# Optimal Unemployment Insurance with Worker Profiling*

[Click here for the latest version]

Sergio Cappellini†

**ABSTRACT**

Efficient unemployment assistance is tailored to workers' human capital. Since human capital is difficult to infer, assistance is provided on the basis of its expected level. Alternatively, workers can be profiled and their actual level of human capital be detected. A profiling program establishes (i) whom to profile, (ii) at what stage of the program and (iii) what benefits to transfer to them, depending on the new information obtained. The paper identifies the determinants of optimal profiling along these three dimensions in a dynamic principal-agent framework with non-contractible effort and two-sided uncertainty about workers' human capital. There are two main findings. First, workers with higher expectations on human capital are incentivized to search for a job, thanks to larger returns on search effort. They are profiled only at a successive stage of the unemployment spell, once the gains from optimal matching between policies and workers outweigh the cost of profiling. Second, since the incentive cost is increasing in the generosity of benefits promised to workers, profiling is used also to lower promised benefits for those workers who are incentivized to search after it.

*JEL classification*: D82, D83, J65, J68

*Keywords*: Bayesian Persuasion, Employability, 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 and participants at the 19th Annual EEFS Conference and at the 16th Annual Symposium on Economic Theory, Policy and Applications (ATINER).

†PhD student, Bocconi University, e-mail: sergio.cappellini@phd.unibocconi.it

# 1 Introduction

A renovated interest in optimal design of active labor-market policies (ALMPs) started in 2007, with the financial crisis. And 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, and results in a push for public spending optimization. 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. Claimants' profiling is present in the welfare programs of most OECD countries[1] 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) are the two Federal-funded programs that profile welfare claimants. Yet, the two programs have a different target. WPRS profiles claimants with the aim of "identifying and ranking or scoring unemployment insurance claimants [...] for referral to appropriate reemployment services" (US Dept. of Labor, 2007), supplied by State Welfare Agencies (SWAs). On the other hand, the purpose of REA is to "[...] identify existing and eliminate potential overpayments, and realize cost savings for UI trust funds" (Poe-Yamagata, 2011). Therefore, US profiling programs generate savings for the provider in two distinct ways. First, by reallocating workers to different policy instruments which better fit their characteristics. Second, by designing transfers based on recipients' characteristics and needs during the unemployment spell. 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. These traits contribute to the formation of reemployment expectations, based on the statistical evidence provided by historical data on claimants and their unemployment history. Thus, the observable individual characteristics lead to an *early assessment* of claimants, and constitute the rationale for the match of each claimant with their

[1] Some 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).

personalized assistance program. In addition, WPRS and REA contemplate the possibility of implementing a further *in-depth assessment* of claimants, in the form of one-on-one interviews and/or skill tests, to detect the actual level of human capital of every single worker with a given level of accuracy.

In-depth assessment generates an additional cost for the provider, which narrows its scope of adoption only to a fraction of claimants. To select participants, WPRS and REA adopt distinct criteria. Indeed, enrollment into WPRS is based on the expectations formed in the early assessment case, while REA randomly chooses the welfare recipients to profile. Such difference mirrors the distinct target of the two programs, that is, referral of poorly employable workers to job-search assistance under WPRS, and exclusion of highly employable ones from too generous incentives under REA. However, the two programs have a common spirit, that is, granting to lowly employable workers a better support, in terms of more effective means for the job search and more generous transfers. In particular, REA pursues this objective by inflicting a 'punishment' to workers who enjoy a higher level of human capital and thus have a stronger appeal on potential employers, while WPRS rewards low-skilled workers by enhancing the level of assistance. Either program thus activates different leverages to achieve cost efficiency.

WPRS and REA only deal with information detection via profiling of workers. The two programs are thus used as ancillary tools to welfare programs, which instead support income *stricto sensu*. Unemployment benefits are paid under four distinct programs, which are activated in succession, depending on the current labor market situation of each US State. Unemployment Insurance benefits last up to 26 weeks in all States. If the worker is still unemployed at the end of the 26th week, he/she is entitled to additional 53 weeks under the Emergency Unemployment Compensation (EUC) program. Moreover, States with unemployment rates exceeding 8.5% pay claimants up to additional 20 weeks of benefits under the Extended Benefits (EB) program. Exhaustees of UI, EUC and EB are finally referred to the Supplemental Nutrition Assistance Program (SNAP), which consists of an constant allowance for purchasing food and constitutes the typical example of a purely income-support policy, with no access requirement, time limit or eligibility assessment. Transfers decline over time, as claimants move from one program to another. WPRS and REA only assess workers' skills during one among UI, EUC and EB, and require profilees to continue their search in the meantime. Any form of inactivity of the worker is not allowed and in fact constitutes a reason for exclusion from any benefit other than SNAP subsidy.

Welfare assistance is funded partly by the Federal government and partly by each State, while the organization and design is mainly deferred to the latter. Profiling programs thus greatly differ along many dimensions, namely (i) when and (ii) how accurately to profile (iii) whom, (iv) whether to request the profilee to search or rest in the meantime, or rather (v) to assist her in the search, and finally (vi) whether the job search should be conducted after profiling, based on the new information obtained, and (vii) what transfer scheme should accompany it.

The main objective of this paper is to develop a suitable framework to analyze the main complementarities between profiling and income-support policies. Optimal welfare provision arises as a solution to 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 stream 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 *per se* informative about hidden reemployment perspectives, and cause a revision of expectations. The paper also extends the analysis to the case where worker's effort is unobservable. Assuming worker's job search to be unobservable by the government would lead to possible mismatches in expectations between it and the worker, if the latter deviates from recommended search effort. Consequently, unemployment benefits would include learning rents and be indexed to the prospective duration of worker's search.

Likewise 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, which can be implemented jointly. The optimal program arises as a sequence of policies over time.

The paper delivers a number of results. First, the cost of search incentives and effort compensation is decreasing in the expected level of employability and increasing in the level of program's generosity toward recipients. This makes the government delegate the job-search to workers when expected chances of re-employment are higher and program's generosity is lower. In particular, workers' search is replaced by assisted search in more generous programs. And, if the current level of promised utility to the worker is allowed to decline over time, a recipient who is initially assisted in the search is possibly referred to two policy sequences, depending on the level of promised generosity. For high-end generosities, the worker does not experience any decrease in promised utility, and her consumption is insured against any risk of search failure. On the contrary, for upper intermediate generosities, continuation utility declines over time and the worker is entrusted with the search at some point of the unemployment spell, where her consumption declines after any failed search. This finding suggests that WPRS-like programs that use profiling as a way to allocate extra services, like assisted job-search, only to needy workers are also more generous toward recipients at the optimum. REA-like programs, instead, also contemplate forms of 'punishment' as a way to ease incentive provision for the welfare provider and are better implemented when the implicit generosity of the program is lower. This second type of programs is characterized by an extensive appeal to incentive provision via transfers, which decline along the unemployment spell and are larger in case of successful job search.

The second finding is that at the optimum profiling of individual characteristics may not be fully accurate. The aim of it is to boost confidence also to a fraction of lowly-employable workers and delegating the job search to them, as well as to the highly-employable ones. This result is due to incentive costs being convex decreasing in expectations, whenever the search cost enters linearly in the utility function of workers, which in turn delivers a concave return of worker's search and a negative value of information when such concavity is prominent. For profiling to improve on government's return, it must be that the savings from reallocation of highly-employable workers to an active policy with a lower promised transfers ahead outweigh the extra cost of both delivering larger welfare support to poorly-employable workers, and adopting the technology. Therefore, likewise assisted search, no worker is ever profiled when her expectations are extremely high or low, that is, when profiling costs outweigh prospective gains. In addition, high employability is never fully detected in profiled workers when concavity of returns from worker's search is prominent, as the government delegates job search also to a fraction of lowly-employable workers, rather than minimize the size of transfers by delegating it only to the highly-employable ones. The paper also states sufficient conditions about the curvature of the first-order derivative of worker's inverse utility that make profiling accuracy positively related to program's generosity.

The third finding is that the optimal program, as well as the actual REA program, features a more prominent decline in transfers for any job-seeking worker who is profiled as highly employable and keeps searching after it. The reason is that incentive costs to job-search being increasing in worker's promised utility advocate for making incentive-compatible transfers less expensive by promising a lower utility to workers who are requested to search upon profiling.

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 and solves for their optimal design, while Section 5 presents the main features of the welfare program. Section 6 conducts a quantitative analysis on REA program in the US and estimates the welfare gain from switching to the optimal benchmark. Section 7 concludes.

# 2 Literature Review

The main contribution of this paper is the development of an amenable framework to study workers' profiling policy within a welfare program toward the jobless. The paper provides an encompassing analysis of the gains and costs of profiling, when implemented jointly with others labor-market policies, in a context where workers' employability is not *ex-ante* observable and agents learn about it either via job-search failures or via profiling.

The existence of an agency problem in the contractual relationship between the welfare provider and the recipients, has long been acknowledged by the literature (Shavell and Weiss, 1979). The provider has the possibility to tackle it either by providing recipients with incentives, or by conducting the search on their behalf. In both cases, the job search produces an extra cost, which sometimes 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 job opportunities are allowed to deteriorate during the unemployment spell, workers who have first been assigned to active policies are later reassigned to passive ones (Pavoni, 2009), as well as those who have first been entrusted with the job search, are monitored or replaced in the search by the government in a subsequent stage of the program (Pavoni and Violante, 2013). Analogously, in this paper any transition to a different policy follows the deterioration in expected employability upon failed job search, independently of who conducts it. Yet, such a deterioration stems from a learning process and does not involve any depreciation of physical human capital. Differently from physical depreciation, expectation revision can be the object of manipulations by the government via an *ad-hoc* profiling strategy.

Profiling is one of the two channels for agents of acquiring information, the other one being job search failures. The difference between the two is that, while learning from failures is intrinsic to the process of job search, profiling instead is under the control of the government, who plays also the role of information designer and sender. A vast and rapidly growing strand of the literature deals with the design of an optimal signaling strategy from a principal/sender to an agent/receiver. The peculiarity of the present framework is the 'hybrid' nature of the problem, which mixes the design of information with that of an effort-incentivizing contract scheme. Consequently, the standard result typical of Bayesian persuasion (as first outlined by the seminal contribution by Kamenica and Gentzkow (2011)), about the optimal signaling delivering the concave closure of the pre-signal envelope function, changes in two directions. First, the cost for implementing profiling lowers its return for every level of expectation and program's generosity. In particular, the cost outweighs the gain from allocative efficiency and incentive cost minimization when expectations are very low or very high. The result differs from Kamenica and Gentzkow (2014), which provides a sufficient condition on the cost of signals for the concavification result to hold even in contexts of costly persuasion. Second, information acquisition can be used to relax the principal's incentive constraint by making agent's utility dependent on the profiling outcome. Indeed, whenever the incentive cost increases in the agent's continuation utility, the principal finds it optimal to lower the utility of the agent who is found highly re-employable and is therefore recommended to search. To the best of my knowledge and notwithstanding the existence of some works embedding incentive provision and Bayesian signaling about hidden characteristics of the agent/receiver into the same framework (for ex., Rodina, 2017; Boleslavsky and Kim, 2018), this paper constitutes the first attempt to embed the two aspects in a dynamic framework.

## 3 Economic Environment

**Players' Interaction.** A risk-neutral government (principal, it) and a risk-averse worker (agent, she) populate the economic environment in discrete time. Each player is infinitely-lived and discount future utility at rate $\beta \in (0, 1)$. The worker can be employed or not, and the government observes her employment status. In period 0, (i) the two players are uncertain about the worker's human capital and hold common expectations about it, (ii) the worker is unemployed, and (iii) the government offers her a contract contingent on any possible future employment status 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 a given expected discounted 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 status 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 $\pi_H$, 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 $v(c, a) = u(c) - e \cdot a$.

**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 is requested by the government to report personal information (social background, past working experiences, education, etc.). According to this initial information, the government makes a first assessment of her level of human capital based on statistical data. Highly-educated and more experienced workers, for instance, are statistically more likely to exit unemployment than workers with less experience and/or lower educational attainment. The assessment attaches to the worker a probability $\mu$

of being high skilled ($\mu = Prob(h = H)$), which is henceforth referred to as *expectation*.[2]

**Assisted-search technology.** Conditional on payment of per-capita cost $\kappa^{ja}$, the government can search on behalf of the worker. 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.** Conditional on payment of cost[3] $\kappa^{wp}$, profiling detects human capital with some accuracy, and returns a publicly observable outcome. Profiling can be thought of as a lottery that returns a binary outcome -'Pass' ($r = p$) or 'Fail' ($r = f$)-, with predetermined odds. 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 $\mu$ and program's generosity $U$

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

# 4 Policies

Any policy arises as the composition of (i) recommended search effort, (ii) consumption contract, and (iii) technology/ies implemented (if any). Combinations of search effort levels and technologies gives 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.

[2] By 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.

[3] The cost includes administrative expenses, as in the case of assisted search.

| 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 $\mu$ and the program's generosity $U$ jointly determine the choice of the policy instrument. The government chooses the optimal policy $i(\mu, U)$ by solving

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

In the following, I introduce the problem of the welfare provider in case of re-employment and for all six instruments during unemployment. First, I define welfare-oriented policies (SA, JS and UI) and later the profiling ones (AP, SP and IP). The definition of each policy is subject to a No-Stick constraint (hereafter, (NS)) on the planner, that prevents agent's continuation utility contingent on any possible state realization from falling below current utility ($U_i^{s,r} \geq U$). This way, the welfare program is designed so not to create punishments for workers who remain unemployed, via a progressive reduction of welfare benefits over time. Such a policy design has been labeled 'soft' by Pavoni et al. (2016), who first highlighted that 'soft' programs are robust to hidden saving.

## 4.1 Welfare Policies

**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

$$
\begin{aligned}
W(h, U) = \max_{\tau, U^w} \tau + \beta W(h, U^w) &= \max_{c^w, U^w} \omega_h - c^w + \beta W(h, U^w) \\
\text{sub:} \quad U &= u(c^w) + \beta U^w && \text{(PK)} \\
U^w &\geq U && \text{(NS)}
\end{aligned}
$$

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 $\tau$ is the wedge between gross ($\omega_h$) 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 ($U^w = U$).[4] Hence from (PK) one can obtain the closed-form expression for consumption $c^w = u^{-1}\big((1-\beta)U\big)$, which compensates for work effort. The expression for labor tax/subsidy thus is

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

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

$$
\begin{aligned}
V^{SA}(\mu, U) &= \max_{c^{sa}, U^u} -c^{sa} + \beta V(\mu, U^u) \\
\text{sub:} \quad U &= u(c^{sa}) + \beta U^u && \text{(PK)} \\
U^u &\geq U && \text{(NS)}
\end{aligned}
$$

The planner transfers $c^{sa}$ and pledges continuation utility $U^u$, without requiring the worker

[4] (NS) is superfluous in this case, as the solution to the problem without (NS) is given by

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

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.** *Social Assistance is an absorbing policy and its continuation utility equals current utility ($U^u = U$).*

*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 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 politically hard to 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 to workers with low expected employability. Current consumption solves (PK) with $U^u = U$.[5] The value of SA is independent of $\mu$ and has a closed-form expression

$$V^{SA}(U) = -\frac{u^{-1}\big((1-\beta)U\big)}{1-\beta} \tag{2}$$

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 their initial expectation $\mu$, according to the formula

$$\mu' := \frac{\mu\big(1-\pi_H\big)}{\mu\big(1-\pi_H\big) + (1-\mu)\big(1-\pi_L\big)} \leq \mu \tag{3}$$

where $\mu'$ is the revised probability that worker's human capital is $h = H$. $\mu'$ is lower than the initial one, with equality holding only if human capital was known already ($\mu \in \{0, 1\}$).

[5] (NS) is superfluous also in this case, as

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

The reason lies in the unbiasedness of $\mu$, that is equal to the actual share of high-skilled workers among those who hold that expectation. Thus, a fraction $\pi(\mu) := \mu\pi_H + (1-\mu)\pi_L$ of them manages to find a new employment, which implies that the high-skilled who remained unemployed after one 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 realization $h = L$.[6]

**Job-Search assistance (JS).** When resorting to assisted search, the government looks for employment on worker's behalf, an activity that costs him $\kappa^{ja}$. The value of JS reads

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

$$\text{sub:} \quad U = u(c^{js}) + \beta\big[\mu\pi_H U^w_H + (1-\mu)\pi_L U^w_L + (1-\pi(\mu))U^u\big] \qquad \text{(PK)}$$

$$U^w, U^u \geq U \quad \text{(NS)}$$

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 connected to human capital realization, conditional on finding a new job for the worker. Given that the worker is risk-averse, under JS the government finds optimal to insure her against both ($U^w_H = U^w_L = U^u$).[7] The value of JS therefore reads

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

with

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

[6] If the failed attempts to exit unemployment are $t$, one for each period, then initial expectation $\mu$ is updated $t$ times according to the formula

$$\mu^{(t)} = \mu^{(t-1)\prime} = \frac{\mu\big(1-\pi_H\big)^t}{\mu\big(1-\pi_H\big)^t + (1-\mu)\big(1-\pi_L\big)^t} \tag{4}$$

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;
- $\lim_{t\to\infty} \mu^{(t)} = 0$, if $\mu^{(0)} < 1$.

[7] The proof is reported in Appendix A: Properties of SA, JS and UI.

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.

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

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

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

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

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

The dispersion in utilities causes a cost of incentives, defined by the difference in cost between the cases of non-contractible and contractible effort.[8] Incentive costs are increasing in the cost of effort and decreasing in the level of patience ($\beta$) and confidence ($\mu$). Indeed, 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.

The problem of the planner reads

$$V^{UI}(\mu, U) = \max_{c^{ui}, U^{w}, U^{u}} -c^{ui} + \beta\big[\pi(\mu)W(\mu, U^{w}) + (1-\pi(\mu))V(\mu', U^{u})\big]$$
$$\text{sub:} \quad \text{(PK) - (IC)}, \qquad U^{w}, U^{u} \geq U \quad \text{(NS)}$$

[8] On the same lines of Pavoni and Violante (2007), one can imagine the existence of a policy which delegates job search to workers whenever effort is contractible. If that is the case, the government will only need to compensate for the worker's effort with the contract

$$c^{jm} = u^{-1}((1-\beta)U + e), \; U^{w} = U^{u} = U$$

If so, the incentive cost is defined as the difference in cost of contract between UI and this new policy.

The No-Stick constraint binds at the optimum in case of search failure, which delivers the following result.

**Proposition 2 (Optimal UI Contract).** *At the optimum, the consumption contract in Unemployment Insurance reads*[9]

$$c^{ui} = u^{-1}((1-\beta)U),\ U^w = U + \frac{e}{\beta\pi(\mu)},\ U^u = U \tag{5}$$

*Proof.* See Appendix A: Properties of SA, JS and UI. ∎

The incentive cost depending negatively on expectations through the utility dispersion generates a comparative advantage of UI for high-end expectations.

**Lemma 1 (Slopes of the value functions with respect to $\mu$ and $U$).** *$V$ is concave increasing in $\mu$ (possibly in the weak sense), and concave decreasing in $U$. Moreover*

$$0 = V_{\mu}^{SA}(U) < V_{t,\mu}^{JS}(\mu, U) < V_{t+1,\mu}^{JS}(\mu, U) < V_{\mu}^{UI}(\mu, U), \quad \forall t \geq 1 \tag{6}$$

*If $1/u'$ is convex,*

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

*and*

$$0 = W_{\mu U}(\mu, U) = V_{\mu U}^{JS}(\mu, U) < V_{\mu U}^{UI}(\mu, U) \tag{8}$$

*Proof.* See Appendix A: Properties of SA, JS and UI. ∎

The marginal value of $\mu$ is increasing in the level of search intensity, duration and effort by the worker. Therefore, fixing generosity and spanning the space of expectations, one observes that policies with higher (resp., lower) marginal returns are optimal for higher (reps., lower) expectations. Hence, SA is only implemented as a policy of last resort, when the worker has little expected chances of exiting unemployment and no search is worth conducting. On the

[9] Differently from policies like SA and W, (NS) constraint plays a role in the solution. Indeed, its absence would cause the optimal contract to be a mix of 'reward' in case of re-employment, and 'punishment' in case of search failure (see Prop. 5).

other end of the spectrum, UI is optimal when the return from search is large and incentive costs are low. JS, instead, displays increasing return in expectations, but constant search cost and is therefore optimal for intermediate $\mu$'s. Furthermore, the duration of any active policy is increasing in expectations. Any worker who enters with a high expectation moves from UI to JS and lastly to SA, unless she is reemployed beforehand. The upper envelope $V$ thus displays a tendency toward between-policy convexity, as the marginal value of expectations is larger in policies that are optimal in the high end of the space of expectations. Such shape of $V$ has deep implications on the choice of optimal profiling.

The second part of Lemma 1 provides a sufficient condition, namely convexity of $1/u'$, for costs of incentive-provision and effort-compensation to be increasing in $U$. Convexity of $1/u'$ can be interpreted as the condition guaranteeing that workers' dislike for risky consumption lotteries (present in UI and absent in SA and JS) is increasing in the expected utility of the lottery.[10] As anticipated by Pavoni and Violante (2007), more generous programs are mainly focused on assistance provision, which in the framework of this paper occurs in the form of income support and assisted search.

[10] It can also be shown that incentive costs are convex in expectations, due to $U^w - U^u$ being a hyperbolic function of $\mu$, and that their sensitivity to an increase in $U$ is decreasing in $\mu$. Concavity of $V^{UI}$ in $\mu$ follows from convexity of incentive costs and the linearity of returns. For a more detailed explanation, see Pavoni and Violante (2007).

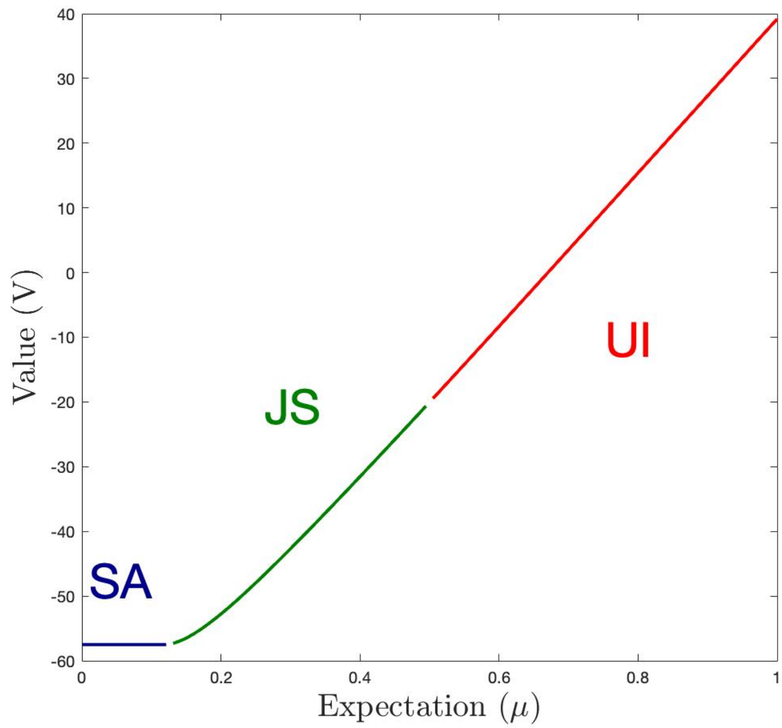


Figure 1: Value of the welfare program with no profiling, for consumption-equivalent generosity of $c = \exp((1-\beta)U) = \$575$.[11]

Fig. 1 shows the value of SA, JS and UI in the space of expectations (on the $x$-axis) and for constant generosity level. Any worker who enters the program with a certain expectation is referred to the policy which is most valuable to the government for that level of expectations. For example, expected employability of claimants with a college degree is different from the one of claimants with a high school diploma. For this reason, the government finds it optimal to initially refer these two groups of claimants to different policy instruments. To interpret the picture, assume that a worker with expectation $\mu$ is first referred to UI for $t$ periods. If she does not manage to find a job in the next period, she revises her initial expectation to $\mu'$ and keeps searching with higher incentives, which is equivalent to move from $\mu$ to $\mu'$ on the $x$-axis. In case she failed to find a job after $t$ periods, she then enters JS with a $t$-fold revised expectation $\mu^{(t)}$. After another series of unsuccessful attempts, she is eventually referred to SA, and once there, she stops searching (and revising expectation, as well).

[11] The parameter values and functional forms used in this Section are: $u(.) = \log(.)$, $\beta = 0.9$, $e = 0.53$, $\kappa^{ja} = 6$, $\kappa^{wp} = 1.5$, $\omega_H = 20$, $\omega_L = 5$, $\pi_H = 0.27$, $\pi_L = 0.14$. All monetary values are divided by 100.

## 4.2 Profiling Policies

Profiling publicly discloses worker's human capital, up to a level of accuracy chosen by the government. For this reason, profiling can be read as the worker undergoing a test with two possible outcomes, 'Pass' ($r = p$) or 'Fail' ($r = f$). Conditional on the outcome of the test, the worker is referred to a different policy. The probability schedule brings both the worker (i.e., the profilee) and the government (i.e., the profiler) to revise expectation $\mu$ upon observation of the public outcome $r$ according to the formula

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

A necessary and sufficient condition for profiling to induce a change in expectations is to avoid returning either outcome with the same probability, irrespective of 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).

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

(MP) can be interpreted as a restriction requiring profiling to be credible. Indeed, considering all workers who share the same expectation $\mu$, inducing any of them to revise their expectation up to $\mu_p$ comes at the cost of inducing an expectation revision down to $\mu_f$ for someone else.[12]

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

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

The government finds it optimal to insure the worker against the risk connected to the profiling outcome, by pledging constant continuation utility under both cases ($U_p^u = U_f^u$, see

[12] Without loss of generality, the posterior upon 'Fail' ($\mu_f$) is set to be lower than the posterior upon 'Pass' ($\mu_p$).

Prop. 2). As far as the profiling strategy is concerned, monotonicity of $V$ in $\mu$ makes the return for the government be larger upon 'Pass', and profiling signal low human capital with full accuracy (i.e., $\mu_f = 0$). Intuitively, reducing the likelihood of 'Fail' outcome increases the frequency of 'Pass' ($q$) more than one to one. This fact, joint with the linearity of $V$ in $\mu$ in low-end expectations (see Fig. 1), makes convenient to limit 'Fail' only to low-skilled people and induce as many workers as possible to upward revise expectations. By a similar reasoning, one may be tempted to guess that, in order to achieve maximization of returns, 'Pass' outcome only targets high-skilled workers (e.g. $\mu_p = 1$) at the optimum. However, this is not always the case, due to concavity of $V$ in $\mu$ for high-end expectations. Indeed, while concave returns cause the marginal gain of 'Pass' informativeness about high human capital to decline in the level of informativeness itself, reducing the frequency of 'Pass' and 'Fail'-ing more workers cause a loss at the margin. Therefore, the planner trades off informativeness of 'Pass' against its frequency up to the point where the gain of higher informativeness equals the cost of lower frequency. In case the marginal gain exceeds the marginal cost for every $\mu$, the test fully discloses high human capital. Otherwise, the internal solution satisfies

$$\frac{V(\mu_p, U) - V(0, U)}{\mu_p} = V_\mu(\mu_p, U) \tag{9}$$

Eq. 9 shows that the upper posterior 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 $\mu_p$, irrespective of the administrative cost of profiling, as the low-skilled ones among them are mistaken in a direction favorable to the government. Hence, disclosing any information about their actual human capital causes it a loss that outweighs the gain of informing high-skilled workers.

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

$$V^{SP}(\mu,U)=\max_{c^{sp},(U^w_r,U^u_r,\mu_r)_{r=\{p,f\}}} -c^{sp}-\kappa^{wp}-\kappa^{ja}+\beta\big[q\big(\pi(\mu_p)W(\mu_p,U^w_p)+(1-\pi(\mu_p))V(\mu'_p,U^u_p)\big)+$$
$$+(1-q)\big(\pi(\mu_f)W(\mu_f,U^w_f)+(1-\pi(\mu_f))V(\mu'_f,U^u_f)\big)\big]$$
$$\text{sub:}\quad U=u(c^{sp})+\beta\big[q\big(\pi(\mu_p)U^w_p+(1-\pi(\mu_p))U^u_p\big)+(1-q)\big(\pi(\mu_f)U^w_f+(1-\pi(\mu_f))U^u_f\big)\big]\quad \text{(PK)}$$
$$U^w_r,U^u_r\geq U,\ \ \forall r\in\{p,f\}\ \ \text{(NS)},\quad \text{(MP)}$$

The planner prefers to insure the worker against the risks related to job search and profiling outcomes, as well as human capital realization, by committing to a constant continuation utility (see Prop. 2). About the informativeness of the profiling strategy, the posterior expectation $\mu_p$ induced by 'Pass' outcome, is either 1 or solves

$$\frac{V(\mu'_p,U)-V(0,U)}{\mu'_p}=V_\mu(\mu'_p,U) \tag{10}$$

Indeed, if in case of AP the randomization in the space of expectations occurs over the upper envelope $V$, now instead the randomization only occurs conditional on job-search failure. Therefore, optimal profiling in SP (net of cost $\kappa^{wp}$) delivers the concave closure of $(1-\pi(\mu))V(\mu',U)$ in the space of expectations $\mu\in[0,1]$.

**Insurance-and-Profiling (IP).** When profiling is implemented jointly with delegated search, the planner's problem reads

$$V^{IP}(\mu,U)=\max_{c^{ip},(U^w_r,U^u_r,\mu_r)_{r=\{p,f\}}} -c^{ip}-\kappa^{wp}+\beta\big[q\big(\pi(\mu_p)W(\mu_p,U^w_p)+(1-\pi(\mu_p))V(\mu'_p,U^u_p)\big)+$$
$$+(1-q)\big(\pi(\mu_f)W(\mu_f,U^w_f)+(1-\pi(\mu_f))V(\mu'_f,U^u_f)\big)\big]$$
$$\text{sub:}\quad U=u(c^{ip})-e+\beta\big[q\big(\pi(\mu_p)U^w_p+(1-\pi(\mu_p))U^u_p\big)+(1-q)\big(\pi(\mu_f)U^w_f+(1-\pi(\mu_f))U^u_f\big)\big]\quad \text{(PK)}$$
$$U\geq u(c^{ip})+\beta\big[qU^u_p+(1-q)U^u_f\big]\quad \text{(IC)},\quad U^w_r,U^u_r\geq U,\ \forall r\in\{p,f\}\ \ \text{(NS)},\quad \text{(MP)}$$

Similarly to SP, profiling delivers the concave closure of $(1-\pi(\mu))V(\mu',U)$ by selecting a posterior upon 'Pass' which equal 1 or solves (10).

## 4.3 Optimal Welfare Program

Under each profiling policy, there exists the possibly of referring also a fraction of low-skilled workers to active policies upon 'Pass'. If so, any worker who receives a 'Pass' and is referred to any 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 and worker's human capital is detected, she does not revise her expectation henceforth. In other words, the policy she is assigned to under either profiling outcome is absorbing, as worker's expectation and promised utility remain the same upon job search failure. The following result about optimal profiling policies in the space of expectations holds.

**Proposition 3.** *Fix generosity. No profiling policy is optimal for very high or very low expectations. Assistance-and-Profiling (AP) is the preferred profiling policy over low-end expectations, Search-assistance-and-Profiling (SP) over intermediate, and Insurance-and-Profiling (IP) over high-end ones.*

*Proof.* See Appendix B: Properties of AP, SP and IP. ■

The first part of the proposition can be explained through gains and losses of profiling. Profiling generates savings for the government by delegating search to high-skilled workers with a lower cost of incentives. The losses are of two types. First, the government incurs administrative expenses. Second, it suffers a loss by passing any information to low-skilled workers who are overconfident about their human capital. Therefore, for very high and very low expectations, workers are on average efficiently matched with policies, and the gains from reallocation and/or transfer reduction are outweighed by the losses. The second part of Prop. 3 outlines the existence of 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 dominant welfare policy is the one implementing (or not) the job search with the same method.

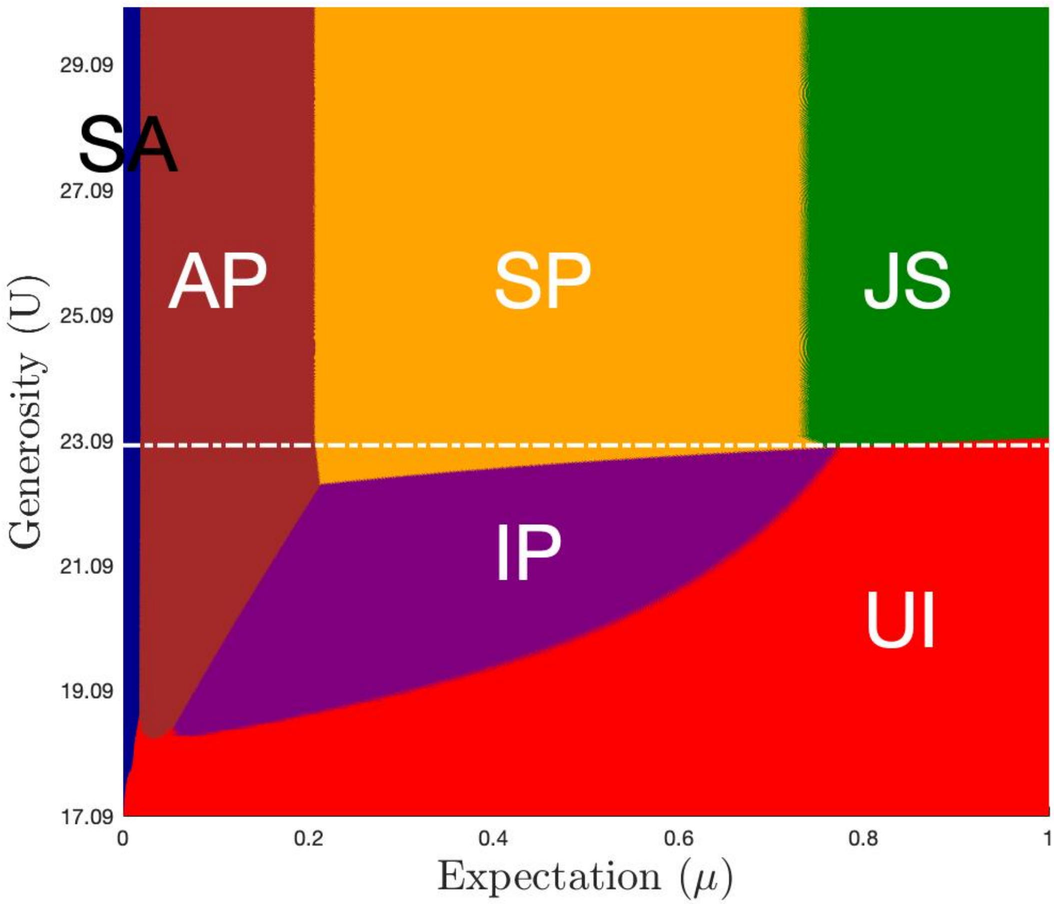


Figure 2: Optimal Policies in the Space of Expectation and Generosity.

Fig. 2 displays the optimal policies in the space of expectations and generosity levels. Moving horizontally from right to left over $x$-axis and holding constant $y$-intercept, one observes the policy sequence of a worker whose expectations are revised downward after failed job-finding attempts. If the worker is profiled, she either discovers to be low-skilled and enters SA ever after, or upward revises her expectations and is referred to an active welfare policy.

The complementarity of search effort and expectations is mirrored also in the best profiling policy adopted. In particular, SA and AP, none of which contemplates any form of search, are optimal for lower-end expectations. JS and SP, which implement assisted search, are optimal for intermediate expectations. And UI and IP, which delegate search to the worker, are optimal for higher-end expectations. As generosity rises (moving vertically from bottom to top of Fig. 2), the return of job search decreases due to higher costs of search-effort compensation and incentive provision (under assumption that $1/u'$ is convex[13]), and worker's search (UI and IP) is replaced by assisted search (JS and SP) for high-end expectations, or by no search (SA and AP) for low-end expectations.

Fig. 2 also proves delegated and assisted search to be substitutes, when worker's utility does not decline over time, as one replaces the other at different program's generosities. This finding mirrors the one in Pavoni et al. (2016), where UI and JS are found to never coexist within the same 'soft' program. In addition, in 'soft' programs profiling pursues different objectives, depending on the level of program's generosity. For high generosity, indeed, the information detected by profiling allows to efficiently allocate Reemployment Services only to high-skilled workers (like in WPRS), while for low generosity, information on human capital is used also to fine-tune transfers (like in REA). This finding does not hold entirely when (NS) constraint is removed and worker's utility is allowed to decline along the unemployment spell (see Section 5).

**Proposition 4** (**Optimal Policy Sequence**). *No optimal policy sequence ever refers workers to profiling for two subsequent periods. Furthermore, profiling is fully accurate for high generosity levels and can therefore occur only once. On the contrary, for low generosity levels profiling does not entirely detect high human capital and delegates the job search also to a fraction of low-skilled workers. In such a case, any worker possibly undergoes profiling more than once.*

*Proof.* See Appendix B: Properties of AP, SP and IP. ■

Prop. 4 sheds light on some features of the optimal sequence of policy instruments, clarifying that it is never optimal to profile workers for two successive periods. The reason is simple: the level of accuracy that results at the end of a two-tier profiling could be achieved by a one-shot profiling, at a lower cost (by avoiding double payment of $\kappa^{wp}$). Second, workers undergo profiling multiple times only when their human capital is not profiled completely. And for this reason the planner prefers to delegate the job-search also to a fraction of low-skilled workers.

[13] See Lemma 1.

# 5 Decreasing Utility

The presence of (NS) constraint prevents workers' utility from falling along the unemployment spell. However, this is not usual in existing programs, which use negative duration dependence of promised utility as an additional incentive to induce the agent to search. Thus, 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.

**Proposition 5** (**Welfare Policies with Decreasing Utility**)**.** *Fix $\mu$ and move $U$. Then, Unemployment Insurance (UI), Jod-Search assistance (JS) and Social Assistance (SA) are optimal for low, intermediate and high $U$, respectively. Now, fix $U$ and move $\mu$. SA, JS and UI are optimal for low, intermediate and high $\mu$, respectively. 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 C: Decreasing Utility. ■

Prop. 5 sheds light on the possible policy patterns that can arise as a function of worker's initial expectation and program's generosity. While the policy location in the $(\mu, U)$ space is the same as in the case of constant promised utility, optimal policy sequences are not. The main difference is that workers in JS may now be followed by UI. The intuition behind this new result is that the cost of incentive provision and effort compensation is increasing in the level of agent's promised utility, and this fact triggers a decrease of promised utility along the unemployment spell. Indeed, allowing for worker's utility to fall over time eases incentive provision and makes worker's job-search more appealing in the eyes of the planner. Fig. 3 shows two instances of optimal policy sequences, for same initial expectation ($\mu_0 = 0.9$) and different levels of generosity. When generosity is higher ($\overline{U}_0 = 25.7$), the worker remains in JS and eventually enters SA with the same utility level and consumption as the entry ones.[14]

[14]In this case, the (NS) constraint has no impact on the design of policies, hence its removal causes no change in the optimal program.

When generosity is lower ($\underline{U}_0 = 23.8$), instead, utility decreases over time and the worker is referred to UI after JS.

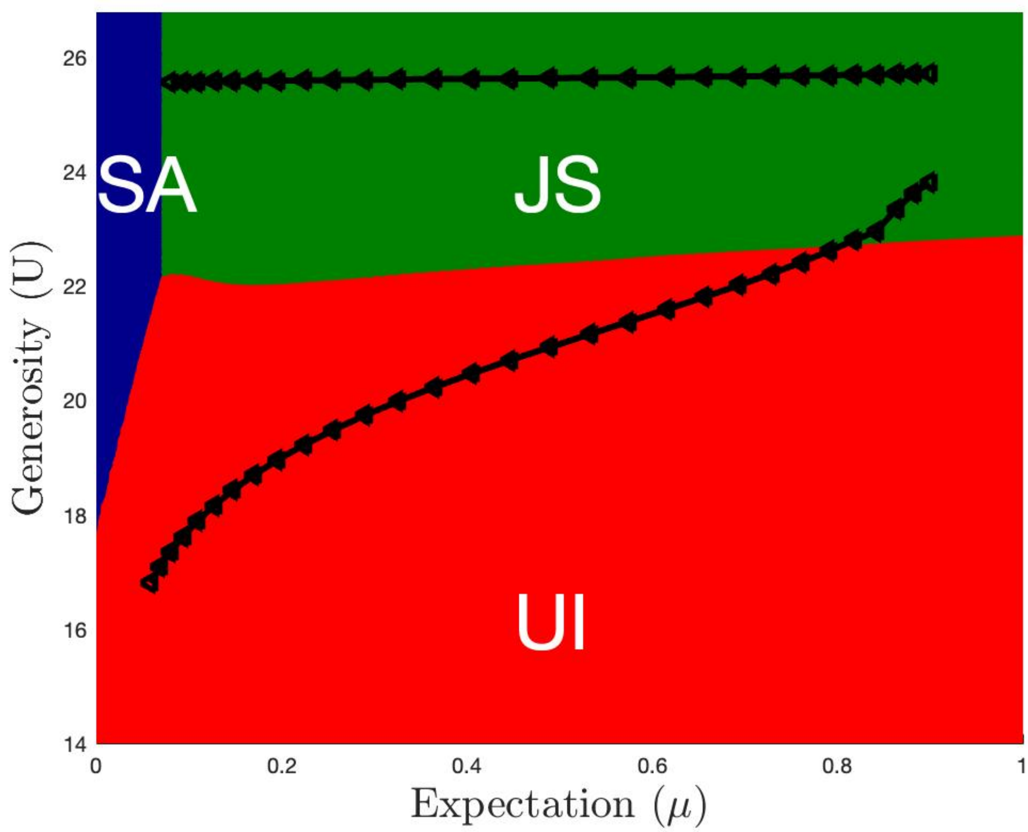


Figure 3: Optimal Income-Support Policies in the Space of Expectation and Generosity.

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 during worker's 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 6** (**Optimal Profiling with Decreasing Utility**). *Assume that worker's continuation utility is allowed to fall along the unemployment spell. Then, when profiling refers workers to JS (for higher generosities), the signal is fully accurate. When, instead, profiling refers workers to UI (for lower generosities), the 'Pass' posterior is either 1 or* *solves*

$$V_\mu(\hat{\mu}, U_p^u) = \frac{V(\hat{\mu}, U_p^u) - V(0, U_f^u) + V_U(0, U_f^u)(U_f^u - U_p^u)}{\hat{\mu}} \tag{11}$$

*with $\hat{\mu} = \mu_p$ in AP and $\hat{\mu} = \mu_p'$ in SP/IP. As with present (NS), $\hat{\mu}$ is increasing in generosity.*

*Proof.* See Appendix C: Decreasing Utility. ∎

The government sets different continuation utilities according to the profiling outcome. As shown in Prop. 5, the cost of incentive provision and effort compensation is increasing in generosity also when no (NS) constraint is imposed, which makes the marginal loss of higher generosity larger in UI than in SA. Hence, the government finds it optimal to lower the net discounted value of future payments upon 'Pass'. 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. 11 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.[15] 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 7 (Non-Soft Profiling Contracts).** *Assume that $1/u'$ is convex and (NS) constraint is absent. Then,*

- *if 'Pass' refers to a policy sequence that contemplates UI at some point of the spell, promised utility upon 'Pass' is lower than current utility, while it remains equal to it when workers are referred to JS/SA ever after;*

[15] The first two forces where already at play in the problem with (NS) constraint (see Section 4).

- *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' (which always refers to UI);*

- *in IP (resp., SP), the net wage upon reemployment is larger than (resp., equal to) current unemployment benefits.*

*Proof.* See Appendix C: Decreasing Utility. ■

Fig. 4 plots the patterns of policies, expectation, utility and unemployment benefits for a worker who enters the program with initial expectation of $\mu_0 = 0.85$ and promised utility of $U_0 = 24.07$. The worker is initially assisted in the search and profiled after 5 months. If she is found low-skilled, she is referred to SA with constant transfers. If, instead, she is found high-skilled, she is requested to search autonomously with transfers declining over time. As profiling is fully accurate and human capital entirely detected, any policy under either profiling outcome is absorbing.

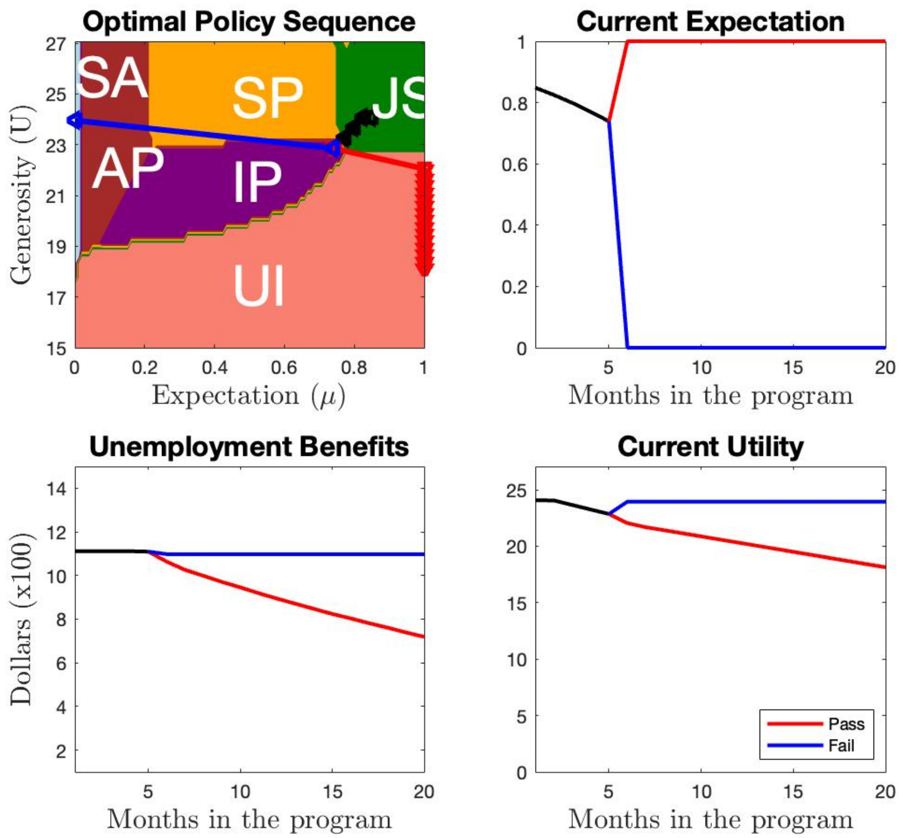


Figure 4: Consumption pattern upon profiling in a program with decreasing utility and $\mu_0 = 0.85$ and $c_0 = 100 \times exp((1-\beta)U_0) = \$1,110$.[16]

Fig. 5 plots the wage tax upon re-employment for the high-skilled worker. During JS and IP, constant net consumption pledged to the re-employed worker and declining expected productivity produce a declining tax pattern. After profiling, wage tax soars, as labor productivity (resp., job-finding probability) is constant and equal to $\omega_H$ (resp., $\pi_H$), whereas incentive costs, and worker's consumption thereof, keep falling over time.

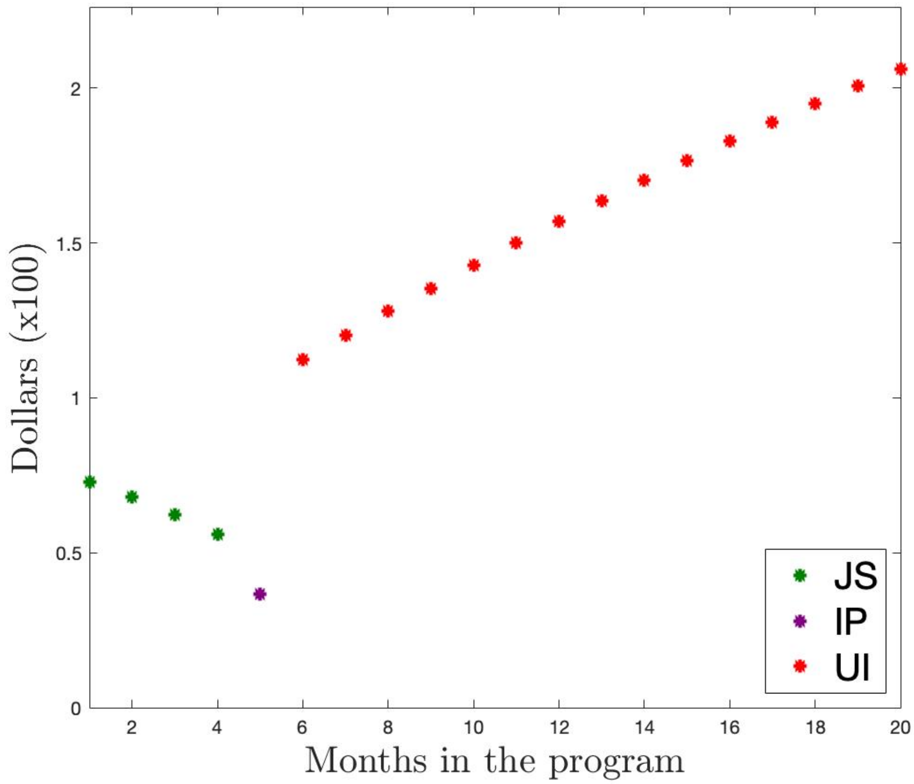


Figure 5: Wage tax upon re-employment for a high-skilled worker with initial expectation and promised utility equal to $\mu_0 = 0.85$ and $c_0 = 100 \times exp((1-\beta)U_0) = \$1,110$, respectively.

# 6 Quantitative Analysis

## 6.1 Parameterization

As anticipated in Section 1, 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. 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

| Parameter | Symbol | Value | Source |
|---|---|---|---|
| **Preferences** | | | |
| Discount Factor | $\beta$ | 0.99 | |
| Search Effort Cost | $e$ | 0.27 | various sources |
| **Labor Market** | | | |
| Job Search Hazard | $\{\pi_H, \pi_L\}$ | {0.27, 0.14} | basic monthly CPS, y. 2019 |
| Net Wage | $\{c^w_H, c^w_L\}$ | {$2,498, $1,128} | Poe-Yamagata et al. (2012) |
| Wage Tax | $\{\tau_H, \tau_L\}$ | {$178, -$224} | EIC, FICA |
| **Worker Profiling** | | | |
| Administrative Cost | $\kappa^{wp}$ | $50 | Poe-Yamagata et al. (2012) |
| **REA programs (FL, ID, IL, NV)** | $c^i$, | | |
| Generosity (consumption equivalent) | $i = FL, ID, IL, NV$ | [$1,350,$2,301] | Nicholson and Needels (2011) |

Table 2: Choice of Parameters Value

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 an employment or 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%.[17]

Turning to to the choice of parameters (see Table 2), the parameters to be chosen are: the functional form of period utility ($u(.)$), the discount factor ($\beta$), 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 ($\kappa^{wp}$). The unit of time is set to one month.

I use a logarithmic specification of utility and set the monthly discount factor equal to $\beta = 0.99$. 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 $\overline{e}^m = 0.67$ and $\overline{e}^w =$

[17] The 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., 2012). The report does not distinguishes 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.

0.97 given the logarithmic specification[18]. 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. (2012) 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 three quarters before displacement and steadily declines later on. The decline is consistent with the *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 last 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.[19] 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. Given that there are two unknowns and three States, I compute $\{c_H, c_L\}$ as the pair that minimizes the loss function

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

with $\varphi_i$ being the fraction of all 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

[18] Logarithm 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.

[19] Poe-Yamagata et al. does not report the percentage of high-skilled recipients in Illinois, which makes their data on wages useless for the estimation of $\{c_H, c_L\}$.

counterpart, I reverse engineer the gross labor income by computing the tax and deductibles that led to net amounts. In 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 with an AGI lower than a certain amount, that depends on their marital status and number of children, are entitled to an Earned Income Credit (EIC). 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 $\$29,976$.[20] 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 at that time. 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 D: Estimation of hazard rates).

Passing to the choice of $\kappa^{wp}$, 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, $\kappa^{wp} = \$50$.

The generosity of any program depends both on the amount of flow endowments and the

[20] The net annual income of high- and low-skilled workers is $\$2,498 \times 12 = \$29,976$ and $\$1,128 \times 12 = \$13,535$, 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 $\$2,139$.

duration. Poe-Yamagata et al. (2012) 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 generosity of State programs. In short, the succession of welfare programs is: Unemployment Insurance (UI) for 26 weeks, Emergency Unemployment Compensation (EUC) for additional 53 weeks, and Extended Benefits (EB) for other 20 weeks, with the last one only applying to States with unemployment rates exceeding 8.5%, which was the case for all four States in 2009. Lastly, after claimants have exhausted all three programs, they are granted unlimited access to Supplemental Nutrition Assistance Program (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), which replaced the Food Stamps Program in 2008. Average total payment was \$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 was receiving a \$501 monthly benefit from SNAP.[21] 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 when she receives regular UI benefits. Worker's utility in SNAP with no search is[22]

$$U_{e=0}^{SNAP} = \frac{u(c^{SNAP})}{1-\beta} = \frac{\log(5.01)}{1-0.99} = 161.1$$

while the utility of reemployment in case she is high-(resp., low-)skilled amounts to

$$U_H^w = \frac{u(c_H^w)}{1-\beta} = \frac{\log(29.16)}{1-0.99} = 337.3 \quad U_L^w = \frac{u(c_L^w)}{1-\beta} = \frac{\log(18.76)}{1-0.99} = 293.2$$

Condition $U_L^w > U_{e=0}^{SNAP} + \frac{e}{\beta \pi_L}$ implies that the worker always finds it convenient to search also during SNAP. Hence, the value of SNAP can be rewritten as function of entry

[21] See SNAP Data Tables at the following link: https://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap.

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

expectation $\mu$

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

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 the last month, the following recursion is implemented

$$U_{i,j,t} = u(c_t^j) - e + \beta[\mu_i^{t-1} \pi_H U_H^w + (1 - \mu_i^{t-1}) \pi_L U_L^w + (1 - \pi(\mu_i^{t-1})) U_{i,j,t+1}], \ 1 \leq t \leq 25,$$
$$j = \{FL, ID, IL, NV\}, \ i = \{< HS, HS, < CD, CD, GD\}$$

with $U_{i,j,26} = U^{SNAP}(\mu_i^{26})$, $j$ indexing States and $i$ indexing education. The initial probability of being high-skilled, $\mu_i^0$, equals the share of high-skilled individuals with same educational attainment, $\theta_i$. The generosity levels of each program and educational attainment, expressed in consumption-equivalent terms,[23] 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, Illinois (resp., Idaho) is the most (resp., least) generous one for all levels of education.

| States | Less Than HS | HS Diploma | Some College | College | Graduate |
|---|---|---|---|---|---|
| Florida | $1,350 | $1,536 | $1,580 | $1,748 | $1,811 |
| Idaho | $1,141 | $1,282 | $1,315 | $1,440 | $1,487 |
| Illinois | $1,666 | $1,920 | $1,981 | $2,212 | $2,301 |
| Nevada | $1,362 | $1,550 | $1,595 | $1,763 | $1,827 |

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

## 6.2 Optimal REA Program

The downward pattern of unemployment benefits displayed by the succession UI, EUC, EB and SNAP programs is inconsistent with the 'soft' design of Section 4, as the No Stick

[23]Consumption equivalent of utility $U$ is given by

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

constraint would prescribe a constant consumption. It would thus make no sense to compare existing programs in Florida, Idaho, Illinois and Nevada to a 'soft' optimal benchmark. I therefore remove the No Stick constraint from the planner's problem and allow current transfers, as well as their net present value, to decline over time.

Given that the REA program does not contemplate the possibility of referral to reemployment services, I restrict government's choice to policies which do not perform any form of search assistance, and let it choose among SA, UI, AP and IP.[24]

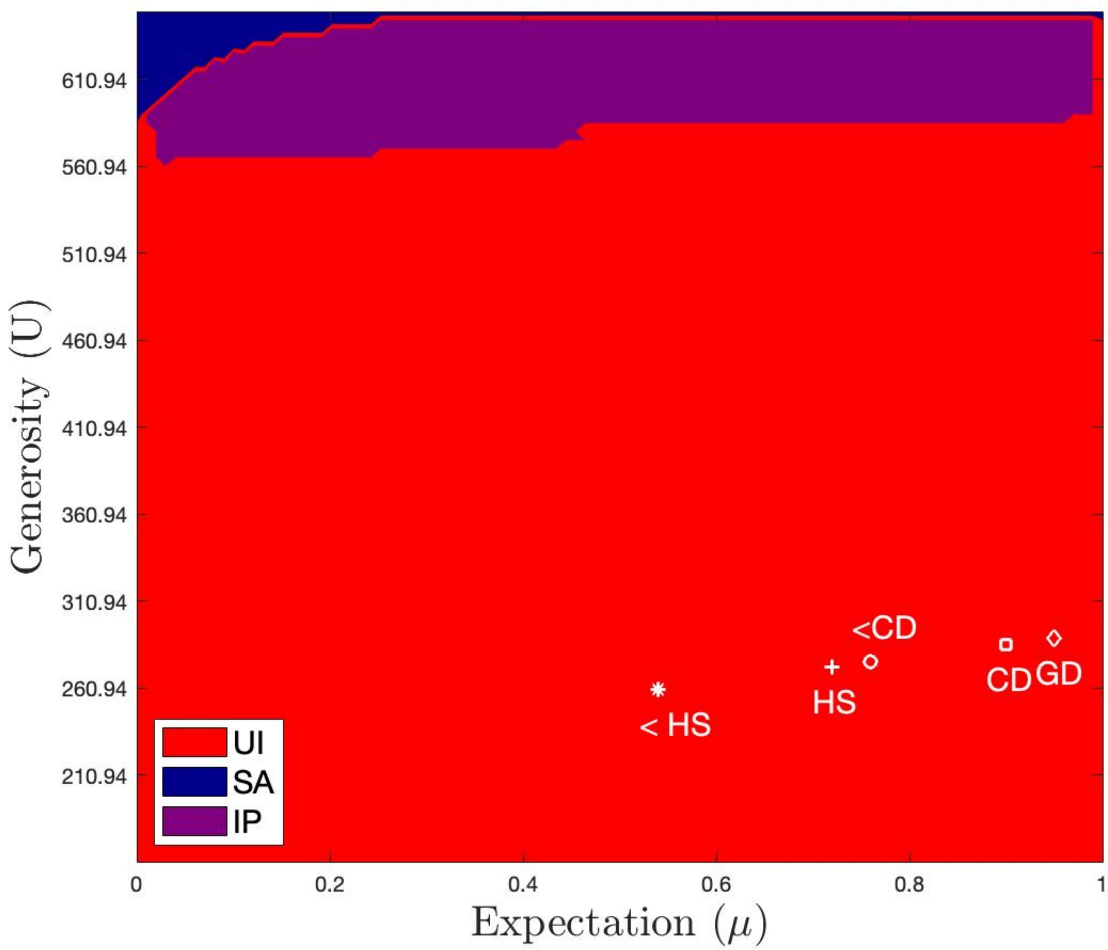


Figure 6: Optimal Policies in the Space of expectations and Generosities.
*Note*: *<HS=Less Than High School, +HS=High School Diploma, o<CD=Some College, □CD=College Degree, ◇GD=Graduate Degree

Fig. 6 reports the optimal policies in the state-space of programs' generosity and initial

[24]Given the initial low generosity of actual programs and assisted search being optimal for high-end generosities only, this exclusion restriction has no bite.

expectation, and locates the REA program implemented in Florida for each educational group.[25] The figure shows that the generosity level of actual programs is so low, to make UI be optimally implemented for the whole unemployment spell. This finding is consistent with the actual program, which delegates search to workers until they are employed and never switches to any passive labor-market policy in the meantime.

The eligibility assessment of REA excludes from UI, EUC or EB benefits and leaves solely on SNAP those workers who turn out to be highly re-employable. In principle, this aspect features also optimal profiling. Indeed, as shown by Prop. 7, payments after IP are lower for high-skilled workers, so as to lower their incentive cost. However, at the optimum no worker is ever profiled, as her promised utility is too low for IP to be optimal at any point of the spell.

Fig. 7 compares the actual and optimal patterns of promised utility, unemployment benefits and wage taxes/subsidies for Florida's jobseekers with a college degree, whose initial expectation and promised utility are $\mu_0 = 0.9$ and $U_0 = 309$, respectively. Optimal benefits (solid line) display a wider variability than their actual counterpart (dashed line). The initial transfer to recipients amounts to $\$2,200$ at the optimum, as opposed to $\$936$ of the actual UI benefit, and declines all along the spell. The actual wage tax, instead, declines over time and becomes negative (i.e., a subsidy) after 17th months, while displaying a hump-shaped pattern at the optimum. The more downward sloped pattern of current transfers, together with a larger wage tax/lower net wage (at least from month 4 on), suggests that the optimal program relies more on the 'stick' and less on the 'carrot' for incentive provision, compared to actual REA. The result is a more rapid decline in worker's promised utility at the optimum, as shown by the top right panel of Fig. 7.

[25]The initial generosity of REA programs in the other States is quite similar.

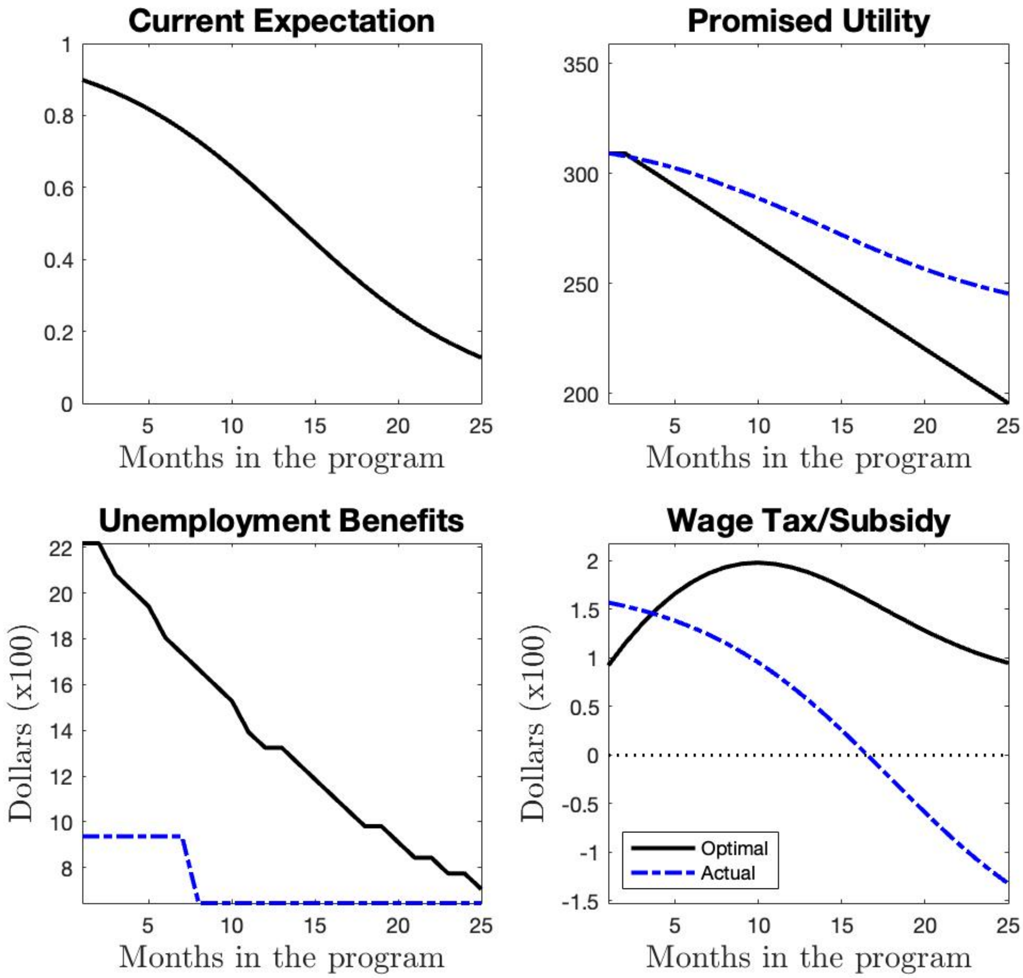


Figure 7: Optimal REA program of Florida for recipients with a college degree over a 25-month horizon (UI+EUC+EB). Initial expectation and generosity are $\mu_0 = 0.9$ and $U_0 = 309$, respectively.

# 7 Conclusions

This paper analyzes the efficiency gains in assisting unemployed workers that can be obtained with profiling. The rationale for embedding profiling into a welfare program stems from the difficulty of inferring recipients' job-finding skills and on-the-job productivity. At the optimum, active labor-market policies and workers' expectations about their skills and productivity are shown to be 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, which come in the form of lower wage taxes or higher wage subsidies.

The effects of implementing worker profiling within the program divide into gains and losses. The gains from workers' profiling stem from incentive alignment between workers and the government. Indeed, rather than pooling into the same policy and contract both high- and low-skilled workers with equal expectations, profiling allows to refer them to the proper job-search method so to minimize the cost of the program. The losses caused by profiling are its implementation cost, and the referral to passive policies of low-skilled workers who are positively mistaken (i.e., overconfident) about their human capital, and would be otherwise assigned to an active policy. This second argument may be conducive to partial detection of hidden skills aimed at strategic persuasion in the sense of Kamenica and Gentzkow (2011).

Second, under profiling the government finds it optimal to randomize not only over expectations, but also over continuation utilities. In particular, an optimal program promises a lower utility to recipients who are profiled as high-skilled and required to search for job, since the incentive and effort compensation costs are increasing in the level of utility that is promised to them. This result, which already features actual REA programs, should be accompanied with a decreasing-in-time pattern of unemployment benefits, as opposed to the constant subsidy under SNAP.

Some questions remain unanswered. First, Michaelides and Muser (2017) study the impact of WPRS and REA programs on the probability of successful job search, and outline three channels that account for it. A services effect, since workers are given proper instruments to conduct the search. A monitoring effect, since workers who do not conduct the search or any other activity that is requested to them are disqualified from the benefits. And a threat effect, since compliance with requirements is verified along the spell. While the first and second effect can be accounted for by the mechanics of the model -the services effect stems from the allocative efficiency achieved via worker profiling, and the non-contractibility of effort can be read as effort monitoring by the government with a limited liability constraint to limit punishment in case of worker's defection-, the threat effect is not present in the paper as information is symmetric. If worker's search effort were not observed by the government/principal, revision of expectations along the unemployment spell could differ between the two parties and the threat effect could be formalized as an additional policy instrument that aligns expectations and curbs worker's rents. Likewise, the paper assumes that both parties share the same initial expectation, as claimants truthfully report their personal data to the provider at the beginning of the program. However, claimants with high level of confidence in reemployment may anticipate being requested to search and choose to misreport their personal data,[26] so as to benefit from larger incentives. If this is the case, the government would need to make search-incentivizing contracts robust to information misreporting and this would further exacerbate the problem of incentive provision. A further shortcoming of the paper is constituted by (i) the cost of profiling being constant and independent of the change in expectations that ensues it, and (ii) the accuracy of profiling being unrestricted.[27] Allowing (i) for a varying cost in accordance to the change induced on the initial expectation, and (ii) for an upper bound on accuracy of profiling, could possibly lead to different profiling strategies, in terms of informativeness of the test, job-search activity to be conducted simultaneously, and policy sequence.

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

[27] Any actual profiling program, as well as any sort of tests aimed at detecting a hidden characteristic, contains a given amount of noise that impedes a fully precise detection.

# References


(1) Addison J., Portugal P., 1989, "Job displacement, relative wage changes and duration of unemployment", Journal of Labor Economics, 7(3), 281-302;

(2) Ashenfelter O., 1978, "Estimating the Effect of Training Programs on Earnings", Review of Economics and Statistics, 60, 47-50;

(3) Bhaskar V., 2012, "Dynamic Moral Hazard, Learning and Belief Manipulation", CEPR Discussion Paper No. DP8948;

(4) Boleslavsky R., Kim K., 2018, "Bayesian persuasion and moral hazard", *mimeo*;

(5) Cappellini S., 2020a, "Worker Profiling and Reemployment Services", *mimeo*;

(6) 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;

(7) Fallick, B. C., 1996, "A review of the recent empirical literature on displaced workers", Industrial and Labor Relations Review, 50:1, 5-16;

(8) Gonzalez F. M., Shi S., 2010, "An Equilibrium Theory of Learning, Search, and Wages", Econometrica, 78(2), 509-537;

(9) Hopenhayn H. A., Nicolini J. P., 1997, "Optimal Unemployment Insurance", Journal of Political Economy, 105(2), 412-438;

(10) Kamenica E., Gentzkow M., 2011, "Bayesian persuasion", American Economic Review, 101(6), 2590-2615;

(11) Kamenica E., Gentzkow M., 2014, "Costly Persuasion", American Economic Review: Papers and Proceedings, 104(5), 457-462;

(12) Keane, M. P. and Wolpin K. I., 1997, "The Career Decisions of Young Men", Journal of Political Economy, 105 (3), 473–522;

(13) 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;

(14) 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;

(15) Manoli D., Michaelides M., Patel A., 2018, "Long-Term Effects of Job-Search Assistance: Experimental Evidence Using Administrative Tax Data", *mimeo*;

(16) Meyer B. D., 1990, "Unemployment insurance and unemployment spells", Econometrica, 58(4), 757-782;

(17) Michaelides M., Mueser P., 2017, "The Labor Market Effects Of US Reemployment Policy: Lessons From An Analysis Of Four Programs During The Great Recession", *mimeo*;

(18) Neal D., 1995, "Industry-Specific Human Capital: Evidence from Displaced Workers", Journal of Labor Economics, 13 (4), 653–677;

(19) Nicholson W. and Needels K., 2011, "The EUC08 Program in Theoretical and Historical Perspective", Mathematica Policy Research, n. 06863.320;

(20) Pavoni N., 2007. "On optimal unemployment compensation", Journal of Monetary Economics, 54, 1612-1630;

(21) Pavoni N., 2009, "Optimal unemployment insurance, with human capital depreciation and duration dependence", International Economic Review, 50(2), 323-362;

(22) Pavoni N., Setty O., Violante G. L., 2013, "Search and Work in Optimal Welfare Programs", NBER Working Paper No. 18666;

(23) Pavoni N., Setty O., Violante G. L., 2016, "The design of 'soft' welfare-to-work programs", Review of Economic Dynamics, 20, 160-180;

(24) Pavoni N., Violante G. L., 2007, "Optimal welfare-to-work programs", Review of Economic Studies, 74(1), 283-318;

(25) Poe-Yamagata E., Benus J., Bill N., Carrington H., Michaelides M., Shen T., 2011, "Impact of the Reemployment and Eligibility Assessemnt (REA) Initiative", IMPAQ International, Task Order No. DOLF091A21507;

(26) Rodina D., 2017, "Information design and career concerns", *mimeo*;

(27) Shavell S., Weiss L. (1979), "The Optimal Payment of Unemployment Insurance Benefits over Time", Journal of Political Economy, 87, 1347-1362;

(28) 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.

# APPENDIX

## Appendix A: Properties of SA, JS and UI

### Proof of Prop. 1

*Proof.* For the moment, assume no randomization over policy histories is possible and ignore the (NS) constraint. Therefore, Envelope Theorem and first-order conditions imply

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

Now, given that SA is optimal in $(\mu, U)$, then $V_U(\mu, U) = V_U^{SA}(\mu, U) = V_U(\mu, U^u)$, and concavity of $V$ in $U$ in turn implies that $U^u = U$.[28] Therefore, (NS) is satisfied with equality and the state space $(\mu, U)$ is equal in the next period, proving that SA to be optimal forever after.

Allow now for randomization over policy histories and assume that the plan

$$\mathcal{X} = \left\{SA, \mathcal{X}'\right\}$$

is optimal for some expectation $\mu$ and generosity $U$, where $\mathcal{X}'$ can possibly contain randomizations over different policy histories. I will show that such a program is equivalent to prescribing SA forever on.

The payoff of $\mathcal{X}$ equals

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

Define an alternative program $\tilde{\mathcal{X}}$, prescribing the randomization between two branches $B_1$ and $B_2$ that are so defined. In $B_1$ SA is prescribed forever after $t$, while the branch $B_2$ prescribes $\mathcal{X}'$. If the first branch is assigned with weight $1-\beta$ and the second with weight $\beta$, then the payoff of $\tilde{\mathcal{X}}$ is equal to the payoff of $\mathcal{X}$, as it looks

$$(1-\beta)V^{SA}(U) + \beta V(\mu, U) = (1-\beta)\frac{-c^{sa}}{1-\beta} + \beta V(\mu, U)$$

[28] $V$ can be shown to be concave over the whole space of $U$ (see Pavoni et al. (2016)).

Now, $\mathcal{X}$ being optimal implies that the return of $B_1$ and $B_2$ is the same

$$V^{SA}(U) = V(\mu, U)$$

Indeed, if the two returns were different, one could assign a weight equal to 1 to the more rewarding branch, so to obtain a higher payoff, therefore contradicting the optimality of $\mathcal{X}$. But then

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

and so $\mathcal{X}$ can be replaced by a program prescribing SA forever on. ■

## Proof of Prop. 2

**Proposition (Optimal Contracts).** *At the optimum, (IC) upon search failure is binding in UI and IP. The contracts therefore are:*

$$\begin{aligned}
&U_i^u = U \\
&c^{ui} = c^{ip} = u^{-1}((1-\beta)U),\ U_{ui}^w = U_{ip}^w = U + \frac{e}{\beta\pi(\mu)} \\
&c^{ap} = u^{-1}((1-\beta)U) \\
&c^{sp} = c^{js} = u^{-1}((1-\beta)U),\ U_{sp}^w = U_{js}^w = U
\end{aligned}$$

*Proof.* The proof is more involved than someone may expect, and requires to first show that

$$V_U(\mu, U) \leq -\frac{1}{u'(u^{-1}((1-\beta)U))} \tag{12}$$

Given that $V$ is the fixed-point solution of a contraction, I guess and verify that (12) holds. In particular, one can already observe that it holds with equality for SA. The value of $i \in \{JS, AP, SP\}$ reads

$$
\begin{aligned}
V^i(\mu, U) = & \max_{c^i, (U_r^w, U_r^u, \mu_r)_{r\in\{p,f\}}} -c^i - \kappa^i + \beta \sum_{r\in\{p,f\}} q_r \left[p^i(\mu_r) W(\mu_r, U_r^w) + (1-p^i(\mu_r)) V(\mu_r', U_r^u)\right] \\
& \text{sub:} \quad U = u(c^i) + \beta \sum_{r\in\{p,f\}} q_r \left[p^i(\mu_r) U_r^w + (1-p^i(\mu_r)) U_r^u\right] \quad (\lambda) \\
& \qquad U_r^w, U_r^u \geq U \quad (\xi_r^w, \xi_r^u)
\end{aligned}
$$

The FOCs and Envelope Condition are:

$$
\begin{aligned}
& \lambda = \frac{1}{u'(c^i)} \\
& \beta q_r p^i(\mu_r)[W_U(\mu_r, U_r^w) + \lambda] + \xi_r^w = 0 \\
& \beta q_r (1-p^i(\mu_r))[V_U(\mu_r', U_r^u) + \lambda] + \xi_r^u = 0 \\
& V_U^i(\mu, U) = -(\lambda + \sum_{r\in\{p,f\}} (\xi_r^w + \xi_r^u))
\end{aligned}
$$

From the second FOC, one already concludes that $U_r^w = U$. Indeed, *per contra*, if $U^w > U$:

$$
\begin{aligned}
& u(c^i) = U - \beta \sum_{r\in\{p,f\}} q_r \left[p^i(\mu_r) U_r^w + (1-p^i(\mu_r)) U_r^u\right] < (1-\beta)U < (1-\beta)U_r^w \\
\Longrightarrow & W_U(\mu_r, U_r^w) < -\frac{1}{u'(c^i)} \Longrightarrow \xi_r^w > 0
\end{aligned}
$$

and the slackness condition guarantees the result. Merging FOC and Envelope Condition yields

$$
V_U^i(\mu, U) = -\frac{1}{u'(c^i)} - \sum_{r\in\{p,f\}} (\xi_r^w + \xi_r^u)
$$

Now, either all $U_r^u = U$, so that $u(c^i) = (1-\beta)U$ and $\sum_{r\in\{p,f\}}(\xi_r^w + \xi_r^u) > 0$, and the result holds, or $U_r^u > U$ and $\xi_r^u = 0$ for some $r = \{p, f\}$, by slackness condition. In the latter case,

$$
V_U^i(\mu, U) \leq -\lambda = V_U(\mu_r', U_r^u) \leq -\frac{1}{u'(u^{-1}((1-\beta)U_r^u))} < -\frac{1}{u'(u^{-1}((1-\beta)U))}
$$

where the second inequality holds by the guess hypothesis. This leads to a contradiction, since $u(c^i) < (1-\beta)U$ by assumption, hence showing that $U_r^u = U$, $\forall r = \{p, f\}$, in $i \in \{JS, AP, SP\}$. And the result follows.

The value of $i \in \{UI, IP\}$ reads

$$
\begin{aligned}
V^i(\mu, U) = & \max_{c^i, (U_r^w, U_r^u, \mu_r)_{r\in\{p,f\}}} -c^i - \kappa^i + \beta \sum_{r\in\{p,f\}} q_r \left[p^i(\mu_r) W(\mu_r, U_r^w) + (1 - p^i(\mu_r)) V(\mu_r', U_r^u)\right] \\
\text{sub:} \quad & U = u(c^i) - e + \beta \sum_{r\in\{p,f\}} q_r \left[p^i(\mu_r) U_r^w + (1 - p^i(\mu_r)) U_r^u\right] \quad (\lambda) \\
& U \geq u(c^i) + \beta \sum_{r\in\{p,f\}} q_r U_r^u \quad (\chi) \\
& U_r^w, U_r^u \geq U \quad (\xi_r^w, \xi_r^u)
\end{aligned}
$$

The FOCs and Envelope Condition are:

$$
\begin{aligned}
& \lambda - \chi = \frac{1}{u'(c^i)} \\
& \beta q_r p^i(\mu_r)[W_U(\mu_r, U_r^w) + \lambda] + \xi_r^w = 0 \\
& \beta q_r \left[(1 - p^i(\mu_r))\big(V_U(\mu_r', U_r^u) + \lambda\big) - \chi\right] + \xi_r^u = 0 \\
& V_U^i(\mu, U) = \chi - \lambda - \sum_{r\in\{p,f\}} (\xi_r^w + \xi_r^u))
\end{aligned}
$$

First notice that the (IC) constraint can be rewritten

$$
-e + \beta \sum_r q_r p^i(\mu_r)(U_r^w - U_r^u) \geq 0
$$

Hence, there is at least one profiling outcome (say $r = p$, wlog) such that $U_p^w > U$ and $W_U(\mu_p, U_p^w) + \lambda = \xi_r^w = 0$. Assume that $U_f^w = U$. Then

$$
W_U(\mu_f, U) + \lambda \leq 0 = W_U(\mu_p, U_p^w) + \lambda
$$

which is absurd as

$$
W_U(\mu_f, U) \leq W_U(\mu_p, U_p^w) = W_U(\mu_f, U_p^w) < W_U(\mu_f, U)
$$

with the equality following by separability and the inequality by concavity of $W$ in the second argument. Therefore, it must be that

$$W_U(\mu_p, U_p^w) + \lambda = W_U(\mu_f, U_f^w) + \lambda \Longrightarrow U_p^w = U_f^w > U$$

In addition, (IC) is binding, since

$$0 = W_U(\mu_r, U^w) + \lambda = W_U(\mu_r, U^w) + \frac{1}{u'(c^i)} + \chi \Longrightarrow \chi > 0$$

as $u(c^i) \leq U - \beta \sum_{r\in\{p,f\}} q_r U_r^u \leq (1-\beta)U < (1-\beta)U^w$.
Merging FOCs and Envelope Condition yields

$$V_U^i(\mu, U) = -\frac{1}{u'(c^i)} - \sum_{r\in\{p,f\}} \xi_r^u$$

Therefore, if all $U_r^u = U$, the same reasoning applies as in the case of $i \in \{JS, AP, SP\}$ and the result follows. Assume *per contra* that $U_r^u > U$ for some outcome $r$, which implies that

$$\xi_r^u = 0 \Longrightarrow (1 - p^i(\mu_r))\Big(V_U(\mu_r', U_r^u) + \frac{1}{u'(c^i)}\Big) = \chi p^i(\mu_r) > 0 \tag{13}$$

which leads to a contradiction as $u(c^i) = U - \beta \sum_{r\in\{p,f\}} q_r U_r^u < (1-\beta)U$ and

$$-\frac{1}{u'(c^i)} < V_U(\mu_r', U_r^u) \leq -\frac{1}{u'(u^{-1}((1-\beta)U_r^u))} < -\frac{1}{u'(u^{-1}((1-\beta)U))} < -\frac{1}{u'(c^i)}$$

with the first inequality following from (13) and the second by the guess hypothesis on $V$. All policy instruments satisfy the initial guess (12), which is thus verified. ■

## Proof of Lemma 1

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

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

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 $\mu$. Moreover, if $V$ is concave in either argument, then so is $V^i$.*

*Proof.* To prove concavity of $V^i$ in $U/\mu$, it suffices to show that:

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

Simply notice that $g=u^{-1}$ is convex, and that $W$ and $V$ are concave in $U^w/\mu$ and $U^u/\mu'$, respectively. 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 $\Gamma^i_\mu$ (i.e., for constant $\mu$) defined as

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

is convex. Same applies to the graph of $\Gamma^i_U$ (i.e., for constant $U$), since PK and IC are linear in $\mu$.

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

- (negative) monotonicity of the objective function in $U$;
- (negative) monotonicity of the feasibility set $\Gamma^i_\mu$ in $U$, i.e.

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

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 and NS constraints are tightened by an increase of $U$, which thus leads to a shrinkage of $\Gamma^i_\mu$.

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

- (positive) monotonicity of the objective function in $\mu$;
- (positive) monotonicity of the feasibility set $\Gamma^i_U$ in $\mu$, i.e.

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

The objective function is always monotone in $\mu$, as so are $W$ and $V$ in their first argument, $\mu'$ is an increasing function of $\mu$ and $W(\mu, U^w) \geq V(\mu^i, U^u)$. Monotonicity of $\Gamma^i(., U)$, instead, holds as an increase of $\mu$ leads to a relaxation of (IC)[29]. Indeed, (IC) is more slack since

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

■

**Slopes with respect to $\mu$**

The next step consists of showing that there exists $\mu_{sa,js}$ and $\mu_{js,ui}$ such that

- SA is optimal for $\mu \leq \mu_{sa,js}$,
- JS is optimal for $\mu \in (\mu_{sa,js}, \mu_{js,ui}]$, and
- UI is optimal for $\mu > \mu_{js,ui}$.

SA is known to be an absorbing policy, and its value is therefore given by (2). Hence $V^{SA}_\mu = 0$. As a consequence, there exists a point where JS switches to SA and $\mu_{sa,js}$ represents the threshold at which the government is indifferent between administering JS for one period and SA forever after $\mathcal{X} = \{JS, SA, SA, ...\}$, and administering SA right from the first period $\mathcal{X}' = \{SA, SA, ...\}$. The derivative of $V$ with respect to $\mu$ when JS is first implemented is

$$V^{JS}_\mu(\mu, U) = \beta\Big[(\pi_H - \pi_L)\big(W(\mu, U) - V(\mu', U)\big) + \pi(\mu)W_\mu(\mu, U) + (1 - \pi(\mu))V_\mu(\mu', U)\frac{\partial \mu'}{\partial \mu}\Big]$$

with $V(\mu', U) = \max_{i=SA,JS} V^i(\mu', U)$

[29](PK) is always relaxed by an increase in $\mu$ (recall that $U^w \geq U^u$ at the optimum).

which is positive at the optimum. Indeed, a necessary condition for JS to be preferable to SA in $\mu$ is that

$$W(\mu, U) \geq V(\mu, U) > V(\mu', U)$$

Therefore, when implementing $\mathcal{X}$, $V$ is positive increasing in $\mu$. And the derivative is increasing in the duration of JS, by an induction argument. Linearity of JS in $\mu$ follows from linearity of $\pi(\mu)W(\mu, U)$ in $\mu$. Thus,

$$V^{JS}(\mu, U) = -c - \kappa^{ja} + \beta\big[\pi(\mu)W(\mu, U) + (1 - \pi(\mu))V^{SA}(U)\big]$$

is linear and

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

is linear as well, as $(1-\pi(\mu))\mu' = \mu(1-\pi_H)$ is linear in $\mu$ and so is any linear transformation of $\mu'$. Moreover,

$$V^{JS}_{t,\mu}(\mu^{JS}_t, U) < V^{JS}_{t+1,\mu}(\mu^{JS}_t, U), \quad \text{with } \; V^{JS}_t(\mu^{JS}_t, U) = V^{JS}_{t+1}(\mu^{JS}_t, U), \; \forall t \geq 0$$

Passing to UI, consider the derivative of $V^{UI}$:

$$\begin{aligned} V^{UI}_\mu(\mu, U) =& \beta\Big[(\pi_H - \pi_L)\big(W(\mu, U^w) - V(\mu', U) - W_U(\mu, U^w)(U^w_1 - U)\big) + \\ & + \pi(\mu)W_\mu(\mu, U^w) + (1 - \pi(\mu))V_\mu(\mu', U)\frac{\partial \mu'}{\partial \mu}\Big] \end{aligned}$$

Clearly, $V^{UI}_\mu(\mu, U) > V^{JS}_\mu(\mu, U)$. Indeed:

- $W_\mu(\mu, U) = W_\mu(\mu, U^w)$;
- $W(\mu, U^w) + W_U(\mu, U^w)(U - U^w) > W(\mu, U)$ due to concavity of $W$ in $U$.

Therefore, $V^{UI}$ crosses JS only once in $\mu_{js,ui}$ from below.

**Slopes with respect to $U$**

I can now prove (7). First, notice that

$$V_U^{SA}(U) = -\frac{1}{u'(c^{sa})} = -\frac{1}{u'(c^w)} = W_U(\mu, U) \quad \text{and} \quad c^{js} = u^{-1}((1-\beta)U) = c^{sa}$$
$$\Longrightarrow V_{1,U}^{JS}(\mu, U) = -\frac{1-\beta}{u'(c^{js})} + \beta\big[\pi(\mu)W_U(\mu, U) + (1-\pi(\mu))V_U^{SA}(U)\big] = -\frac{1-\beta}{u'(c^{sa})} + \beta V_U^{SA}(U) = V_U^{SA}(U)$$

and, by induction,

$$\begin{aligned} V_{t+1,U}^{JS}(\mu, U) &= -\frac{1-\beta}{u'(c^{js})} + \beta\big[\pi(\mu)W_U(\mu, U) + (1-\pi(\mu))V_{t,U}^{JS}(\mu', U)\big] \\ &= -\frac{1-\beta}{u'(c^{js})} + \beta\big[\pi(\mu)W_U(\mu, U) + (1-\pi(\mu))V_{t-1,U}^{JS}(\mu', U)\big] = V_{t,U}^{JS}(\mu, U) \end{aligned}$$

Therefore, I can conclude that $V_{t+1,U}^{JS}(\mu, U) = V_{t,U}^{JS}(\mu, U) = V_U^{SA}(U) = W_U(\mu, U) < 0$.

I will now prove that

$$V_U^{UI}(\mu, U) < V_U^{JS}(\mu, U)$$

$$\begin{aligned} V_U^{JS}(\mu, U) &= -\frac{1-\beta}{u'(c^{js})} + \beta\big[\pi(\mu)W_U(\mu, U) + (1-\pi(\mu))V_U(\mu', U)\big] \\ V_U^{UI}(\mu, U) &= -\frac{1-\beta}{u'(c^{ui})} + \beta\big[\pi(\mu)W_U(\mu, U^w) + (1-\pi(\mu))V_U(\mu', U)\big], \quad U^w = U + \frac{e}{\beta\pi(\mu)} \end{aligned}$$

Now, given that $W$ is concave in $U$, it holds that $W_U(\mu, U) > W_U(\mu, U^w)$, while $c^{js} = u^{-1}((1-\beta)U) = c^{ui}$ and so the inequality is shown.

**Supermodularity**

$W_{\mu U}(\mu, U) = 0$ follows from separability of $W$ in both arguments.

$$\begin{aligned} V_{\mu U}^{JS}(\mu, U) &= \beta\big[(\pi_H - \pi_L)(W_U(\mu, U) - V_U(\mu', U)) + (1-\pi(\mu))\frac{\partial \mu'}{\partial \mu}V_{\mu U}(\mu', U)\big] = 0 \\ V_{\mu U}^{UI}(\mu, U) &= \beta\Big[(\pi_H - \pi_L)(W_U(\mu, U^w) - V_U(\mu', U) + (U - U^w)W_{UU}(\mu, U^w)) + (1-\pi(\mu))\frac{\partial \mu'}{\partial \mu}V_{\mu U}(\mu', U)\Big] \\ &> \beta\Big[(\pi_H - \pi_L)(W_U(\mu, U) - V_U(\mu', U) + (1-\pi(\mu))\frac{\partial \mu'}{\partial \mu}V_{\mu U}(\mu', U)\Big] \geq 0 \end{aligned}$$

where the equality follows from the fact that $V^{SA}$ is independent of $\mu$ and from (7), and the first inequality follows from concavity of $W_U$ in $U$ and the second one again from (7).

**Concavity in $U$**

Concavity in $U$ follows from a recursive argument, as

$$V_{UU}^{SA}(U) = (1-\beta)\frac{u''(c^{sa})}{u'(c^{sa})^3} < 0$$

And moving upward in the space of expectations, concavity follows by Lemma 2, as the hypothesis about concavity of $V$ in $U$ is met.

**Concavity in $\mu$**

Concavity in $\mu$ descends from linearity of $V^{SA}$ and concavity of $W$ in $\mu$, by Lemma 2. ■

# Appendix B: Properties of AP, SP and IP

## Proof of Prop. 3

*Proof.* Define $\mu_{i,j}$ the threshold in the space of expectations where the planner is indifferent between policies $i$ and $j$.

For $\mu \leq \mu_{js,ui}$, IP is dominated by SP as

$$\begin{aligned}
V^{IP}(\mu, U) = & - c^{ui} - \kappa^{wp} + \beta[\pi(\mu)W(\mu, U_{ui}^{w}) + q(1-\pi(\mu_p))V(\mu_p', U) + (1-q)(1-\pi(\mu_f))V(\mu_f', U)] \\
\leq & - c - \kappa^{ja} - \kappa^{wp} + \beta\big[\pi(\mu)W(\mu, U) + q(1-\pi(\mu_p))V(\mu_p', U) + (1-q)(1-\pi(\mu_f))V(\mu_f', U)] \\
\leq & - c - \kappa^{ja} - \kappa^{wp} + \beta\big[\pi(\mu)W(\mu, U) + \\
& + \max_{\hat{q}, \hat{\mu}_f, \hat{\mu}_p} \big\{\hat{q}(1-\pi(\hat{\mu}_p))V(\hat{\mu}_p', U) + (1-\hat{q})(1-\pi(\hat{\mu}_f))V(\hat{\mu}_f', U)\big\}\big] = V^{SP}(\mu, U)
\end{aligned}$$

where the first inequality follows from $V^{UI}(\mu, U) \leq V^{JS}(\mu, U)$ when $\mu \leq \mu_{js,ui}$. Thus, IP crosses SP in $\mu_{sp,ip} \geq \mu_{js,ui}$.

For $\mu \leq \mu_{sa,js}$, SP is dominated by AP as

$$\begin{aligned}
V^{SP}(\mu, U) = & - c - \kappa^{ja} - \kappa^{wp} + \beta\big[\pi(\mu)W(\mu, U) + q(1-\pi(\mu_p))V(\mu_p', U) + (1-q)(1-\pi(\mu_f))V(\mu_f', U)] \\
\leq & - c - \kappa^{wp} + \beta\big[\pi(\mu)V^{SA}(U) + q(1-\pi(\mu_p))V(\mu_p', U) + (1-q)(1-\pi(\mu_f))V(\mu_f', U)] \\
\leq & - c - \kappa^{wp} + \beta \max_{\hat{q} \in \Delta^*([0,1])} \int_0^1 \hat{q}(x)V(\hat{\mu}(x), U) = V^{AP}(\mu, U)
\end{aligned}$$

where the first inequality follows from $V^{JS}(\mu, U) \leq V^{SA}(U)$ when $\mu \leq \mu_{sa,js}$, while the second from the fact that the distribution $(q_1, q_2, q_3) = (\pi(\mu), q(1-\pi(\mu_p)), (1-q)(1-\pi(\mu_f)))$

over posteriors $(\mu_1, \mu_2, \mu_3) = ((\mu\pi_H)/\pi(\mu), \mu'_p, \mu'_f)$ belongs to $\Delta^*([0,1])$.[30] Thus, SP crosses AP in $\mu_{sp,jp} \geq \mu_{sa,js}$.

■

## Proof of Prop. 4

*Proof.* The value of SP/AP can be written as

$$\begin{aligned} V^{SP}(\mu,U) =& -u^{-1}((1-\beta)U) - \kappa^{wp} - \kappa^{ja} + \beta[\pi(\mu)W(\mu,U) + q(1-\pi(\mu_p))V(\mu'_p,U)+ \\ &+ (1-q)(1-\mu(\mu_f))V(\mu'_f,U)] = qV^{JS}(\mu_p,U) + (1-q)V^{JS}(\mu_f,U) - \kappa^{wp} \\ V^{AP}(\mu,U) =& -u^{-1}((1-\beta)U) - \kappa^{wp} + \beta[qV(\mu_p,U) + (1-q)V(\mu_f,U)] \\ =& qV^{SA}(\mu_p,U) + (1-q)V^{SA}(\mu_f,U) - \kappa^{wp} \end{aligned}$$

Assuming *per contra* that IP referred to SP, the return would be

$$\begin{aligned} V^{IP}(\mu,U) &= -c_{ip} - \kappa^{wp} + \beta\big[\pi(\mu)W(\mu,U^w_{ip}) + q(1-\pi(\mu_p))V(\mu'_p,U) + (1-q)(1-\pi(\mu_f))V^{SP}(\mu'_f,U)\big] \\ &< -c_{ip} - \kappa^{wp} + \beta\big[\pi(\mu)W(\mu,U^w_{ip}) + q(1-\pi(\mu_p))V(\mu'_p,U)+ \\ &\qquad + (1-q)(1-\pi(\mu_f))\big(q'_fV^{JS}(\mu_{fp},U) + (1-q'_f)V^{JS}(\mu_{ff},U)\big)\big] \\ &= -c_{ip} - \kappa^{wp} + \beta\big[\pi(\mu)W(\mu,U^w_{ip}) + q(1-\pi(\mu_p))V(\mu'_p,U)+ \\ &\qquad + (1-q)\big(q_f(1-\pi(x))V^{JS}(\mu_{fp},U) + (1-q_f)(1-\pi(y))V^{JS}(\mu_{ff},U)\big)\big] \\ &\leq -c_{ip} - \kappa^{wp} + \beta\big[\pi(\mu)W(\mu,U^w_{ip}) + \max_{q\in\Delta^*([0,1])}\int_0^1 q(x)(1-\pi(\mu(x))V(\mu(x)',U)dx\big] \end{aligned}$$

where the first inequality follows as $\kappa^{wp} > 0$, and the second one as the distribution $(q_1, q_2, q_3) = (q, (1-q)q_f, (1-q)(1-q_f))$ over posteriors $(\mu_1, \mu_2, \mu_3) = (\mu_p, x, y)$ with $x = \frac{\mu_{pf}(1-\pi_L)}{\mu_{pf}(1-\pi_L)+(1-\mu_{pf})(1-\pi_H)}$, $y = \frac{\mu_{ff}(1-\pi_L)}{\mu_{ff}(1-\pi_L)+(1-\mu_{ff})(1-\pi_L)}$ and $q_fx + (1-q_f)y = \mu_f$ is such that:

- $q(1-\pi(\mu_p)) + (1-q)q_f(1-\pi(x)) + (1-q)(1-q_f)(1-\pi(y)) = 1-\pi(\mu) = \int_0^1 q(x)(1-\pi(\mu(x))dx$;

[30] $\mu_1$ solves
$$\pi(\mu)\mu_1 + q(1-\pi(\mu_p))\mu'_p + (1-q)(1-\pi(\mu_f)\mu'_f = \mu$$

- $q_f(1-\pi(x)) = \frac{\mu_f - y}{x-y}\frac{(1-\pi_H)(1-\pi_L)}{(1-\mu_{pf})(1-\pi_H)+\mu_{pf}(1-\pi_L)} = \frac{\mu_f - y}{\mu_{pf}-\mu_{ff}}\left[(1-\mu_{ff})(1-\pi_H)+\mu_{ff}(1-\pi_L)\right] = \frac{\mu_f(1-\pi_H)-\mu_{ff}(1-\pi(\mu_f))}{\mu_{pf}-\mu_{ff}} = q'_f(1-\pi(\mu_f))$

But then then initial randomization was not optimal. An analogous argument can be used to show that IP never refers workers to AP. Moreover, the whole proof revolves around the fact that the value of SP and AP is linear in $\mu$, which allows to write them as a convex combination of their no-profiling policy counterparts. Hence, given that IP does not play any role, one can argue that the inequalities above show that no profiling policy is ever followed by either SP or AP. What is left to prove is that the same applies to IP. The difference with respect to SP and AP is that IP is concave in $\mu$ and so

$$V^{IP}(\mu, U) > qV^{UI}(\mu_p, U) + (1-q)V^{UI}(\mu_f, U) - \kappa^{wp}$$

Assume *per contra* that AP refers workers to IP. Given that UI is optimal in $\mu = 1$, it must be the case that the 'Pass' posterior is strictly lower than 1 and satisfies

$$V_\mu(\mu_p, U) = \frac{V(\mu_p, U) - V(0, U)}{\mu_p}$$

Now consider the posteriors induced by IP in $\mu_p$. Given that it is not optimal to refer any worker to either SP or AP, it must refer worker to either UI, JS or SA. Moreover, JS has convex envelope, while SA is constant in the space of expectations. Hence, the 'Fail' posterior is zero ($\mu_{fp} = 0$). The 'Pass' posterior $\mu_{pp}$ is lower than 1 and solves

$$V_\mu(\mu'_{pp}, U) = \frac{V(\mu'_{pp}, U) - V(0, U)}{\mu'_{pp}}$$

Indeed, the concave closure of the upper envelope $V$ uniquely defines posteriors, since the function

$$D(\mu, U) = V_\mu(\mu, U) - \frac{V(\mu, U) - V(0, U)}{\mu}$$

is strictly decreasing in $\mu$. Hence, $\mu'_{pp} = \mu_p$, which means that IP refers workers with 'Pass' to the optimal policy in $(\mu_p, U)$, that is, IP, by hypothesis. But then, we have reached a contradiction, since IP can only refer workers to either UI, JS or SA. An analogous argument can be used to prove that SP never refers workers to IP either.

The second part of the proposition deals with non-successive multiple profiling. If 'Pass'

outcome is associated to JS, it means that UI is never optimal for that generosity level, and that profiling is fully revealing. Therefore, if JS is the optimal policy upon 'Pass', it means that $\mu = 1$, and that no second transition from any profiling policy is possible.
On the contrary, when 'Pass' refers workers to UI, $\mu_p$ is increasing in $U$, as $D_{\mu U}(\mu_p, U) > 0$.

$$D_\mu(\mu_p, U) = V_{\mu\mu}(\mu_p, U) - \mu_p^{-1} D(\mu_p, U) = V_{\mu\mu}(\mu_p, U) < 0$$
$$D_{\mu U}(\mu_p, U) = V_{\mu U}(\mu_p, U) - \frac{V_U(\mu_p, U) - V_U(0, U)}{\mu_p} > 0$$
$$\implies \frac{\partial \mu_p}{\partial U} = -\frac{D_{\mu U}}{D_{\mu\mu}}(\mu_p, U) > 0$$

where the first inequality follows from concavity of $V^{UI}$ in $\mu$, and the second one from supermodularity of $V^{UI}$ and $V_U^{UI} < V_U^{SA}$ (see Lemma 1).

■

# Appendix C: Decreasing Utility

**Lemma 3.** *$V^{UI}$ is supermodular (i.e., $V_{\mu U}^{UI}(\mu, U) > 0$). Instead, $V^{JS}$ is supermodular if and only if $U_{JS}^u < U < \frac{u(c_{JS})}{1-\beta} = U_{JS}^w$.*

*Proof.* The problem defining $V^{UI}$ can be rewritten replacing (IC) constraint

$$U \geq u(c_{UI}) + \beta U_{UI}^u$$

with

$$U_{UI}^w - U_{UI}^u \geq \frac{e}{\beta\pi(\mu)} \quad \text{(IC ')}$$

Therefore the derivative of $V^{UI}$ wrt $U$ is

$$V_U^{UI}(\mu, U) = -\lambda'^{UI}$$

where $\lambda'^{UI}$ is the Lagrange multiplier of the (PK) constraint. Now, an increase of $\mu$ makes easier to satisfy (PK), since $U_{UI}^w - U_{UI}^u > 0$, hence $\lambda_\mu'^{UI} < 0$ and so

$$V_{\mu U}^{UI}(\mu, U) = -\lambda_\mu'^{UI} > 0$$

Similarly, the mixed derivative of $V^{JS}$ reads:

$$V^{JS}_{\mu U}(\mu, U) = -\lambda^{JS}_{\mu}$$

while FOCs lead to

$$V^{JS}_{U}(\mu, U) = W_U(\mu, U^w_{JS}) \Longrightarrow u(c_{JS}) = u(c^w_{JS}) = (1-\beta)U^w_{JS}$$

and so

$$U = \frac{1-\beta+\beta\pi(\mu)}{1-\beta}u(c_{JS}) + \beta(1-\pi(\mu))U^u_{JS} \tag{14}$$

Therefore, three are the possible cases:

- if $\frac{u(c_{JS})}{1-\beta} = U^u_{JS} = U$, then an increase of $\mu$ leaves the JS contract intact, as the (PK) constraint is always satisfied. Hence $\lambda^{JS}_{\mu} = -V^{JS}_{\mu U}(\mu, U) = 0$
- if $\frac{u(c_{JS})}{1-\beta} < U < U^u_{JS}$, then it is harder for the planner to satisfy the (PK) constraint when $\mu$ is larger. Hence $\lambda^{JS}_{\mu} = -V^{JS}_{\mu U}(\mu, U) > 0$
- if $U^u_{JS} < U < \frac{u(c_{JS})}{1-\beta}$, then it is easier for the planner to satisfy the (PK) constraint when $\mu$ is larger. Hence $\lambda^{JS}_{\mu} = -V^{JS}_{\mu U}(\mu, U) < 0$

■

## Proof of Prop. 5

*Proof.* First, $V^{SA}$ and $W$ do not change, as (NS) constraint does not play any role in their definition (see Section 4). Therefore, unemployment benefits and utility remain constant in SA.

Second, with no (NS) the derivative of the value of each policy $i$ with respect to $U$ is

$$V^{i}_{U}(\mu, U) = -\frac{1}{u'(c_i)} \tag{15}$$

which can be obtained by applying the envelope theorem to the problem of each policy.
Third, first-order conditions in UI and JS impose

$$V_U(\mu', U_{UI}^u) - V_U^{UI}(\mu, U) = V_U(\mu', U_{UI}^u) + \lambda^{UI} - \chi^{UI} = \frac{\pi(\mu)}{1-\pi(\mu)}\chi^{UI} > 0$$
$$V_U^{UI}(\mu, U_{UI}) = -(\lambda^{UI} - \chi^{UI}) = W_U(\mu, U_{UI}^w) + \chi^{UI} > W_U(\mu, U_{UI}^w)$$
$$V_U^{JS}(\mu, U) = -\lambda^{JS} = W_U(\mu, U_{JS}^w) = V_U(\mu', U_{JS}^u)$$

where $\lambda_i$ (resp., $\chi_i$) is the Lagrange multiplier associated to (PK) (resp., (IC)) constraint.

**Unemployment Benefits**

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

$$V_U(\mu', U_{UI}^u) > V_U^{UI}(\mu, U) \Longrightarrow c_{UI}^u < c_{UI}$$
$$V_U^{UI}(\mu, U_{UI}) > W_U(\mu, U_{UI}^w) \Longrightarrow c_{UI} < c_{UI}^w$$
$$V_U^{JS}(\mu, U) = W_U(\mu, U_{JS}^w) = V_U(\mu', U_{JS}^u) \Longrightarrow c_{JS} = c_{JS}^w = c_{JS}^u$$

where the implications follow from (15). Moreover, if UI lasts for more than one period, then

$$V_U^{UI}(\mu', U_{UI}^u) > V_U^{UI}(\mu, U) > V_U^{UI}(\mu', U) \Longrightarrow U_{UI}^u < U$$

where the first inequality holds by FOC and the second by supermodularity of $V^{UI}$.

**Optimal Policies in the $U$ Space**

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

$$V_U^{UI}(\mu, U) \leq V_U^{JS}(\mu, U) \leq V_U^{SA}(U) = W_U(\mu, U) \tag{16}$$

In UI the continuation utility upon job-search failure is lower than $U$ (since it was equal to it with (NS) constraint). Hence,

$$u(c_{UI}) = U - \beta U_{UI}^u \geq (1-\beta)U \Longrightarrow V_U^{UI}(\mu, U) \leq -\frac{1}{u'(u^{-1}((1-\beta)U))}$$

If JS never refers to UI, then payments remain constant and so

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

What is left to show is that, if JS refers to UI in $(\mu, U)$, then the statement still holds true. Consider a program that implements UI in the first and second period, and label its value $\hat{V}^{UI}$. *Per contra*, assume that $\hat{V}_U^{UI}(\mu, U) > V_U^{JS}(\mu, U)$. First, notice that the inequality

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

holds and implies that $(1-\beta)U \leq u(c_{UI}) < u(c_{JS})$. And from (14), it must be that $U_{JS}^u < U < \frac{u(c_{JS})}{1-\beta}$.

Now, by FOCs in the problem of UI and JS, it holds that

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

Since UI is followed by UI in $(\mu, U)$, then by concavity of UI in $U$ and assumption on slopes of $V^{UI}$ and $V^{JS}$ in the space of $U$, it must hold that $U_{UI}^u < U_{JS}^u$. But this is impossible as

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

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

Therefore, it has been shown that $\hat{V}_U^{UI}(\mu, U) \leq V_U^{JS}(\mu, U)$, also when both policies are followed by UI, which means that $\hat{V}^{UI}$ crosses $V^{JS}$ from above in the $U$ space. Hence, UI dominates JS for low-end generosity levels.

**Optimal Policies in the $\mu$ 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 $\mu$

$$\begin{aligned}V_\mu^{JS}(\mu,U) =&\beta(\pi_H-\pi_L)\big[W(\mu,U_{JS}^w)-V(\mu',U_{JS}^u)-\lambda^{JS}(U_{JS}^u-U_{JS}^w)\big]+\\&+\beta\Big[\pi(\mu)W_\mu(\mu,U_{JS}^w)+(1-\pi(\mu))\frac{\partial\mu'}{\partial\mu}V_\mu(\mu',U_{JS}^u)\Big] \qquad (17)\\V_\mu^{UI}(\mu,U) =&\beta(\pi_H-\pi_L)\big[W(\mu,U_{UI}^w)-V(\mu',U_{UI}^u)-\lambda^{UI}(U_{UI}^u-U_{UI}^w)\big]+\\&+\beta\Big[\pi(\mu)W_\mu(\mu,U_{UI}^w)+(1-\pi(\mu))\frac{\partial\mu'}{\partial\mu}V_\mu(\mu',U_{UI}^u)\Big] \qquad (18)\end{aligned}$$

Consider a program that prescribes JS first and SA later. Then, using FOC $-\lambda^{JS}=W_U(\mu,U_{JS}^w)$, it holds that

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

where the first inequality follows from concavity of $W$ in $U$ and the second by definition of $W$ and $V^{SA}$. Therefore, at the optimum JS crosses SA from below, and the inequality $V_\mu^{SA}(U)<V_\mu^{JS}(\mu,U)$ is shown.

Assume JS is followed by policy $i$, and consider a program that prescribes UI first and policy $i$ later with the additional constraint that $U_{UI}^u\geq U_{JS}^u$, and label its value $\hat{V}^{UI}$. Moreover,

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

where the first inequality follows from the statement proved above and the second one from FOC of UI.[31] In addition, one can easily show that $V^{JS}$ is supermodular and that $U_{JS}^u\leq U\leq U_{JS}^w$. Indeed:

$$V_U^i(\mu',U_{JS}^w)\leq W_U(\mu',U_{JS}^w)=W_U(\mu,U_{JS}^w)=V_U^i(\mu',U_{JS}^u) \qquad (19)$$

where the inequality holds since $V_U^i\leq W_U$ at any point of the $(\mu,U)$ space, the first equality as $W_{\mu U}=0$ and the second equality from FOC of JS. Thus, by concavity of $V^i$ in $U$, it holds that $U_{JS}^u\leq U\leq U_{JS}^w$. Supermodularity follows from Lemma 3. There are two possible cases:

- $U_{JS}^u<U_{UI}^u$;

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

- $U^u_{JS} = U^u_{UI}$.

Consider the first case, where the constraint does not bind. Hence, derivatives (17) and (18) can be rewritten

$$
\begin{aligned}
V^{JS}_{\mu}(\mu, U) =& \beta(\pi_H - \pi_L)\big[W(\mu, U^w_{JS}) - V^i(\mu', U^u_{JS}) + W_U(\mu, U^w_{JS})(U^u_{UI} - U^w_{JS}) + V^i_U(\mu', U^u_{JS})(U^u_{JS} - U^u_{UI})\big] + \\
&+ \beta\Big[\pi(\mu)W_{\mu}(\mu, U^w_{JS}) + (1 - \pi(\mu))\frac{\partial \mu'}{\partial \mu} V^i_{\mu}(\mu', U^u_{JS})\Big] \\
\hat{V}^{UI}_{\mu}(\mu, U) =& \beta(\pi_H - \pi_L)\big[W(\mu, U^w_{UI}) - V^i(\mu', U^u_{UI}) + W_U(\mu, U^w_{UI})(U^w_{JS} - U^w_{UI}) + W_U(\mu, U^w_{UI})(U^u_{UI} - U^w_{JS})\big] + \\
&+ \beta\Big[\pi(\mu)W_{\mu}(\mu, U^w_{UI}) + (1 - \pi(\mu))\frac{\partial \mu'}{\partial \mu} V^i_{\mu}(\mu', U^u_{UI})\Big]
\end{aligned}
$$

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

$$
\begin{aligned}
&W(\mu, U^w_{JS}) + W_U(\mu, U^w_{JS})(U^u_{UI} - U^w_{JS}) - V^i(\mu', U^u_{JS}) + V^i_U(\mu', U^u_{JS})(U^u_{JS} - U^u_{UI}) < \\
<& W(\mu, U^w_{UI}) + W_U(\mu, U^w_{UI})(U^w_{JS} - U^w_{UI}) + W_U(\mu, U^w_{UI})(U^u_{UI} - U^w_{JS}) - V^i(\mu', U^u_{UI})
\end{aligned}
$$

and

$$
\pi(\mu)W_{\mu}(\mu, U^w_{JS}) + (1 - \pi(\mu))\frac{\partial \mu'}{\partial \mu} V^i_{\mu}(\mu', U^u_{JS}) < \pi(\mu)W_{\mu}(\mu, U^w_{UI}) + (1 - \pi(\mu))\frac{\partial \mu'}{\partial \mu} V^i_{\mu}(\mu', U^u_{UI})
$$

The first inequality holds since:

- $W(\mu, U^w_{JS}) < W(\mu, U^w_{UI}) + W_U(\mu, U^w_{UI})(U^w_{JS} - U^w_{UI})$, by concavity of $W$ in $U$;
- $W_U(\mu, U^w_{JS})(U^u_{UI} - U^w_{JS}) < W_U(\mu, U^w_{UI})(U^u_{UI} - U^w_{JS})$, as $U^u_{UI} < U < U^w_{JS} < U^w_{UI}$;
- $V^i(\mu', U^u_{UI}) < V^i(\mu', U^u_{JS}) + V^i_U(\mu', U^u_{JS})(U^u_{UI} - U^u_{JS})$, by concavity of $V^i$ in $U$.

The second inequality holds since $W_{\mu U} = 0$ and $V^i_{\mu}(\mu', U^u_{JS}) < V^i_{\mu}(\mu', U^u_{UI})$, by assumption $U^u_{JS} < U^u_{UI}$ and supermodularity of $V^i$.

Consider now the second case. (17) and (18) can be rewritten

$$\begin{aligned}
V_\mu^{JS}(\mu,U) =&\beta(\pi_H-\pi_L)\big[W(\mu,U^w_{JS})-V^i(\mu',U^u_{JS})+W_U(\mu,U^w_{JS})(U^u_{JS}-U^w_{JS})\big]+\\
&+\beta\Big[\pi(\mu)W_\mu(\mu,U^w_{JS})+(1-\pi(\mu))\frac{\partial\mu'}{\partial\mu}V^i_\mu(\mu',U^u_{JS})\Big]\\
\hat{V}_\mu^{UI}(\mu,U) =&\beta(\pi_H-\pi_L)\big[W(\mu,U^w_{UI})-V^i(\mu',U^u_{JS})+W_U(\mu,U^w_{UI})(U^w_{JS}-U^w_{UI})+W_U(\mu,U^w_{UI})(U^u_{JS}-U^w_{JS})\big]+\\
&+\beta\Big[\pi(\mu)W_\mu(\mu,U^w_{UI})+(1-\pi(\mu))\frac{\partial\mu'}{\partial\mu}V^i_\mu(\mu',U^u_{JS})\Big]
\end{aligned}$$

and the result follows as $U^u_{JS}<U^w_{JS}<U^w_{UI}$ and

$$W(\mu,U^w_{JS})<W(\mu,U^w_{UI})+W_U(\mu,U^w_{UI})(U^w_{JS}-U^w_{UI})$$

Therefore, it has been shown that $\hat{V}^{UI}$ crosses $V^{JS}$ from below in the $\mu$ space, and so does $V^{UI}$, which implies that UI dominates JS for high expectations.

**Continuation Utility**

Passing to the last part of the statement, utility upon failed job search being lower in UI (and strictly so when UI is implemented also in the successive period) has been shown already, by noticing that $U^u_{UI}=U$ when (NS) is imposed. Also $U^u_{JS}\le U$ has been shown by (19). If JS refers to UI, then

$$V_U^{UI}(\mu,U)\le V_U^{JS}(\mu,U)=V_U^{UI}(\mu',U^u_{JS})<V_U^{UI}(\mu,U^u_{JS})$$

where the first inequality follows from (16), the second from FOC of JS and the third from supermodularity of $V^{UI}$. And it implies that $U^u_{JS}<U$. ■

## Proof of Prop. 6 and 7

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

**AP**

At the optimum,

$$V_U^{AP}(\mu, U) = V_U(\mu_p, U_{AP}^{u,p}) = V_U^{SA}(U_{AP}^{u,f}) \tag{20}$$

$$-\frac{1}{u'(c_{AP})} = V_U^{AP}(\mu, U) = V_U^{SA}(U_{AP}^{u,f}) = -\frac{1}{u'(c_{SA})} \tag{21}$$

which implies that $c_{AP} = c_{AP}^{u,p} = c_{SA}$, and $U_p^u \leq U \leq U_f^u$. Indeed, by (21), it follows

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

where the passage follows from $u(c_{AP}) = u(c_{SA}) = (1-\beta)U_{AP}^{u,f}$. If referred to JS -which is optimal only for high-end generosities-, then $\mu_p = 1$ given the linearity of JS in $\mu$. 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 where (NS) is imposed, 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 $V_U(\mu, U) < V_U^{SA}(U)$

$$V_U^{SA}(U_{AP}^{u,f}) = V_U(\mu_p, U_{AP}^{u,p}) \Longrightarrow U_{AP}^{u,p} < U < U_{AP}^{u,f}$$

Moreover, $V^{UI}$ is concave in each argument and supermodular (see Lemma 3).

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

The Hessian matrix of second-order derivatives reads

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

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

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

Differentiating the two FOCs wrt $U$ yields

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

and solving the system yields

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

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

**SP**

At the optimum

$$\begin{aligned} & V_U^{SP}(\mu, U) = W_U(\mu_p, U_{SP}^{w,p}) = W_U(\mu_f, U_{SP}^{w,f}) = V_U(\mu'_p, U_{SP}^{u,p}) = V_U(\mu'_f, U_{SP}^{u,f}) \\ \Longrightarrow & c_{SP} = c_{SP}^w = c_{SP}^{u,p} = c_{SP}^{u,f}, \quad U_{SP}^{u,p} \leq U \leq U_{SP}^{u,f} = U_{SP}^{w,p} = U_{SP}^{w,f} \end{aligned}$$

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

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

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.

**IP**

The optimal IP contract satisfies

$$V_U(\mu_p', U_{IP}^{u,p}) - V_U^{IP}(\mu, U) = \frac{\pi(\mu_p)}{1-\pi(\mu_p)}\chi^{IP} > \frac{\pi(\mu_f)}{1-\pi(\mu_f)}\chi^{IP} = V_U(\mu_f', U_{IP}^{u,f}) - V_U^{IP}(\mu, U) > 0$$

$$\implies U_{IP}^{u,p} < U_{IP}^{u,f}$$

$$V_U^{IP}(\mu, U) = -(\lambda^{IP} - \chi^{IP}) > W_U(\mu, U_{IP}^{w}) \implies c_{IP}^{u,p} < c_{IP}^{u,f} < c_{IP} < c_{IP}^{w}$$

The 'Pass' posterior in IP satisfies

$$V_\mu(\mu_p', U_{IP}^{u,p}) = \frac{V(\mu_p', U_{IP}^{u,p}) - V^{SA}(U_{IP}^{u,f}) + V_U^{SA}(U_{IP}^{u,f})(U_{IP}^{u,f} - U_{IP}^{u,p})}{\mu_p'}$$

with $U_{IP}^{u,f} > U_{IP}^{u,p}$. Moreover, $U_{IP}^{u,p} < U$, as

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

where the first inequality follows from (16), the second one from $V_U(\mu_p', U_{IP}^{u,p}) > V_U^{IP}(\mu, U)$, and the last one from $U_{IP}^{u,p} < U_{IP}^{u,f}$. Therefore, the 'Pass' in IP refers workers to UI. Indeed, it would be suboptimal to refer workers to JS, with $U_{IP}^{u,p} < U$ and $\mu_p = 1$, as IP being optimal in $(\mu, U)$ with $\mu < 1$ implies that UI is optimal in $(1, \underline{U})$ with $\underline{U} \leq U$. ■

# Appendix D: Estimation of hazard rates

In order to infer the hazard rates $\{\pi_H, \pi_L\}$, I proceed as follows. First, from the basic monthly Current Population Survey (CPS), I derive the fraction of high- and low-skilled workers for each level of educational attainment $\theta_i$, $i \in \{LessHighSc., HighSc., SomeCollege, College, Graduate\}$.[32] 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,

[32]High-skilled workers are defined as those who earn a wage higher than the mean of $\omega_H$ and $\omega_L$, that is, $2,527.

using the following formulas

$$\pi_1 = 1 - Prob(t > 1) = 1 - \frac{\# jobless\, for\, t > 1}{\# jobless}$$
$$\pi_1 + (1 - \pi_1)\pi_2 = 1 - Prob(t > 2) = 1 - \frac{\# jobless\ for\ t > 2}{\# jobless}$$
$$....$$

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, $\psi_{i,t}$. Lastly, I compute $\{\pi_H, \pi_L\}$ that minimize

$$\{\pi_H, \pi_L\} = \arg\min_{\hat{\pi}_H, \hat{\pi}_L} \sum_t \big( \sum_i \psi_{i,t}(\theta_i \hat{\pi}_H + (1 - \theta_i)\hat{\pi}_L) - \pi_t \big)^2$$

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 $a_t = \sum_i \psi_{it}\theta_i,\ b_t = \sum_i \psi_{it}(1 - \theta_i) = 1 - a_t$.[33] The results are reported in Table 4. The hazard rate $\pi_t$ is quite stable over time, as well as the share of any education level among all jobless people with same duration of unemployment spell, $\psi_{it}$. The estimated hazard rates are $\pi_H = 0.27$ and $\pi_L = 0.14$.

[33]First-order conditions for $\pi_H$ and $\pi_L$ return the minimizers of the convex objective function.

| | Total | < HS | HS Diploma | Some College[34] | College D. | Graduate D. | |
|---|---|---|---|---|---|---|---|
| $\theta_i$ | 39,333 | 0.54 | 0.72 | 0.76 | 0.9 | 0.95 | |
| Horizon | Total | $\psi_{it}$=Pr(Education=i \| Horizon $\geq t$) | | | | | Haz. Rate ($\pi_t$) |
| 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 4: Education-cohort size for any unemployment spell duration.

[34]'Some College' item includes workers with an Associate Degree, which is a post-secondary course of study lasting 2 or 3 years.