Rising UI Benefits over Time♣

Tomer Blumkin* **Efraim Sadka****

July, 2008

Revised: November, 2009

Abstract

We re-examine a key result in the optimal UI literature that benefits should decline over time. We show that when the population is heterogeneous, Pareto-efficiency may call for multiple payment schedules, some with benefits that fall over time and some with benefits that rise over time.

JEL Classification: H2, D6

Key Words: Optimal Taxation, Re-distribution, Unemployment Insurance, Inequality

♣ We wish to thank Jan Van-Ours, Yoram Weiss, Eran Yashiv and an anonymous referee for useful comments and suggestions.

* Department of Economics, Ben-Gurion University, Beer-Sheba 84105, Israel, CesIfo, IZA. E-mail: tomerblu@bgumail.bgu.ac.il.

** The Eitan Berglas School of Economics, Tel Aviv University, Tel-Aviv 69978, Israel, CesIfo, IZA. E-mail: sadka@post.tau.ac.il (Corresponding Author)

1. Introduction

In 1911, Britain introduced the first publicly financed unemployment insurance (UI) system. Since then, UI programs of one sort or another have become key policy tools in virtually all industrial economies. The primary goal of these programs is to provide consumption smoothing over periods of employment and unemployment.[1] The drawback of UI programs stems from the moral hazard effect. Indeed, the main line of research on the optimal design of UI benefits has focused on issues of the trade-off between consumption smoothing and moral hazard [see, for instance, Karni (1999), for a broad survey].

The main insight provided by the early models that appeared in the late 70's [Baily (1978), Flemming (1978) and Shavell and Weiss (1979)] was the desirability of a declining schedule, that is, benefits should decline over the spell of unemployment so as to mitigate the moral hazard effect.[2] These early models have been extended in several directions. Two such examples are Hopenhayn and Nicolini (1997) who extended the set of fiscal instruments by allowing for a wage tax after re-employment; and Fredrikson and Holmlund (2001), that consider a general equilibrium framework with endogenous wage determination (through bargaining between firms and workers). Notably, these models preserve the declining pattern featured by the early contributions.[3]

---

[1] UI programs also serve to enhance the efficiency of job search and matching in the labor market.

[2] A declining time profile is fairly prevalent. Typically, UI benefits are offered for a limited duration and then replaced by lower benefits categorized as social or income assistance (the latter is often means tested to further mitigate the moral hazard issue). According to OECD (2004) UI duration in OECD countries ranges between 6 and 60 months (Belgium being an exception with an unlimited duration in some cases).

[3] The literature emphasized the role of a declining UI schedule as a means to mitigate the tradeoff between consumption smoothing and moral hazard. Another strand in the literature demonstrates how a declining schedule can mitigate a different form of tradeoff between sorting (providing workers with incentives to wait for jobs that are more suitable) and unemployment [see Cremer, Merchand and Pestieu (1996) and, more recently, Blumkin, Hadar and Yashiv (2005)].

A common assumption in the literature on optimal UI programs, which is the focus of this study, is the homogeneity of workers.[4] Naturally, there may be many dimensions of heterogeneity one may consider. Wage rates or job opportunities are clearly among them. However, individuals with the same wage rate (or job opportunities) may still have different personal attributes that affect their re-employment prospects. These attributes include; *inert-alia*, health, education and social networking. Typically, these attributes are difficult and very costly to observe, or endogenous. For instance, education may be an observable characteristic; however, it is clearly not exogenous. If the government attempts to employ it as a 'tagging' device [á la Akerlof (1978)], individual decisions on investment in human capital (education) will be distorted. We refer to this broad class of attributes as search ability. Petrongolo (2001) provides empirical evidence for the existence of this dimension of heterogeneity. She estimates re-employment chances, and shows that the latter are adversely affected by health problems and positively affected by educational attainment, for both male and female, carefully controlling for differences in job opportunities and a large number of personal attributes, including; inter alia, marital status, age, homeownership, union membership and past employment.

In this paper, we focus on this unobservable dimension of heterogeneity and study the time profile of UI benefits. Indeed, this kind of heterogeneity lends itself to the study of the time profile of UI benefits, because other dimensions of

[4] Another common assumption in the literature is that the unemployed have no access to the capital market. This assumption implies that consumption during unemployment spells equals UI benefits. UI arrangement thus plays a dual role by providing both insurance and liquidity. A recent paper by Shimer and Werning (2008) challenges the early studies, by relaxing this assumption, and allowing workers to borrow and save. Shimer and Werning demonstrate that when workers have sufficient liquidity (perfect access to a risk-less asset), a constant benefit schedule of unlimited duration is optimal [with constant absolute risk aversion (CARA) preferences] and nearly optimal [with constant relative risk aversion (CRRA) preferences]. Notably, the optimal net unemployment subsidy is rising very slowly over time with CRRA preferences and, consequently, the welfare gain relative to a constant subsidy schedule is miniscule.

heterogeneity, such as wage heterogeneity, may be addressed by wage-dependent UI or other means. To render the analysis more tractable, without sacrificing the gist of our approach, we measure the search ability by a single (composite) index.

Relaxing the assumption of homogenous workers has several important implications. Naturally, with heterogeneous agents, there is no longer one Pareto-efficient UI program; but there will rather be many Pareto-efficient UI programs. Furthermore, a UI program need no longer consist of a single schedule, but may well consist of several time-dependent schedules which are incentive compatible. For instance, a UI program may offer one schedule with benefits rising over time and another one with benefits diminishing over time, and with individuals self-selecting between them. Specifically, allowing for individuals to differ in their search ability, we show that there may well exist Pareto-efficient programs that consist of schedules that offer rising benefits over time. In particular, we show that Pareto-efficient UI programs that favor individuals with low search ability offer the latter schedules with rising benefits over time, whereas individuals with high search ability are still offered schedules with declining benefits over time (as suggested by the literature).[5]

The organization of the paper will be as follows. In the following section we introduce the model. In section 3 we derive the properties of the optimal UI system. We conclude in section 4.

2. The Model

We construct a simple bare-bone framework with just the key ingredients necessary to demonstrate our point. Consider a two-type economy where each

[5] Shimer and Werning (2006) analyze the time-profile of UI benefits with heterogeneous workers. They focus on wage heterogeneity, whereas we choose to focus on differences in search abilities. Our papers also differ in some other aspects: first, we assume that the unemployed are credit-constrained; second, there are no lump-sum transfers in our setting; finally and most importantly, we allow for a menu of UI schedules.

individual (i=1,2) lives for two periods. We assume a continuum of individuals and normalize to unity the number of individuals for each type, with no loss in generality. Job search is conducted in the beginning of each period. Each individual engages in search for a job which offers her (for each working period) a wage rate denoted by w>0. If the individual finds a job in the beginning of the first period of life, she works both in the first and in the second period. Otherwise she engages in a second round of search in the beginning of the second period, and provided that she finds a job, she works in the second period.

All individuals share the same instantaneous utility from consumption given by $u(c)$, where $u'>0, u''<0, u'(c)\to 0$ as $c\to\infty$ and $u'(c)\to\infty$ as $c\to 0$.[6]

We assume that the probability of finding a job is a function of both the type of the individual (her ability), denoted by a^i, and the search effort exerted, denoted by e, measured in utility terms. Specifically, we let $p^i(e)=a^i\cdot p(e)$ denote the probability that an individual of type i finds a job, conditional on exerting an effort e, where $p'>0, p''<0, p'(e)\to 0$ as $e\to\infty$ and $p'(e)\to\infty$ as $e\to 0$. We further assume that $d[-p''(e)/p'(e)]/de\le 0$. That is, the marginal effort curve declines at a (weakly) falling rate (note the analogy to the non-increasing absolute risk aversion feature). The latter assumption seems reasonable, as $p(e)$ is naturally bounded from above by unity. The economic implication of the assumption is that the disincentive effect of the UI system (diminished search incentives in response to more generous benefits and/or higher payroll taxes) would be (weakly) stronger for the high search-ability individual.[7] It is straightforward to find functional forms for $p(e)$ that satisfy all

[6] Note that without being excessively unrealistic, in light of the empirical evidence [see the discussion in Saez (2002)] suggesting that labor supply elasticity (conditional on participation) is fairly low, we simplify by focusing on the extensive margin (participation choice in the labor market) while ignoring the intensive margin (labor-leisure choice), by dropping leisure from the utility function.

[7] The assumption simplifies our analysis but a weaker assumption would suffice for the arguments.

the properties specified above; for instance, the exponential case, given by $p(e)=1-\exp(-\gamma e)$, where $\gamma>0$, used by Hopenhayn and Nicolini (1997), amongst others, for numerical analysis. For concreteness, we let a type-2 individual be more able in searching for a job ($a^2>a^1>0$), that is, for a given search effort, she is more likely to find a job. Moreover, other things equal, she faces stronger incentives to search (because $\partial p^2(e)/\partial e>\partial p^1(e)/\partial e$ for all *e*), and hence will exert higher search efforts.

A standard assumption in the UI literature is that search effort is unobserved by the government. In our setup we further supplement this assumption by supposing, á la Mirrlees (1971) that the individual search ability (type) is also a private information, unobserved by the government.

Suppose that the government offers an UI program of the following form. An individual is entitled to UI benefits, denoted by b_1 and b_2, during the first and second period of unemployment, respectively; and pays a payroll tax, denoted by τ, at any working period. We let $V_1^i(b_1,b_2,\tau)$ denote the maximal level of utility derived by an individual of type *i* faced with the UI system $\langle b_1,b_2,\tau\rangle$. Thus,

(1) $$V_1^i(b_1,b_2,\tau)\equiv\max_e\left\{p^i(e)\cdot 2\cdot u(w-\tau)+[1-p^i(e)]\cdot[u(b_1)+V_2^i(b_2,\tau)]-e\right\},$$

where $V_2^i(b_2,\tau)\equiv\max_e\left\{p^i(e)\cdot u(w-\tau)+[1-p^i(e)]\cdot u(b_2)-e\right\}$.

We denote by $e_1^i(b_1,b_2,\tau)$ and $e_2^i(b_2,\tau)$, the optimal choice of efforts in period 1 and 2, respectively, by type-*i* individuals. We henceforth omit the arguments of e_1 and e_2 for notational simplicity.

Two remarks are in order. First, we simplify by assuming, with no loss of generality, that the individuals have no time preference (that is, the subjective discount rate is zero). Second, in order to stay in line with the early literature on the optimal design of UI benefits [see, for instance, the seminal contribution of Shavell and Weiss (1979)], we make the following assumptions: (i) individuals have no other sources of income, (ii) an unemployed individual cannot borrow (or lend); and (iii) the consumption good is non-storable.[8]

We turn next to the derivation of the Pareto-efficient UI programs.[9] As there are two types of individuals, the government can possibly offer two schedules, each of which chosen by a different type of individual (a separating equilibrium).[10] We denote by (b_1^1, b_2^1, τ^1) and (b_1^2, b_2^2, τ^2) the two benefit-tax schedules designed for the low-ability individual (type 1) and high-ability individual (type 2), respectively. To be incentive-compatible, these schedules must satisfy the following self-selection constraints (which state that each type has no incentive to mimic the other type):

(2) $$V_1^i(b_1^i, b_2^i, \tau^i) \geq V_1^i(b_1^j, b_2^j, \tau^j);\ i, j = 1,2;\ j \neq i,$$

These schedules must satisfy also a revenue constraint, which by virtue of the law of large numbers, requires that expected net revenues are non-negative:[11]

[8] The assumption that individuals have no other sources of income is made to simplify the analysis and can be relaxed by assuming a constant exogenous (type independent) source of income. The assumption that individuals are credit-constrained seems plausible as the unemployed find it often difficult to borrow due to moral hazard issues (a primary reason for government provision of UI benefits). Ruling out saving and borrowing is quite standard in the repeated moral hazard literature. A rare exception is Fudenberg, Holmstrom and Milgrom (1990); see also the related discussion in Rogerson (1985). For recent studies that consider the optimal UI system while allowing workers to save and borrow freely, see Shimer and Werning (2006) and (2008). See also our discussion in footnote 4 above.

[9] Evidently, our concept of Pareto-efficiency is in a constrained (second-best) sense, in light of the asymmetry in information between the individuals and the government.

[10] In a recent paper, Luttmer and Zeckhauser (2008) consider a model where individuals are imperfectly informed about their types and acquire information about the latter over time. In this case offering a menu of schedules can generate welfare gains relative to offering a single schedule which would suffice to attain the optimum in the perfect information case. Luttmer and Zeckhauser show, however, that in the special two-type case, offering just a single schedule would suffice to attain the optimum. Our model differs in that individuals are perfectly informed about their type right from the outset. We show that offering two schedules would be optimal. While this might seem to be inconsistent at a first blush, it should be noted, that our model could be re-formulated as suggesting only a single schedule.

[11] This specification assumes that the interest rate is zero and that the government has no revenue needs.

(3)

$$\sum_i \left[2 \cdot p^i(e_1^i) + [1 - p^i(e_1^i)] \cdot p^i(e_2^i)\right] \cdot \tau^i - \sum_i \left[[1 - p^i(e_1^i)] \cdot b_1^i + [1 - p^i(e_1^i)] \cdot [1 - p^i(e_2^i)] \cdot b_2^i\right] \geq 0$$

The set of Pareto-efficient UI programs consists of the set of pairs of 3-tuples, $[(b_1^1, b_2^1, \tau^1), (b_1^2, b_2^2, \tau^2)]$, which maximize a weighted average of the two utilities:

(4) $$W \equiv (1-\alpha) \cdot V_1^1(b_1^1, b_2^1, \tau^1) + \alpha \cdot V_1^2(b_1^2, b_2^2, \tau^2),$$

subject to the self-selection constraints in (2) and the revenues constraint in (3), where $0 \leq \alpha \leq 1$ is the social welfare weight assigned to type-2 individuals.

3. Properties of the Pareto-Efficient UI Programs

An interesting preliminary question is whether a Pareto-efficient UI program must offer two distinct schedules (a separating equilibrium) or just a single schedule (a pooling equilibrium). Our first result establishes indeed that a Pareto-efficient UI program necessarily involves a separating equilibrium. Furthermore, we show, that when the social welfare function [in equation (4)] is sufficiently egalitarian (that is, α is small enough), then the corresponding Pareto-efficient UI program has a binding incentive compatibility constraint for the high-search ability (type 2) individual. Formally,

Lemma: (i) A Pareto-efficient UI program must have a separating equilibrium.

(ii) For a sufficiently egalitarian social welfare function, the incentive compatibility constraint of the high search-ability individual is binding in the Pareto-efficient UI program.

Proof: See Appendix A. ■

The intuition for these results is as follows. In general the government would prefer to tailor-make a distinct policy for each type of individuals. As this is infeasible in practice, the government can employ observable differences across individuals in order to partially tailor-make policies for individuals. In our case this takes the form of offering two different time sequences of benefits (i.e., two separate UI benefit schedules) employing the differences in time-preferences across the two individuals. The second part of the lemma follows from the assumed egalitarianism: if the said incentive compatibility constraint is not binding, the government can take some more from the rich and give to the poor.

Now, we turn to our main point: under some conditions a Pareto-efficient UI program which favors the low-ability individual must offer this individual a schedule with rising benefits over time; whereas the high-ability individual is still offered a schedule with declining benefits over time.[12] More specifically, two conditions are required. First, the social welfare weight (α) assigned to the high search-ability individual (type 2) is small. Second, the search disincentive effect on the low search-ability individual (type 1) in period 2, caused by the UI system, is rather small. Formally, we measure this disincentive effect by the term $\left|\Delta^1(b_2^1)\right|$, where $\Delta^1(b_2^1) \equiv \partial p^1 \big/ \partial e_2^1 \cdot \partial e_2^1 \big/ \partial b_2^1$ denotes the effect of a small increase in the benefit (b_2^1) offered to a type-1 individual during her second period of unemployment on her probability to find a job, evaluated at the individual optimum. In the exponential case, given by $p(e) = 1 - \exp(-\gamma e)$, where $\gamma > 0$, this would amount to assuming that γ is sufficiently large. Formally,

Proposition: Let the weight assigned to the high-ability individual (that is, α) be sufficiently small. Let (b_1^1, b_2^1, τ^1) and (b_1^2, b_2^2, τ^2) constitute a Pareto-efficient UI program associated with this α. Suppose further, that the second-period distortion of

---

[12] It is worth noting that when UI programs are restricted to a single schedule (pooling equilibrium), one can show that Pareto-efficiency would imply that benefits should decline over time.

the low-ability individual search incentives caused by the UI program [that is, $\Delta^1(b_2^1)$] is sufficiently small. Then, $b_1^1 < b_2^1$ and $b_1^2 > b_2^2$.

Proof: See Appendix A. ■

The rationale for this result is as follows. In general, bearing on the optimal tax literature, one would like to reduce the distortion as much as possible for the high-ability individual [see Sadka (1976)]. The reason being the fact that it is the constraint that the high-ability individual will not mimic her low-ability counterpart (and, thereby, enjoy the transfer accorded to her) that is binding, and not the converse one (recall, that in general the government redistributes from the rich to the poor). Thus, there is no reason to distort (at the margin) the choice of the high-ability individual. In contrast, distorting the choice of the low-ability individual is required in order to dis-incentivize mimicking by the high-ability individual. In the UI context, reducing the distortion amounts to mitigating the moral hazard problem. As the UI literature suggests, this can be achieved by offering a schedule with declining benefits over time for the high-ability individual. A declining schedule for the low-ability individual would also serve the purpose of mitigating the moral hazard problem. However, for the low-ability individual there is also another consideration. In analogy to the optimal tax literature, allowing benefits to rise over time enables the government to mitigate the binding incentive constraint of the high-ability type (discouraging her from choosing the schedule designed for the low-ability individual); see part (ii) of the lemma.[13] The reason for this derives from the fact that the marginal rate of substitution between the benefit levels in the two periods is higher for the high-ability individual, as she is less likely to remain unemployed in the second period,

[13] When this second consideration is absent, as is indeed the case when the two individuals are identical, our setting replicates the standard result in the literature of declining UI benefits over time [e.g., Shavell and Weiss (1979)]; for a proof see Appendix A.

conditional on being unemployed in the first period. By mitigating this constraint, the government can raise the utility of the low-ability individual at the expense of the high-ability individual, which is desirable when α is sufficiently small. Thus, when the distortion of the search incentive of the low-ability individual caused by a rising schedule is small enough, the "re-distribution" motive dominates, and it is Pareto-efficient to offer the low-ability individual a schedule with rising benefits over time.[14] Note that we focus on the distortion of the search-effort incentive of the low-ability individual in the second period, because the distortion is more pronounced then. To see this, note that the first period effort is affected by the benefit levels in both periods; thus, as the UI benefit for this individual is lowered in the first period and raised in the second period, the combined effect on the first period effort is rather small. In contrast, the second period effort is affected only by the benefit level in the second period (which was raised). This clearly induces the individual to exert lower search efforts in the second period.

4. Conclusions

The primary goal of UI programs is to provide a tool for consumption smoothing over periods of employment and unemployment. A key result in the literature suggests that so as to mitigate the moral hazard effects inherent to such programs, due to the asymmetry in information, benefits should decline over time. Indeed, UI programs in most industrial countries follow this pattern.

The result that benefits should decline over time dwells on several assumptions; notably, that individuals are homogenous. In this paper, we re-examine

[14] This idea of searching for additional policy tools aimed at mitigating incentive compatibility constraints in order to enhance welfare lies at the core of the second-best policy design literature that followed the seminal works of Diamond and Mirrlees (1971) and Mirrlees (1971). For instance, Diamond and Mirrlees (1978) show that taxing the returns to saving can mitigate the incentive compatibility constraint, thereby improving the disability insurance system.

this result in the context of heterogeneous individuals.[15] We suggest that Pareto-efficient UI programs generally offer a variety of benefit-contribution schedules (separating equilibria). More importantly, Pareto-efficient UI programs which favor the low search-ability workers may offer the latter schedules with rising benefits over time.

[15] Our main goal was to show that this result hinges critically on the assumption of homogenous individuals. For this purpose, we chose a setting which deviates from the homogenous benchmark in the simplest possible way; namely, a setting with two types of individuals (and two periods only).

References

Akerlof, G. (1978). "The Economics of Tagging as Applied to the Optimal Income Tax," American Economic Review 68, 8-19.

Baily, M. (1978), "Some Aspects of Optimal Unemployment Insurance", Journal of Public Economics, 10, 379-402

Blumkin, T., Y. Hadar and E. Yashiv (2005), "Firm Productivity Dispersion and the Matching Role of UI Policy", IZA Discussion Paper # 1733

Cremer, H., M. Marchand and P. Pestieau (1996), "The Optimal Level of Unemployment Insurance Benefits in a Model of Employment Mismatch", *Labor Economics*, 2, 407-420

Dimaond, P. and J. Mirrlees (1971), "Optimal Taxation and Public Production I: Production Efficiency", *American Economic Review*, 61, 8-27

Dimaond, P. and J. Mirrlees (1971), "Optimal Taxation and Public Production II: Tax Rules", *American Economic Review*, 61, 261-278

Dimaond, P. and J. Mirrlees (1978), "Model of Social Insurance with Variable Retirement ", Journal of Public Economics, 10, 295-336

Flemming, J. (1978), "Aspects of Optimal Unemployment Insurance: Search, Leisure, Savings and Capital Market Imperfections", *Journal of Public Economics*, 10, 403-425

Fredriksson, P. and B. Holmlund (2001), "Optimal Unemployment Insurance in Search Equilibrium", Journal of Labor Economics, 19, 370-399

Fudenberg, D., B. Holmstrom and P. Milgrom (1990), "Short-Term Contracts and Long-Term Agency Relationships", Journal of Economic Theory, 51, 1-31

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

Karni, E. (1999), "Optimal Unemployment Insurance – A Survey", Southern Economic Journal, 66, 442-465

Luttmer, E. and R. Zeckhauser (2008), "Schedule Selection by Agents: From Price Plans to Tax Tables", NBER Working Paper # 13808

Mirrlees, J. (1971) "An Exploration in the Theory of Optimum Income Taxation", Review of Economic Studies, 38, 175-208

Petrongolo, B. (2001), "Re-employment Probabilities and Returns to Matching", Journal of Labor Economics, 19, 716-741

Rogerson, W. (1985), "Repeated Moral Hazard", Econometrica, 53, 69-76

Sadka, E. (1976), "On Income Distribution, Incentive Effects and Optimal Income Taxation", Review of Economic Studies, 43, 261-268

Saez, E. (2002), "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses", Quarterly Journal of Economics, 117, 1039-1073

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

Shimer, R. and I. Werning (2006), "On the Optimal Timing of Benefits with Heterogeneous Workers and Human Capital Depreciation", MIT Working Paper 06-12

Shimer, R. and I. Werning (2008), "Liquidity and Insurance for the Unemployed", American Economic Review, forthcoming

Appendix A: Proofs

1. Proof of the Lemma

(i) We first turn to establish a simple claim which will be used in the proof. For this purpose we introduce a new piece of notation: we denote by $\Delta^i(b_2) \equiv \partial p^i / \partial e_2^i \cdot \partial e_2^i / \partial b_2$ the effect of a small increase in the benefit (b_2) offered to a type-*i* individual during her second period of unemployment on her probability to find a job, evaluated at the individual optimal choice.

Claim: $\Delta^1(b_2) \geq \Delta^2(b_2)$.

Proof of the Claim: The first order condition for the type-*i* individual optimization (with respect to search effort in the second period) implies:

(A1) $a^i \cdot p'(e_2^i) \cdot [u(w-\tau) - u(b_2)] - 1 = 0$.

Fully differentiating the expression in (A1) with respect to b_2, employing (A1) and re-arranging yields:

(A2) $\Delta^i(b_2) = u'(b_2) / [u(w-\tau) - u(b_2)]^2 \cdot [p'(e_2^i) / p''(e_2^i)]$.

Strict concavity of the function $p(e)$ immediately implies that $e_2^2 > e_2^1$, as $a^2 > a^1$. The claim follow then by virtue of our assumption that $d[-p''(e)/p'(e)]/de \leq 0$. ■

We turn next to prove the first part of the lemma. The proof will be by way of contradiction. Let the optimal solution be a pooling equilibrium. Denote the optimal schedule by (b_1, b_2, τ). Differentiating the indirect utility in (1) with respect to b_1 and b_2, employing the envelope theorem, implies that the marginal rate of substitution between the two benefit levels (fixing the payroll tax, τ) is given by: $\left|MRS^i\right| = \dfrac{u'(b_1)}{[1 - p^i(e_2^i)] \cdot u'(b_2)}$. By virtue of the properties of the probability function, $p^i(e)$, $p^2(e) > p^1(e)$, hence, $\left|MRS^2\right| > \left|MRS^1\right|$. Now, suppose that the government offers a second schedule (b_1', b_2', τ), where: $b_1' = b_1 + \varepsilon_1, b_2' = b_2 + \varepsilon_2$, $\varepsilon_1 > 0 > \varepsilon_2$, ε_1 and ε_2 are small and $-\varepsilon_2 / \varepsilon_1 = \left|MRS^2\right|$. That is we offer a second schedule which lies on the indifference curve of the high-ability type going through the original schedule by moving (slightly) along her indifference curve in the southeast direction (see Figure 1 in the Appendix B). By construction, the high-ability individual (type 2) will be indifferent between the two schedules, whereas the low-ability individual (type 1) will strictly prefer the original (presumably optimal) schedule to the new schedule, as her indifference curve is flatter than that of the high-ability type. Thus, the incentive constraints are still satisfied. Moreover, as we decrease the benefit to which the high-ability individual is entitled during the second period, she will increase her search effort (and hence her chances to find a job) during this period. By construction of the new schedule, the effort exerted during the first period remains the same. Differentiating the revenue constraint in (3), denoting by $\Omega(\varepsilon_1, \varepsilon_2)$ the total effect of introducing the new schedule on the net tax revenues, it follows that:

(A3)

$$\begin{aligned}\Omega(\varepsilon_1,\varepsilon_2) &= \varepsilon_2 \cdot [1-p^2(e_1^2)] \cdot \Delta^2(b_2) \cdot (\tau + b_2) - \varepsilon_1 \cdot [1-p^2(e_1^2)] - \varepsilon_2 \cdot [1-p^2(e_1^2)] \cdot [1-p^2(e_2^2)] \\ &= \varepsilon_2 \cdot [1-p^2(e_1^2)] \cdot \left[\Delta^2(b_2) \cdot (\tau + b_2) + [1-p^2(e_2^2)] \cdot [u'(b_2)/u'(b_1) - 1]\right],\end{aligned}$$

where the equality follows by substituting the term $-\varepsilon_2 / \left|MRS^2\right|$ for ε_1.

A necessary condition for the original schedule to be Pareto-efficient is that $\Omega(\varepsilon_1, \varepsilon_2) \le 0$. Otherwise, offering the new schedule will maintain the incentive constraints, attain the same level of utility (for both types) as in the original schedules, but result in a positive fiscal surplus. This surplus can be utilized to attain a Pareto improvement, by raising the level of utility at all states for both types by the same arbitrarily small amount, which does not change the search and the mimicking incentives. This necessary condition implies in particular that:

(A4) $\Delta^2(b_2)\cdot(\tau+b_2)+[1-p^2(e_2^2)]\cdot[u'(b_2)/u'(b_1)-1]\geq 0$.

Now, suppose that the government is offering an alternative schedule (in addition to the original presumably Pareto-efficient schedule). Denote this new schedule by (b_1'',b_2'',τ), where $b_1''=b_1+\varepsilon_1, b_2''=b_2+\varepsilon_2$, $\varepsilon_2>0>\varepsilon_1$, ε_1 and ε_2 are small and $-\varepsilon_2/\varepsilon_1=\left|MRS^1\right|$. This time the new schedule lies on the indifference curve of the low-ability type going through the original schedule, by moving along her curve in the north-west direction (see Figure 2 in Appendix B). By construction, the low-ability individual (type 1) will be indifferent between the two schedules, whereas the high-ability individual (type 2) will strictly prefer the original (presumably optimal) schedule to the new one, as her indifference curve is steeper than that of the low-ability type. Thus, the two incentive constraints are satisfied. Moreover, as we increase the benefit to which the low-ability individual is entitled during the second period, she will decrease her search effort (and hence her chances to find a job) during this period. By construction of the new schedule, the effort exerted during the first period remains the same. Differentiating the revenue constraint in (3), denoting by $\Psi(\varepsilon_1,\varepsilon_2)$ the total effect of introducing the new schedule on the net tax revenues, it follows that:

(A5)

$$\begin{aligned}\Psi(\varepsilon_1,\varepsilon_2)&=\varepsilon_2\cdot[1-p^1(e_1^1)]\cdot\Delta^1(b_2)\cdot(\tau+b_2)-\varepsilon_1\cdot[1-p^1(e_1^1)]-\varepsilon_2\cdot[1-p^1(e_1^1)]\cdot[1-p^1(e_2^1)]\\&=\varepsilon_2\cdot[1-p^1(e_1^1)]\cdot\left[\Delta^1(b_2)\cdot(\tau+b_2)+[1-p^1(e_2^1)]\cdot[u'(b_2)/u'(b_1)-1]\right],\end{aligned}$$

where the equality follows by substituting the term $-\varepsilon_2/\left|MRS^1\right|$ for ε_1.

A necessary condition for the original schedule to be Pareto-efficient is that $\psi(\varepsilon_1,\varepsilon_2)\leq 0$. Otherwise, offering the new schedule will maintain the incentive constraints, attain both types the same level of utility as in the original schedules, but result in a positive fiscal surplus, which can attain a Pareto improvement. This necessary condition implies in particular that:

(A6) $\Delta^1(b_2)\cdot(\tau+b_2)+[1-p^1(e_2^1)]\cdot[u'(b_2)/u'(b_1)-1]\leq 0$.

Comparing the two necessary conditions (A4) and (A6) yields a contradiction, noting that by the claim $\Delta^1(b_2)\geq\Delta^2(b_2)$; $p^1(e_2^1)<p^2(e_2^2)$, by virtue of the properties of the probability function; and, $u'(b_2)/u'(b_1)-1>0$, by virtue of condition (A4). This completes the proof of part (i). ■

(ii) We turn to prove that when the social welfare function is sufficiently egalitarian (that is, α is small enough), then the corresponding Pareto-efficient UI program has a binding incentive compatibility constraint for the high-search ability (type 2) individual. By the first part of the lemma we know that any Pareto-efficient UI program must offer two distinct benefit-tax schedules to the two types of individuals. Fix some small α and let the corresponding Pareto-efficient pair of schedules be given by (b_1^1,b_2^1,τ^1) and (b_1^2,b_2^2,τ^2). Now suppose by way of contradiction that the incentive compatibility constraint of the high-ability (type 2) individual is not binding. By virtue of continuity one can slightly reduce the benefit levels offered to the high-ability individual in both periods (b_1^2,b_2^2), without affecting her mimicking incentives. This slight modification creates a fiscal surplus. The reason for this is twofold: (i) a mechanical effect associated with the cut in benefit levels; (ii) a behavioral effect associated with inducing the high-ability individual to exert higher search efforts in both periods. The surplus can be used to raise the utility level of both agents without violating the two incentive constraints [e.g., this goal can be achieved by raising the utility of both individuals in each state of nature (employed and unemployed) and each period, by the same amount). All in all (following the combined change in the two original schedules) we obtain an increase in the utility derived by low-ability (type-1) individual; and, a decrease in the utility derived by the high-ability (type-2). For low enough values of α (the weight assigned to the high-ability type), this implies an increase in social welfare. Note, that for high values of α, this modification will rather give rise to a decrease in social welfare. This completes the proof.■

2. Proof of the Proposition

We first prove that in the Pareto-efficient UI program, $b_1^1 < b_2^1$, for sufficiently small values of α.

Let (b_1^1, b_2^1, τ^1) and (b_1^2, b_2^2, τ^2) denote the two Pareto-efficient benefit-tax schedules designed for the low-ability individual (type 1) and high-ability individual (type 2), respectively. Now consider the following two alternative schedules, denoted by $(\tilde{b}_1^1, \tilde{b}_2^1, \tilde{\tau}^1)$ and $(\tilde{b}_1^2, \tilde{b}_2^2, \tilde{\tau}^2)$, obtained by small perturbations around the original schedules, where:

(i) $\tilde{b}_1^1 = b_1^1 + \varepsilon_{11}, \tilde{b}_2^1 = b_2^1 + \varepsilon_{12}$ and $\tilde{\tau}^1 = \tau^1$;
(ii) $\tilde{b}_1^2 = b_1^2 + \varepsilon_{21}, \tilde{b}_2^2 = b_2^2 + \varepsilon_{22}$ and $\tilde{\tau}^2 = \tau^2 + \varepsilon_2$,
(iii) $\varepsilon_{11} < 0, \varepsilon_{12} > 0, \varepsilon_{21} < 0, \varepsilon_{22} < 0$ and $\varepsilon_2 > 0$,
(iv) $-u'(w - \tau^2) \cdot \varepsilon_2 = u'(b_2^2) \cdot \varepsilon_{22} = u'(b_1^2) \cdot \varepsilon_{21}$,
(v) $u'(b_1^1) \cdot \varepsilon_{11} + u'(b_2^1) \cdot [1 - p^1(e_{12}^1)] \cdot \varepsilon_{12} = 0$,
(vi) $-2 \cdot u'(w - \tau^2) \cdot \varepsilon_2 = [1 - p^2(e_{11}^2)] \cdot \left[u'(b_1^1) \cdot \varepsilon_{11} + u'(b_2^1) \cdot [1 - p^2(e_{12}^2)] \cdot \varepsilon_{12}\right]$,

where $e_{1,t}^i$ denotes the optimal effort exerted at time *t* by type *i*, when faced with the schedule designed for the low-ability individual (type 1). In words, we slightly shift the schedule designed for the low-ability individual along her indifference curve in the north-west direction, thereby creating a slack in the incentive constraint of the high-ability individual. To see this, observe that the benefit level in period 1 for type 1 is reduced ($\varepsilon_{11} < 0$), whereas the benefit level in period 2 for this type is increased ($\varepsilon_{12} > 0$) and the tax level does not change ($\tilde{\tau}^1 = \tau^1$). Equation (v) then ensures that the change in the UI schedule of type 1 does not alter her utility level. Note further that this change creates a slack in the incentive compatibility constraint of type 2. This derives from the fact that the indifference curve of type 2, in the (b_1, b_2)-space, is steeper than that of type 1 (see figures 1 and 2). The size of the slack in utility terms is given by the expression on the right-hand side of equation (vi). We then use this slack to reduce the total expected utility (V_1^2) of the high-ability individual, by reducing the benefit levels in both periods ($\varepsilon_1^2, \varepsilon_2^2 < 0$) and by increasing the tax ($\varepsilon_2 > 0$). Equation (iv) ensures that the utility derived by type 2 falls by the same amount in each state of nature (employed and unemployed) and each period. Therefore, type 2 will not change her effort level in either period. As the individual lives for two periods and the discount rate is assumed to be zero, the decline in the total expected utility of type 2 (V_1^2) is twice the size of each term in equation (iv). Equation (vi) then ensures that the incentive compatibility constraint of type 2 remains binding.

Differentiating the revenue constraint in (3), employing (iv)-(vi) to substitute for $\varepsilon_{11}, \varepsilon_{12}, \varepsilon_{21}$ and ε_{22}, yields, after re-arrangement, the total effect of introducing the two new schedules on the net tax revenues as a function of ε_2, denoted by $\Xi(\varepsilon_2)$:

$$\text{(A7)} \quad \Xi(\varepsilon_2) = \varepsilon_2 \cdot \left[\begin{array}{l} 2 \cdot p^2(e_{21}^2) + [1 - p^2(e_{21}^2)] \cdot p^2(e_{22}^2) + [1 - p^2(e_{21}^2)] \cdot u'(w - \tau^2) / u'(b_1^2) \\ + [1 - p^2(e_{21}^2)] \cdot [1 - p^2(e_{22}^2)] \cdot u'(w - \tau^2) / u'(b_2^2) \\ + [1 - p^1(e_{11}^1)] \cdot \dfrac{2 \cdot u'(w - \tau^2)}{u'(b_2^1)} \cdot \dfrac{1}{[1 - p^2(e_{12}^2)]} \cdot \dfrac{1}{[p^2(e_{12}^2) - p^1(e_{12}^1)]} \\ \times \left[\Delta^1(b_2^1) \cdot (\tau^1 + b_2^1) + [1 - p^1(e_{12}^1)] \cdot [u'(b_2^1) / u'(b_1^1) - 1] \right] \end{array} \right]$$

A necessary condition for the original schedules to be Pareto-efficient is that $\Xi(\varepsilon_2) \leq 0$. Otherwise, offering the modified schedules will maintain the two incentive constraints, will maintain the utility derived by the low-ability individual and result in a fiscal surplus. This surplus can be utilized to raise the utility derived by the low-ability individual; hence, to raise social welfare, for sufficiently small values of α (notice that this will be the case although the suggested modification will result in, all in all, a decrease in the utility level of the high-ability individual, as long as the weight assigned to the low-ability agent is high enough). By virtue of the properties of the probability function, it follows that $p^2(e_{12}^2) - p^1(e_{12}^1) > 0$. Moreover, the term $(\tau^1 + b_2^1)$ is bounded from above ($w - \tau^1 > b_2^1$, as we assume an interior solution for the individual optimization problem). Thus, it follows from (A7), that when $\Delta^1(b_2^1)$ is small, a necessary condition for the original schedule to be Pareto-efficient is that $u'(b_2^1)/u'(b_1^1) - 1 < 0$. By virtue of the strict concavity of the utility function, it thus follows that $b_1^1 < b_2^1$. This completes the first part of the proof.

We turn next to prove that $b_1^2 > b_2^2$. Our proof will be in two steps. First we show that in the Pareto-efficient solution $\tau^2 > 0$. Then we prove that, provided that $\tau^2 > 0$, it follows that $b_1^2 > b_2^2$. Consider first the first step. Suppose, by way of contradiction, that $\tau^2 \leq 0$. By virtue of the revenue constraint in (3) it follows that the government obtains expected net revenues (net fiscal surplus) from the low-type individual to offset the negative surplus (in expected terms) derived from the high-ability individual (type 2). In particular, $\tau^1 > 0$. Now suppose that the government offers the high-ability individual the same schedule designed for the low-ability individual (rather than the presumably Pareto-efficient schedule). That is, the government implements a pooling equilibrium, where both types face the same schedule, (b_1^1, b_2^1, τ^1). It suffices to show that such a modification of the presumably efficient schedules results in a fiscal surplus (in expected terms) for the high-ability individual as well, to obtain the desired contradiction. The aggregate surplus (from both types) may be utilized to raise the utility of the low-ability individual, hence to raise social welfare for sufficiently small values of α.

To prove that a positive surplus is obtained for the high-ability type, it suffices to show that her search effort (hence employment chances) is higher (in each period) than that of the low-ability type, when faced with the same schedule. The reason being that in such a case, the expected net surplus for the high-ability type would exceed that obtained for the low-ability type (which is by presumption strictly positive). To see that search efforts of the high-ability type are indeed higher, denote by $e_t(a, b_1, b_2, \tau)$; t=1,2, the optimal efforts exerted by an individual with ability a, faced with a schedule (b_1, b_2, τ), in periods 1 and 2, respectively. Consider first the second period. Re-formulating the indirect utility given in (1), yields (omitting the tax parameters for notational convenience):

(A8) $\quad V_2(a, b_2, \tau) \equiv \{a \cdot p[e_2(a)] \cdot u(w-\tau) + [1 - a \cdot p[e_2(a)]] \cdot u(b_2) - e_2(a)\}$.

The first order condition for the individual optimization (with respect to search effort in the second period) implies:

(A9) $\quad a \cdot p'[e_2(a)] \cdot [u(w-\tau) - u(b_2)] - 1 = 0$.

Strict concavity of the function $p(e)$ immediately implies that $\partial e_2(a) / \partial a > 0$.

We turn next to the first period. Re-formulating the indirect utility given in (1), yields:

(A10)

$$V_1(a, b_1, b_2, \tau) \equiv \{a \cdot p[e_1(a)] \cdot 2 \cdot u(w-\tau) + [1 - a \cdot p[e_1(a)]] \cdot [u(b_1) + V_2(a, b_2, \tau)] - e_1(a)\}.$$

The first order condition for the individual optimization (with respect to search effort in the first period) implies:

(A11) $a \cdot p'[e_1(a)] \cdot [2 \cdot u(w-\tau) - u(b_1) - V_2(a, b_2, \tau] = 1$.

Fully differentiating the first-order condition in (A11) with respect to *a*, yields:

(A12)

$$p''[e_1(a)] \cdot \frac{\partial e_1}{\partial a} \cdot \left[2 \cdot u(w-\tau) - u(b_1) - V_2(a, b_2, \tau\right] - p'[e_1(a)] \cdot p[e_2(a)] \cdot [u(w-\tau) - u(b_2)] = -\frac{1}{a^2}$$

By virtue of the concavity of the probability function $p(e)$, it follows that,

(A13) $Sign\left[\frac{\partial e_1}{\partial a}\right] = -Sign\left[p'[e_1(a)] \cdot p[e_2(a)] \cdot [u(w-\tau) - u(b_2)] - \frac{1}{a^2}\right]$.

Now, substituting from (A8) and (A11) into (A13) and re-arranging, implies that:

(A14)

$$Sign\left[\frac{\partial e_1}{\partial a}\right] =$$

$$-Sign\left[\frac{a \cdot p[e_2(a)] \cdot [u(w-\tau) - u(b_2)]}{a^2 \cdot \left[\left[2 - a \cdot p[e_2(a)]\right] \cdot u(w-\tau) - u(b_1) - \left[1 - a \cdot p[e_2(a)]\right] \cdot u(b_2) + e_2(a)\right]} - \frac{1}{a^2}\right].$$

To prove that $\frac{\partial e_1}{\partial a} > 0$, it suffices to show that:

(A15) $$\frac{a \cdot p[e_2(a)] \cdot [u(w-\tau) - u(b_2)]}{\left[\left[2 - a \cdot p[e_2(a)]\right] \cdot u(w-\tau) - u(b_1) - \left[1 - a \cdot p[e_2(a)]\right] \cdot u(b_2) + e_2(a)\right]} < 1,$$

which holds if-and-only-if,

(A16) $[u(w-\tau) - u(b_1)] + \left[1 - 2a \cdot p[e_2(a)]\right] \cdot [u(w-\tau) - u(b_2)] + e_2(a) > 0.$

By definition of the probability function, it follows that $ap[e_2(a)] \leq 1$. Thus, the condition in (A16) follows if $u(b_2) > u(b_1)$. However, this follows from the first part of the proof. We thus yield a contradiction and establish that $\tau_2 > 0$.

We turn next to prove that when $\tau_2 > 0$, it follows that $b_1^2 > b_2^2$. Let the Pareto-efficient schedule offered to the type-2 individual be given by (b_1^2, b_2^2, τ^2). Suppose by way of contradiction that $b_1^2 \le b_2^2$ Consider now a small perturbation to the (presumably) efficient schedule. Denote this schedule by $(b_1^2{}', b_2^2{}', \tau^2)$, where: $b_1^2{}' = b_1^2 + \varepsilon_1, b_2^2{}' = b_2^2 + \varepsilon_2$, $\varepsilon_1 > 0 > \varepsilon_2$, and $-\varepsilon_2 / \varepsilon_1 = \left|MRS^2\right|$. By the same reasoning used in the proof of the lemma [part (i)], it is easy to verify that the perturbed system satisfies the two incentive constraints and hence maintains the same level of utility for the low-ability individual. Differentiating the revenue constraint in (3), denoting by $\Omega(\varepsilon_1, \varepsilon_2)$ the total effect of the perturbation on the net tax revenues, it follows that:

(A17)

$$\Omega(\varepsilon_1, \varepsilon_2) = \varepsilon_2 \cdot [1 - p^2(e_1^2)] \cdot \Delta^2(b_2^2) \cdot (\tau^2 + b_2^2) - \varepsilon_1 \cdot [1 - p^2(e_1^2)] - \varepsilon_2 \cdot [1 - p^2(e_1^2)] \cdot [1 - p^2(e_2^2)]$$
$$= \varepsilon_2 \cdot [1 - p^2(e_1^2)] \cdot \left[\Delta^2(b_2^2) \cdot (\tau^2 + b_2^2) + [1 - p^2(e_2^2)] \cdot [u'(b_2^2) / u'(b_1^2) - 1]\right],$$

where the equality follows by substituting the term $-\varepsilon_2 / \left|MRS^2\right|$ for ε_1.

By virtue of our presumption that $b_1^2 \le b_2^2$ and the strict concavity of the utility function, it follows that $\Omega(\varepsilon_1, \varepsilon_2) > 0$. Thus, we obtain a fiscal surplus that can attain a Pareto improvement. We obtain the desired contradiction. This concludes the proof. ■

3. The Identical Individual Case

In this part we demonstrate that when individuals are identical ($a^1 = a^2 = a$), we replicate the standard result in the literature [see e.g., Shavell and Weiss (1979)]; namely, showing that in the optimal UI benefit-tax schedule, benefits should decline over time. Maintaining the notation used above, the government is seeking to maximize the total expected utility derived by the representative agent, $V_1(b_1, b_2, \tau)$, subject to a revenue constraint:

(A18) $\left[2 \cdot p(e_1) + [1 - p(e_1)] \cdot p(e_2)\right] \cdot \tau - \left[[1 - p(e_1)] \cdot b_1 + [1 - p(e_1)] \cdot [1 - p(e_2)] \cdot b_2\right] \ge 0$.

Let the Pareto-efficient schedule be given by the triplet (b_1, b_2, τ). Suppose by way of contradiction that $b_1 \leq b_2$ Consider now a small perturbation to the (presumably) efficient schedule. Denote this modified schedule by (b_1', b_2', τ), where: $b_1' = b_1 + \varepsilon_1, b_2' = b_2 + \varepsilon_2$, $\varepsilon_1 > 0 > \varepsilon_2$, and $-\varepsilon_2 / \varepsilon_1 = |MRS|$. By the same reasoning used in the proof of the lemma [part (i)], it is easy to verify that the perturbed system maintains the same level of utility for the individual. Moreover, by construction of the new schedule, the effort exerted during the first period remains the same. Differentiating the revenue constraint in (A18), denoting by $\Omega(\varepsilon_1, \varepsilon_2)$ the total effect of the perturbation on the net tax revenues, it follows that:

(A19)

$$\Omega(\varepsilon_1,\varepsilon_2) = \varepsilon_2 \cdot [1-p(e_1)] \cdot \Delta(b_2) \cdot (\tau + b_2) - \varepsilon_1 \cdot [1-p(e_1)] - \varepsilon_2 \cdot [1-p(e_1)] \cdot [1-p(e_2)]$$
$$= \varepsilon_2 \cdot [1-p(e_1)] \cdot \left[\Delta(b_2) \cdot (\tau + b_2) + [1-p(e_2)] \cdot [u'(b_2)/u'(b_1) - 1]\right],$$

where the equality follows by substituting the term $-\varepsilon_2 / |MRS|$ for ε_1, and where $\Delta(b_2) = \partial p / \partial e_2 \cdot \partial e_2 / \partial b_2$.

By virtue of our presumption that $b_1 \leq b_2$ and the strict concavity of the utility function, it follows that $\Omega(\varepsilon_1, \varepsilon_2) > 0$. Thus, we obtain a fiscal surplus that can be used to raise the utility of the agent. We obtain the desired contradiction. This concludes the proof. ■

Appendix B: Figures

Figure 1

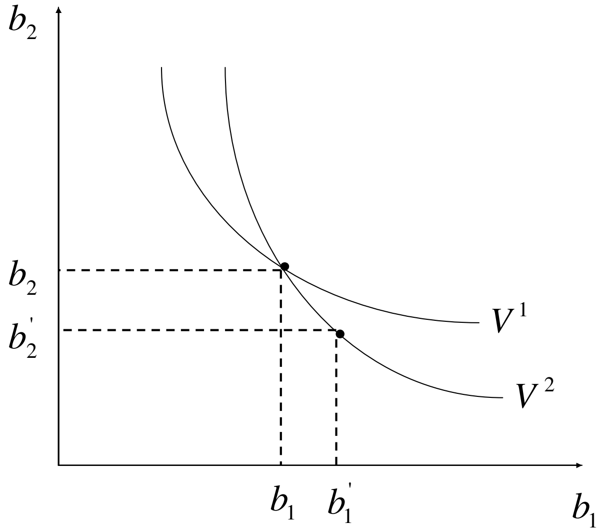

Figure 2