Social Insurance

Prepared for

The New Palgrave Dictionary of Economics and Law

Stefania Albanesi*

Columbia University

October 31, 2005

1. Introduction

Individuals face uncertainty about their labor income. Chronic medical conditions can arise that reduce or remove their ability to work. Alternatively, unemployment spells can occur, temporarily reducing their earnings. The individual burden of these risks can be eased by private insurance and by social insurance schemes. A social insurance scheme is a government transfer program whereby individuals who claim a condition

*I wish to thank Claudia Olivetti, Nicola Pavoni and Bruce Preston for very helpful comments and discussions. Contact information: 1022 International Affains Building, 420 West 118th Street, New York, NY 10027. E-mail: sa2310@columbia.edu.

or state that reduces their labor income, such as disability or unemployment, obtain a transfer from the government for the duration of this state. Social insurance can also be carried out through income taxation.

The design of optimal social insurance is a classic problem in economics. This problem is interesting because neither private insurance or public arrangements can fully distinguish between low income by choice or low income by necessity. For example, the diagnosis of medical factors is often subjective and symptoms for certain conditions, such as back pain or recurrent migraines, are hard to verify. If full insurance for chronic medical conditions were available, individuals could exploit this by reporting fake symptoms to claim insurance benefits and stop working. Similarly, an unemployed worker's ability to find new employment depends on the effort she exerts in the job search. If full unemployment insurance is available, an unemployed worker would have no incentive to search for a new job, if her search effort cannot be monitored. Private information on earning ability implies that full insurance might defeat itself by removing the incentive to work. Hence, the essential trade-off in the design of optimal social insurance schemes is the one between risk-sharing and incentives.

2. The trade-off between insurance and incentives

Consider the following simple social insurance problem[1]. Individuals live for two periods. In the first period of their life, they are endowed with one unit of consumption and they consume. In the second period, they consume and they may work if they are able to. The probability of being able to work is π, a number between zero and one. Work is publicly observable and the probability distribution of ability is known, but ability is private information. Preferences over consumption in each period are represented by $u(c)$, where $c \geq 0$ is consumption and the function u is strictly increasing and strictly concave. The utility cost of working if able is γ, a number strictly greater than zero. An individual produces one unit of output if she works, zero otherwise.

The government wishes to induce able agents to work in period two and to maximize their lifetime expected utility, given by:

$$u(c_1) + \pi u(c_2^w) + (1-\pi) u(c_2^n) - \pi\gamma,$$

where c_1 is consumption in period 1 and c_2^w and c_2^n are consumption conditional on work and no work in period two.

There is no private insurance, but the government is able to pool risk by distributing consumption, conditional on work in period two, so that the expected value of lifetime

[1] The set up for this example is adapted from Diamond and Mirrlees (1978).

consumption equals the expected value of lifetime work for any individual. Individuals cannot save but the government has access to a storage technology with return $R > 0$. Hence, an individual's lifetime consumption profile must satisfy the constraint:

$$Rc_1 + \pi c_2^w + (1 - \pi) c_2^n \leq R + \pi, \tag{2.1}$$

Since the government cannot observe ability, the consumption profile must also satisfy an incentive compatibility constraint:

$$u(c_2^w) - \gamma \geq u(c_2^n), \tag{2.2}$$

to ensure that able individuals will work.

The optimal social insurance scheme satisfies the following two inequalities when the incentive compatibility constraint (2.2) is binding:

$$\frac{u'(c_2^w)}{u'(c_2^n)} < 1, \tag{2.3}$$

$$\frac{u'(c_1)}{Eu'(c_2)} < R, \tag{2.4}$$

where $Eu'(c_2)$ is the expected marginal utility of consumption in period two.

It immediately follows from (2.3) that full insurance is not possible and $c_2^w > c_2^n$. Hence, private information implies that the first best optimum is not attainable. Equa-

tion (2.4) is derived from the government's Euler equation and uncovers another dimension of the incentive problem. It implies that the consumption path associated with the optimal insurance scheme displays a wedge between the marginal intertemporal rate of substitution and the intertemporal rate of transformation, R. This intertemporal wedge indicates that the marginal cost of transferring consumption to period two is greater than the value of foregone consumption at time one. The additional cost stems from the need to maintain incentive compatibility. This requires consumption to be conditional on work and therefore stochastic in the second period. Since utility from consumption is strictly concave, a given increase in expected utility requires more resources when consumption is stochastic rather than deterministic.

The intertemporal wedge implies that individuals would prefer to reduce their consumption in the first period. Thus, in general, the optimal intertemporal consumption profile cannot be obtained without preventing access to the capital market or imposing a tax on savings. For example, Golosov and Tsyvinski (2004a) study optimal disability insurance in competitive equilibrium and show that the intertemporal wedge requires disability benefits to be asset-tested.

Private information on individuals' ability to work gives rise to adverse selection in the social insurance problem. The resulting partial insurance and the presence of an intertemporal wedge (2.4) hold generally with adverse selection (Golosov, Kocherlakota and Tsyvinski, 2003) and also characterize social insurance under moral hazard (Roger-

son, 1985), as in the unemployment example described in the introduction. Moreover, (2.4) exemplifies a general feature of social insurance with private information, that is, that the government would want to control trade in related commodities (Atkinson and Stiglitz, 1976, and Varian, 1980).

3. Long Run Inequality

Social insurance models with infinite horizon generate normative implications for long run consumption inequality. The key step in deriving these implications is to formulate the social insurance problem recursively. The pioneering work of Green (1987) provides an early example of a recursive solution to a dynamic social insurance problem with infinite horizon and constant discounting. The main insight is to restate the optimal social insurance scheme as a contract between the government, the principal, and individuals, the agents. Such a contract retains memory of the past history of outcomes and assigns current transfers and promises of future transfers based on that history and the current outcome. Despite their large dimensionality, such histories can be summarized by agents' *promised value*[2]. This one dimensional object, which corresponds to an agent's expected lifetime utility at a point in time, encodes an agent's history and

[2]These results can be found in Spear and Srivastava (1987), Thomas and Worral (1990), Abreu, Pierce and Stacchetti (1990), and Phelan and Townsend (1991).

permits a recursive formulation of the problem. Incentive compatibility implies that current transfers and promises of future transfers will depends on promised utility and on the current outcome. The government's promises of future transfers can be represented as *continuation values* that determine agents' promised utility in the subsequent period.

For example, in an infinite horizon version of the earlier disability insurance example, the individual history for some period $t > 1$ is the sequence of work outcomes from period 1 to period $t - 1$. This can be summarized by an individual's expected lifetime utility at the beginning of time t, her promised value. The optimal consumption allocation at time t and the continuation value will depend on the promised value and on the current work outcome. The continuation value corresponds to an individual's expected lifetime utility in the subsequent period. Hence, individuals' future consumption depends on current work and on the past history of work.

The history dependence of the consumption path resulting from an optimal insurance scheme implies that the trade-off between insurance and incentives shapes the evolution of the consumption distribution. Specifically, incentive compatibility implies that, not only current transfers, but also continuation values will be conditional on current outcomes. An immediate consequence is that the degree of consumption inequality tends to continually increase. An additional implication can be derived from the intertemporal wedge, which implies that individuals will face a downward sloping path of promised

lifetime utility under the optimal social insurance scheme.

Taken together, these results give rise to an extreme conclusion. Consumption inequality should grow without bound, with all individuals in the population converging to their minimum promised lifetime utility, except for a vanishing fraction converging to their bliss point. This *immiseration* property is robust. It obtains in partial (Green, 1987, and Thomas and Worrall, 1990) and general (Atkeson and Lucas, 1992) equilibrium, under weak assumptions on preferences (Phelan, 1998), and holds in adverse selection and, in somewhat weaker form (Pavoni, 2004), in moral hazard environments.

The radical implications for consumption inequality generated by optimal social insurance models have prompted research on alternative normative criteria for the government's problem, based on an intergenerational interpretation of the infinite horizon framework. In standard models of social insurance, future generations are only considered indirectly, via the altruism of the earlier ones. Phelan (2005) proposes a government objective with equal weight on all future generations and shows that this implies a finite amount of inequality in the limit. Farhi and Werning (2005) explore a class of social insurance allocations that take into account individuals currently alive, as well as future generations, by assigning the latter a vanishingly small weight in the government's objective. They find that long-run inequality remains bounded and all individuals avoid misery.

The recursive principal-agent approach to social insurance problems underlies most macroeconomic applications. Two recently prominent areas of interest are optimal income taxation with unobserved skills and optimal unemployment insurance with hidden effort.

4. Optimal Income Taxation

Income taxes are an important instrument for social insurance. Hence, the basic trade-off between insurance and incentives also underlies the design of optimal tax systems. This intuition drives Mirrlees's (1971) seminal study of optimal income taxes. The main assumption is that labor income is observable but it depends on individual effort and skills, which are private information. Taxes are restricted to depend on labor income only, but conditional on this, the government is relatively unconstrained. Lump-sum taxes and arbitrarily progressive or regressive tax schemes can all be part of the armory. Mirrlees studies a static economy and finds that optimal marginal income taxes are low and slightly declining in income. Diamond (1998) and Saez (2001) find that marginal income taxes are high and sharply increasing in income at low income levels. Diamond and Saez's results can be interpreted as a prescription for a rapid phase out of social benefits for low income individuals, which is consistent with the US system. The properties of the optimal marginal income tax are very sensitive on the assumed properties of the skill distribution, which explains the difference in findings.

Albanesi and Sleet (2005) and Kocherlakota (2005a) apply Mirrlees's approach to dynamic economies and derive implications for optimal capital and labor income taxes. The optimal taxes depend on the agents' past history. This is achieved by conditioning the tax payments on the entire history of labor income, or, equivalently, on outstanding wealth when skill shocks are i.i.d. The properties of optimal capital income taxes stem from the intertemporal wedge associated with the optimal consumption path. As noted earlier, this wedge implies that individuals would like to save more for two reasons: their lifetime consumption path is downward sloping and they face consumption risk, given that the optimal scheme only provides partial insurance. Capital income must then be taxed to prevent this excessive saving. The optimal marginal asset tax, however, has a very specific form: it is decreasing in labor income. Hence, it is stochastic and negatively correlated with consumption. Excessive saving is discouraged by making after tax returns on assets correlate positively with labor income, and thus reducing the hedging value of holding assets[3]. Albanesi and Sleet (2005) also study the properties of marginal labor income taxes and find that they should be high at low income and decreasing in wealth. Kocherlakota (2005b) provides an extensive review of these findings.

[3] Albanesi (2005) studies optimal capital income taxes in economies with moral hazard and idiosyncratic capital returns. She finds that the intertemporal wedge can be negative in this class of economies. In this case, the optimal marginal capital income taxes are increasing in income.

5. Unemployment Insurance

Hopenhayn and Nicolini (1997) consider the design of optimal unemployment insurance under moral hazard[4]. The probability of finding a new job depends on the unemployed worker's search effort, which is private information. The unemployed worker is risk-averse and cannot borrow or save. Upon finding a new job, the worker will be employed forever at a fixed wage. The optimal unemployment insurance scheme is self-financing and comprises two elements: a sequence of unemployment benefits and a lump sum tax levied when the worker finds new employment. The presence of moral hazard implies that the replacement ratio, that is the fraction of previous salary transferred to the worker in the form of unemployment benefits, must be strictly smaller than one when the worker is searching for a new job. Hence, the optimal scheme only provides partial insurance against unemployment.

The main results are that unemployment benefits should be decreasing over the course of the unemployment spell and the employment tax is increasing with the length of the unemployment spell, under mild conditions on preferences. The decreasing benefits result[5] is a manifestation of the intertemporal wedge in this setting. The intertemporal wedge implies that promised utility should be decreasing as long as the worker is unemployed, which requires unemployment benefits to decrease over time. The em-

[4] Shimer and Werning (2005) analyze optimal unemployment insurance in a search model with adverse selection.

[5] This result was first derived in the seminal paper of Shavell and Weiss (1979).

ployment tax result stems from consumption smoothing, since continuation utility rises discretely when a worker becomes employed. By taxing a newly employed worker, the government optimally smooths this jump in consumption. A worker's promised utility declines further for longer unemployment spells, hence, the employment tax is increasing with the length of unemployment.

Pavoni (2004) enriches Hopenhayn and Nicolini's model by introducing a realistic feature: a worker's human capital depreciates over the unemployment spell. The optimal unemployment insurance program displays two novel features. First, if human capital depreciates rapidly enough during unemployment, benefits are bounded below by a minimal assistance level. Second, the optimal employment tax should decrease with the length of the unemployment spell. These new findings are a consequence of the fact that a worker's wage upon employment depends positively on her human capital. As human capital depreciates over the length of the unemployment spell, expected utility upon employment declines. This reduces and eventually eliminates the government's need to decrease benefits over time to induce a decline in promised lifetime utility over the unemployment spell. Similarly, it reduces the value of the employment tax required to smooth consumption in the transition from unemployment to employment.

More recently, Pavoni and Violante (2005) study the optimal design of welfare-to-work programs. These programs are a mix of government expenditures on passive policies, such as unemployment insurance and social assistance, and active policies, such as

job search monitoring, training and wage taxes or subsidies, targeted to the unemployed. Most governments in fact use a combination of passive and active policies in dealing with unemployment. There are several novel findings. First, the optimal welfare to work program endogenously generates a permanent policy of last resort, which resembles a social assistance program. The unemployed worker is given a constant lifetime benefit and is not active in job search or training. Second, optimal unemployment benefits are generally decreasing or constant during unemployment, but they must increase after a successful spell of training. These findings result from the fact that utility conditional on employment decreases with the length of the unemployment spell and increases after job training. The central assumption is once again human capital depreciation over the unemployment spell. Indeed, it is shown that human capital depreciation is necessary for policy transition to be part of an optimal welfare-to-work program. Pavoni and Violante find that, by providing more insurance to skilled workers and more incentives to unskilled workers, the optimal welfare-to-work program delivers significant welfare gains with respect to the existing U.S. system.

6. Concluding Remarks

Most studies of optimal social insurance exclude the presence of private insurance contracts. Does the government have a special role in the provision of insurance with private information? The seminal work of Prescott and Townsend (1984) demonstrates

that the First Welfare Theorem holds for a large class economies with private information. Golosov and Tsyvinski (2004) study an adverse selection economy and allow for private insurance provision. They show that, if all trades are observable and individuals can sign binding contracts with private insurers ex ante, the only effect of public provision of social insurance is crowding out of private insurance. On the other hand, if certain trades are not observable, privately provided insurance is not Pareto optimal and government policies can increase welfare. Bisin and Guatioli (2004) examine a moral hazard economy and show that if agents' contractual relationships with competiting insurance providers cannot be monitored, the competitive equilibrium allocation is not Pareto optimal.

The observability of trades allows insurance providers to restrict participation in additional contractual relationships with other agents. Absent this, individuals can undo the incentive effects of an insurance contract by purchasing additional insurance or engaging in trades that provide a limited amount of self-insurance. These results suggest that if exclusivity of private insurance contracts cannot be enforced, government provision of insurance plays a critical role.

References

[1] Abreu, Dilip, David Pierce and Ennio Stacchetti. 1990. Toward a Theory of Discounted Repeated Games with Imperfect Monitoring. *Econometrica* 58: 1041-1063.

[2] Albanesi, Stefania. 2005. Optimal Taxation of Risky Productive Capital with Private Information. Manuscript, Columbia University.

[3] Albanesi, Stefania and Christopher Sleet. 2005. Dynamic Optimal Taxation with Private Information. Forthcoming, *The Review of Economic Studies*.

[4] Atkeson, Andrew and Robert E. Lucas. 1992. On Efficient Distribution with Private Information. *The Review of Economic Studies* 59.3:427-453.

[5] Atkinson, A.B. and J.E. Stiglitz. 1976. The Design of the Tax Structure: Direct versus Indirect Taxation. *Journal of Public Economics* 6: 55-75.

[6] Bisin, Alberto and Danilo Guatioli. 2004. Moral Hazard and Non-Exclusive Contracts. *Rand Journal of Economics* 35: 306-328.

[7] Diamond, Peter. 1998. Optimal taxation: An example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review* 88:83-95.

[8] Diamond, Peter and James Mirrlees. 1978. A model of social insurance with variable retirement. *Journal of Public Economics* 10:295-336.

[9] Farhi, Emanuel and Ivan Werning. 2005. Inequality, Social Discounting and Progressive Estate Taxation. NBER WP 11408.

[10] Golosov, Mikhail, Narayana Kocherlakota, Aleh Tsyvinski. 2003. Optimal Indirect and Capital Taxation. *The Review of Economic Studies* 70.3: 569.

[11] Golosov, Mikhail and Aleh Tsyvinski. 2004a. Designing Disability Insurance: A Case for Asset Testing. NBER Working Paper 10792.

[12] Golosov, Mikhail and Aleh Tsyvinski. 2004b. Optimal Taxation with Endogenous Insurance Markets. NBER Working Paper 11185.

[13] Green, Edward J. 1987. Lending and the Smoothing of Uninsurable Income. In Prescott E. and Wallace N. (eds.), *Contractual Arrangements for Intertemporal Trade* (Minneapolis, University of Minnesota Press), 3-25.

[14] Hopenhayn, Hugo A. and Juan Pablo Nicolini. 1997. Optimal Unemployment Insurance. *The Journal of Political Economy* 105.2: 412-438.

[15] Kocherlakota, Narayana. 2005a. Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation. Forthcoming, *Econometrica*.

[16] Kocherlakota, Narayana. 2005b. Advances in Dynamic Optimal Taxation. Paper presented at the World Congress of the Econometric Society, London.

[17] Mirrlees, James. 1971. An exploration in the theory of optimum income taxation. *The Review of Economic Studies* 38:175-208.

[18] Pavoni, Nicola. 2004. Optimal Unemployment Insurance, with Human Capital Depreciation, and Duration Dependence. Manuscript, University College London.

[19] Pavoni, Nicola and Giovanni L. Violante. 2005. Optimal Welfare to Work Programs. Manuscript, New York University.

[20] Phelan, Christopher. 1998. On the Long Run Implications of Repeated Moral Hazard. *Journal of Economic Theory* 79: 174-191.

[21] Phelan, Christopher. 2005. Opportunity and Social Mobility. Forthcoming, *The Review of Economic Studies*.

[22] Phelan, Christopher and Robert Townsend. 1991. Computing multi-period, information constrained optima. *The Review of Economic Studies* 58:853-881.

[23] Prescott, Edward C. and Robert Townsend. 1984. Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard. *Econometrica* 52: 21-45.

[24] Rogerson, William. 1985. Repeated moral hazard. *Econometrica* 53:69-76.

[25] Saez, Emmanuel. 2001. Using elasticities to derive optimal income tax rates. *The Review of Economic Studies* 68:205-229.

[26] Shavell, Stephen and Laurence Weiss. 1979. The Optimal Payment of Unemployment Insurance Benefits over Time. *The Journal of Political Economy* 87.6: 1347-1362.

[27] Shimer, Robert and Ivan Werning. 2005. Liquidity and Insurance for the Unemployed. NBER WP 11689.

[28] Spear, Stephen and Sanjay Srivastava. 1987. On Repeated Moral Hazard with Discounting. *The Review of Economic Studies* 54: 599-617.

[29] Thomas, Jonathan and Tim Worrall. 1990. Income Fluctutations and Asymmetric Information: An Example of a Repeated Principal Agent Problem. *Journal of Economic Theory* 51: 367-390.

[30] Varian, Hal R. 1980. Redistributive taxation as social insurance. *Journal of Public Economics*, 14:49-68.