

Homburg, Stefan; Lohse, Tim

Working Paper

Optimal Tax-Transfer-Schemes under Partial Information

Diskussionsbeitrag, No. 298

Provided in Cooperation with:

School of Economics and Management, University of Hannover

Suggested Citation: Homburg, Stefan; Lohse, Tim (2004) : Optimal Tax-Transfer-Schemes under Partial Information, Diskussionsbeitrag, No. 298, Universität Hannover, Wirtschaftswissenschaftliche Fakultät, Hannover

This Version is available at:

<http://hdl.handle.net/10419/22410>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Optimal Tax-Transfer-Schemes under Partial Information

By Stefan Homburg and Tim Lohse

DISCUSSION PAPER NO. 298

MAY 2004

ISSN: 0949-9962

Chair of Public Finance, School of Economics and Management, University of Hannover, Königsworther Platz 1, D-30167 Hannover, Germany. www.fwi.uni-hannover.de.

ABSTRACT: According to a widely held belief, “all who are able to work, should work”. We consider this statement within a framework of non-linear taxation. The crucial difference between our model and the standard model is that the government can distinguish between productive persons and the disabled. A general proposition regarding the design of tax-transfer-schemes under such partial information is derived. Moreover, it is shown that unemployment on the side of the productive poor may still be optimal.

KEYWORDS: Optimal Taxation, Employment, Poverty, Welfare.

JEL-CLASSIFICATION: H21, I38

1. INTRODUCTION

During the last decades, persistently high unemployment rates have produced a lot of theoretical and empirical research. Among the established “stylized facts” is the finding that unemployment rates in most industrialized countries tend to be higher at the bottom of the income scale than at the top. Relatively higher unemployment rates of the less qualified cannot be explained by wage rigidities alone, at least unless one offers a compelling reason why such rigidities should be more pronounced in case of the less qualified. Therefore, an alternative explanation maintains that the welfare state itself induces high unemployment at the bottom of the income scale: In granting a basic allowance to every person in need, which becomes reduced upon acceptance of a job, the welfare state implicitly defines a certain reservation wage rate. If the reservation wage rate exceeds the productivity of the less qualified, these become unemployed even in a competitive labor market. In order to have a convenient short-hand, we may refer to this occurrence as induced unemployment.

Pure theorists do not find induced unemployment particularly alarming. This is because Mirrlees (1971) pointed out early that it can be socially optimal to have an interval of less productive persons who do not work. For this sort of voluntary unemployment, Seade (1977, p. 215) coined the term “bunching at the bottom”. The economic intuition of why bunching at the bottom may be optimal is straightforward: If the government cannot observe people’s inherent productivities, which equal the wage rates in a competitive equilibrium, but can only observe realized incomes, it must make sure that the more productive have no incentive to mimic the less productive by reducing their effort. A suitable device for preventing the rich from mimicking the poor are high marginal tax rates. Therefore, optimal tax-transfer-schemes typically have high marginal tax rates at the bottom of the income scale (Tuomala 1990). Distorting a person’s labor-leisure-choice using high marginal tax rates effectively means reducing his consumption and income along an indifference curve. Unless the indifference curves become horizontal in the neighborhood of zero effort, optimality may require a negative income which, however, is impossible because of the non-negativity constraint. In such a case, several persons with different productivities are bunched together at the corner solution of zero income.

Many applied economists, not to mention the public, seem to disagree strongly with such a kind of reasoning. There is a widely held belief that “all who are able to work, should work” and that basic allowances should be confined to the truly needy. In this spirit, Chambers (1989), Cuff (2000) and in particular Besley and Coate (1995) considered the efficacy of workfare as an additional instrument, and Diamond (1980) as well as Saez (2002) analysed negative marginal tax rates at the bottom of the income scale. However, Besley and Coate showed workfare to be non-optimal if the government has a welfare objective. Saez used a model where all persons have the same productivity but different preferences. His results do not carry over to the Mirrlees framework, where optimal marginal tax rates are always non-negative, even if the labour-leisure choice is a pure participation decision (Homburg 2002).

The present paper proceeds from the postulate that “all who are able to work, should work” which in itself presupposes a government’s capability to detect the disabled. Hence, we assume *partial information* in the sense that the government can distinguish the disabled from

the productive, but cannot distinguish among the different productive types. This assumption deviates both from the usual second-best approach, where the government has no information whatsoever about the types, and from the first-best approach, where the government has full information. It appears important to trace out the analytical consequences of this “first and a half-best” setting because several policy recommendations proceed implicitly from the premise that the government in fact has partial information.

A good example is the report of the Academic Advisory Council at the German Federal Ministry of Economics and Labor (2002). The Council suggested i) leaving the basic allowance (*Sozialhilfe*) unchanged in case of disabled persons, ii) cutting it substantially in case of productive persons and iii) lowering the implicit marginal tax rates on low incomes in order to induce less productive types to join the labor force. In sum, persons with zero income are treated differently, depending on whether or not they are disabled, and marginal tax rates on low incomes are reduced. In the following section we investigate whether such a policy is indeed optimal under partial information.

2. THE MODEL

We consider the finite variant of the standard optimal taxation model. There are several types $h = 0 \dots H$ ($H > 1$), whose exogenous wage rates and fractions are denoted as w^h and f^h , respectively. We assume $0 = w^0 < w^1 < \dots < w^H$, so that type 0 persons are completely disabled whereas all others are productive. A person consuming c^h and earning labor income y^h enjoys utility $u(c^h, y^h/w^h)$, where y^h/w^h represents effort. The utility function is strictly increasing in consumption, strictly decreasing in effort and strictly concave. Moreover, we assume for simplicity that its cross derivative vanishes and that all consumptions are strictly positive at the optimum. The government's objective reads:

$$(1) \quad \begin{aligned} \max_{(c^h, y^h)_{h=0 \dots H}}! \quad EU &= \sum_{h=0}^H u(c^h, \frac{y^h}{w^h}) f^h \\ \text{s.t.} \quad i) \quad &\sum_{h=0}^H (y^h - c^h - g) f^h = 0, \\ \quad \quad \quad ii) \quad &u(c^k, \frac{y^k}{w^k}) \geq u(c^h, \frac{y^h}{w^k}) \quad \text{for all } k, h > 0. \end{aligned}$$

Thus the government maximizes expected utility of a person choosing a tax-transfer-scheme from behind a veil of ignorance, subject to the budget constraint i), where g represents an exogenous revenue requirement, and subject to the self-selection constraints ii) which ensure that no person can make himself better off by mimicking somebody else. The difference between the standard model and the present one is that the self-selection constraints refer to types $k, h > 0$ only, rather than to all types. This means that the productive cannot mimic the disabled, and vice versa. Solutions of this problem induce truthful reporting of abilities. Optimal taxes or transfers follow implicitly as $T(y^h) = y^h - c^h$, and the discrete marginal tax rates are defined as

$$(2) \quad m^h = \frac{T(y^h) - T(y^{h-1})}{y^h - y^{h-1}}.$$

The only general result of the standard model which holds even without the single-crossing property states that the marginal tax rates are strictly less than one. This is easy to see: By the very definition of m^h , a marginal tax rate beyond one hundred per cent implies that $c^h - c^{h-1}$ and $y^h - y^{h-1}$ have opposite signs, but no person $h > 1$ would accept a pair containing less consumption and more income (effort) as compared to another. This, as well as the two following features, is also true under partial information:

$$(3) \quad y^h \geq y^{h-1} \quad \text{and} \quad c^h \geq c^{h-1} \quad \text{and} \quad u(c^h, \frac{y^h}{w^h}) = u(c^{h-1}, \frac{y^{h-1}}{w^h}) \quad \text{for all } h > 1.$$

The inequalities state that income and consumption increase weakly in productivity. Bunching at strictly positive incomes is possible but will be ruled out in what follows. The equality states that the downward adjacent self-selection constraints are binding at an optimum, which is due to the fact that the government wishes to redistribute from top to bottom. Together with the monotonicity, this ‘‘chain property’’ implies that all remaining self-selection constraints are automatically satisfied and hence can be neglected without loss of generality. It should be clear that the government’s budget constraint also holds with equality at an optimum.

As in the standard approach, we can now maximize the objective in (1) subject to the constraints (3). Comparing the two approaches immediately reveals that the use of partial information pays: Referring to type 1 persons as the *productive poor*, a constraint preventing the productive poor from mimicking the disabled is missing under partial information. As this constraint binds in the standard case, the outcome of the present model strictly dominates the second-best optimum. The social value of partial information has already been pointed out by Akerlof (1978). Yet, owing to the complexity of the Mirrlees model, neither he nor the literature following him (e.g. Immonen, Kanbur, Keen and Tuomala 1998) could derive general results. The only insights in the field stem from examples and simulations. Using a slightly different technique, we are able to identify a general property of optimal tax-transfer-schemes:

Proposition: Every optimum under partial information satisfies $c^0 > c^1$.

Proof: Since the government’s budget constraint and the downward adjacent self-selection constraints hold as equalities and are linearly independent (Homburg 2003), we can infer necessary first-order conditions using the Lagrangean

$$(4) \quad \mathcal{L} = \sum_{h=0}^H u(c^h, \frac{y^h}{w^h}) f^h + \lambda \sum_{h=0}^H (y^h - c^h - g) f^h + \sum_{h=2}^H \mu^h \left(u(c^h, \frac{y^h}{w^h}) - u(c^{h-1}, \frac{y^{h-1}}{w^h}) \right).$$

Differentiating with respect to consumption yields:

$$(5) \quad \frac{\partial \mathcal{L}}{\partial c^0} = \frac{\partial u}{\partial c^0} f^0 - \lambda f^0 = 0,$$

$$(6) \quad \frac{\partial \mathcal{L}}{\partial c^h} = \frac{\partial u}{\partial c^h} f^h - \lambda f^h + \mu^h \frac{\partial u}{\partial c^h} - \mu^{h+1} \frac{\partial u}{\partial c^h} = 0, \quad h > 0,$$

where $\mu^1 = \mu^{H+1} = 0$. Rearranging terms gives

$$(7) \quad f^0 = \frac{\lambda f^0}{\partial u / \partial c^0},$$

$$(8) \quad f^h + \mu^h - \mu^{h+1} = \frac{\lambda f^h}{\partial u / \partial c^h}, \quad h > 0.$$

Adding these equations over all h , the variables f^h sum up to one, the variables μ^h cancel out each other, and solving for λ shows that the shadow price of the budget constraint is the harmonic mean of the marginal utilities of consumption:

$$(9) \quad \lambda = \frac{1}{\sum_{h=0}^H \frac{f^h}{\partial u / \partial c^h}}.$$

According to (5), the marginal utility of the disabled equals λ . As consumption is monotonically increasing in h , the marginal utility of the productive poor exceeds λ , implying $c^0 > c^1$. ■

The proposition states that the disabled should have more consumption than the productive poor. This may appear surprising but has a simple explanation. Partial information virtually separates the population into two subsets of persons who cannot mimic one another. Given this segregation, resources should be used so as to equalize the marginal utilities of the persons belonging to the two subsets, implying that the marginal utility of the disabled should equal the average marginal utility of the productive. As an alternative explanation, increasing c^0 and c^1 marginally induces the same resource costs per person, but increasing c^1 entails an additional incentive cost, in that it tightens the self-selection constraint which prevents type 2 persons from mimicking the productive poor. Owing to the chain property, increasing c^1 makes it necessary to increase consumption, or to reduce effort, of all persons with higher productivities, so that the rich become better off at the expense of the poor.

Under the assumptions made, our proposition holds perfectly general, irrespective of whether or not the optimal policy succeeds in eliminating bunching at the bottom. But if it does, we obtain an even more irritating outcome.

Corollary: If the optimal policy succeeds in getting the productive poor to work, the marginal tax rate at the bottom exceeds one hundred per cent.

Proof: From the proposition, the productive poor have less consumption and, if working, they have higher income than the disabled, implying $m^1 > 1$. ■

This invalidates the received wisdom which says that marginal tax rates can never exceed one hundred per cent, and also runs counter to the notion that marginal tax rates should be low at the bottom of the income scale. While the standard model, as mentioned above, often entails high marginal tax rates at the bottom, these are now complemented by a tax rate even beyond one hundred per cent. However, the corollary only holds if the productive poor are on work. This is not necessarily the case because, contrary to a further widespread view, treating the disabled and the productive differently does not suffice to rule out bunching at the bottom. It may still be optimal to have some productive persons idle and then, following the usual logic, the marginal tax rate at the bottom will fall short of one hundred per cent.

3. AN EXAMPLE

Consider an economy with five types, a utility function $u(c^h, y^h/w^h) = \ln c^h + \ln(500 - y^h/w^h)$, a uniform productivity distribution, and per-capita revenue $g = 100$. Table 1 depicts the standard second-best optimum. Marginal tax rates are high at the bottom of the income scale and decrease monotonically. The “implicit” marginal rate – the difference between the marginal rate of substitution and the marginal rate of transformation – vanishes at the very top of the income scale, as is well known, but the discrete marginal rate defined in (2) is positive (and converges to zero as the income distribution grows dense). More importantly, there is induced unemployment since the productive poor are bunched together with the disabled at zero income.

w	c	y	m	u·f
0	543	0	--	2,50
2	543	0	--	2,50
4	728	509	64%	2,50
8	1433	2.226	59%	2,53
10	1994	3.006	28%	2,58
Σ				12,62

Table 1: Standard Optimum.

Table 2 shows the corresponding optimum under partial information. The government’s ability to distinguish between the disabled and the productive induces it to let the former enjoy average consumption (this is true only for log-linear utility functions). As a consequence, expected utility rises from 12,62 to 12,67, but the policy change does not present a Pareto-improvement: the utilities of the productive poor and of the middle classes have been diminished. Marginal tax rates are roughly as high as in the standard case, and have been supplemented by a horrific marginal tax rate of 516 per cent at the bottom of the income scale. The policy’s aim to get the productive to work has been reached in that induced unemployment does no longer exist.

w	c	y	m	u·f
0	1137	0	--	2,65
2	511	151	516%	2,46
4	706	661	62%	2,47
8	1394	2.309	58%	2,52
12	1937	3.063	28%	2,57
Σ				12,67

Table 2: Optimum under Partial Information.

The final Table 3 warns, however, that the last outcome does not hold generally. The table has been constructed using the same assumptions as above, the only difference being that the smallest wage rate has been reduced from two to one. Again, the productive poor are bunched together with the disabled at income zero and obtain less consumption, because their consumption, in the sense pointed out above, is more costly. Therefore, even if the government can separate the disabled from the productive, pooling some of them at zero income, rather than separating them with respect to working hours, may still be optimal. It is no accident that the marginal tax rate at the bottom falls short of one hundred per cent: Whenever there exists an index $\hat{h} > 1$ such that $y^h = 0$ for all $h < \hat{h}$ and $y^h > 0$ for all $h \geq \hat{h}$, the marginal tax rate at the smallest positive income is given by $m^{\hat{h}}$ and is below one hundred per cent, since type \hat{h} must be prevented from mimicking his left-hand neighbor.

w	c	y	m	u·f
0	1113	0	--	2,65
1	463	0	--	2,47
4	693	665	65%	2,47
8	1377	2.322	59%	2,51
10	1920	3.080	28%	2,56
Σ				12,66

Table 3: Bunching under Partial Information.

4. CONCLUSION AND EVALUATION

This paper has analyzed the structure of optimal tax-transfer-systems if the government has partial information about people's abilities. The upshot is that an optimal scheme grants more consumption to the disabled than to the productive poor. This finding may explain why persons with certain characteristics signaling disability – such as age, illness, or dependent children – can expect a better treatment in many welfare states as compared to unemployed persons without such characteristics. The intuition behind the result is not that idle productive are more expensive than idle disabled persons. Rather, increasing consumption of the productive entails an incentive cost which is absent when increasing consumption of the disabled.

Our analysis reinforces the case for high marginal tax rates at the bottom of the income scale which, under partial information, may even exceed one hundred per cent. On the other hand, bunching at the bottom can still be optimal. Hence, the model sustains the policy recommendation cited in the introduction, as far as it concerns lower transfers to the productive poor; but it does not support the further recommendation which aims at eliminating bunching at the bottom using low marginal tax rates.

In evaluating the model, three caveats are in order. Firstly, we have assumed that the disabled can be distinguished from the productive perfectly and costlessly. If the distinction involves administrative costs, other aspects become important which have been analyzed by Boadway, Marceau and Sato (1999). Secondly, our discrete setting conveys a clear-cut conceptual dis-

inction between the disabled and the productive. Considering a limiting process, where the productivity distribution becomes dense, will obscure the borderline because, then, persons with an arbitrarily low productivity are likely to exist in the right-hand neighborhood of the disabled. The distinction between the disabled and the productive is no more a positive issue but a normative one; the government must decide which persons qualify as disabled. Nevertheless, our above analysis can easily be generalized to cover this case: Suppose the government fixes a threshold \hat{h} such that all types with productivities below $w^{\hat{h}}$, which can be identified by hypothesis, qualify as disabled. At the optimum, these types obtain the common bundle $(c^0, 0)$, and type \hat{h} obtains $(c^{\hat{h}}, y^{\hat{h}})$, where $c^{\hat{h}} < c^0$. Again, the marginal tax rate at the bottom of the income distribution will be strictly positive and will exceed one hundred per cent if and only if there is no bunching at the bottom.

Thirdly, and most importantly, exogenous productivities have been assumed. This premise seems innocuous in the standard model – which can be extended to include educational decisions – but not in case of partial information. Given the harsh treatment of the productive poor, these may be apt to become disabled. Strategies as drinking or drug addiction come into mind. Perhaps in a generalized model, which takes account of endogenous productivities, a lower tax pressure at the bottom would be optimal.

REFERENCES

- Academic Advisory Council (2002) *Reform des Sozialstaats für mehr Beschäftigung im Bereich gering qualifizierter Arbeit*. Bundesministerium für Wirtschaft, Dokumentation Nr. 512.
- Akerlof, G. A. (1978) The Economics of “Tagging” as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning, *American Economic Review* **68**, pp. 8–19.
- Besley, T. und St. Coate (1995) The Design of Income Maintenance Programmes. *Review of Economic Studies* **62**, pp. 187–221.
- Boadway, R., N. Marceau and M. Sato (1999) Agency and the design of welfare systems. *Journal of Public Economics* **73**, S. 1–30.
- Cuff, K. (2000) Optimality of Workfare with Heterogeneous Preferences. *Canadian Journal of Economics* **33**, pp. 149–174.
- Chambers, R. G. (1989) Workfare or Welfare? *Journal of Public Economics* **40**, pp. 79–97.
- Diamond, P. A. (1980) Income Taxation with Fixed Hours of Work. *Journal of Public Economics* **13**, pp. 101–110.
- Homburg, St. (2002) Optimal Marginal Tax Rates for Low Incomes: Positive, Negative, or Zero? Discussion Paper No. 255, University of Hannover.
- Homburg, St. (2003) An Axiomatic Proof of Mirrlees’ Formula. *Public Finance/Finances Publiques* **53** (3-4), 1998/2003, pp. 285–295.
- Immonen, R., R. Kanbur, M. Keen M and M. Tuomala (1998) Tagging and taxing: The Optimal Use of Categorical and Income Information in Designing Tax/transfer Schemes. *Economica* **65**, pp. 179–192.

Mirrlees, J. A. (1971) An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies* **38**, pp. 175–208.

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

Seade, J. K. (1977) On the Shape of Optimal Tax Schedules. *Journal of Public Economics* **7**, pp. 203–235.

Tuomala, M. (1990) *Optimal Income Tax and Redistribution*. Oxford: Clarendon Press.