Using a Microeconometric Model of Household Labour Supply to Design Optimal Income Taxes

Published version:

DOI:10.1111/sjoe.12015

(Article begins on next page)

Using a Microeconometric Model of Household Labour Supply to Design Optimal Income Taxes*

Rolf Aaberge° and Ugo Colombino#

Abstract

Empirical applications of optimal taxation theory have typically adopted analytical expressions for the optimal taxes and then imputed numerical values to their parameters by calibration or using previous estimates. We aim at avoiding the restrictive assumptions and the possible inconsistencies of that approach. By contrast, we identify optimal taxes by iteratively running a microeconometric model based on 1994 Norwegian data until a given social welfare function is maximized given the public budget constraint. The optimal rules envisage monotonically increasing marginal rates (negative on very low incomes) and – compared to the current rule – a lower average rate, lower marginal rates on low incomes and higher marginal rates on very high incomes.

Keywords: Optimal taxation, Microsimulation, Random Utility Model.

JEL classification: H21, H31, J22.

* We thank Tom Wennemo for skilful programming assistance and a referee for very detailed and useful comments and suggestions. Ugo Colombino received financial support from the Italian Ministry of University and Research and the Compagnia di San Paolo. ICER is gratefully acknowledged for providing financial support and excellent working conditions for Rolf Aaberge.

°Rolf Aaberge, Research Department, Statistics Norway, Oslo, Norway, rolf.aaberge@ssb.no.

#Ugo Colombino, Department of Economics "Cognetti De Martiis", University of Torino, Italy, ugo.colombino@unito.it

I. Introduction

This paper presents an empirical analysis of optimal income taxation. The purpose is not new, but the exercise illustrated here differs in many important ways from previous attempts to empirically compute optimal taxes. The standard procedure adopted in the literature starts with some version of the optimal taxation framework originally set up in the seminal paper by Mirrlees (1971). The next step typically consists of imputing numerical values – either determined by calibration or taken from previous empirical analysis – to the parameters (e.g. wage elasticities of labour supply) appearing in the formulas produced by the theory. This literature is surveyed by Tuomala (1990). A recent strand of research adopts a similar approach to address the inverse optimal taxation problem, i.e. retrieving the social welfare function that makes optimal a given tax rule (Bourguignon and Spadaro, 2005). There are two main problems with the optimal taxation literature: 1) The theoretical results become amenable to an operational interpretation only by adopting rather restrictive assumptions concerning the preferences, the composition of the population and the structure of the tax rule; 2) The empirical measures used as counterparts of the theoretical concepts are usually derived from previous estimates obtained under assumptions different from those used in the theoretical model. As a consequence the consistency between the theoretical model and the empirical measures is dubious and the significance of the numerical results remains uncertain. The typical outcome of these exercises envisages a lump-sum transfer which is progressively taxed away by very high marginal tax rates (MTRs) on lower incomes (i.e. a negative income tax mechanism). Beyond the "break-even point" (i.e. the income level where the transfer is completely exhausted), the MTRs are close to constant. Tuomala (2010) suggests, however, that these results are essentially forced by the restrictive assumptions made upon preferences, labour supply elasticities and distribution of productivities (or wage rates). Interestingly, when Tuomala (2010) adopts a more flexible specification of the utility function he finds that the optimal system is progressive with monotonically increasing MTRs.

While most of the studies mentioned above were essentially illustrative numerical exercises, several recent contributions have attempted to use optimal taxation results in the empirical evaluation or design of tax-transfer reforms. Diamond (1980, 1998), Revesz (1989) and Saez (2001) make Mirrlees's results more easily interpretable by reformulating them in terms of labour (or income) supply elasticities in order to provide a direct link between theoretical results and empirical measures. Saez (2002) develops a model amenable to empirical implementation that focuses on the relative magnitude of the labour supply elasticities at the extensive and intensive margin. Immervoll et al. (2007) adopt Saez's (2002) approach to evaluate alternative income support policies in European countries. Blundell et al. (2009) and Haan and Wrohlich (2010) also use Saez (2002) to evaluate taxes and transfers for lone mothers in Germany and UK, whereas Kleven et al. (2009) provide results on the taxation of couples. Although these new contributions are interesting attempts to advance towards the empirical implementation of theoretical optimal taxation results, they still rely on restrictive assumptions and suffer from a possible inconsistency between the theoretical model and the empirical measures used to implement it. For example, the basic version (adopted in the empirical exercises mentioned above) of the model proposed by Saez (2002) does not account for income effects[1] and moreover relies on rather restrictive assumptions upon the way the households respond to changes in the relative attractiveness of the opportunities in the budget set.[2] When it comes to empirical applications (as in Immervoll et al. (2007), Blundell et al. (2009) and Haan and Wrohlich (2010)), the parameters of the theoretical models are given numerical values estimated with microeconometric models that do not adopt the same restrictive assumptions as Saez (2002). Of course, some of those limitations and potential inconsistencies might be overcome in the future, but analytical solutions of the optimal income taxation problem will likely never be fully consistent with flexible structural labour supply models. We follow here a different and possibly complementary approach. We do not start from theoretical results dictating conditions for optimal tax rules under various assumptions.

---

[1] Income effects can be accounted for, as in Saez (2001), at the cost of notable analytical and computational complications.

[2] In Saez (2002) each individual has only three opportunities to choose from: non-participation and two adjacent labour income brackets.

Instead we use a microeconometric model of labour supply in order to identify by simulation the tax rule that maximizes a social welfare function under the constraints that the households maximize their own utility and the total net tax revenue remains constant. The microeconometric simulation approach is common in evaluating tax reforms, but has not been much used in empirical optimal taxation studies. The closest examples adopting a similar approach are represented by Fortin et al. (1993), Aaberge and Colombino (2006), Colombino et al. (2010) and Blundell and Shephard (2011).[3] Analytical solutions are still crucial for understanding the “grammar” of the problem and for suggesting promising directions of reform. By contrast, microeconometric models and computational solutions allow for the introduction of less restrictive specifications of preferences and opportunity sets and the evaluation of more complex tax-transfer rules. The estimated model we use here represents the choices of both couples and singles, it adopts a flexible specification of the preferences, it accounts for quantity constraints in labour supply choices and it can accommodate a detailed representation of complex tax-transfer systems. The optimal tax rules turn out to envisage an average tax rate lower than the current one, a modest lump sum tax (interpretable as a property tax), a negative tax on low incomes (close to mechanisms such as the Earned Income Tax Credit or the In-Work Benefit policies) and a progressive MTR profile culminating to a 100 per cent MRT on very high incomes (about 1.5 per cent of the sample). This scenario contrasts sharply with respect to the results obtained by the numerical exercises inspired by the seminal contribution of Mirrlees (1971). It is closer to the picture that typically comes out of empirical applications adopting the theoretical results of Saez (2002). However, using a flexible microeconometric model as a computational tool, we are able to explore a larger variety of tax-transfer rules and to perform a more articulated analysis of the effects of the various rules upon different segments of the population. Obviously, the results of our computational exercise cannot claim similar generality as the analytical solutions. While the latter

[3] Fortin et al. (1993) use a calibrated (not estimated) model with rather restrictive (Stone-Geary) preferences and focus on alternative income support schemes rather than on the whole tax rule. Aaberge and Colombino (2006) report on preliminary results of a simpler version of the exercise illustrated in this paper. Colombino et al. (2010) analyse basic income support mechanisms in some European countries. Blundell and Shephard (2011) identify the optimal design of a specific UK policy addressed to low income families with children. They do not treat the problem of interpersonal comparability, which, however, in their case might be less important given the smaller and less heterogeneous target population.

establish an explicit relationship between the fundamentals of the economy (preferences, skill distribution etc.), the former are application-specific (in this paper, Norway-specific): this is the price of accounting for a more detailed and flexible representation of preferences and opportunity sets. In principle, however, this limitation of our computational exercise could be overcome. By performing similar exercises on many different economies, one should again be able to identify – empirically – a more general relationship between the fundamentals of the economy and the optimal income tax rule. As explained in Section II, the microeconometric model used in this study contains 78 parameters capturing the heterogeneity of preferences and opportunities among households and individuals. The estimated model is used to simulate the choices given a particular tax rule. Those choices are therefore generated by preferences and opportunities that vary across the decision units. However, since preferences are heterogeneous and some individuals live as singles whereas others form families and live together, when it comes to social evaluation it does not make sense to treat the estimated utility functions as comparable individual welfare functions. To solve the interpersonal comparability problem we adopt a method that consists of using a common utility function in order to produce interpersonally comparable individual welfare measures to be used as arguments of the social welfare function. The common utility function is justified as a normative standard where the social planner treats individuals symmetrically and it is only used to compute and compare the individual welfare levels that provide the basis for the social welfare evaluation of tax reforms; it is not used for simulating household behaviour (where instead the estimated individual utility functions are used). This procedure, which circumvents the problem of interpersonal comparability of heterogeneous preferences, is well-established in the empirical public economics literature. It is proposed in Deaton and Muellbauer (1980) and in Hammond (1991), and it forms the basis for the definition and measurement of a money-metric measure of utility in King (1983) and in Aaberge et al. (2004). Moreover, it has been applied for example by Fortin et al. (1993), Colombino (1998) and Colombino et al. (2010). As a practical matter, an average of the estimated individual utility functions or an estimated utility function (individual welfare function) with common parameters (as in our case) is typically used.

A brief description of the microeconometric model is presented in Section II, while the empirical specification of the model and the estimates of the model parameters are provided by Aaberge and Colombino (2011). In order to illustrate the behavioural implications of the estimates, Section II reports wage elasticities of labour supply, whereas income elasticities are presented by Aaberge and Colombino (2011). Since the microeconometric model, once estimated, is used for a rather ambitious purpose – simulating choices in view of identifying optimal tax rules – it is important to check its reliability: ultimately, the model should be judged on its ability to do the job it is built for, i.e. predicting the outcomes of policy changes. In Section II we therefore perform an out-of-sample prediction exercise where we use the model (estimated on 1994 data) to predict household-specific distributions of income in Norway in 2001. In Section III we present the social welfare evaluation method and the computational procedure for solving the optimal taxation problem. The resulting optimal rules are presented in Section IV. Section V contains the final comments.

II. The modelling framework

The microeconometric labour supply model

In this section we present a sketch of the microeconometric model. A full description is given in Aaberge and Colombino (2011). The model can be considered as an extension of the standard multinomial logit model, and differs from the traditional models of labour supply in several respects.[4] First, it accounts for observed as well as unobserved heterogeneity in preferences and constraints. Second, it includes both single person households and married or cohabiting couples making joint labour supply decisions. A proper model of the interaction between spouses in their labour supply decisions is important as most of the individuals are married or cohabiting. Third, by taking all the details of the tax system into account, the budget sets become complex and non-convex in certain intervals. For expository simplicity we consider in this section only the behaviour of a single person household. The extension to couples is fully explained in Aaberge and Colombino (2011). The agents

[4] Examples of previous applications of this approach are found in Aaberge et al. (1995, 1999, 2000, 2004).

choose a job from alternatives characterized by the wage rate *w,* hours of work *h* and other characteristics *j*. The problem solved by the agent is the following:

$$\begin{aligned} &\max_{(w,h,s,j)\in B} U\left(c,h,s,j\right) \\ &\text{s.t.} \\ &c = f(wh,I) \end{aligned} \tag{2.1}$$

where

h = hours of work,

w = the pre-tax wage rate,

s = observed job characteristics (besides h and w, e.g. occupational sector),

j = unobserved (by the analyst) job and/or household characteristics,

I = the pre-tax non-labour income (exogenous),

c = net disposable income,

f = tax rule that transforms gross pre-tax incomes (wh,I) into net disposable income c,

B = the set of all opportunities available to the household (including non-market opportunities, or "leisure" activities, i.e. a "jobs" with $w = 0$ and $h = 0$).

Agents can differ not only in their preferences and in their wage (as in the traditional model) but also in the number of available jobs of different types. Moreover, for the same agent, wage rates (unlike in the traditional model) can differ from job to job. Let $p(h,w,s)$ denote the density of available jobs of type (h,w,s). By representing the choice set B by a probability density p we can, for example, allow for the fact that jobs with hours of work in a certain range are more or less likely to be found, possibly depending on agents' characteristics; or for the fact that for different agents the relative number of market opportunities may differ. We assume that the utility function can be factorised as

$$U\left(f(wh,I),h,s,j\right) = v\left(f(wh,I),h,s\right)\varepsilon(j) \tag{2.2}$$

where v and $\varepsilon(j)$ are respectively the systematic and the random component. The term $\varepsilon(j)$ is a random variable that accounts for the effect on utility of all the characteristics of the household–job match that are observed by the household but not by us. Moreover, we assume that $\varepsilon(j)$ is i.i.d. according to the Type III Extreme Value distribution. We observe the chosen *h*, *w* and *s*. We can therefore specify the probability that the agent chooses a job with observed characteristics (*h,w,s*). It can be shown that, under the assumptions (2.1) – (2.2) and given the extreme value distribution for *ε*, we can write the probability density function of a choice (*h*,*w,s*) as[5]

$$\varphi(h,w,s)=\frac{v(f(wh,I),h,s)p(h,w,s)}{\iiint v(f(xy,I),y,z)p(x,y,z)dxdydz}. \qquad (2.3)$$

The density (2.3) is the contribution of an observation (*h, w, s*) to the likelihood function, which is then maximized in order to estimate the parameters of $v(f(hw,I),h,s)$ and $p(h,w,s)$. The intuition behind expression (2.3) is that the probability of a choice (*h*,*w*,*s*) can be expressed as the relative attractiveness – weighted by a measure of "availability" $p(h,w,s)$ – of jobs of type (*h, w, s*). Given convenient parametric specifications of the functions *v* and *p*, the 78 parameters of the model can be estimated by maximizing the likelihood function formed on the basis of expression (2.3). The estimation is based on 1994 data collected by the 1995 Norwegian Survey of Level of Living, which includes detailed income data from tax reported records.[6] We have restricted the ages of the individuals to between 20 and 62 in order to minimize the inclusion in the sample of individuals who in principle are eligible for retirement, since analysis of retirement decisions is beyond the scope of this study. Moreover, self-employed as well as individuals receiving permanent disability benefits are excluded from the sample. The sample contains 1,842 couples, 309 single females and 312 single males. The estimates are reported in Aaberge and Colombino (2011).

When interpreting the random utility model illustrated above it is important to stress that in the model household members choose from a set of jobs (characterized by *h*, *w* and other characteristics *s* and *j*), not just from jobs that differ in hours of work *h*. Households' responses

[5] For the derivation of the choice density (2.3), see Aaberge et al. (1999). Note that (2.3) can be considered as a special case of the more general framework developed by Dagsvik (1994). A more specialized type of continuous multinomial logit was introduced by Ben-Akiva and Watanatada (1981).

[6] At the time of performing the exercise presented in this paper, the 1994 sample was chosen due to the relatively stable macroeconomic conditions.

therefore include many dimensions: hours, wage rates and non-pecuniary job characteristics. Theoretical optimal taxation models typically consider effort as the agents' choice variable. Effort does not coincide with hours of work; it might include searching for jobs of better quality, putting more effort into each hour of work or even configuring reported incomes in a more favourable way in view of taxation. A related concept – taxable income – has been used, among others, by Feldstein (1995) and Gruber and Saez (2002). The idea is that in evaluating the effects of changes in taxes, one should not just look at hours of work (and participation), since households' responses include many other dimensions. At least some of these multi-dimensional responses are reflected in taxable income. However, structural empirical models of labour supply used for tax reform evaluations have traditionally considered hours of work as the sole choice variable, implicitly equating hours of work and effort. An exception is provided by Bourguignon and Spadaro (2005), who under rather special assumptions are able to impute to each agent an effort value. Our model does not strictly equate effort to hours of work, since households – as a response to a change in the tax system – might choose a new job that differs from the previous one not only with respect to hours of work but also wage rate and other job characteristics. However, while we account for the disutility of hours of work and choice of sector, we only implicitly account (through the random utility component) for the fact that the other dimensions of effort may also bear a utility cost. Therefore, we cannot claim that our model is completely consistent with the "effort-taxable income" approach. We return to this limitation in the Conclusions (Section V).

Behavioural implications

In this section we illustrate some of the behavioural implication of the estimates. First, we report the wage elasticities of labour supply because they are useful for understanding and interpreting the optimal taxation results to be presented in Section IV. Second, since the model will be used for a rather ambitious operation (i.e. computing optimal tax-transfer rules) we illustrate the prediction performance of the model with an out-of-sample exercise.

The wage elasticities reported in Table 2.1 are computed by means of stochastic simulation. Note that the households face exogenous opportunity joint density functions of h, w and s. Since many individuals in this labour supply model of discrete choice will not react to small exogenous changes, the elasticities in Table 2.1 have been computed as an average of the percentage changes in labour supply from a 10 per cent increase in the means of the wage densities. Given the simulated responses of each individual, we aggregate them to compute the aggregate elasticities. We find that the overall wage elasticity is equal to 0.12, which suggests rather low behavioural responses. However, by looking behind the overall elasticity, the picture changes substantially. The major features of the estimated labour supply elasticities can be summarized as follows: (a) labour supply of married women is far more elastic than for married men; (b) individuals belonging to low-income households are much more elastic than individuals belonging to high-income households. As demonstrated, for example, by Røed and Strøm (2002) and by Meghir and Phillips (2008), these results are consistent with the findings in many recent studies. The profile of the wage elasticity across the income deciles is related to the hours worked. Households belonging to the higher income brackets on average participate more and work longer hours, which – other things being equal – tends to lower the wage elasticity of labour supply. Table 2.1 also reveals that cross-elasticities are relevant in many cases, which supports the importance of modelling joint household's decisions.

[Table 2.1]

In principle, elasticities such as those illustrated above might be used to compute optimal tax-transfer rules as is done for example in Saez (2001), Saez (2002), Immervol et al. (2007) and Blundell et al. (2009). As we explained in Section I, we do not think that this procedure is totally satisfactory, due to the possible inconsistencies between the assumptions adopted by the theoretical optimal taxation model and the assumptions adopted in producing the elasticities. For example, the empirical exercises such as those mentioned above typically adopt theoretical results that ignore cross-elasticities and income effects.[7] Microeconometric models are based on assumptions that are much more flexible and general than those leading to the theoretical results. The approach we adopt in this paper exploits this greater flexibility and guarantees consistency in the assumptions by obtaining the optimal tax-transfer

[7] Aaberge and Colombino (2006, 2011) report more detailed results on wage and income elasticities. Income effects are not negligible: their order of magnitude is close to the order of magnitude of wage elasticities and they vary substantially across gender, household type, location in the income distribution and composition of income.

rules computationally, i.e. we iteratively run the microeconometric model of household behaviour until the social welfare function is maximized under the constraint that the total net tax revenue remains constant.

For the out-of-sample prediction exercise we used the model estimated on the 1994 sample and the 2001 data (exogenous variables) from the 2002 Norwegian Survey of Level of Living, in order to predict the choices made in 2001 under the new 2001 tax rules.[8] Tables 2.2 and 2.3 describe some of the characteristics of the 1994 and 2001 tax regimes.

[Tables 2.2 and 2.3]

The basic features of the 1994 Norwegian tax system were determined by a major tax reform of 1992, which introduced a so-called dual income tax system characterized by a 28 per cent flat tax rate on capital income in combination with progressive tax rates on labour income plus 7.8 per cent social insurance contribution. Further measures broadened the tax base of business income substantially and removed various previous tax credits and deductions. In order to reduce incentives for taxpayers to classify labour income as capital income, the reform established rules for mandatory income splitting for dividing business income into capital and labour income, with the resulting imputed wage income taxed according to a two-bracket progressive surtax. The associated top marginal tax rates for wage earners and owners of small businesses were 49.5 per cent and 52.4 per cent in1994. Between1992 and 2001, marginal rates as well as the threshold of the highest bracket of the surtax increased, resulting in the statutory tax rates for 2001 shown in Table 2.3. Disposable income is the variable used for comparing predicted outcomes to observed outcomes. The predictions are obtained individual by individual, evaluating the utility function – including the stochastic component drawn from the Type III extreme value distribution – at each alternative and identifying the selected alternative as the one with the highest utility level. The individual predictions are then aggregated into the 10 means of the 10 income deciles. Table 2.4 reports the results of an out-of-sample prediction exercise. The model turns out to be rather successful in reproducing the income distribution.

---

[8] Both 1994 and 2001 data are characterized by relatively stable macroeconomic conditions.

[Tables 2.4]

III. The design of optimal income taxes

The framework of the social planner

The literature on optimal taxation relies on the maximization of social welfare functions defined as summary measures of the distribution of individual utility levels, where utility levels are assumed to be interpersonally comparable. In this section we explain the method adopted for allowing the aggregation of comparable individual welfare levels and the computational procedure used to solve the optimal taxation problem.

Individual welfare functions

Since the microeconomic labour supply model used in this study allows heterogeneous preferences for leisure and consumption and, moreover, some individuals live as singles and others as a couple, it makes no sense to treat the estimated utility functions as comparable individual welfare functions. Thus, it is necessary to introduce measures of individual welfare that permit interpersonal comparisons.[9] Here we follow an approach similar to the ones advocated by Deaton and Muellbauer (1980), King (1983) and Hammond (1991). Specifically, we use a common utility function to evaluate the bundles chosen by the households according to their own preferences.[10] The common utility function (or individual welfare function) is to be interpreted just as the input of a social welfare function. It is not used to simulate behaviour, only to evaluate – in a comparable way – the results of choices made according to the actual individual utility functions.

The individual welfare function V used by the social planner is specified as follows:

$$\log V(y,h) = \gamma_2 \left(\frac{y^{\gamma_1} - 1}{\gamma_1} \right) + \gamma_4 \left(\frac{L^{\gamma_3} - 1}{\gamma_3} \right) \qquad (3.1)$$

[9] See Boadway et al. (2002) and Fleurbaey and Maniquet (2006) for a discussion of interpersonal comparability of utility when preferences for leisure differ between individuals.

[10] An alternative approach has been recently proposed and illustrated by Decoster and Haan (2010).

where L is leisure, defined as $L = 1 - (h/8736)$, and y is the equivalent individual's income after tax defined by

$$y = \begin{cases} c = f(wh, I) & \text{for singles} \\ \frac{c}{\sqrt{2}} = \frac{1}{\sqrt{2}} f(w_F h_F, w_M h_M, I) & \text{for married/cohab. individuals.} \end{cases} \tag{3.2}$$

The Box-Cox functional form of expression (3.1) is the same as that adopted for specifying the utility functions in the microeconometric model (Aaberge and Colombino, 2006, 2011). By dividing the couple income by $\sqrt{2}$ we transform incomes of couples into comparable single individual incomes.[11] In order to estimate the parameters of the individual welfare function (3.1) we use expression (2.3) with the systematic part of the utility function (v) replaced by the individual welfare function (V) and conditional on the estimated opportunity densities p. The sample is the same as that used for estimating the model of Section II. Table 3.1 displays the parameter estimates of V.

[Table 3.1]

A different way to circumvent the interpersonal comparability problem consists in avoiding interpersonal welfare level comparisons altogether and basing the social evaluation exclusively on ordinal comparisons. We provide an example of this method in Table 4.5, where we present the number of "winners" under the optimal tax rules.[12]

Social Welfare Functions

When evaluating the distribution of individual welfare effects of a tax system and/or a tax reform it is necessary to summarize the gains and losses by a social welfare function. The simplest welfare function is the one that adds up the comparable welfare gains over individuals. The objection to the linear additive welfare function is that the individuals are given equal welfare weights, irrespective of whether they are poor or rich. Concern for distributive justice requires, however, that poor individuals

[11] The "square root scale" is one of the equivalence scales commonly used in OECD publications. The number of household members, including children, is taken into account in the specification of the utility function, where it affects the marginal utilities of income and leisure (Aaberge and Colombino 2006, 2011).

[12] This is just an illustration; a proper application of the ordinal criterion would require defining the optimal tax in a different way, for example the tax rule that maximizes the number of winners.

are assigned larger welfare weights than rich individuals. This structure is captured by the family of rank-dependent welfare functions,[13]

$$W = \int_0^1 q(t)F^{-1}(t)dt \tag{3.3}$$

where F^{-1} is the left inverse of the cumulative distribution function of the individual welfare levels V with mean μ, and $q(t)$ is a positive weight-function defined on the unit interval. The social welfare functions (3.3) can be given a similar normative justification as is underlying the "expected utility" social welfare functions introduced by Atkinson (1970). Given suitable continuity and dominance assumptions for the preference ordering $\succeq$ defined on the family of distributions ***F***, Yaari (1987, 1988) demonstrated that the following axiom,

Axiom (Dual independence). *Let F_1, F_2 and F_3 be members of **F** and let $\alpha \in [0,1]$ Then $F_1 \succeq F_2$ implies*

$$\left(\alpha F_1^{-1} + (1-\alpha)F_3^{-1}\right)^{-1} \succeq \left(\alpha F_2^{-1} + (1-\alpha)F_3^{-1}\right)^{-1},$$

characterizes the family of rank-dependent measures of social welfare functions (4.3) where $q(t)$ is a positive non-decreasing function of *t*. We refer to Yaari (1987, 1988) for a discussion of the difference between the dual independence axiom and the conventional independence axiom used to justify the "expected utility" social welfare functions. In this paper we use the following specification of $q(t)$,

$$q_k(t) = \begin{cases} -\log t, & k = 1 \\ \frac{k}{k-1}\left(1 - t^{k-1}\right), & k = 2,3,.... \end{cases} \tag{3.4}$$

Note that the inequality aversion exhibited by the social welfare function W_k (associated with $q_k(t)$) decreases with increasing *k*. As $k \to \infty$, W_k approaches inequality neutrality and coincides with the linear additive welfare function defined by

---

[13] Several other authors have discussed the rationale for rank-dependent measures of inequality and social welfare, see e.g. Sen (1974), Hey and Lambert (1980), Donaldson and Weymark (1980, 1983), Weymark (1981), Ben Porath and Gilboa (1992) and Aaberge (2001).

$$W_\infty = \int_0^1 F^{-1}(t)dt = \mu \ . \qquad (3.5)$$

It follows by straightforward calculations that $W_k \leq \mu$ for all *k* and that W_k is equal to the mean μ for finite *k* if and only if *F* is the egalitarian distribution. Thus, W_k can be interpreted as the equally distributed individual welfare level. As recognized by Yaari (1988) this property suggests that C_k, defined by

$$C_k = 1 - \frac{W_k}{\mu}, \ k = 1, 2, ... \qquad (3.6)$$

can be used as a summary measure of inequality.[14] As noted by Aaberge (2000, 2007), C_1 is actually equivalent to a measure of inequality that was proposed by Bonferroni (1930), whilst C_2 is the Gini coefficient. Aaberge (2000, 2007) demonstrates that C_1 exhibits strong downside inequality aversion and is particularly sensitive to changes that concern the poor part of the population, whilst C_2 normally pays more attention to changes that take place in the middle part of the income distribution. The C_3 coefficient exhibits upside inequality aversion and is thus particularly sensitive to changes that occur in the upper part of the income distribution. Due to the close relationship between C_1, C_2 and C_3 Aaberge (2007) proposed treating them as a group and calling them Gini's Nuclear Family of inequality measures. In order to ease the interpretation of the inequality aversion profiles exhibited by W_1, W_2, W_3 and W_∞, Table 3.2 provides the ratios of the corresponding weights – defined by (3.4) – of the median individual and the 1 per cent poorest, the 5 per cent poorest, the 30 per cent poorest and the 5 per cent richest individual for different social welfare criteria. As can be observed from the weight profiles provided by Table 3.2, W_1 will be particular sensitive to changes in policies that affect the welfare of the poor, whereas the inequality aversion profile of W_3 is rather moderate and W_∞ exhibits neutrality with respect to inequality.

[14] Note that Aaberge (2001) provides an axiomatic justification for using the C_k – measures as criteria for ranking Lorenz curves. Thus, the justification of the social welfare function $W_k = \mu(1 - C_k)$ defined by (3.3) (and (3.6)) can also be made in terms of a value judgement of the trade-off between the mean and (in)equality in the distribution of welfare.

[Table 3.2]

The Optimal Taxation Problem

We strictly consider only personal income taxation. Following the tradition of the optimal income tax literature, all the other dimensions of the wider tax system (VAT, consumption taxes, payroll taxes, social assistance etc.) are kept constant as of 1994 in Norway. The optimal taxation problem considered in this exercise can be formulated as follows:

$$
\begin{aligned}
&\max_{\vartheta} W\left(V\left(y_{1F},h_{1F}\right),V\left(y_{2F},h_{2F}\right),...,V\left(y_{nF},h_{nF}\right),V\left(y_{1M},h_{1F}\right),V\left(y_{2M},h_{2M}\right),...,V\left(y_{nM},h_{nM}\right)\right)\\
&\text{s.t.}\\
&y_{iF}=y_{iM}=c_i\big/\sqrt{2},\ i=1,2,...,n\\
&\sum_{i=1}^{n}\left(w_{iF}h_{iF}+w_{iM}h_{iM}+I_i-f(w_{iF}h_{iF},w_{iM}h_{iM},I_i;\vartheta)\right)\geq G\\
&\left(c_i,h_{iF},h_{iM},s_{iF},s_{iM},j_i\right)=\arg\max_{(w,h,s,j)\in B_i} U_i\left(c,h_F,h_M,s_F,s_M,j\right)\text{s.t. } c_i=f(w_{iF}h_{iF},w_{iM}h_{iM},I_i;\vartheta),\ i=1,2,...,n.
\end{aligned}
\tag{3.7}
$$

where n is the number of households and G denotes the total net tax revenue required (set equal to the current one in our exercise). For simplicity of exposition, expression (3.7) assumes that the n households are couples, while in fact we consider both couples and singles. In (3.7) each couple contributes to the social welfare functions with two terms corresponding to the individual welfare functions of the two partners. For singles, we have just one term $V\left(y_i,h_i\right)$ and $y_i=c_i$ (according to expression (3.2)). All the variables are the same as those appearing in expression (2.1) in Section II. The function $c_i=f(w_{iF}h_{iF},w_{iM}h_{iM}I_i;\vartheta)$, which transforms gross incomes $(w_{iF}h_{iF},w_{iM}h_{iM},I_i)$ into net available income c_i, denotes a class of tax rules defined up to a vector of parameters ϑ. We will consider a class of piecewise-linear tax rules with a (positive or negative) lump-sum transfer and five income brackets. The parameters will therefore be the amount of the lump-sum transfer, the lower and upper limits of the income brackets and the MTR applied to the income brackets. Household i maximizes its own utility given the tax rule $f(w_{iF}h_{iF},w_{iM}h_{iM}I_i;\vartheta)$ by choosing the "job" $\left(c_i,h_{iF},h_{iM},s_{iF},s_{iM},j_i\right)$. Taking the individual utility-maximizing choices into account as a constraint (i.e. the incentive-compatibility constraint), the social planner searches for the tax rule – i.e.

the parameter vector ϑ – that maximizes the social welfare function *W*, subject to the constraint that the total net tax revenue must at least be as large as *G*. The social welfare function *W* takes as arguments the evaluations of the chosen "jobs" according to the individual welfare function *V*. Given the very flexible and general specifications adopted for the random utility functions and of the opportunity sets (Aaberge and Colombino, 2011), problem (3.7) cannot be solved analytically. The maximization of *W* is performed by a global maximization procedure that efficiently scans the parameter space. At each run of the iterative procedure, the maximization of the individual utility function is simulated by the microeconometric model described in Section II. The search for the optimal tax rule is limited to the class of piecewise-linear rules, with five brackets:

$$c = \begin{cases} z + d - \tau_1 z \quad \text{if } 0 < z \le z_1 \\ z + d - \sum_{i=1}^{l-1} \tau_i \left(z_i - z_{i-1} \right) - \tau_l \left(z - z_{l-1} \right) \quad \text{if } z_{l-1} < z \le z_l,\ l = 2,3,4 \\ z + d - \sum_{i=1}^{4} \tau_i \left(z_i - z_{i-1} \right) - \tau_5 \left(z - z_4 \right) \quad \text{if } z > z_4 \end{cases} \tag{3.8}$$

where c is net available income, z is the sum of gross market income (earnings plus capital income) and taxable public transfers, $\left(\tau_1, \tau_2, ..., \tau_5\right)$ are the marginal tax rates applied to the five income brackets, z_i is the upper limit of the i-th bracket (i = 1, 2, 3, 4), and d is a lump-sum that can be positive (i.e. a lump-sum transfer) or negative (i.e. a lump-sum tax). Thus, each particular tax rule is characterized by 10 parameters $\left(\tau_1, \tau_2, \tau_3, \tau_4, \tau_5, z_1, z_2, z_3, z_4, d\right)$. The tax rule is quite flexible since the MTRs are allowed to take positive or negative values ≤ 1 and the bracket-limits are allowed to take any positive value only subject to the constraints $z_i \ge z_{i-1}$. The tax rule specified by expression (3.8) replaces the current rule as of 1994, whose main characteristics are illustrated by the examples of Table 2.2 and also belongs to the class of piece-wise linear tax rules.[15] The dataset is the same as the

[15] Taxes include the part of social security contributions paid by the employee.

one used for the estimation of the model (Section II). The identification of the optimal tax rules consists of four steps:

1. For each household we simulate the opportunity set, which contains the observed job plus 199 market and non-market alternatives drawn from the estimated *p* densities defined in Section II. For each household and each alternative in the opportunity set we then draw a value ε from the Type III extreme value distribution. Next, the new tax rule is applied to individual earners' gross incomes in order to obtain disposable incomes (income after tax) corresponding to each alternative in the choice set. For each household, a new choice $(c, h_F, h_M, s_F, s_M, j)$ for couples or (c, h, s, j) for singles – is given by the alternative that maximizes the household-specific utility functions defined by (2.2).[16]
2. To each decision maker (wife or husband or single) an equivalent income *y* is imputed according to expression (3.2). The purpose of this procedure is to convert the distribution of incomes (*c*) across heterogeneous families into a distribution of (equivalent) incomes (*y*) across adult individuals.
3. As a result of the previous steps, we now have for each individual a simulated pair (*y, h*). We then compute the individual welfare levels by applying to the chosen (*y, h)* the individual welfare function (3.1).
4. We then compute the social welfare function W_k for $k = 1,2,3,\infty$.

The optimization is performed by iterating the steps 1–4 in order to find the tax rule in the class (3.8) that produces the highest value of W_k for each value of *k*, under the constraint of constant total tax revenue.[17]

[16] Colombino (1998), Colombino et al. (2010) and Blundell and Shephard (2011) use a different method, where the maximum utility is not found by simulation but is measured instead by the expected maximum utility (McFadden 1978). The two methods are asymptotically equivalent, but the method adopted in this paper turns out to be more flexible and robust for producing disaggregated results.

[17] The optimal tax-transfer parameters are determined by an iterative grid-search procedure developed by Tom Wennemo at the Research Department of Statistics Norway. Each optimization requires the evaluation of approximately 200 000 tax-transfer rules.

IV. The optimal tax-transfer schedules

The results of our exercise are reported in Tables 4.1–4.5. Table 4.1 displays the optimal tax rules. In order to ease the comparability of the behavioural responses to the 1994 tax system and the various optimal tax systems we report proportions of individuals by family status in specific tax income brackets in Table 4.2. Tables 4.3 and 4.4 provide additional information of the behavioural implications of the optimal tax rules. Table 4.5 displays the percentages of winners under the optimal rule by income deciles of the 1994 income distribution.

[Tables 4.1 – 4.5]

Under any social welfare function, the MTRs are continuously increasing for all levels of income. Clearly the pattern of elasticities – sharply decreasing with respect to income – illustrated in Table 2.1 contributes to the profile of the optimal MTR. The most striking results are represented by the negative MTR on the first bracket and by the 100 per cent MTR on the last bracket. These results are obviously driven by the pattern of the elasticities displayed in Table 2.1. From each of the panels of Table 2.1 we observe that the labour supply of the 10 per cent poorest are very responsive to changes in economic incentives whereas the 10 per cent richest are inelastic. Moreover, by comparing the fourth and fifth panel of Table 2.1 we see for married/cohabitating females that hours supplied (given participation), in particular for those belonging to the poorest couples, are far more responsive than participation.[18] A negative MTR on low incomes – in fact a subsidy or a tax credit on the wage rate – is close to policies actually implemented, such as the Working Families Tax Credit in the UK, the Earned Income Tax Credit in the USA and the In-Work Tax Credit in Sweden. The 100 per cent MTR on the last bracket – despite the fact that it could hardly be realistically implemented – does make sense within the limits of our model. As shown in Table 2.1, people in the richest decile exhibit on average a wage elasticity of labour supply equal, or very close, to zero (with the exception of married

[18] Except when income effects are assumed to be zero, the relationship between the optimal MTR and the wage elasticities is complicated. Other things being equal, a large (small) compensated wage elasticity contributes to a low (high) optimal MTR. Table 2.1 reports uncompensated wage elasticities. However, even when accounting for the incomes elasticities (reported in Aaberge and Colombino 2011) the resulting compensated wage elasticities are inversely related to the income levels.

women). At least part of this segment of the population (1.5 per cent) is willing to work the same amount of hours despite a reduction of the net wage rate, and non-pecuniary characteristics of the job (captured by the utility random component) may induce them to choose jobs where the marginal net wage is zero: their earnings become a rent which as such is captured by the optimal tax rule.[19] This argument carries over *a fortiori* to unearned income (completely inelastic in our model), which for this segment of the population might represent a very significant part of total income. The overall picture emerging from our exercise is in sharp contrast with most of the results obtained by the numerical exercises based on Mirrlees's optimal tax formulas. The typical outcome of those exercises envisages a positive lump-sum transfer which is progressively taxed away by very high marginal tax rates on lower incomes (i.e. a negative income tax mechanism); after the income level where the transfer is exhausted the tax rule is close to proportional. Tuomala (2010) suggests, however, that these results are essentially forced by the restrictive assumption typically made upon preferences, elasticities and distribution of productivities (or wage rates).

Table 4.1 shows that the more egalitarian the social welfare criterion, the more progressive the optimal tax rule. For example, the optimal rule according to Bonferroni is more progressive than the optimal rule according to Gini, which in turn is more progressive than the optimal utilitarian rule.

The lump-sum *d* turns out to be a tax. This result can be explained by the fact that households with small and medium high incomes are particularly sensitive to changes in marginal taxes (see Table 2.1). Thus MTR on low and average incomes are kept low both for minimizing distortions and for fulfilling distributive goals. However, since the total net tax revenue must be kept unchanged, the optimal tax rule envisages a universal lump-sum tax. A possible practical implementation close to a lump-sum tax might be represented by a tax on wealth or on property (e.g. on owner-occupied houses). According to this interpretation, the optimal tax rules would imply – with respect to the 1994 rule – a lower taxation on earnings complemented by a property tax.

[19] An anonymous referee correctly argued that this same segment of the population might be willing to accept a marginal tax rate above 100 per cent (i.e. a negative marginal net wage rate). However we decided to constrain the optimal marginal tax rates to be less than or equal to 100 per cent since official rates above 100 per cent would not be realistically considered anyway.

All the optimal rules imply a higher income after tax for most levels of gross income (Table 4.3). In other words, the optimal rules are able to extract the same total tax revenue from a larger total gross income (i.e. under a lower average tax rate). This result, together with those commented upon at point (a) above, provides a controversial perspective in view of the tax reforms implemented in many developed countries during the last decades. In most cases those reforms embodied the idea of improving efficiency and labour supply incentives through a lower average tax rate and lower MTRs on the highest incomes.[20] Our results give clear support to the first part: lowering the average tax rate; as to the second part, the picture is less clear-cut. Our results suggest that a lower average tax rate should be mainly obtained by lowering the marginal and average tax rates particularly on low and average incomes (and also on a substantial part of high incomes) and by sharply increasing them on very high income levels.[21]

Table 4.4 shows that the strongest labour supply response comes from households in the lower income deciles, who are those who show a more elastic labour supply. While females in couples receive a stronger incentive to work under the Bonferroni regime than under the Utilitarian regime, the opposite is the case for the males. This happens because the wife faces on average lower wages than the husband and the more relevant tax brackets for her are the lower ones, those where the Bonferroni regime imposes much lower MTRs than the Utilitarian regime (and than the current regime). On the other hand, the Utilitarian regime is especially favourable (also compared to the current regime) for those who decide to locate themselves in high tax brackets, where husbands are more likely to be found. The implication is that a more egalitarian criterion also involves stronger work incentives for married women (and especially those in the lower income deciles), and therefore also a more egalitarian inter-gender distribution of income.

---

[20] For example Blundell (1996) reports that during the 80s and early 90s in some countries the top marginal tax rates were cut from 70–80 per cent down to about 40–50 per cent. On these issues the discussion in Røed and Strøm (2001) is especially relevant.

[21] A second important difference between our exercise and the implemented reforms referred to in the main text, is that those reforms typically envisaged a reduction of the total tax revenue together with the reduction in the average tax rate, while in our simulations we keep the total tax revenue unchanged.

Table 4.5 shows the percentage of winners under the optimal rules, by marital status, gender and household income decile under the current 1994 rule. An individual is defined as a winner if her/his welfare is higher under the new tax rule than under the current 1994 rule. All the optimal rules would largely "win the referendum" against the current rule, since they all imply a strong majority of winners. The percentage of winners, however, varies substantially across the different subgroups and especially across income deciles. Single women in the IX and X income deciles are the only ones who would "vote against" all the optimal tax rules. The current (1994) tax system provides important deductions that favour in particular the group of relatively well-off single women with children. The deductions are removed in the class of tax-transfer rule we optimize upon. As a consequence, a majority of those women turn out to be losers under the optimal rules.

V. Conclusions

We have performed an exercise in designing optimal income taxes which – unlike what is typically done in the literature – does not rely on *a priori* theoretical optimal taxation results, but instead employs a microeconometric model of labour supply in order to maximize a social welfare function with respect to a parametrically defined income tax rule. Modern microeconometric models of labour supply are based on very general and flexible assumptions. They can accommodate many realistic features such as general structures of heterogeneous preferences, simultaneous decisions of household members, complicated (non-convex, non continuous, non-differentiable etc.) constraints and opportunity sets, multidimensional heterogeneity of both households and jobs, quantitative constraints etc. It is simply not feasible (at least so far) to obtain analytical solutions for the optimal income taxation problem in such environments. Yet those features are very relevant and important especially in view of evaluating or designing reforms. Analytical solutions remain indispensable for understanding the grammar of the problem and for suggesting promising classes of tax-transfer systems that can then be more deeply investigated with the microeconometric model.[22] The microeconometric model adopted in this paper and fully described in Aaberge and Colombino (2006, 2011) is designed to allow for a detailed description of complex choice sets and budget constraints. The model is used to identify by simulation the tax rule that maximizes a social welfare function. We keep fixed the current (1994) system of transfers, income support and social assistance policies, but allow for a lump-sum that can be positive (i.e. a transfer) or negative (i.e. a tax). We explore a variety of different social welfare criteria. The MTRs always turn out to be monotonically increasing with income. More egalitarian social welfare functions tend to imply more progressive tax rules. For all the social welfare functions used, the optimal bottom MTR is negative and the optimal top MRT always turns out to be 100 per cent for sufficiently high gross income levels (depending on the social welfare function, approximately above 720 000 – 790 000 NOK 1994), which concerns not more than 2 per cent of the tax payers. The negative MTR on the lowest income bracket suggests a mechanism close to policies like the WFTC in the UK or the EITC in the USA. The 100 per cent top MTR can be mainly explained by the inelastic labour supply at the top of the income distribution (Table 2.1) and by non-pecuniary characteristics that may make a job attractive even though it carries a 100 per cent marginal tax rate. All the optimal tax rules imply an average tax rate lower than the current 1994 one and imply – with respect to the current rule – lower marginal rates on low and/or average income levels and a higher marginal rate on very high income levels. The pattern of wage elasticities of labour supply illustrated in Section II helps explain the profile of the optimal tax rules. Our results are partially at odds with the tax reforms that took place in many countries during the last decades. While those reforms embodied the idea of lowering average tax rates, the way they were implemented typically consisted in reducing the top marginal rates. Our results instead suggest lowering average tax rates by reducing marginal rates except for very high income levels.

[22] The philosophy inspiring this approach is similar to the one adopted long ago in engineering and, recently and successfully, also in other applications of mechanism design (auctions, negotiation procedures, matching markets etc.) where analytical solutions are complemented by computational simulations or experiments that account for a host of realistic features that cannot be included in the theoretical model. Roth (2002) provides a very inspired survey of this approach.

Even though we think the approach illustrated here can usefully complement theoretical work and analytical solutions and actually improve upon them concerning the representation of preferences, constraints and policies, clearly there are many dimensions of the tax-transfer rules that are relevant for their evaluation (e.g. implementation and administrative costs) but are beyond the purpose of our exercise. Moreover, some of the results illustrated in Section IV might change with the inclusion in the behavioural model of features that are currently not accounted for. A candidate for further refinements is the modelling of the choice by households at the top of the income distribution. For example, the optimal top MTR might turn out to be lower than 100 per cent if we were able to fully account for other dimensions of households' response, such as inter-country mobility and taxable income response.[23]

[23] See Feldstein (1995) and Gruber and Saez (2002). However, based on previous exercises where we constrained the top MTR to be lower than 100 per cent we expect the overall qualitative features of the optimal tax rule to remain unaffected.

References

Aaberge, R. (2000): Characterizations of Lorenz Curves and Income Distributions, *Social Choice and Welfare*, 17, 639-653.

Aaberge, R. (2001): Axiomatic Characterization of the Gini Coefficient and Lorenz Curve Orderings, *Journal of Economic Theory*, 101, 115-132.

Aaberge, R. (2007): Gini's Nuclear Family, *Journal of Economic Inequality*, 5, 305-322.

Aaberge, R. and U. Colombino (2006): Designing Optimal Taxes with a Microeconometric model of Household Labour Supply, IZA Discussion Paper No. 2468.

Aaberge, R. and U. Colombino (2011): Empirical Optimal Income Taxation: A Microeconometric Application to Norway, CHILD Working Paper No. 16/2011.

Aaberge, R., Colombino, U. and S. Strøm (1999): Labour Supply in Italy: An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints, *Journal of Applied Econometrics*, 14, 403-422.

Aaberge, R., Colombino, U. and S. Strøm (2000): Labour Supply Responses and Welfare Effects from Replacing Current Tax Rules by a Flat Tax: Empirical Evidence from Italy, Norway and Sweden, *Journal of Population Economics,* 13, 595-621.

Aaberge, R., Colombino, U. and S. Strøm (2004): Do More Equal Slices Shrink the Cake? An Empirical Investigation of Tax-Transfer Reform Proposals in Italy, *Journal of Population Economics*, 17, 767-785.

Aaberge, R., Dagsvik J. K. and S. Strøm (1995): Labour Supply Responses and Welfare Effects of Tax Reforms, *Scandinavian Journal of Economics*, 97, 635-659.

Atkinson, A.B. (1970): On the measurement of inequality, *Journal of Economic Theory*, 2, 244-263.

Ben-Akiva, M. and T. Watanatada (1981): Application of a Continuous Spatial Choice Logit Model, in Manski, C. F. and McFadden D. (eds.) *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press.

Ben Porath, E. and I. Gilboa (1994): Linear Measures, the Gini Index, and the Income-Equality Trade-off, *Journal of Economic Theory*, 64, 443-467.

Blundell, R. (1996): Labour Supply and Taxation. In Devereux, M. (ed.): *The Economics of Tax Policy*, Oxford University Press.

Blundell, R., M. Brewer, P. Haan and A. Shephard (2009): Optimal Income Taxation of Lone Mothers: An Empirical Comparison of the UK and Germany, *Economic Journal*, 119, F101-F121.

Blundell, R. and A. Shephard (2011): Employment, Hours of Work and the Optimal Taxation of Low Income Families, *Review of Economic Studies* (forthcoming).

Boadway, R., M. Marchand, P. Pestieau and M. Racionero (2002): Optimal Redistribution with Heterogeneous Preferences for Leisure, *Journal of Public Economic Theory*, 4, 475 – 498.

Bonferroni, C. (1930): *Elementi di Statistica Generale*, Seeber, Firenze.

Bourguignon F. and A. Spadaro (2005): Tax-benefit revealed social preferences, PSE Working Paper 2005-22.

Colombino U. (1998): Evaluating the effects of new telephone tariffs on residential users' demand and welfare. A model for Italy, *Information Economics and Policy*, 10(3), 283-303.

Colombino. U., Locatelli M., Narazani E. and C. O'Donoghue (2010): Alternative Basic Income Mechanisms: An Evaluation Exercise with a Microeconometric Model, *Basic Income Studies*, 5(1).

Dagsvik, J.K. (1994): Discrete and Continuous Choice, Max-Stable Processes and Independence from Irrelevant Attributes, *Econometrica,* 62, 1179-1205.

Deaton, A. And J. Muellbauer (1980): *Economics and Consumer Behavior*, Cambridge University Press, Cambridge.

Decoster, A. and P. Haan (2010): Empirical Welfare Analysis in Random Utility Models of Labour Supply, IZA Discussion Paper 5301.

Diamond, P. (1980): Income Taxation with Fixed Hours of Work, *Journal of Public Economics*, 12, 101-110.

Diamond, P. (1998): Optimal Income Taxation: An Example with a U-Shaped pattern of Optimal Marginal Tax Rates, *American Economic Review*, 88, 83-95.

Donaldson, D. and J.A. Weymark (1980): A Single Parameter Generalization of the Gini Indices of Inequality, *Journal of Economic Theory*, 22, 67-86.

Donaldson, D. and J.A. Weymark (1983): Ethically flexible Indices for Income Distributions in the Continuum, *Journal of Economic Theory*, 29, 353-358.

Feldstein, M. (1995): The effect of marginal tax rates on taxable income: a panel study of the 1986 Tax Reform Act, *Journal of Political Economy*, 103(3), 551–572.

Fleurbaey, M. and F. Maniquet (2006): Fair Income Tax, *Review of Economic Studies*, 73, 55-84.

Fortin, B., Truchon, M. and L. Beauséjour (1993): On Reforming the Welfare System: Workfare meets the Negative Income Tax, *Journal of Public Economics*, 31, 119-151.

Gruber, J. and E. Saez (2002): The Elasticity of Taxable Income: Evidence and Implications, *Journal of Public Economics*, 84, 1-32.

Hammond, P.J. (1991): Interpersonal Comparisons of Utility: Why and How They Are and Should Be Made?. In Elster, J. and J. E. Roemer (eds.): *Interpersonal comparisons of well-being. Studies in Rationality and Social Change*, Cambridge; New York and Melbourne: Cambridge University Press in collaboration with Maison des Sciences de l'Homme, 1991; 200-254.

Haan, P. and K. Wrohlich (2010): Optimal Taxation: The Design of Child-Related Cash and In-Kind Benefits, *German Economic Review*, 11, 278-301.

Hey, J. D. and P. J. Lambert (1980): Relative Deprivation and the Gini Coefficient: Comment, *Quarterly Journal of Economics*, 94, 567-573.

King, M. (1983): Welfare Analysis of Tax Reforms Using Household Data, *Journal of Public Economics*, 21, 183-214.

Kleven, H. J., Kreine, C, T. and E. Saez (2009): The Optimal Income Taxation of Couples, *Econometrica*, 77, 537-560.

Immervoll H., Kleven H. J., Kreiner C. T., and E. Saez (2007): Welfare Reforms in European Countries: A Microsimulation Analysis, *Economic Journal,* 117, 1-44.

McFadden, D. (1978): Modeling the Choice of Residential Location, in (eds.) A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull *Spatial Interaction Theory and Planning Models*, North Holland, 75–96.

Meghir, C. and D. Phillips (2008): Labour Supply and Taxes, IZA Discussion Paper n. 3405, IZA, Bonn.

Mirrlees, J. A. (1971): An Exploration in the Theory of Optimal Income Taxation, *Review of Economic Studies*, 38, 175-208.

Revesz, J. T. (1989): The Optimal Taxation of Labour Income, *Public Finance*, 44, 453 – 475.

Roth A. E. (2002): The Economist as Engineer: Game Theory, Experimental Economics and Computation as Tools of Design Economics, *Econometrica*, 70, 1341-1378.

Røed, K. and S. Strøm (2002): Progressive Taxes and the Labour Market - Is the Trade-Off between Equality and Efficiency Inevitable?, *Journal of Economic Surveys*, 16, 77-100.

Saez, E. (2001): Using Elasticities to Derive Optimal Income Tax Rates, *Review of Economic Studies*, 68, 205-229.

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

Sen, A. (1974): Informational Bases of Alternative Welfare Approaches, *Journal of Public Economics*, 3, 387-403.

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

Tuomala, M. (2010): On optimal non-linear income taxation: numerical results revisited, *International Tax and Public Finance*, 17, 259-270.

Weymark, J. (1981): Generalised Gini Inequality Indices, *Mathematical Social Sciences*, 1, 409-430.

Yaari, M.E. (1987): The dual theory of choice under risk, *Econometrica*, 55, 95-115.

Yaari, M.E. (1988): A Controversial Proposal Concerning Inequality Measurement, *Journal of Economic Theory*, 44, 381-397.

Table 2.1. Labour supply elasticities with respect to wage for single females, single males, married females and married males by deciles of household disposable income*. Norway 1994

Family status	Type of elasticity		Female elasticities		Male elasticities	
		Income decile under the 1994 tax system	*Own wage elasticities*	*Cross elasticities*	*Own wage elasticities*	*Cross elasticities*
Single females and males	Elasticity of the probability of participation	I	0.59		0.00	
		II	0.45		0.00	
		III-VIII	0.06		0.06	
		IX	0.00		0.00	
		X	0.00		0.00	
		All	0.12		0.04	
	Elasticity of the conditional expectation of total supply of hours	I	-0.17		0.77	
		II	-0.04		0.00	
		III-VIII	-0.08		-0.08	
		IX	-0.07		0.00	
		X	0.00		0.00	
		All	-0.09		-0.02	
	Elasticity of the unconditional expectation of total supply of hours	I	0.42		0.77	
		II	0.42		0.00	
		III-VIII	-0.02		-0.02	
		IX	-0.07		0.00	
		X	0.00		0.00	
		All	0.02		0.02	
Married/cohabitating females and males	Elasticity of the probability of participation	I	1.03	-0.28	0.90	-0.23
		II	0.35	-0.14	0.79	0.00
		III-VIII	0.14	-0.23	0.13	-0.10
		IX	0.12	-0.12	0.06	-0.06
		X	0.07	0.00	0.06	-0.19
		All	0.21	-0.19	0.23	-0.11
	Elasticity of the conditional expectation of total supply of hours	I	1.51	-0.01	0.87	0.11
		II	0.62	-0.53	0.38	-0.08
		III-VIII	0.27	-0.24	0.18	-0.14
		IX	0.08	-0.22	0.02	-0.09
		X	0.19	-0.10	-0.02	-0.23
		All	0.31	-0.25	0.16	-0.13
	Elasticity of the unconditional expectation of total supply of hours	I	2.54	-0.29	1.77	-0.12
		II	0.97	-0.67	1.17	-0.08
		III-VIII	0.41	-0.47	0.31	-0.24
		IX	0.20	-0.34	0.08	-0.14
		X	0.26	-0.10	0.05	-0.42
		All	0.52	-0.42	0.39	-0.23

Table 2.2.Current tax rule in Norway as of 1994 for singles without children and couples without children and with two wage earners(*)

Gross earnings (NOK 1994)	Tax
(0 – 17000)	0
(17000 – 24709)	0.25x - 4250
(24709 – 28250)	0.078x
(28250 – 140500)	0.302x - 6328
(140500 – 208000)	0.358x - 14196
(208000 – 234500)	0.453x - 33956
(234500 –)	0.495x - 43804

(*) x denotes annual earnings. Taxes include the part of social security contributions paid by the employee. 10 000 NOK ≃ 1 250 Euros.

Table 2.3. The 2001 tax function for singles without children and couples without children and with two wage earners(*)

Gross earnings (NOK 2001)	Tax
[0 – 22200)	0
[22200 – 32267)	0.25x – 5550
[32267 – 60600)	0.078x
[60600 – 144545)	0.358x – 16968
[144545 – 183182)	0.296x – 8064
[183182 – 289000)	0.358·x – 19 348
[289000 – 793200)	0.493x – 58 363
[793200 –)	0.553x – 105 955

(*) x denotes annual earnings. Taxes include the part of social security contributions paid by the employee. 10 000 NOK ≃ 1 250 Euros.

Table 2.4. Observed and predicted *relative* distributions of disposable income in 2001. Mean decile income in percent of mean income

Deciles	Couples		Single females		Single males	
	Observed	*Simulated*	*Observed*	*Simulated*	*Observed*	*Simulated*
1	50	49	45	47	41	42
2	68	64	56	61	54	55
3	77	74	68	71	65	67
4	83	83	79	79	76	76
5	89	90	90	88	87	86
6	95	98	101	98	97	97
7	102	107	111	108	107	108
8	111	117	123	121	119	121
9	125	131	139	138	137	141
10	199	187	189	188	218	207
9	129	128	142	136	150	135
10	159	151	177	166	178	161

Table 3.1. Estimates of the parameters of the individual welfare function, Norway 1994

Variable	Parameter	Estimate	Stand.dev.
Income after tax (y)			
	γ_1	-0.649	0.086
	γ_2	3.026	0.138
Leisure (L)			
	γ_3	-12.262	0.556
	γ_4	0.045	0.011

Table 3.2. Distributional weight profiles of four different social welfare functions

	W_1 (Bonferroni)	W_2 (Gini)	W_3	W_∞ (Utilitarian)
q(.01)/q(.5)	6.64	1.98	1,33	1
q(.05)/q(.5)	4,32	1,90	1,33	1
q(.30)/q(.5)	1,74	1,40	1,21	1
q(.95)/q(.5)	0,07	0,10	0,13	1

Table 4.1 Optimal tax rules according to alternative social welfare criteria(*)

	Social welfare function			
	W_1 (Bonferroni)	W_2 (Gini)	W_3	W_∞ (Utilitarian)
τ_1	-0.30	-0.80	-0.70	-0.80
τ_2	0.06	0.20	0.22	0.24
τ_3	0.29	0.26	0.26	0.29
τ_4	0.39	0.38	0.37	0.33
τ_5	1.00	1.00	1.00	1.00
d	-13 600	-7 500	-5 200	-5 800
z_1	10 000	10 000	10 000	10 000
z_2	120 000	130 000	140 000	230 000
z_3	220 000	230 000	240 000	290 000
z_4	730 000	720 000	720 000	790 000

(*) d, z_1, z_2, z_3 and z_4 are in 1994 NOK (10 000 NOK $\simeq$ 1250 Euros).

Table 4.2 Percentage of individuals by income intervals under different tax systems.

	Proportions located in various gross income segments			
Income intervals	**1994 tax system**			
	Couples (Males)	Couples (Females)	Single Males	Single Females
0 –30 000	5	16	0	0
30 000 – 130 000	11	33	26	24
130 000 – 230 000	31	35	41	51
230 000 – 730 000	52	16	33	24
730 000 ->	2	0	0	0
	W_1 – optimal tax system			
0 –30 000	2	10	0	0
30 000 – 130 000	9	32	22	22
130 000 – 230 000	29	41	42	51
230 000 – 730 000	58	17	35	27
730 000 ->	2	0	0	0
	W_2 – optimal tax system			
0 –30 000	3	12	0	0
30 000 – 130 000	9	32	23	22
130 000 – 230 000	28	39	41	50
230 000 – 730 000	59	17	36	28
730 000 ->	1	0	0	0
	W_3 – optimal tax system			
0 –30 000	3	12	0	0
30 000 – 130 000	8	31	23	21
130 000 – 230 000	27	39	41	50
230 000 – 730 000	60	17	36	28
730 000 ->	2	0	0	0
	W_∞ – optimal tax system			
0 –30 000	3	13	0	0
30 000 – 130 000	8	31	22	20
130 000 – 230 000	25	37	40	50
230 000 – 730 000	62	18	38	30
730 000 ->	2	0	0	0

Table 4.3 Percentage changes in participation rates, annual hours of work and disposable income under the optimal tax rules

		Social welfare function			
		W_1 (Bonferroni)	W_2 (Gini)	W_3	W_∞ (Utilitarian)
Single males	Participation rates	2.3	2.3	2.3	2.3
	Annual hours	4.9	5.0	5.1	6.0
	Disposable income	10.0	9.7	10.0	11.9
Single females	Participation rates	4.0	4.4	4.4	4.8
	Annual hours	6.0	6.6	6.6	9.0
	Disposable income	4.7	4.6	4.5	6.6
Couples	Participation rates, M	2.6	2.3	2.3	2.7
	Participation rates, F	5.8	4.3	3.9	3.3
	Annual hours, M	5.9	6.2	6.6	9.4
	Annual hours, F	10.6	8.1	7.0	6.3
	Disposable income	9.2	9.4	9.9	13.3

Table 4.4 Percentage changes in labour supply (total hours) by household income decile under the optimal tax rules

		Social welfare function							
		W_1 (Bonferroni)		W_2 (Gini)		W_3		W_∞ (Utilitarian)	
	Income decile under the 1994 system	Male	Female	Male	Female	Male	Female	Male	Female
Singles	I	60.5	64.7	57.3	54.4	62.8	54.4	62.8	60.3
	II	18.6	17.9	18.6	21.3	18.6	21.3	20.3	29.3
	III-VIII	0.9	3.0	1.2	4.5	1.1	4.5	1.7	7.2
	IX	0.0	0.0	0.0	0.0	0.0	0.0	2.6	-0.4
	X	1.3	0.0	1.3	0.0	1.3	0.0	1.3	0.0
	All	*4.9*	*6.0*	*5.0*	*6.6*	*5.1*	*6.6*	*6.0*	*9.0*
Couples	I	50.4	74.4	43.8	60.7	42.8	56.0	49.5	60.8
	II	22.2	22.9	23.3	19.5	22.9	20.5	32.7	18.6
	III-VIII	2.6	7.7	3.5	5.1	4.2	3.6	6.7	2.7
	IX	0.7	0.5	0.7	1.0	0.7	0.5	1.2	-0.3
	X	-3.3	0.0	-2.9	0.2	-2.9	0.9	-1.4	0.1
	All	*5.9*	*10.6*	*6.2*	*8.1*	*6.6*	*7.0*	*9.4*	*6.3*

Table 4.5. Percentage of winners under optimal tax rules

		Social welfare function							
		W_1 (Bonferroni)		W_2 (Gini)		W_3		W_∞ (Utilitarian)	
	Income decile under the 1994 system	Male	Female	Male	Female	Male	Female	Male	Female
Singles	**I**	79	76	83	72	83	72	79	69
	II	66	62	66	55	62	55	55	55
	III-VIII	86	68	85	68	81	68	77	66
	IX	79	45	83	45	83	45	83	48
	X	76	34	79	38	79	38	86	41
	All	*82*	*63*	*82*	*62*	*79*	*62*	*76*	*61*
Couples	**I**	62	64	64	67	62	65	61	64
	II	70	72	72	73	73	73	70	73
	III-VIII	84	85	84	86	84	87	83	87
	IX	85	87	86	88	88	90	88	91
	X	71	69	72	70	74	72	79	78
	All	*79*	*80*	*70*	*81*	*80*	*82*	*79*	*83*