, U_f^u) + \frac{\partial \mu_f^i}{\partial \mu_f} \frac{1 - p^i(\mu_f)}{1 - p^i(\mu_p)} (\mu_f - \mu_p) \mathbf{V}_\mu(\mu_p^i, U_f^u)$$

The second inequality follows from concavity of \mathbf{V} in the second argument

$$\mathbf{V}(\mu_p^i, U_f^u) < \mathbf{V}(\mu_p^i, U_p^u) + \mathbf{V}_U(\mu_p^i, U_p^u)(U_f^u - U_p^u)$$

The second-to-last passage follows from the slack constraint $\xi(U_p^u - U_f^u) = 0$ and first-order condition on U_p^u ,

$$\beta q \{ (1 - p^i(\mu_p)) [V_U(\mu_p^i, U_p^u) + \lambda] - \chi \} + \xi = 0$$

And the third (and last) inequality follows as (i) profiling outcomes refer workers to different policies,⁴⁵ (ii) $\mu_f^i < \mu_p^i$, and (iii) condition (8) of Prop. 2.

Therefore, it must be that $\mu_f = 0$. Referral to SA follows from the fact that SA is optimal (if any) at low-end expectations. \blacksquare

Proof of Prop. 6

Proof. Consider the two first-order conditions of the AP problem

$$\begin{aligned} \mathbf{V}_\mu(\mu_p, U_{AP}^u) - \frac{\mathbf{V}(\mu_p, U_{AP}^u) - \mathbf{V}(0, U_{AP}^u)}{\mu_p} &= 0 \\ q \mathbf{V}_U(\mu_p, U_{AP}^u) + (1 - q) \mathbf{V}_U(0, U_{AP}^u) + \lambda^{AP} &= 0 \end{aligned}$$

For the pair (μ_p, U_{AP}^u) to be a point of maximum, it must be that the Hessian matrix H of second-order derivatives has positive determinant.

$$H = \begin{bmatrix} q \mathbf{V}_{UU}(\mu_p, U_{AP}^u) + (1 - q) \mathbf{V}_{UU}(0, U_{AP}^u) & q \left(\mathbf{V}_{\mu U}(\mu_p, U_{AP}^u) - \frac{V_U(\mu_p, U_{AP}^u) - V_U(0, U_{AP}^u)}{\mu_p} \right) \\ \mathbf{V}_{\mu U}(\mu_p, U_{AP}^u) - \frac{V_U(\mu_p, U_{AP}^u) - V_U(0, U_{AP}^u)}{\mu_p} & \mathbf{V}_{\mu\mu}(\mu_p, U_{AP}^u) \end{bmatrix}$$

Differentiating the two conditions by U yields

$$\begin{aligned} \mathbf{V}_{\mu\mu}(\mu_p, U_{AP}^u) \frac{\partial \mu_p}{\partial U_{AP}^u} + \mathbf{V}_{\mu U}(\mu_p, U_{AP}^u) - \frac{\mathbf{V}_U(\mu_p, U_{AP}^u) - \mathbf{V}_U(0, U_{AP}^u)}{\mu_p} &= 0 \\ \left\{ q \left[\mathbf{V}_{\mu U}(\mu_p, U_{AP}^u) - \frac{V_U(\mu_p, U_{AP}^u) - V_U(0, U_{AP}^u)}{\mu_p} \right] \frac{\partial \mu_p}{\partial U_{AP}^u} + q \mathbf{V}_{UU}(\mu_p, U_{AP}^u) + (1 - q) \mathbf{V}_{UU}(0, U_{AP}^u) \right\} \frac{\partial U_{AP}^u}{\partial U} &+ \\ &+ \lambda_U^{AP} = 0 \end{aligned}$$

⁴⁵Since each policy is concave in the first argument, referring workers to the same policy, irrespective of the profiling outcome, would produce a net loss.

Plugging the expression of $\frac{\partial \mu_p}{\partial U_{AP}^u}$ from the first equation into the second one, the term in curly brackets becomes

$$\Delta := -\frac{q}{\mathbf{V}_{\mu\mu}(\mu_p, U_{AP}^u)} \left[\mathbf{V}_{\mu U}(\mu_p, U_{AP}^u) - \frac{\mathbf{V}_U(\mu_p, U_{AP}^u) - \mathbf{V}_U(0, U_{AP}^u)}{\mu_p} \right]^2 + q\mathbf{V}_{UU}(\mu_p, U_{AP}^u) + (1-q)\mathbf{V}_{UU}(0, U_{AP}^u)$$

with $\Delta < 0$, as $\det(H) > 0$ and $\mathbf{V}_{\mu\mu}(\mu_p, U_{AP}^u) < 0$. Therefore the second equation becomes

$$\frac{\partial U_{AP}^u}{\partial U} \Delta + \lambda_U^{AP} = 0$$

which shows that $\frac{\partial U_{AP}^u}{\partial U} > 0$, since $V_U^{AP}(\mu, U) = -\lambda_U^{AP} < 0$. From the first equation, local supermodularity of \mathbf{V} and $\mathbf{V}_U(\mu_p, U) < \mathbf{V}_U(0, U)$ (see Prop. 2) yield $\frac{\partial \mu_p}{\partial U_{AP}^u} > 0$, which deliver the result.

The proof of μ_p being monotone increasing in U also under SP and IP follows accordingly. Finally, full accuracy of profiling whenever ‘Pass’ refers to JS and no incentive is provided along the spell follows from linearity of V^{JS} in μ (see proof of Lemma 2). ■

Appendix C: Statistical Profiling

Proof of Prop. 7 and 8

Proof. $\mu_f = 0$ descends from the linearity of SA and JS in μ . In addition, any worker who receives a ‘Pass’ (resp., ‘Fail’) is referred to UI/JS (resp., SA).

Assistance-and-Profiling (AP)

At the optimum,

$$V_U^{AP}(\mu, U) = \mathbf{V}_U(\mu_p, U_{AP}^u) = \mathbf{V}_U(0, U_{AP}^u) \quad (26)$$

$$-\frac{1}{u'(c_{AP})} = V_U^{AP}(\mu, U) = \mathbf{V}_U(0, U_{AP}^u) = -\frac{1}{u'(c_{AP}^{u,f})} \quad (27)$$

which implies that $c_{AP} = c_{AP}^{u,p} = c_{AP}^{u,f}$, and $U_p^u \leq U \leq U_f^u$. Indeed, by (27), it follows

$$U = u(c_{AP}) + \beta(qU_{AP}^{u,p} + (1-q)U_{AP}^{u,f}) = (1-\beta)U_{AP}^{u,f} + \beta qU_{AP}^{u,p}$$

where the passage follows from $u(c_{AP}) = u(c_{AP}^{u,f}) = (1-\beta)U_{AP}^{u,f}$, and the expression of consumption in SA (see Prop. 1). If referred to JS -which is optimal only for high-end generosities-, then $\mu_p = 1$ given the linearity of JS in μ . Moreover, for U high enough, JS never refers to UI, and

so $U_{JS}^w = U = U_{JS}^u$, which in turn implies that $u(c_{JS}) = (1 - \beta)U_{JS}^w = (1 - \beta)U = u(c_{SA})$ and

$$V_U^{JS}(\mu, U) = -\frac{1}{u'(c_{JS})} = -\frac{1}{u'(c_{SA})} = V_U^{SA}(U)$$

Therefore, if referred to JS/SA forever after, then $U_{AP}^{u,p} = U_{AP}^{u,f} = U_{AP}^u$. So, nothing changes with respect to the case with ND constraint, whenever AP refers workers to SA and JS forever after, that is, for higher generosities.

Assume, instead, AP refers to UI directly, or to JS which later refers to UI. Then, by Prop. 2,

$$V_U^{SA}(U_{AP}^{u,f}) = \mathbf{V}_U(\mu_p, U_{AP}^{u,p}) \implies U_{AP}^{u,p} < U < U_{AP}^{u,f}$$

I now show that the ‘Pass’ posterior μ_p in AP is increasing in U .

$$\begin{aligned} \mathbf{V}_U(\mu_p, U_{AP}^{u,p}) + \lambda^{AP} &= 0 \\ -\frac{\mathbf{V}(\mu_p, U_{AP}^{u,p}) - \mathbf{V}(0, U_{AP}^{u,f}) + \lambda^{AP}(U_{AP}^{u,p} - U_{AP}^{u,f})}{\mu_p} + \mathbf{V}_\mu(\mu_p, U_{AP}^{u,p}) &= 0 \end{aligned}$$

The Hessian matrix of second-order derivatives reads

$$H = \begin{bmatrix} \mathbf{V}_{UU}(\mu_p, U_{AP}^{u,p}) & \mathbf{V}_{\mu U}(\mu_p, U_{AP}^{u,p}) \\ \mathbf{V}_{\mu U}(\mu_p, U_{AP}^{u,p}) & \mathbf{V}_{\mu\mu}(\mu_p, U_{AP}^{u,p}) \end{bmatrix}$$

$(\mu_p, U_{AP}^{u,p})$ are a point of maximum of the objective function if and only if

$$\mathbf{V}_{UU}(\mu_p, U_{AP}^{u,p}) < 0, \quad \det(H) > 0$$

The first condition holds as \mathbf{V} is concave in each argument (see proof of Lemma 2). Differentiating the two FOCs wrt U yields

$$\begin{aligned} \mathbf{V}_{\mu U}(\mu_p, U_{AP}^{u,p}) \frac{\partial \mu_p}{\partial U} + \mathbf{V}_{UU}(\mu_p, U_{AP}^{u,p}) \frac{\partial U_{AP}^{u,p}}{\partial U} &= -\frac{\partial \lambda^{AP}}{\partial U} \\ \mathbf{V}_{\mu\mu}(\mu_p, U_{AP}^{u,p}) \frac{\partial \mu_p}{\partial U} + \mathbf{V}_{\mu U}(\mu_p, U_{AP}^{u,p}) \frac{\partial U_{AP}^{u,p}}{\partial U} &= \frac{U_{AP}^{u,p} - U_{AP}^{u,f}}{\mu_p} \frac{\partial \lambda^{AP}}{\partial U} \end{aligned}$$

and solving the system, one obtains

$$\begin{bmatrix} \frac{\partial \mu_p}{\partial U} \\ \frac{\partial U_{AP}^{u,p}}{\partial U} \end{bmatrix} = \det(H)^{-1} \begin{bmatrix} \mathbf{V}_{\mu\mu}(\mu_p, U_{AP}^{u,p}) & -\mathbf{V}_{\mu U}(\mu_p, U_{AP}^{u,p}) \\ -\mathbf{V}_{\mu U}(\mu_p, U_{AP}^{u,p}) & \mathbf{V}_{UU}(\mu_p, U_{AP}^{u,p}) \end{bmatrix} \begin{bmatrix} \frac{U_{AP}^{u,p} - U_{AP}^{u,f}}{\mu_p} \\ -1 \end{bmatrix} \frac{\partial \lambda^{AP}}{\partial U}$$

Both derivatives are positive, since $U_{AP}^{u,p} - U_{AP}^{u,f} < 0$ and an increase in U makes it harder for the planner to satisfy (PK) constraint (i.e., $\partial \lambda^{AP} / \partial U > 0$).

Search-assistance-and-Profiling (SP)

At the optimum

$$\begin{aligned} V_U^{SP}(\mu, U) &= W_U(\mu_p, U_{SP}^{w,p}) = W_U(0, U_{SP}^{w,f}) = \mathbf{V}_U(\mu'_p, U_{SP}^{u,p}) = \mathbf{V}_U(0, U_{SP}^{u,f}) \\ \implies c_{SP} &= c_{SP}^w = c_{SP}^{u,p} = c_{SP}^{u,f}, \quad U_{SP}^{u,p} \leq U \leq U_{SP}^{u,f} = U_{SP}^{w,p} = U_{SP}^{w,f} \end{aligned}$$

since

$$\frac{u(c_{SP})}{1-\beta} = \frac{u(c_{SP}^w)}{1-\beta} = U_{SP}^w = \frac{u(c_{SP}^{u,f})}{1-\beta} = U_{SP}^{u,f}$$

where the last equality follows from referral to SA upon ‘Fail’. So

$$U = (1 - \beta + \beta\pi(\mu) + \beta(1 - q)(1 - \pi_L))U_{SP}^{u,f} + \beta q(1 - \pi(\mu_p))U_{SP}^{u,p}$$

and the same argument in AP applies, meaning that the continuation utility upon ‘Pass’ falls if and only if the outcome refers workers directly or indirectly to UI.

Insurance-and-Profiling (IP)

The optimal IP contract satisfies

$$\begin{aligned} \mathbf{V}_U(\mu'_p, U_{IP}^{u,p}) - V_U^{IP}(\mu, U) &= \frac{\pi(\mu_p)}{1 - \pi(\mu_p)}\chi^{IP} > \frac{\pi_L}{1 - \pi_L}\chi^{IP} = \mathbf{V}_U(0, U_{IP}^{u,f}) - V_U^{IP}(\mu, U) \implies U_{IP}^{u,p} < U_{IP}^{u,f} \\ \mathbf{V}_U(\mu'_p, U_{IP}^{u,p}) > \mathbf{V}_U(0, U_{IP}^{u,f}) > V_U^{IP}(\mu, U) &= -(\lambda^{IP} - \chi^{IP}) > W_U(\mu, U_{IP}^w) \implies c_{IP}^{u,p} < c_{IP}^{u,f} < c_{IP} < c_{IP}^w \end{aligned} \quad (28)$$

Moreover, $U_{IP}^{u,p} < U$, as

$$(1 - \beta)U_{IP}^{u,p} \leq u(c_{IP}^{u,p}) < u(c_{IP}) = U - \beta[qU_{IP}^{u,p} + (1 - q)U_{IP}^{u,f}] < U - \beta U_{IP}^{u,p}$$

where the first inequality follows from (23), the second one from (28) and the last one from $U_{IP}^{u,p} < U_{IP}^{u,f}$.

Passing to the equation that determines the ‘Pass’ posterior in IP, the first-order condition of μ_p reads

$$\begin{aligned} \frac{1}{\mu_p} [(1 - \pi_L)\mathbf{V}(0, U_{IP}^{u,f}) - (1 - \pi(\mu_p))\mathbf{V}(\mu'_p, U_{IP}^{u,p})] - (\pi_H - \pi_L)\mathbf{V}(\mu'_p, U_{IP}^{u,p}) &+ \frac{(1 - \pi_H)(1 - \pi_L)}{1 - \pi(\mu_p)}\mathbf{V}_\mu(\mu'_p, U_{IP}^{u,p}) + \\ + \lambda^{IP} \left[\frac{1}{\mu_p} [(1 - \pi_L)U_{IP}^{u,f} - (1 - \pi(\mu_p))U_{IP}^{u,p}] - (\pi_H - \pi_L)U_{IP}^{u,p} \right] &+ \frac{\chi^{IP}}{\mu_p} [U_{IP}^{u,p} - U_{IP}^{u,f}] = 0 \end{aligned}$$

Rearranging the terms and using the first order condition on $U_{IP}^{u,f}$

$$(1 - \pi_L)[\mathbf{V}_U(0, U_{IP}^{u,f}) + \lambda^{IP}] = \chi^{IP}$$

it yields

$$\mathbf{V}_\mu(\mu'_p, U_{IP}^{u,p}) = \frac{\mathbf{V}(\mu'_p, U_{IP}^{u,p}) - \mathbf{V}(0, U_{IP}^{u,f}) + \mathbf{V}_U(0, U_{IP}^{u,f})(U_{IP}^{u,f} - U_{IP}^{u,p})}{\mu'_p}$$

with $U_{IP}^{u,f} > U_{IP}^{u,p}$. ■

Appendix D: Moral Hazard

Proof of Prop. 9

Proof. The first part of the proof is contained in the [Technical Appendix](#). It shows that multiple deviations can be accounted for by single one-shot deviations, that is, deviations from recommended action lasting only one period. Now consider the recursion (15)

$$\begin{aligned} \Lambda(1, \mu^{T-1}) &= 0 \\ \Lambda(2, \mu^{T-2}) &= u(c_{T-1}) - e + \beta\pi(\mu^{T-2})U_T^w + \beta(1 - \pi(\mu^{T-2}))U_T^u - U_{T-2}^u \\ &= U_{T-2}^u + \beta[\pi(\mu^{T-2}) - \pi(\mu^{T-1})](U_T^w - U_T^u) - U_{T-2}^u = \beta[\pi(\mu^{T-2}) - \pi(\mu^{T-1})] \frac{e}{\beta\pi(\mu^{T-1})} \\ \Lambda(T-j, \mu^j) &= u(c_{j+1}) - e + \beta\pi(\mu^j)U_{j+2}^w + \\ &\quad + \beta(1 - \pi(\mu^j)) \underbrace{[u(c_{j+2}) - e + \beta\pi(\mu^{j+1})U_{j+3}^w + \beta(1 - \pi(\mu^{j+1}))[\dots]]}_{\Lambda(T-j-1, \mu^{j+1}) + U_{j+2}^u} - U_{j+1}^u \\ &= U_{j+1}^u + \beta[\pi(\mu^j) - \pi(\mu^{j+1})](U_{j+2}^w - U_{j+2}^u) + \beta(1 - \pi(\mu^j))\Lambda(T-j-1, \mu^{j+1}) - U_{j+1}^u \\ &= \beta[\pi(\mu^j) - \pi(\mu^{j+1})] \frac{\beta\Lambda(T-j-1, \mu^{j+1}) + e}{\beta\pi(\mu^{j+1})} + \beta(1 - \pi(\mu^j))\Lambda(T-j-1, \mu^{j+1}) \\ &= \left(\frac{\pi(\mu^j)}{\pi(\mu^{j+1})} - 1 \right) e + \beta\pi(\mu^j) \left(\frac{1}{\pi(\mu^{j+1})} - 1 \right) \Lambda(T-j-1, \mu^{j+1}), \quad 0 \leq j \leq T-1 \end{aligned}$$

And notice that the constraint (\hat{IC}, t) , defined as

$$U_s(\mathcal{W}, \mu^s, \sigma^s) = u(c_s(\sigma^s)) + \beta[U_{s+1}(\mathcal{W}, \mu^{s+1}, (\sigma^s, u)) + \Lambda(T-s, \mu^s)]$$

makes the contract robust against any possible deviation after period t , thanks to the recursive definition of Λ . In particular,

$$(\hat{IC}, t) \iff (IC, s), \quad \forall s \geq t$$

Hence the whole set of IC constraints can be expressed by

$$U = U_0(\mathcal{W}, \mu, \sigma_0) = u(c) + \beta[U^u + \Lambda(T, \mu)], \quad (\hat{IC}, 0)$$

Λ is defined by the recursion in (15), and is independent of U^u . In addition, $\Lambda(t+1, \mu) \geq \Lambda(t, \mu)$, with inequality being strict for $\mu \in (0, 1)$.⁴⁶ Indeed, taking the difference between $\Lambda(t+1, \mu)$ and $\Lambda(t, \mu)$, it holds:

$$\begin{cases} \Lambda(2, \mu) - \Lambda(1, \mu) = \left(\frac{\pi(\mu)}{\pi(\mu')} - 1\right)e > 0 \\ \Lambda(t+1, \mu) - \Lambda(t, \mu) = \beta\pi(\mu)\left(\frac{1}{\pi(\mu')} - 1\right)(\Lambda(t, \mu') - \Lambda(t-1, \mu')) > 0, \quad \forall t \geq 2 \end{cases}$$

■

Lemma 3. *The value of $(UI, t)_{t \geq 1}$ is increasing in μ .*

Proof. The problem of policy $(UI, t)_{t \geq 1}$ reads

$$\begin{aligned} V_t^{UI}(\mu, U) &= \max_{(z, U^w, U^u) \in \Gamma(\mu, U)} -g(z) + \beta[\pi(\mu)W(\mu, U^w) + (1 - \pi(\mu))V_{t-1}^{UI}(\mu', U^u)] \\ \text{sub: } \Gamma(\mu, U) &= \left\{ (z, U^w, U^u) : U = z - e + \beta[\pi(\mu)U^w + (1 - \pi(\mu))U^u], \right. \\ &\quad \left. U \geq z + \beta[U^u + \Lambda(t, \mu)] \right\} \end{aligned}$$

V_1^{UI} is monotone increasing in μ (see Lemma 1). By induction, assume that V_{t-1}^{UI} is increasing in μ . Positive monotonicity of V_t^{UI} in μ follows from:

- (positive) monotonicity of the objective function in μ ;
- (positive) monotonicity of the feasibility set Γ_U^i in μ , i.e.

$$\mu < \tilde{\mu} \implies \Gamma_U^i(\mu) \subseteq \Gamma_U^i(\tilde{\mu})$$

The objective function is always monotone in μ , as so are W and V_{t-1}^{UI} in their first argument, μ' is an increasing function of μ and $W(\mu, U) \geq V_{t-1}^{UI}(\mu', U)$ at the optimum. Monotonicity of Γ_U , instead, holds whenever an increase of μ leads to a relaxation of (IC).⁴⁷ Now, if $\Lambda(t, \cdot)$ is constant or decreasing, this always holds. Indeed, (IC) is more slack if $\Lambda(t, \cdot)$ is decreasing as

$$\mu < \tilde{\mu} \implies U^w - U^u \geq \frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)} > \frac{e/\beta + \Lambda(t, \tilde{\mu})}{\pi(\tilde{\mu})}$$

To prove that monotonicity holds also when $(\Lambda(t, \cdot))_{t \geq 1}$ is increasing in μ , I prove that the RHS

⁴⁶In $\mu \in \{0, 1\}$, no learning occurs and learning rents are null.

⁴⁷(PK) is always relaxed by an increase in μ (recall that $U^w \geq U^u$ in optimum).

is decreasing in μ .

From the definition of Λ in (15), I can rewrite

$$\frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)} = \frac{e}{\pi(\mu)} \left(\frac{1}{\beta} - 1 \right) - \beta \Lambda(t-1, \mu') + \beta \left(\frac{e/\beta + \Lambda(t-1, \mu')}{\pi(\mu')} \right) \quad (29)$$

Define $f(\mu) := \frac{\pi(\mu)}{\pi(\mu')}$, and notice that it is concave in μ . Indeed:

$$f_\mu(\mu) = (\pi_H - \pi_L)^2 \frac{(1-\mu)^2 \pi_L (1-\pi_L) - \mu^2 \pi_H (1-\pi_H)}{[(1-\pi_H)\pi_H \mu + (1-\pi_L)\pi_L(1-\mu)]^2}$$

$$f_{\mu\mu}(\mu) = -\frac{2(\pi_H - \pi_L)^2 \pi_H \pi_L (1-\pi_H)(1-\pi_L)}{[(1-\pi_H)\pi_H \mu + (1-\pi_L)\pi_L(1-\mu)]^3} < 0$$

Thus, the derivative of $\Lambda(t, \mu)$ by μ reads

$$\Lambda_\mu(t, \mu) = f_\mu(\mu)e + \beta [f_\mu(\mu) - (\pi_H - \pi_L)] \Lambda(t-1, \mu') + \beta [f(\mu) - \pi(\mu)] \frac{\partial \mu'}{\partial \mu} \Lambda_\mu(t-1, \mu') \quad (30)$$

Two cases are possible:

1. $f_\mu(\mu) \geq \pi_H - \pi_L$
2. $f_\mu(\mu) < \pi_H - \pi_L$

If the first case applies, then

$$\Lambda_\mu(t, \mu) > 0 \implies \Lambda_\mu(t-1, \mu') > 0$$

Assume *per contra* that $\Lambda_\mu(t-1, \mu') < 0$. But then by (strict) concavity of f , $f_\mu(\mu') > \pi_H - \pi_L$. Which, coupled with the expression of the derivative in (30), implies that for the assumption to be true, it must be that $\Lambda_\mu(t-2, \mu'') < 0$, and so on, until

$$f_\mu(\mu^{(t-2)})e = \Lambda_\mu(2, \mu^{(t-2)}) < 0 < \pi_H - \pi_L < f_\mu(\mu^{(t-2)})$$

Therefore, I have reached a contradiction.

Now, I am ready to prove by induction that

$$\Lambda_\mu(t, \mu) > 0 \wedge f_\mu(\mu) \geq \pi_H - \pi_L \implies \frac{\partial}{\partial \mu} \left(\frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)} \right) < 0$$

Base Step ($t = 2$)

Notice that the result is always true for $t = 2$, as the expression reads

$$\frac{e/\beta + \Lambda(2, \mu)}{\pi(\mu)} = \frac{e}{\pi(\mu)} \left(\frac{1}{\beta} - 1 \right) + \frac{e}{\pi(\mu')}$$

Induction Step

Assume *per contra* that

$$\Lambda_\mu(t, \mu) > 0 \wedge f_\mu(\mu) \geq \pi_H - \pi_L \wedge \frac{\partial}{\partial \mu} \left(\frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)} \right) > 0$$

Since the first two addends of (29) have been shown to be decreasing in μ , for it to be true it must be that $\frac{\partial}{\partial \mu'} \left(\frac{e/\beta + \Lambda(t-1, \mu')}{\pi(\mu')} \right) > 0$. However,

$$\Lambda(t, \mu) > 0 \wedge f_\mu(\mu) \geq \pi_H - \pi_L \implies \Lambda_\mu(t-1, \mu') > 0 \wedge f_\mu(\mu') > \pi_H - \pi_L \implies \frac{\partial}{\partial \mu'} \left(\frac{e/\beta + \Lambda(t-1, \mu')}{\pi(\mu')} \right) < 0$$

where the second implication follows by induction hypothesis. Hence the contradiction.

What is left to be shown is the following:

$$\Lambda_\mu(t, \mu) > 0 \wedge f_\mu(\mu) < \pi_H - \pi_L \implies \frac{\partial}{\partial \mu} \left(\frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)} \right) < 0$$

Base Step ($t = 2$)

Same as in the case above, as the thesis always applies.

Induction Step

The derivative by μ has the following expression

$$\frac{\partial}{\partial \mu} \left(\frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)} \right) = \frac{1}{\pi(\mu)} \left[\Lambda_\mu(t, \mu) - (\pi_H - \pi_L) \frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)} \right]$$

So assume *per contra* that it is positive. Then this means that $\Lambda_\mu(t, \mu) > (\pi_H - \pi_L) \frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)}$.

Moreover, by (29), it either means that $\Lambda_\mu(t-1, \mu') < 0$ or that

$$\frac{\partial}{\partial \mu'} \left(\frac{e/\beta + \Lambda(t-1, \mu')}{\pi(\mu')} \right) > 0$$

The first case can not apply, as (30) would imply that

$$(\pi_H - \pi_L) \frac{e}{\beta \pi(\mu)} < (\pi_H - \pi_L) \frac{e/\beta + \Lambda(t, \mu)}{\pi(\mu)} < \Lambda_\mu(t, \mu) < f_\mu(\mu) e < (\pi_H - \pi_L) e$$

which is impossible, as $\frac{1}{\beta \pi(\mu)} > 1$. Therefore, it must be the case that

$$\frac{\partial}{\partial \mu'} \left(\frac{e/\beta + \Lambda(t-1, \mu')}{\pi(\mu')} \right) > 0 \implies \Lambda_\mu(t-1, \mu') > (\pi_H - \pi_L) \frac{e/\beta + \Lambda(t-1, \mu')}{\pi(\mu')} > 0$$

Now, if $f_\mu(\mu') \geq \pi_H - \pi_L$, I have reached a contradiction, since I have shown above that

$$\Lambda_\mu(t-1, \mu') > 0 \wedge f_\mu(\mu') \geq \pi_H - \pi_L \implies \frac{\partial}{\partial \mu'} \left(\frac{e/\beta + \Lambda(t-1, \mu')}{\pi(\mu')} \right) < 0$$

If, instead, $f_\mu(\mu') < \pi_H - \pi_L$, then

$$\Lambda_\mu(t-1, \mu') > 0 \wedge f_\mu(\mu') < \pi_H - \pi_L \implies \frac{\partial}{\partial \mu'} \left(\frac{e/\beta + \Lambda(t-1, \mu')}{\pi(\mu')} \right) < 0$$

by induction hypothesis, and a contradiction is reached in this case, too. ■

Proof of Prop. 10

When profiling is adopted jointly with private search, the planner's problem reads

$$\begin{aligned} V^{IP}(\mu, U) &= \max_{c^{ip}, U^w, (U_r^u)_{r=\{p,f\}}, \mu_p} -c^{ip} - \kappa^{wp} + \beta [\pi(\mu)W(\mu, U^w) + q(1 - \pi(\mu_p))V(\mu'_p, U_p^u) + \\ &\quad + (1-q)(1 - \pi_L)V(0, U_f^u)] \\ \text{sub: } U &= u(c^{ip}) - e + \beta [\pi(\mu)U^w + q(1 - \pi(\mu_p))U_p^u + (1-q)(1 - \pi_L)U_f^u] \\ U &\geq u(c^{ip}) + \beta [q(\hat{\Lambda}^{i(\mu_p, U_p^u)}(\mu_p) + U_p^u) + (1-q)(\hat{\Lambda}^{i(0, U_f^u)}(0) + U_f^u)] \\ \text{with: } \hat{\Lambda}^i(\mu) &= \begin{cases} \Lambda(t, \mu), & \text{if } i = (UI, t) \\ 0, & \text{otherwise} \end{cases}, \quad (\text{MP}) \end{aligned}$$

Then, the first-order condition reads

$$\begin{aligned} &\frac{\partial q}{\partial \mu_p} [(1 - \pi(\mu_p))V(\mu'_p, U_p^u) - (1 - \pi_L)V(0, U_f^u)] - q(\pi_H - \pi_L)V(\mu'_p, U_p^u) + \\ &+ q \frac{(1 - \pi_H)(1 - \pi_L)}{1 - \pi(\mu_p)} V_\mu(\mu'_p, U_p^u) + \lambda \left[\frac{\partial q}{\partial \mu_p} [(1 - \pi(\mu_p))U_p^u - (1 - \pi_L)U_f^u] - q(\pi_H - \pi_L)U_p^u \right] - \\ &- \chi \left[\frac{\partial q}{\partial \mu_p} \Lambda(t, \mu_p) + q\Lambda_\mu(t, \mu_p) \right] = 0 \end{aligned}$$

Which can be rewritten as

$$V_\mu(\mu'_p, U_p^u) = \frac{V(\mu'_p, U_p^u) - V(0, U_f^u)}{\mu'_p} + \left(\lambda - \frac{\chi}{1 - \pi_L} \right) \frac{U_p^u - U_f^u}{\mu'_p} + \frac{\chi}{1 - \pi_L} \frac{\mu_p \Lambda_\mu(t, \mu_p) - \Lambda(t, \mu_p)}{\mu'_p}$$

and plugging in $-V_U^{SA}(U_f^u) = -V_U(0, U_f^u) = \lambda - \frac{\chi}{1 - \pi_L}$ and $\lambda = -W_U(\mu, U^w)$ delivers the result.

Appendix E: Estimation of hazard rates

In order to infer the hazard rates $\{\pi_H, \pi_L\}$, I proceed as follows. From the basic monthly Current Population Survey (CPS), I first derive the fraction of high- and low-skilled workers for each level of educational attainment ϑ_i , $i \in \{LessHighSc., HighSc., SomeCollege, College, Graduate\}$, with high-skilled workers being those who earn a wage higher than the mean. Then, I compute

the hazard rate out of unemployment for each time horizon $(\pi_t)_{t \geq 1}$, from the cross-section of jobless workers who report to have been unemployed for t periods of time, using the following formulas

$$\begin{aligned}\pi_1 &= 1 - Prob(t > 1) = 1 - \frac{\# \text{jobless for } t > 1}{\# \text{jobless}} \\ \pi_1 + (1 - \pi_1)\pi_2 &= 1 - Prob(t > 2) = 1 - \frac{\# \text{jobless for } t > 2}{\# \text{jobless}} \\ &\dots\end{aligned}$$

Third, by looking at the same cross-sections, I compute the share of those with same spell duration (at the time the survey is conducted) who also have attained the same educational level, ψ_{it} . Lastly, I compute $\{\pi_H, \pi_L\}$ that minimize

$$\{\pi_H, \pi_L\} = \arg \min_{\hat{\pi}_H, \hat{\pi}_L} \sum_t \left(\sum_i \psi_{it} (\vartheta_i \hat{\pi}_H + (1 - \vartheta_i) \hat{\pi}_L) - \pi_t \right)^2$$

that is,

$$\pi_H = \frac{\sum_t b_t \sum_s \pi_s a_s - \sum_s \pi_s \sum_t a_t b_t}{12 \sum_t a_t^2 - (\sum_t a_t)^2}, \quad \pi_L = \frac{(\sum_t \pi_t)(\sum_s a_s^2) - \sum_s \pi_s a_s \sum_t a_t}{12 \sum_t a_t^2 - (\sum_t a_t)^2}$$

with $a_t = \sum_i \psi_{it} \vartheta_i$, $b_t = \sum_i \psi_{it} (1 - \vartheta_i) = 1 - a_t$.⁴⁸ The results are reported in Table 5. The hazard rate π_t is quite stable over time, as well as the share of any education level among all jobless people with same duration of unemployment spell, ψ_{it} . The estimated hazard rates are $\pi_H = 0.27$ and $\pi_L = 0.14$.

⁴⁸First-order conditions for π_H and π_L return the minimizers of the convex objective function.

	Total	< High Sch.	High Sch. D.	< Col. ⁴⁹	Col. D.	Grad. D.	
ϑ_i	39,333	0.54	0.72	0.76	0.9	0.95	
Horizon	Total	$\psi_{it}=\Pr(\text{Education}=i \mid \text{Horizon} \geq t)$					Haz. Rate (π_t)
t=1	3,481	0.11	0.31	0.29	0.28	0.01	0.22
t=2	2,517	0.11	0.32	0.29	0.27	0.01	0.28
t=3	1,742	0.11	0.32	0.29	0.27	0.01	0.31
t=4	1,316	0.11	0.32	0.29	0.28	0.01	0.24
t=5	1,081	0.11	0.32	0.29	0.28	0	0.18
t=6	815	0.11	0.33	0.28	0.27	0	0.25
t=7	586	0.12	0.33	0.28	0.27	0	0.28
t=8	468	0.12	0.33	0.28	0.27	0	0.2
t=9	356	0.11	0.32	0.29	0.27	0	0.24
t=10	274	0.11	0.31	0.28	0.29	0	0.23
t=11	215	0.11	0.33	0.26	0.29	0	0.22
t=12	167	0.11	0.34	0.25	0.3	0	0.22

Table 5: Education-cohort size for any unemployment spell duration.

⁴⁹'< Col.' item includes workers who attended college, but have not earned a degree, and workers with an Associate Degree, which is a post-secondary course of study lasting 2 or 3 years.

Annex I Missouri Self-Eligibility Assessment Form



**DIVISION OF
EMPLOYMENT
SECURITY**

P.O. Box 3915
Jefferson City, MO 65102-3915
www.labor.mo.gov/DES

Date Mailed: _____

Reemployment Services and Eligibility Assessment

You have been selected to participate in a Reemployment Services and Eligibility Assessment (RESEA) program designed to help you return to work. As a condition of continued eligibility for unemployment benefits, you must participate in an interview at the Missouri Job Center shown below. If you have already returned to work, contact the Missouri Job Center at the phone number listed below.

In the interview, you will receive individual assistance with your reemployment plan, reemployment services and finding training that may be available.

You are required to bring photo identification, your completed Work Search Record, and your completed Self-Evaluation Questionnaire (enclosed) to the interview. For your convenience, a Work Search Record is located on the reverse of the Self-Evaluation Questionnaire. If you have a resume, bring the most recent version with you.

The interview may last up to 90 minutes. You can significantly reduce the time you spend at the Job Center if you pre-register with MoJobs at <https://app-jobs.mo.gov> prior to your appointment.

Participation in this program is mandatory and failure to attend any scheduled RESEA appointment will result in an interruption of your unemployment benefits. If you are unable to attend this interview for any reason, you must call the Missouri Job Center prior to the scheduled appointment date and time and explain why you cannot attend.

Report on _____ at _____ to:

MISSOURI JOB CENTER

Local Phone _____

*IMPORTANT: If needed, call 573-751-9040 for assistance in the translation and understanding of the information in this document.
¡IMPORTANTE!: Si es necesario, llame al 573-751-9040 para asistencia en la traducción y entendimiento de la información en este documento.
Missouri Division of Employment Security is an equal opportunity employer/program. Auxiliary aids and services are available upon request to individuals with disabilities. TDD/TTY: 800-735-2966 Relay Missouri: 711*

**MISSOURI
DEPARTMENT OF LABOR
& INDUSTRIAL RELATIONS**

MODES-4632 (04-18)
Benefits



DIVISION OF EMPLOYMENT SECURITY

MISSOURI DEPARTMENT OF LABOR AND INDUSTRIAL RELATIONS ELIGIBILITY ASSESSMENT

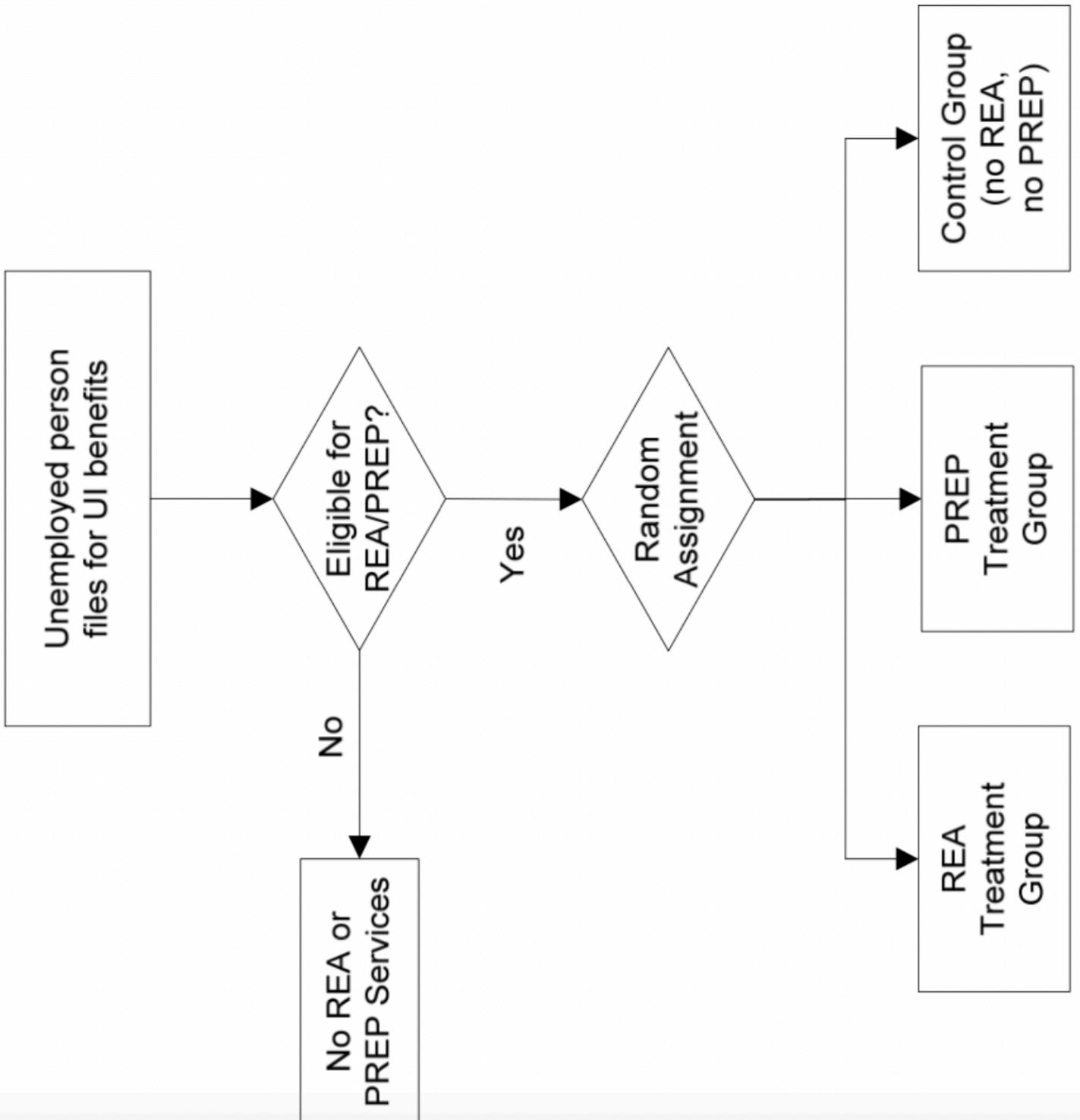
1. Have you worked since you filed for unemployment insurance benefits? This includes full-time work, part-time work, or temporary work. Yes No If Yes, provide dates of employment.
Beginning Employment Date: _____ Ending Employment Date: _____
2. Please provide your rate of pay on your last job.
Hourly wage: \$ _____ or Salary: \$ _____ Weekly Monthly
3. How much experience did you have on that job? (check one)
 Less than 6 months 6 months – 1 year 1 year – 3 years 3 years – 5 years 5+ years
4. Are you looking for: Full-time work Part-time work Both
5. What type of work are you seeking?
 Construction Retail Office Services Management
 Manufacturing Transportation Health Care Other _____
6. What days are you available for work? (check all that apply)
 Sunday Monday Tuesday Wednesday Thursday Friday Saturday
7. What hours are you available for work?
From: _____ a.m. p.m. To: _____ a.m. p.m.
8. What is the lowest pay you will accept for work?
Hourly wage: \$ _____ or Salary: \$ _____ Weekly Monthly
9. What type of transportation do you have to get to a job? (check one)
 Private Vehicle Public Transportation Family/Friend None Other _____
10. How many miles are you willing to travel to a job (one way)? (check one)
 0-5 miles 5-10 miles 10-20 miles 20-30 miles More _____
11. Do you attend or plan to attend school or training? Yes No
If currently attending school or training, provide name of educational or training institution: _____
12. Are you self-employed? Yes No
If Yes, please provide the number of hours worked per week. _____ hours worked per week.
13. Do you have limitations that may keep you from performing the type of work that you are seeking?
 Yes No If Yes, please explain. _____

14. Do you have dependents who require care during work hours? Yes No
If Yes, will you be able to make arrangements for the dependents if you are offered work? Yes No

Name _____

Date _____

Annex II Florida REA Selection Process



Technical Appendix

Setting

- $T < \infty$
- $\sigma_t \in \{u, w\}$ describes the worker status, either unemployed or employed. If $\sigma_t = w$, the worker finds reemployment, which is an absorbing state. Hence $p(\sigma_{t+1} = w | \sigma_t = w) = 1$.
- $\sigma^t = \{\sigma_0, \dots, \sigma_t\}$ is a public history describing the employment status of the worker
- $c_t(\sigma^t)$ is the transfer function, with $c_t(\sigma^t) \geq 0$ for every σ^t . Let $\mathbf{c}(\alpha \setminus \sigma^t)$ be the stream of transfers downstream of node σ^t
- $a_t(\sigma^t)$ is the effort level, with

$$a_t(\sigma^t) \in \begin{cases} \{0, e\}, & \text{if } \sigma_t = u \\ 0, & \text{if } \sigma_t = w \end{cases}$$

The effort is unobservable by the government. Denote by $\mathbf{a}(\alpha \setminus \sigma^t)$ the continuation plan of effort costs downstream of node σ^t , and $\mathbf{a}(\sigma^t)$ its upstream counterpart

- $h \in \{L, H\}$ is the hidden state, which is revealed once the worker finds reemployment
- $\mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1}))$ -with $\sigma^t = (\sigma^{t-1}, \sigma_t)$ - is the expectation held by the worker during unemployment, expressing the probability about state H. This is clearly a non-contractible variable, as the worker can hide it from the government. $\mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1}))$ impacts the probability of future σ_{t+1} . in particular,

$$p(\sigma_{t+1} = w | \sigma_t = u, \mu_t, a_t(\sigma^t) = e) = \pi(\mu_t), \quad p(\sigma_{t+1} = w | \sigma_t = u, \mu_t, a_t(\sigma^t) = 0) = 0$$

where I have dropped dependence of μ_t by $(\sigma^t, \mathbf{a}(\sigma^{t-1}))$ to ease notation. Moreover, $\mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1}))$ undergoes an updating process every time the worker exerts effort in t and remains unemployed in $t + 1$

$$\mu_{t+1}(\sigma^t, \mathbf{a}(\sigma^{t-1}), \sigma_{t+1} = u, a_t(\sigma^t) = e) = \frac{\mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1}))(1 - \pi_H)}{\mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1}))(1 - \pi_H) + (1 - \mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1}))) (1 - \pi_L)} \quad (31)$$

Instead, if no effort is exerted, the worker does not revise expectation⁵⁰

$$\mu_{t+1}(\sigma^t, \mathbf{a}(\sigma^{t-1}), \sigma_{t+1} = u, a_t(\sigma^t) = 0) = \mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1})) \quad (32)$$

⁵⁰Notice that $\mu_{t+1}(\sigma^t, \mathbf{a}(\sigma^{t-1}), \sigma_{t+1} = w, a_t(\sigma^t) = 0)$ is not defined as h is disclosed once $\sigma_{t+1} = w$.

- $r_s(\sigma^s)$ is the return for the government in period s upon history σ^s . If $\sigma_s = w$,

$$r_s(\sigma^s) = \omega(\mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1}))), \quad \text{with } t = \inf_s \{y_s = w\} - 1$$

Otherwise, if $\sigma_t = u$, $r_s(\sigma^s) = 0$.

Worker's Problem in UI

Let $\mathcal{W}(\sigma^t) = (\mathbf{c}, \mathbf{a})(\alpha \setminus \sigma^t) = \{c_s(\sigma^s), a_s(\sigma^s)\}_{s=t}^T$ denote the contract offered by the government to the worker. Worker's expected utility reads

$$\begin{aligned} U_t(\mathcal{W}, \mathbf{a}(\sigma^{t-1}), \sigma^t) &= \mathbf{E} \left\{ \sum_{s=t}^T \beta^{s-t} (u(c_s(\sigma^s)) - a_s(\sigma^s)) \middle| \mathcal{W}(\sigma^t), \mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1})) \right\} + \\ &+ \beta^{T+1-t} \sum_{\sigma^{T+1}} p(\sigma^{T+1} | \sigma^t, \mu_t, \mathbf{a}(\sigma^T)) U_{T+1}(\sigma^{T+1}) \\ &= \sum_{s=t}^T \beta^{s-t} \sum_{\sigma^s} p(\sigma^s | \sigma^t, \mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1})), a_t(\sigma^t)) \left\{ u(c_s(\sigma^s)) - a_s(\sigma^s) \middle| \mathcal{W}(\sigma^t) \right\} \\ &+ \beta^{T+1-t} \sum_{\sigma^{T+1}} p(\sigma^{T+1} | \sigma^t, \mu_t, \mathbf{a}(\sigma^T)) U_{T+1}(\sigma^{T+1}) \\ &= u(c_t(\sigma^t)) - a_t(\sigma^t) + \\ &+ \beta \left[p(\sigma_{t+1} = w | \sigma_t = u, \mu_t, a_t(\sigma^t)) \sum_{s=t+1}^T \beta^{s-(t+1)} \sum_{h \in \{H, L\}} p(h | \sigma_{t+1} = w, \mu_t) \left\{ u(c_s(\sigma^s)) - e \middle| \mathcal{W}'(\sigma^t, w, h) \right\} + \right. \\ &+ p(\sigma_{t+1} = u | \sigma_t = u, \mu_t, a_t(\sigma^t)) \sum_{s=t+1}^T \beta^{s-(t+1)} \times \\ &\quad \times \sum_{\sigma^s} p(\sigma^s | \sigma_{t+1} = u, \mu_{t+1}^u, a_{t+1}(\sigma^t, u)) \left\{ u(c_s(\sigma^s)) - a_s(\sigma^s) \middle| \mathcal{W}'(\sigma^t, u) \right\} \left. \right] + \\ &+ \beta^{T+1-t} \sum_{\sigma^{t+1}} p(\sigma^{t+1} | \sigma^t, \mu_t, \mathbf{a}(\sigma^t)) \sum_{\sigma^{T+1}} p(\sigma^{T+1} | \sigma^{t+1}, \mu_{t+1}, \mathbf{a}(\sigma^T)) U_{T+1}(\sigma^{T+1}) \\ &= u(c_t(\sigma^t)) - e + \\ &+ \beta \sum_{h \in \{H, L\}} p(h | \mu_t) p(\sigma_{t+1} = w | h) \left[\sum_{s=t+1}^T \beta^{s-(t+1)} \left\{ u(c_s(\sigma^s)) - e \middle| \mathcal{W}'(\sigma^t, w, h) \right\} + \beta^{T-t} U_{T+1}(\sigma^t, \mathbf{w}, h) \right] + \\ &+ \beta(1 - \pi(\mu_t)) \left[\sum_{s=t+1}^T \beta^{s-(t+1)} \sum_{\sigma^s} p(\sigma^s | \sigma_{t+1} = u, \mu_{t+1}^u, a_{t+1}(\sigma^t, u)) \left\{ u(c_s(\sigma^s)) - a_s(\sigma^s) \middle| \mathcal{W}'(\sigma^t, u), \mu_{t+1}^u \right\} + \right. \\ &\quad \left. + \sum_{\sigma^{T+1}} p(\sigma^{T+1} | \sigma^{t+1}, \mu_{t+1}^u, \mathbf{a}(\sigma^T)) U_{T+1}(\sigma^{T+1}) \right] \\ &= u(c_t(\sigma^t)) - e + \beta \left\{ \pi(\mu_t) \left[\frac{\mu_t \pi_H}{\pi(\mu_t)} U_{t+1}(\mathcal{W}', \mathbf{a}(\sigma^t), (\sigma^t, w, H)) + \frac{(1 - \mu_t) \pi_L}{\pi(\mu_t)} U_{t+1}(\mathcal{W}', \mathbf{a}(\sigma^t), (\sigma^t, w, L)) \right] + \right. \\ &\quad \left. + (1 - \pi(\mu_t)) U_{t+1}(\mathcal{W}', \mathbf{a}(\sigma^t), (\sigma^t, u)) \right\} \end{aligned}$$

with $\mu_{t+1}^u = \mu_{t+1}(\sigma^t, \sigma_{t+1} = u, \mathbf{a}(\sigma^{t-1}), a_t(\sigma^t))$ ⁵¹.

The IC constraint starting from time t reads

$$U_t(\mathcal{W}, \mathbf{a}(\sigma^{t-1}), \sigma^t) \geq U_t((\mathbf{c}, \hat{\mathbf{a}})(\alpha \setminus \sigma^t), \mathbf{a}(\sigma^{t-1}), \sigma^t), \quad \forall \hat{\mathbf{a}}(\alpha \setminus \sigma^t) \in A_t(\alpha \setminus \sigma^t) \quad (33)$$

Government's Problem in UI

The problem for the government reads

$$V^{UI}(U, \mathbf{a}(\sigma^{t-1}), \sigma^t) = \max_{\mathcal{W}} \mathbf{E} \left\{ \sum_{s=t}^T \beta^{s-t} (r_s(\sigma^s) - c_s(\sigma^s)) \middle| \mathcal{W}(\sigma^t), \mu_t(\sigma^t, \mathbf{a}(\sigma^{t-1})) \right\}$$

sub: (33), $U_s(\mathcal{W}'(\sigma^s), \mathbf{a}(\sigma^{s-1}), \sigma^s) \geq U, \quad \forall s \geq t$

Given that expectation is revised upon (failure and) effort exerted only in the last period, it follows a Markovian process, meaning that expectation in $t+1$ can be predicted by expectation μ_t and effort in t , and realization of σ_{t+1} . Thus, I define $x_t = (\mu_t(\sigma^t, x_{t-1}), a_t(\sigma^t))$ and write

$$\mu_{t+1}(\sigma^t, \sigma_{t+1} = u, \mathbf{a}(\sigma^t)) = \mu_{t+1}(\sigma^t, \sigma_{t+1} = u, \mu_t(\sigma^t, x_{t-1}), a_t(\sigma^t)) = \mu_{t+1}(\sigma^t, \sigma_{t+1} = u, x_t)$$

And given that reemployment is absorbing and discloses the state, there exists an isomorphism between all unemployment histories $(\sigma^s, \sigma_{s+1} = u)$ and terminal realizations $\sigma_{s+1} = u$, as no extra information is contained in σ^s which can not be inferred by observing $\sigma_{s+1} = u$. As a result, next expectation μ_{t+1} only depends on current expectation μ_t and effort a_t and future realization of σ_{t+1} :

$$\mu_{t+1}(\sigma^t, \sigma_{t+1} = u, x_t) = \mu_{t+1}(\sigma_{t+1} = u, \mu_t, a_t) = \begin{cases} \mu_t, & \text{if } a_t = 0 \\ \frac{\mu_t(1-\pi_H)}{1-\pi(\mu_t)}, & \text{if } a_t = e \end{cases}$$

Therefore:

$$\begin{aligned} U_t(\mathcal{W}, \mu_t, \sigma^t) &= u(c_t(\sigma^t)) - e + \\ &+ \beta \left\{ \pi(\mu_t) \left[\frac{\mu_t \pi_H}{\pi(\mu_t)} U_{t+1}(\mathcal{W}', H, \sigma^t, \sigma_{t+1} = w) + \frac{(1-\mu_t)\pi_L}{\pi(\mu_t)} U_{t+1}(\mathcal{W}', L, \sigma^t, \sigma_{t+1} = w) \right] + \right. \\ &\quad \left. + (1-\pi(\mu_t)) U_{t+1}(\mathcal{W}', \mu_{t+1}^u, \sigma^t, \sigma_{t+1} = u) \right\} \\ &= u(c_t(\sigma^t)) - e + \beta \left\{ \mu_t \pi_H U_{t+1}(\mathcal{W}', H, \sigma^t, \sigma_{t+1} = w) + (1-\mu_t) \pi_L U_{t+1}(\mathcal{W}', L, \sigma^t, \sigma_{t+1} = w) + \right. \\ &\quad \left. + (1-\pi(\mu_t)) U_{t+1}(\mathcal{W}', \mu_{t+1}^u, \sigma^t, \sigma_{t+1} = u) \right\} \end{aligned}$$

⁵¹Note that by assumption on absorbing nature of re-employment, $\sigma^s = (\sigma^t, (y_j = w)_{j+1}^s)$, $\forall \sigma^s \succeq (\sigma^t, \sigma_{t+1} = w)$.

Government's problem can be rewritten as

$$\begin{aligned}
V^{UI}(U, \mu_t, \sigma^t) &= \max_{\mathcal{W} \in \Omega(\mu_t, \sigma^t)} \mathbf{E} \left\{ \sum_{s=t}^T \beta^{s-t} (r_s(\sigma^s) - c_s(\sigma^s)) \middle| \mathcal{W}(\sigma^t), \mu_t \right\} \\
&= \max_{c_t(\sigma^t), a_t(\sigma^t) \in \Gamma(\mu_t, \sigma^t)} r_t(\sigma^t) - c_t(\sigma^t) + \\
&+ \beta \left[\pi(\mu_t) \max_{\mathcal{W}' \in \Omega'(\mu_{t+1}, \sigma^t, \sigma_{t+1}=w)} \mathbf{E} \left\{ \sum_{s=t+1}^T \beta^{s-(t+1)} \left(\overbrace{r_s(\sigma^s)}^{=\omega(\mu_t)/\pi(\mu_t)} - c_s(\sigma^s) \right) \middle| \mathcal{W}'(\sigma^t, \sigma_{t+1}=w), \mu_{t+1} \right\} + \right. \\
&\left. + (1 - \pi(\mu_t)) \max_{\mathcal{W}' \in \Omega'(\mu_{t+1}, \sigma^t, \sigma_{t+1}=u)} \mathbf{E} \left\{ \sum_{s=t+1}^T \beta^{s-(t+1)} (r_s(\sigma^s) - c_s(\sigma^s)) \middle| \mathcal{W}'(\sigma^t, \sigma_{t+1}=u), \mu_{t+1} \right\} \right]
\end{aligned}$$

$$\text{sub: } U_t(\mathcal{W}, \mu_t, \sigma^t) \geq U_t((\mathbf{c}, \mathbf{a}')(\alpha \setminus \sigma^t), \mu_t, \sigma^t), \quad \forall \mathbf{a}'(\alpha \setminus \sigma^t) \in A_t(\alpha \setminus \sigma^t) \quad (IC)$$

$$\mu_t := \mu_t(\sigma^t, \mu_{t-1}, a_{t-1}(\sigma^{t-1}))$$

Lemma 4. *The government prefers to insure the worker against the risk of h realization upon reemployment.*

Lemma 5. *Define (IC, s) the constraint that makes contract \mathcal{W} robust to the alternative strategy $\mathbf{a}'(\alpha \setminus \sigma^s) = (0, e \mathbf{1}_{T-s}) \in A_s(\alpha \setminus \sigma^s)$ that shirks in s and sticks to effort from $s+1$ to the final period T*

$$U_s(\mathcal{W}, \mu_s, \sigma^s) \geq U_s((\mathbf{c}, \mathbf{a}')(\alpha \setminus \sigma^s), \mu_s, \sigma^s) = u(c_s(\sigma^s)) + \beta U_{s+1}(\mathcal{W}, \mu_s, (\sigma^s, u))$$

If (IC, s) $_{s=t}^T$ are all binding under contract \mathcal{W} , then contract \mathcal{W} is feasible.

Proof. First, consider that no learning motive or moral hazard problem is present upon reemployment, when state is disclosed, nor there is any chance that any reemployed W falls back into unemployment ($p(y_s = u | y_t = w) = 0, \forall s > t$). Hence in order to verify (33), one can only focus on continuation histories $\sigma^s \succsim \sigma^t$ where $\sigma^s = (\sigma^t, (y_j)_{j=t+1}^s) = (\sigma^t, (u)_{j=t+1}^s)$. For this reason, I therefore adopt the convention that $\sigma^{s+1} = (\sigma^s, u)$. Define also μ_s as the expectation in period s if contract \mathcal{W} is followed.

Notice that continuation utility at time $s > t$ upon reemployment ($y_s = w$) follows

$$U_s(\mathcal{W}, h, \sigma^s) = u(c_s(\sigma^s)) - e + \beta U_{s+1}(\mathcal{W}, h, (\sigma^s, w))$$

while upon failure ($y_s = u$), it follows

$$\begin{aligned}
U_s(\mathcal{W}, \mu_s, \sigma^s) &= u(c_s(\sigma^s)) - e + \beta \left[\mu_s \pi_H U_{s+1}(\mathcal{W}, H, (\sigma^s, w)) + (1 - \mu_s) \pi_L U_{s+1}(\mathcal{W}, L, (\sigma^s, w)) + \right. \\
&\left. + (1 - \pi(\mu_s)) U_{s+1}(\mathcal{W}, \mu_{s+1}^u, (\sigma^s, u)) \right]
\end{aligned}$$

$$\text{with: } \mu_{s+1}^u = \mu_{s+1}(\sigma^s, \sigma_{s+1} = u, \mu_s, e)$$

By Lemma 4, focusing on contracts \mathcal{W} such that

$$\begin{aligned} U_{s+1}(\mathcal{W}, H, (\sigma^s, w)) &= U_{s+1}(\mathcal{W}, L, (\sigma^s, w)) \\ \implies \mu\pi_H U_{s+1}(\mathcal{W}, H, (\sigma^s, w)) + (1 - \mu)\pi_L U_{s+1}(\mathcal{W}, L, (\sigma^s, w)) &= \pi(\mu)U_{s+1}(\mathcal{W}, (\sigma^s, w)) \end{aligned}$$

is without loss of generality.

Second, the following holds true:

$$\begin{aligned} U_s(\mathcal{W}, \mu_s, \sigma^s) &= u(c_s(\sigma^s)) + \beta U_{s+1}(\mathcal{W}, \mu_s, \sigma^{s+1}) \\ \implies U_s(\mathcal{W}, \mu, \sigma^s) &\geq u(c_s(\sigma^s)) + \beta U_{s+1}(\mathcal{W}, \mu, \sigma^{s+1}), \quad \forall \mu : \mu > \mu_s \end{aligned} \quad (34)$$

The proof of (34) will be given by induction, joint with the main statement.

Base Step ($t = T$)

UI contract ends in $t = T$, where the only possible deviation is $\hat{\mathbf{a}}(\alpha \setminus \sigma^T) = \hat{a}_T(\sigma^T) = 0$. Thus, for W to be robust to this deviation, it must be that

$$\begin{aligned} U_T(\mathcal{W}, \mu_T, \sigma^T) &= u(c_T(\sigma^T)) - e + \beta [\mu_T \pi_H U_{T+1}(\sigma^T, w, H) + (1 - \mu_T) \pi_L U_{T+1}(\sigma^T, w, L) + \\ &\quad + (1 - \pi(\mu_T)) U_{T+1}(\sigma^T, u)] \\ U_T(\mathcal{W}, \mu_T, \sigma^T) &\geq u(c_T(\sigma^T)) + \beta U_{T+1}(\sigma^T, u) \end{aligned}$$

Since (IC, T) is binding by assumption, it holds

$$U_T(\mathcal{W}, \mu_T, \sigma^T) = u(c_T(\sigma^T)) + \beta U_{T+1}(\sigma^T, u), \quad U_{T+1}(\sigma^T, w) - U_{T+1}(\sigma^T, u) = \frac{e}{\beta \pi(\mu_T)} > 0$$

and then, for $\mu > \mu_T$,

$$\begin{aligned} &U_T(\mathcal{W}, \mu, \sigma^T) - \beta U_{T+1}(\mathcal{W}, \mu, (\sigma^T, u)) = \\ &= u(c_T(\sigma^T)) - e + \beta \pi(\mu) [U_{T+1}(\mathcal{W}, (\sigma^T, w)) - U_{T+1}(\mathcal{W}, (\sigma^T, u))] > \\ &> u(c_T(\sigma^T)) - e + \beta \pi(\mu_T) [U_{T+1}(\mathcal{W}, (\sigma^T, w)) - U_{T+1}(\mathcal{W}, (\sigma^T, u))] = \\ &= U_T(\mathcal{W}, \mu_T, \sigma^T) - \beta U_{T+1}(\mathcal{W}, \mu_T, (\sigma^T, u)) \end{aligned}$$

which proves (34) for $t = T$.

Induction Step ($t \leq T - 1$)

First, notice that

$$\begin{aligned}
U_t(\mathcal{W}, \mu_t, \sigma^t) &= u(c_t(\sigma^t)) + \beta U_{t+1}(\mathcal{W}, \mu_t, \sigma^{t+1}) \\
\implies -e + \beta\pi(\mu_t) [U_{t+1}(\mathcal{W}, (\sigma^t, w)) - U_t(\mathcal{W}, \mu_{t+1}, \sigma^{t+1})] \\
&= \beta [U_{t+1}(\mathcal{W}, \mu_t, \sigma^{t+1}) - U_{t+1}(\mathcal{W}, \mu_{t+1}, \sigma^{t+1})] \\
&= \beta [-e + \beta\pi(\mu_t) (U_{t+2}(\mathcal{W}, (\sigma^{t+1}, w)) - U_{t+2}(\mathcal{W}, \mu_{t+1}, \sigma^{t+2}))] \\
\implies U_{t+1}(\mathcal{W}, (\sigma^t, w)) - U_{t+1}(\mathcal{W}, \mu_{t+1}, \sigma^{t+1}) &> \beta [U_{t+2}(\mathcal{W}, (\sigma^{t+1}, w)) - U_{t+2}(\mathcal{W}, \mu_{t+1}, \sigma^{t+2})]
\end{aligned} \tag{35}$$

where the equalities hold since (IC, t) and (IC, $t+1$) are binding. Now, by induction hypothesis, $\forall \mu : \mu > \mu_{t+1}$,

$$\begin{aligned}
U_{t+1}(\mathcal{W}, \mu_{t+1}, \sigma^{t+1}) &= u(c_{t+1}(\sigma^{t+1})) + \beta U_{t+2}(\mathcal{W}, \mu_{t+1}, \sigma^{t+2}) \\
\implies U_{t+1}(\mathcal{W}, \mu, \sigma^{t+1}) &\geq u(c_{t+1}(\sigma^{t+1})) + \beta U_{t+2}(\mathcal{W}, \mu, \sigma^{t+2})
\end{aligned}$$

Therefore, take any $\mu > \mu_t$ and

$$\begin{aligned}
U_t(\mathcal{W}, \mu_t, \sigma^t) - \beta U_{t+1}(\mathcal{W}, \mu_t, \sigma^{t+1}) &= u(c_t(\sigma^t)) - \beta u(c_{t+1}(\sigma^{t+1})) - e(1 - \beta) + \\
+ \beta\pi(\mu_t) [U_{t+1}(\mathcal{W}, (\sigma^t, w)) - \beta U_{t+2}(\mathcal{W}, (\sigma^{t+1}, w))] &+ \beta(1 - \pi(\mu_t)) [U_{t+1}(\mathcal{W}, \mu_t^u, \sigma^{t+1}) - \beta U_{t+2}(\mathcal{W}, \mu_t^u, \sigma^{t+2})] \\
< u(c_t(\sigma^t)) - \beta u(c_{t+1}(\sigma^{t+1})) - e(1 - \beta) &+ \\
+ \beta\pi(\mu) [U_{t+1}(\mathcal{W}, (\sigma^t, w)) - \beta U_{t+2}(\mathcal{W}, (\sigma^{t+1}, w))] &+ \beta(1 - \pi(\mu)) [U_{t+1}(\mathcal{W}, \mu_t^u, \sigma^{t+1}) - \beta U_{t+2}(\mathcal{W}, \mu_t^u, \sigma^{t+2})] \\
< u(c_t(\sigma^t)) - \beta u(c_{t+1}(\sigma^{t+1})) - e(1 - \beta) &+ \\
+ \beta\pi(\mu) [U_{t+1}(\mathcal{W}, (\sigma^t, w)) - \beta U_{t+2}(\mathcal{W}, (\sigma^{t+1}, w))] &+ \beta(1 - \pi(\mu)) [U_{t+1}(\mathcal{W}, \mu^u, \sigma^{t+1}) - \beta U_{t+2}(\mathcal{W}, \mu^u, \sigma^{t+2})] \\
= U_t(\mathcal{W}, \mu, \sigma^t) - \beta U_{t+1}(\mathcal{W}, \mu, \sigma^{t+1})
\end{aligned}$$

where the first inequality follows from (35) above, as $\pi(\mu) > \pi(\mu_t)$, while the second inequality follows from induction hypothesis. I can thus conclude that (34) holds also for t .

I now pass to the proof of the main part of the proposition, that is, that binding IC constraints is a sufficient condition to account for all possible deviations occurring from t onward. By induction hypothesis, \mathcal{W} satisfies all $(IC, s)_{s=t}^T$ with equality, and that guarantees robustness to all possible deviations over histories in $\alpha \setminus \sigma^{t+1}$, i.e.

$$U_{t+1}(\mathcal{W}, \mu_{t+1}, \sigma^{t+1}) \geq U_{t+1}((\mathbf{c}, \mathbf{a}')(\alpha \setminus \sigma^{t+1}), \mu_{t+1}, \sigma^{t+1}), \quad \forall \mathbf{a}'(\alpha \setminus \sigma^{t+1}) \in A_{t+1}(\alpha \setminus \sigma^{t+1})$$

What it is to show is that \mathcal{W} is robust also to all possible deviations in $\alpha \setminus \sigma^t$, i.e.

$$U_t(\mathcal{W}, \mu_t, \sigma^t) \geq U_t((\mathbf{c}, \mathbf{a}')(\alpha \setminus \sigma^t), \mu_t, \sigma^t), \quad \forall \mathbf{a}'(\alpha \setminus \sigma^t) \in A_t(\alpha \setminus \sigma^t)$$

First of all, notice that $A_t(\alpha \setminus \sigma^t) = \{0, e\} \times A_{t+1}(\alpha \setminus \sigma^{t+1})$ can be decomposed into:

- all effort histories with positive effort in t , i.e. $A_e = A_t(\alpha \setminus \sigma^t) \cap \{a_t(\sigma^t) = e\}$;
- all effort histories with zero effort in t , i.e. $A_0 = A_t(\alpha \setminus \sigma^t) \cap \{a_t(\sigma^t) = 0\}$;

Second, assumption on robustness to any $\mathbf{a}'(\alpha \setminus \sigma^{t+1}) \in A_{t+1}(\alpha \setminus \sigma^{t+1})$ guarantees robustness of \mathcal{W} to the first set of deviations A_e , since $\mu_{t+1} = \frac{\mu_t(1-\pi_H)}{1-\pi(\mu_t)} = \mu_t^u$. Indeed, pick any $\mathbf{a}'(\alpha \setminus \sigma^t) \in A_e$. Then, it follows

$$\begin{aligned} U_t(\mathcal{W}, \mu_t, \sigma^t) &= u(c_t(\sigma^t)) - e + \beta[\pi(\mu^t)U_{t+1}(\mathcal{W}, \mu_t^w, (\sigma^t, w)) + (1 - \pi(\mu^t))U_{t+1}(\mathcal{W}, \mu_t^u, \sigma^{t+1})] \\ &\geq u(c_t(\sigma^t)) - e + \beta[\pi(\mu^t)U_{t+1}(\mathcal{W}, \mu_t^w, (\sigma^t, w)) + (1 - \pi(\mu^t))U_{t+1}(\mathcal{W}', \mu_t^u, \sigma^{t+1})] = U_t(\mathcal{W}', \mu_t, \sigma^t) \end{aligned}$$

where the inequality follows from robustness to $\mathbf{a}'(\alpha \setminus \sigma^{t+1}) \in A_{t+1}(\alpha \setminus \sigma^{t+1})$.

What is left to show is robustness of \mathcal{W} to A_0 . By assumption, (IC, t) and (IC, $t+1$) are binding, which means that

$$U_t(\mathcal{W}, \mu_t, \sigma^t) = U_t((\mathbf{c}, \tilde{\mathbf{a}})(\alpha \setminus \sigma^t), \mu_t, \sigma^t) = U_t((\mathbf{c}, \hat{\mathbf{a}})(\alpha \setminus \sigma^t), \mu_t, \sigma^t) \quad (36)$$

with $\hat{\mathbf{a}}(\alpha \setminus \sigma^t) = (0, e\mathbf{1}_k)$, $\tilde{\mathbf{a}}(\alpha \setminus \sigma^t) = (e, 0, e\mathbf{1}_{k-1})$. Define $\tilde{\mathcal{W}} = (\mathbf{c}, \tilde{\mathbf{a}})(\alpha \setminus \sigma^t)$ and $\hat{\mathcal{W}} = (\mathbf{c}, \hat{\mathbf{a}})(\alpha \setminus \sigma^t)$.

Thus, by construction

$$U_{t+2}(\tilde{\mathcal{W}}, \mu_t^u, \sigma^{t+2}) = U_{t+2}(\hat{\mathcal{W}}, \mu_t^u, \sigma^{t+2}) \quad (37)$$

Indeed, both alternative strategies prescribe to set effort cost to 0 either at stage t or $t+1$ (but not both), and therefore the expectation at node $\sigma^{t+2} = (\sigma^t, u, u)$ is equal to μ_t^u under both strategies. Moreover, they prescribe positive effort forever after, until the last period T .

Pick any $\mathbf{a}'(\alpha \setminus \sigma^t) \in A_0$. There are two possibilities: $a'_{t+1}(\sigma^{t+1}) = e$ or $a'_{t+1}(\sigma^{t+1}) = 0$. If the first case applies, consider the alternative deviation strategy $\mathbf{a}''(\alpha \setminus \sigma^t) \in A_e$ so constructed:

$$a''_t(\sigma^t) = e, \quad a''_{t+1}(\sigma^{t+1}) = 0, \quad \mathbf{a}''(\alpha \setminus \sigma^{t+2}) = \mathbf{a}'(\alpha \setminus \sigma^{t+2})$$

Likewise, define $\mathcal{W}' = (\mathbf{c}, \mathbf{a}')(\alpha \setminus \sigma^t)$ and $\mathcal{W}'' = (\mathbf{c}, \mathbf{a}'')(\alpha \setminus \sigma^t)$. Hence, by construction,

$$U_{t+2}(\mathcal{W}', \mu_t^u, \sigma^{t+2}) = U_{t+2}(\mathcal{W}'', \mu_t^u, \sigma^{t+2}) \quad (38)$$

for the same reason as in (37), and

$$\begin{aligned} U_t(\mathcal{W}', \mu_t, \sigma^t) &= U_t(\hat{\mathcal{W}}, \mu_t, \sigma^t) + \beta^2(1 - \pi(\mu_t)) [U_{t+2}(\mathcal{W}', \mu_t^u, \sigma^{t+2}) - U_{t+2}(\hat{\mathcal{W}}, \mu_t^u, \sigma^{t+2})] \\ U_t(\mathcal{W}'', \mu_t, \sigma^t) &= U_t(\tilde{\mathcal{W}}, \mu_t, \sigma^t) + \beta^2(1 - \pi(\mu_t)) [U_{t+2}(\mathcal{W}'', \mu_t^u, \sigma^{t+2}) - U_{t+2}(\tilde{\mathcal{W}}, \mu_t^u, \sigma^{t+2})] \end{aligned}$$

which follows from the fact that \mathcal{W}' is identical to $\hat{\mathcal{W}}$ in periods t and $t + 1$, and the same holds true for \mathcal{W}'' and $\tilde{\mathcal{W}}$.

One can easily see that the RHS of the two equations are equal, by (36), (37) and (38), which causes also the LHS to be equal

$$U_t(\mathcal{W}', \mu_t, \sigma^t) = U_t(\mathcal{W}'', \mu_t, \sigma^t)$$

But then, given that \mathcal{W} is robust to any alternative strategy in A_e ,

$$\mathbf{a}''(\alpha \setminus \sigma^t) \in A_e \implies U_t(\mathcal{W}, \mu_t, \sigma^t) \geq U_t(\mathcal{W}'', \mu_t, \sigma^t) = U_t(\mathcal{W}', \mu_t, \sigma^t)$$

proving that \mathcal{W} is robust to $\mathbf{a}'(\alpha \setminus \sigma^t)$, too.

Now, consider the case where $a'(\sigma^{t+1}) = 0$ and the strategies $\hat{\mathbf{a}}$ and $\tilde{\mathbf{a}}$ defined as above, and also $\ddot{\mathbf{a}}(\alpha \setminus \sigma^t) = (0, 0, e\mathbf{1}_{k-1})$. I first show that

$$U_t(\mathcal{W}, \mu_t, \sigma^t) \geq U_t(\ddot{\mathcal{W}}, \mu_t, \sigma^t)$$

under the assumption of (IC, t) being binding

$$U_t(\mathcal{W}, \mu_t, \sigma^t) = U_t(\hat{\mathcal{W}}, \mu_t, \sigma^t) = u(c_t(\sigma^t)) + \beta U_{t+1}(\hat{\mathcal{W}}, \mu_t, \sigma^{t+1})$$

which boils down to prove that

$$\begin{aligned} U_{t+1}(\hat{\mathcal{W}}, \mu_t, \sigma^{t+1}) &\geq U_{t+1}(\ddot{\mathcal{W}}, \mu_t, \sigma^{t+1}) = u(c_{t+1}(\sigma^{t+1})) + \beta U_{t+2}(\ddot{\mathcal{W}}, \mu_t, \sigma^{t+2}) \\ \implies U_{t+1}(\mathcal{W}, \mu_t, \sigma^{t+1}) &\geq u(c_{t+1}(\sigma^{t+1})) + \beta U_{t+2}(\mathcal{W}, \mu_t, \sigma^{t+2}) \end{aligned} \quad (39)$$

where the first inequality follows from the fact that both strategies prescribe no effort in t , and the second inequality follows from the fact that $\hat{\mathcal{W}} = \mathcal{W}$ (resp., $\ddot{\mathcal{W}} = \mathcal{W}$) over $\alpha \setminus t + 1$ (resp., $\alpha \setminus t + 2$). By assumption, (IC, $t + 1$) is binding

$$U_{t+1}(\mathcal{W}, \mu_t^u, \sigma^{t+1}) = u(c_{t+1}) + \beta U_{t+2}(\mathcal{W}, \mu_t^u, \sigma^{t+2})$$

which, jointly with (34) and since $\mu_t^u < \mu_t$, causes (39). Now, \mathbf{a}' and $\ddot{\mathbf{a}}$ prescribe the same action

in periods t and $t+1$. Therefore, in order to prove that \mathcal{W} is robust against $\mathbf{a}'(\alpha \setminus \sigma^t)$, it is enough to show that

$$U_{t+2}(\mathcal{W}, \mu_t, \sigma^{t+2}) = U_{t+2}(\check{\mathcal{W}}, \mu_t, \sigma^{t+2}) \geq U_{t+2}(\mathcal{W}', \mu_t, \sigma^{t+2}) \quad (40)$$

Now, there are two possibilities, $a'_{t+2}(\sigma^{t+2})$ can either be 0 or e . If the first case occurs, in order to prove (40) it is enough to show

$$U_{t+3}(\mathcal{W}, \mu_t, \sigma^{t+3}) \geq U_{t+3}(\mathcal{W}', \mu_t, \sigma^{t+3}) \quad (41)$$

Indeed, (IC, $t+2$) binding and (34) jointly cause

$$U_{t+2}(\mathcal{W}, \mu_t, \sigma^{t+2}) \geq u(c_{t+2}(\sigma^{t+2})) + \beta U_{t+3}(\mathcal{W}, \mu_t, \sigma^{t+3})$$

On the other hand, if $a'_{t+2}(\sigma^{t+2}) = e$, then

$$U_{t+2}(\check{\mathcal{W}}, \mu_t, \sigma^{t+2}) = u(c_{t+2}(\sigma^{t+2})) - e + \beta [\pi(\mu_t)U_{t+3}(\mathcal{W}, (\sigma^{t+2}, w)) + (1 - \pi(\mu_t))U_{t+3}(\check{\mathcal{W}}, \mu_t^u, \sigma^{t+3})],$$

$$\check{\mathcal{W}} = \{\mathcal{W}, \mathcal{W}'\}$$

But then proving (40) boils down to show (41). I have just established the following implication

$$U_{j+1}(\mathcal{W}, \mu'_{j+1}, \sigma^{j+1}) \geq U_{j+1}(\mathcal{W}', \mu'_{j+1}, \sigma^{j+1}) \implies U_j(\mathcal{W}, \mu'_j, \sigma^j) \geq U_j(\mathcal{W}', \mu'_j, \sigma^j), \quad \forall j : t \leq j \leq T$$

where μ'_j is the expectation in period j if strategy \mathbf{a}' is applied. But then the proof is complete, as

$$U_{T+1}(\mathcal{W}, \mu'_{T+1}, \sigma^{T+1}) = U_{T+1}(\mathcal{W}, \sigma^{T+1}) = U_{T+1}(\mathcal{W}', \mu'_{T+1}, \sigma^{T+1}) \implies U_t(\mathcal{W}, \mu'_t, \sigma^t) \geq U_t(\mathcal{W}', \mu'_t, \sigma^t)$$

■

Lemma 5 proves to be useful in light of the following result.

Lemma 6. *In optimum, all (IC, s) $^T_{s=0}$ constraints are binding.*

Proof. By contradiction, assume that $\mathcal{W} = (\mathbf{c}, \mathbf{a})(\alpha \setminus \sigma^0)$ is optimum and that (IC, t) is slack

$$U_t(\mathcal{W}, \mu_t, \sigma^t) > u(c_t(\sigma^t)) + \beta U_{t+1}(\mathcal{W}, \mu_t, (\sigma^t, u))$$

Then there exists $\varepsilon > 0$ such that

$$\begin{cases} c'_{t+1}(\sigma^t, w) = c_{t+1}(\sigma^t, w) - \varepsilon \\ U_t(\mathcal{W}', \mu_t, \sigma^t) = u(c_t(\sigma^t)) + \beta U_{t+1}(\mathcal{W}', \mu_t, \sigma^{t+1}) \end{cases}$$

where $\mathcal{W}' = (\mathbf{c}', \mathbf{a})(\alpha \setminus \sigma^0)$ is defined as

$$c'_s(\sigma^s) = c_s(\sigma^s), \quad \forall \sigma^s \neq (\sigma^t, w), \quad c'_{t+1}(\sigma^t, w) = c_{t+1}(\sigma^t, w) - \varepsilon$$

Now, government's payoff is larger under \mathcal{W}' than under \mathcal{W} , as payment to the worker in history (σ^t, w) is lower in the former case. Moreover, by Lemma 5, \mathcal{W}' is also feasible, since it satisfies all $(\text{IC}, s)_{s=0}^T$ constraints with equality. But this contradicts that \mathcal{W} is optimum. ■

Thus, robustness against all one-shot deviations from the prescribed effort sequence constitutes a necessary condition for a contract to be optimum (by Lemma 6) and sufficient one for it to be robust against any multiple deviation (by Lemma 5). Therefore, focusing on the set of contracts with such characteristic is without loss of generality.