Review of Income and Wealth
Series 50, Number 1, March 2004

MEANS-TESTING THE CHILD BENEFIT

by Tom Kornstad*

and

Thor O. Thoresen

Statistics Norway

Improving the distributional impact of transfers may be costly if it reduces labor supply. In this paper we show how effects of changes in the design of the child benefit program can be examined by employing information from behavioral and non-behavioral simulations on micro data. The direct distributional effects are assessed by tax-benefit model calculations, while female labor supply responses to alternative child benefit schemes are simulated under the assumption that choices are discrete. Distributional effects after labor supply responses are also shown. The study confirms that greater targeting of the child benefit is traded against reductions in female labor supply.

1. Introduction

In almost all industrialized countries one finds a publicly provided cash transfer specifically directed at families with children,[1] a child benefit scheme. There is considerable variation across countries with respect to the structure of benefit systems, cf. for instance Bradshaw *et al.* (1993), and Atkinson (1995). A key distinction is that between systems in which the transfer is conditional on income and programs where the conditioning is done with respect to other variables, as family type, number of children and their age. The Norwegian child benefit scheme is of the latter type, but calls are frequently heard from politicians and experts to direct more resources to the most needy families, i.e. to means-test the benefit. The calls for greater target efficiency have intensified in line with a more generous system, making program expenditures increase by about 25 percent in real terms from 1990 to 1998, reaching about 13 billion Norwegian kroner (NOK) (about $1.7 billion)[2] in 1998, or approximately 7 percent of the revenue from taxation of individuals' income and wealth. Marring the debate, however, has been the lack of comprehensive information on the effects of altering the scheme. For instance, qualified measures of labor supply effects have been lacking. This information is crucial in welfare analyses of changes in the tax and transfer system, not the least since

Note: We particularly thank John Dagsvik and two referees of this journal for many helpful comments and suggestions. Valuable comments have also been received from seminar participants at the 1998 annual meeting for the Economic Research Programme on Taxation, Halvorsbøle (Norway) and at the 55th Congress of the International Institute of Public Finance, Moscow, August 1999. We have benefited from research assistance by Kirsten Hansen, Bård Lian and Ingeborg F. Solli. Financial support from the Norwegian Research Council is gratefully acknowledged.

*Correspondence to: Tom Kornstad, Research Department, Statistics Norway, PO Box 8131 Dep., N-0033 Oslo, Norway (tom.kornstad@ssb.no).

[1]One exception is the United States, where the support for families with children is provided by other modes of assistance.

[2]1994 exchange rates.

maintaining high employment is seen as a key element in efforts to avoid the potentially negative impact of population ageing on living standards.

In this paper we assess the effects of three alternative means-testing devices by analyzing information derived from behavioral and non-behavioral simulations on micro data. The analysis is limited to families with married mothers, as the labor supply decisions of lone mothers depend on special support programs for lone parents. The distributional effects are discussed in terms of effects on equivalent household income both before and after labor supply adjustments, while the labor supply effects are restricted to effects on female labor supply. We choose to focus on female labor supply because mothers are the child benefit recipients and because results from empirical studies suggest that women are more responsive to changes in taxes and transfers than men (Blundell, 1993; Blundell and MaCurdy, 1999). The approach allows us to contrast the distributional gains associated with means-testing with the effects on female labor supply.[3]

Our results apply to the discussions of means-testing in the literature; a selection of recent studies includes Garfinkel (1982), Besley (1990), Atkinson (1995), Immonen *et al*. (1998), and Creedy (1996, 1998). The analysis also relates to the discussions concerning the Family Credit and the Working Families Tax Credit in the U.K. (see Duncan and Giles, 1996; Blundell *et al*., 2000) and the Earned Income Tax Credit in the U.S. (cf. Scholz, 1996; Eissa and Hoynes, 1998), as these in-work benefits phase-out with respect to income levels. However, while the main motivation for reforms in these two tax credits appears to be to strengthen work incentives, or to "make work pay," it is distributional concerns that have actuated the calls for changes in the Norwegian child benefit scheme. On the other hand, it can be argued that these two motives are just two sides of the same coin, as improved incentives to work can weaken the effects of the "poverty trap" and improve living standards for the less well-off.

We address the effects of three alternative modes of means-testing recently discussed in a green paper on transfers to families with children (Ministry of Children and Family Affairs, 1996): taxation of the benefit as wage income for the mother and two ways of testing the transfer against family income. The first alternative of income testing entails a loss of 50 percent in child benefit for families whose incomes exceed NOK400,000 ($56,700) per year and 100 percent for those exceeding NOK600,000 ($85,100) per year. The other alternative involves a deduction of one tenth of a Norwegian krone in child benefit for every krone the family earns in excess of NOK250,000 ($35,500) per year.

The direct (non-behavioral) distributional effects of altering the child benefit transfer system are described by tax-benefit model calculations based on information from income tax returns for about 4,500 households with a married mother eligible for child benefit. In the analyses of the labor supply responses a structural discrete choice labor supply model is applied. The specification of the particular

[3]Of course, there might be other important motives when determining the child benefit scheme. For instance, Sadka, Garfinkel, and Moreland (1982) emphasize the stigma associated with income testing, while Lundberg, Pollak, and Wales (1997) stress the importance of the child benefit on the intra-family distribution of consumption. Another aspect is administrative costs, as emphasized by Atkinson (1995).

discrete choice model is motivated by the following: in many occupations there are non-pecuniary job attributes, such as job satisfaction and other means of self-realization that may be just as important as wage rates and hours of work. The individual then faces a labor market consisting of a finite set of feasible job opportunities where each job is assumed to have fixed working hours and non-pecuniary attributes. In this sense the choice set is discrete. To the analyst the set of feasible job opportunities is, however, unobservable. We make specific assumptions about the distribution of preferences and the latent choice sets, which enables us to derive an empirical model for observed behavior.

We find similar results as found in the analyses of the tax credits in the U.K. (Blundell *et al.*, 2000) and in the U.S. (Scholz, 1996; Eissa and Hoynes, 1998) that also phase-out with respect to income; there are adverse female labor supply effects involved when eligibility is based on family income. Labor supply adjustments also impact on the distribution of income, and one advantage of the micro simulation framework applied here is that the distributional effects after labor supply responses can be considered explicitly. When behavioral adjustments are included in the evaluation of effects on income inequality, the alternative in which the 50 percent and 100 percent tapers are introduced at NOK400,000 and NOK600,000 becomes more attractive from a distributional point of view, as the income losses due to behavioral adjustments, on average, are more pronounced at the high end of the decile income ranking than among low-income families.

This paper is organized as follows: in Section 2 we present the current child benefit scheme and the three means-testing alternatives, while the direct distributional effects of altering the scheme are described in Section 3. Section 4 presents the specification of the discrete choice model for female labor supply and provides estimation results. The model is employed to assess the labor supply responses to means-testing in Section 5. In Section 6 we discuss the distributional effects after labor supply responses, and round up our arguments in Section 7.

2. The Child Benefit Scheme and Methods of Means-Testing

The Norwegian Child Benefit Scheme

In this section the structure of the Norwegian child benefit system is briefly described and we present the three alternatives of means-testing.[4] We address changes in the child benefit scheme for 1994, the year of data collection. As the program structure has basically remained unchanged over the last few years, the main findings will apply to the current system too.

According to Bradshaw *et al.* (1993, p. 35) the Norwegian child benefit scheme is rather generous compared to those of other countries. Child benefit was paid to the mother of children under 16 years of age in 1994. For the first child she received NOK10,416 per year, which is about 1,350 U.S. dollars (according to 1994 exchange rates). The amount per child increased with the number of children; for a fifth child she would receive NOK13,392 ($1,900). In addition, the mother

[4]Since the analysis is restricted to households consisting of married couples (with children), the special scheme for lone mothers is not discussed.

received NOK5,040 ($710) per child per year if the child was younger than 3 years, and NOK3,752 ($530) per child per year if the family lived in the northern parts of the country. Thus, a family with three children, of which one was younger than 3 obtained nearly NOK40,000 in child benefit in 1994, or about $5,700 per year. Average disposable household income including actual child benefit in the sample is NOK325,000 ($46,100).

Consequently, the child benefit represents an important contribution to family welfare. However, the affluent are as eligible to the support as other families. This "inefficiency" in the transfer arrangement has prompted suggestions that the support should be more targeted towards those in greater need. We will consider the effects of three such attempts to satisfy this requirement, recently discussed in a green paper (Ministry of Children and Family Affairs, 1996). These alternative methods raise different amounts of revenue; to make them more comparable we increase child benefit rates in each alternative, thereby eliminating revenue effects of reforms before potential labor supply adjustments. For instance, in the case of taxing the child benefit as wage income, the yearly rates are simultaneously raised by NOK7,000 ($990) in order to ensure that the taxation has no effect on tax revenues, before labor supply adjustments.

Taxation of the Child Benefit

The most straightforward way of means-testing is to include the transfer in the tax base and let the tax system resolve the targeting problem. As shown in Figure 1, Norwegian taxation of wage income is progressive with a two-tier surtax: thresholds are triggered at NOK208,000 ($29,500) and NOK234,500 ($33,200) p.a. for single filers (tax class 1) and NOK252,000 ($35,700) and NOK263,000 ($37,300) for joint filers (tax class 2).[5] Given the statutory tax rates shown in Figure 1, it would be reasonable to assume that the inclusion of the child benefit in taxable income would prove relatively advantageous to the poorest families.

Income Testing

Two methods of income testing are considered, one in which a 50 percent taper applies for (gross) family income above NOK400,000 ($56,700) per year, and the remaining 50 percent is lost when family income exceeds NOK600,000 ($85,100). In the following this method will be referred to as income testing alternative A. The other model of income testing, alternative B, is a linearized version of income testing, where one-tenth of a Norwegian krone in child benefit is deducted for every krone the family earns in excess of NOK250,000 ($35,500) per year. This scheme increases the marginal tax rates for incomes above NOK250,000 by 10 percentage points, gradually reducing the amount received to zero.

Figure 2 displays the effects of the two income testing schemes for a family with three children, one of whom is under 3 years old. As the figure shows, the family will "lose" their child benefit (NOK38,736, $5,500) in the interval between

[5]Most couples are taxed at the individual level (single filers), as joint filing is advantageous only when one spouse has a small or no income, which is not the case for a majority of Norwegian couples.

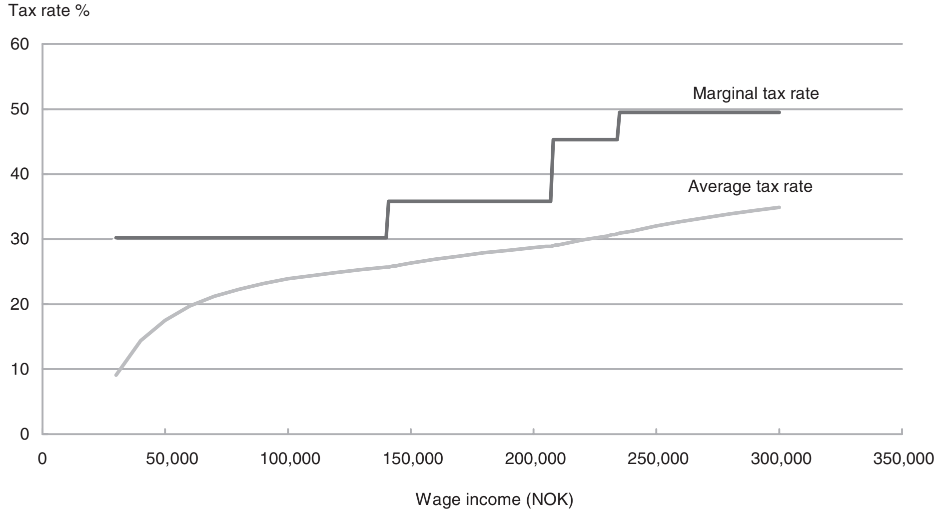

Figure 1. Marginal and Average Tax Rates on Wage Income. Single Filers in 1994

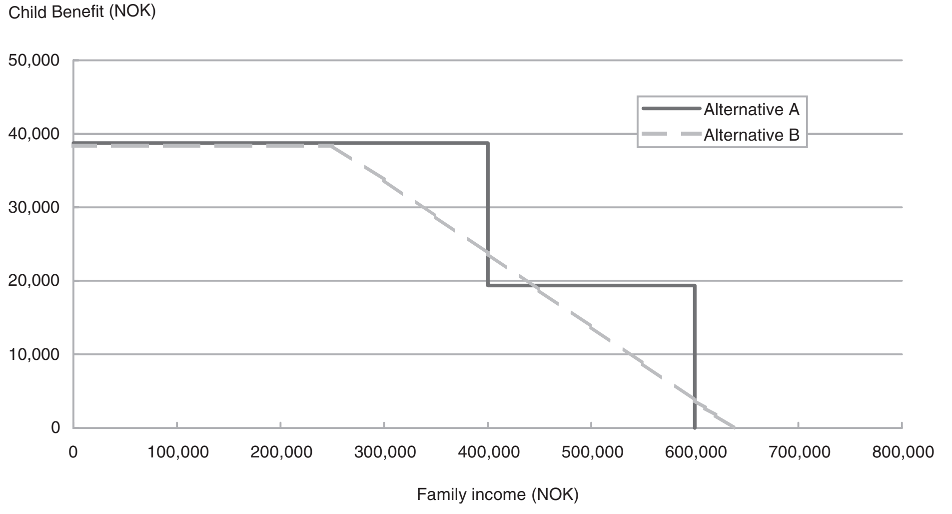

Figure 2. Two Alternative Methods of Testing Child Benefit Against Family Income. Example Based on a Family With Three Children (One Under 3)

NOK250,000 ($35,500) and NOK637,360 ($90,400) in gross family income under the linearized deduction scheme (alternative B).

3. The Direct Distributional Effects

This section discusses the direct distributional effects on equivalent household income of alterations in the child benefit scheme, as described by simulation results from the Norwegian tax-benefit model LOTTE.[6] The tax-benefit model simulations are based on the 1994-wave of the Income Distribution Survey (IDS), which includes about 4,500 households with a married mother eligible for child benefit. IDS is a large representative sample survey consisting of information from administrative and statistical registers, including registers of income tax returns. The household composition is established by interviews.[7]

Empirical evaluations of the direct effect of changes in taxes and transfers are associated with methodological controversies. Income definitions and techniques for the re-weighting of income may impact on the assessment of distributional effects of tax policy changes. An extensive examination of these issues is beyond the scope of this paper, but the following points summarize our main assumptions: Firstly, post-tax income is defined as gross income minus taxes, plus a number of tax-free benefits such as the child benefit, housing benefits, social security benefits, etc. Interest expenses, for instance from house acquisition, are not deducted due to the undervaluation of imputed income from owner-occupied homes in our data.[8] Secondly, post-tax income is aggregated over household members, weighted with an equivalence scale, and each person in the household is represented by household equivalent income. Thirdly, the equivalence scale is defined by the square root of number of household members (Buhmann *et al.*, 1988), where the relative weight given to a child is 0.75 of an adult. This income measure will be referred to as equivalent income in the following.

The effects of changes in the child benefit scheme depend on the distribution of the child benefit in the current system. Figure 3 shows the distribution of the equivalent child benefit (measured in thousand Norwegian kroner) by equivalent income decile. The distributional impact of the child benefit can be said to be favorable since the lower income deciles on average receive more of the benefit than wealthier individuals. This result is mainly due to a negative correlation between equivalent income and number of benefit-eligible children. While the average number of children is above 2 for deciles 1–5, the figure for decile 10 is 1.6.

To what extent there is assortative mating among couples has consequences with respect to the distributional effects from the policy changes. For instance, in the case of including the child benefit in the tax base, the redistributional effect will be more evident when high-income females are found in high-income households. When ranking the couples by male income, we find that female income distributes fairly evenly on male income. Thus, it does not seem to be any strong correlation of income among spouses.

Figure 3 also includes estimates for working hours by deciles. These are imputed from the sample of married mothers used in the behavioral analysis below. The figure shows a positive correlation between equivalent income and

[6]The model is used extensively in the preparation of national budgets in Norway.

[7]In the case of non-response, household information is replaced by information from administrative registers on family composition.

[8]By calculating an average ratio between the market value (only reported by a smaller number of individuals) and imputed value of housing for income tax purposes, we have estimated income from housing and defined an alternative income concept. These results are not shown here, but the distribution of income when including "market profits" from housing and deducting interest rates is not very different from the distribution of income according to the income concept applied.

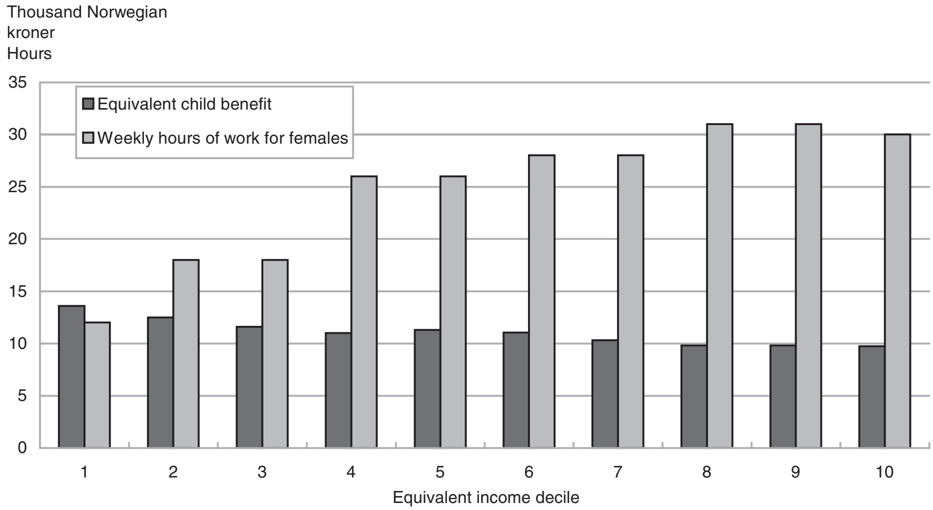

Figure 3. Equivalent Child Benefit and Women's Working Hours Per Week by Equivalent Income Decile

female working hours. Ranking of individual well-being according to equivalent income is thus questionable, but this problem is ignored in the following.

The direct distributional effects of the three alternative ways of means-testing are set out in Figure 4 as deviations in percent from the average pre-reform equivalent income in deciles. The taxation method, i.e. taxation of the child benefit as mothers' wage income and redistributing the collected revenue into an equal-sized increase in all child benefit rates,[9] has very little impact on incomes. The main reason is that the key progressive element in the Norwegian tax system, the surtax (cf. Figure 1), is not very effective here since only a small proportion of mothers are liable to surtax. Another reason is that the high-earning mothers are not necessarily found at the high-end of the equivalent income distribution: both mating patterns (cf. discussion above) and the ambiguous relationship between the ranking by equivalence scales and the ranking by tax progression (see, e.g. Lambert, 1993), mean that persons with high incomes in terms of tax assessment can be found in lower equivalent income deciles. Thus, the redistributional impact of this revenue-neutral tax-law change is rather limited, as seen in Figure 4. In total, the income inequality among married couples with children 0–16 years is reduced by 0.5 percent, as measured by the Gini coefficient. A 0.5 percent increase in the Gini coefficient corresponds to introducing an equal-sized lump-sum tax of 0.5 percent on the mean income and redistributing the collected tax revenue as proportional transfers whereby each unit receives 0.5 percent of its income (Aaberge, 1997). Alternatively, it can be interpreted as a 0.5 percent increase in the expected difference in incomes between two individuals drawn at random from the income distribution (Jenkins, 1991).

[9]As noted in Section 2 this means that yearly child benefit rates are increased by approximately NOK7,000 ($990).

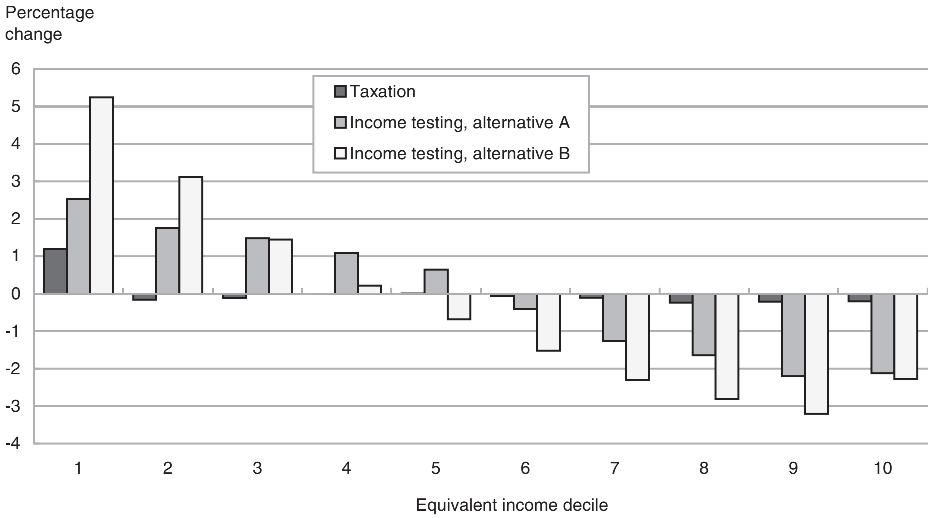

Figure 4. Direct Distributional Effects from Three Methods of Means-Testing. Gain in Percent by Equivalent Income Decile

Compared to the effects from taxing the child benefit, income testing entails more advantageous distributional effects. Especially the alternative in which one-tenth of a Norwegian krone is deducted for every krone the family earns in excess of NOK250,000 or $35,500 (income testing, alternative B) has distinct redistributional effects. In sum, the income inequality is reduced by more than 6 percent among married couples with children. Implementing 50 percent and 100 percent tapers at NOK400,000 ($56,700) and NOK600,000 ($85,100) respectively (alternative A) gives a smaller redistributive effect, a 2.6 percent reduction in overall income inequality as measured by the Gini coefficient.

4. A Model for Female Labor Supply

Discrete Choice Specification

As pointed out in the introduction, effects on labor supply are crucial when discussing reforms in the child benefit scheme. In the following we discuss our empirical approach to the assessment of behavioral responses of mothers. The model is a family decision model, where choices are made with respect to female labor supply, whereas male labor supply decisions are given.[10] Male income is, thus, exogenously determined, and treated as non-labor income when the partner makes her decisions. Assuming that males are unresponsive with respect to the policy changes in question is an approximation, but empirical labor supply studies often find that female labor supply is more responsive to policy changes than male labor supply (Blundell, 1993; Blundell and MaCurdy, 1999).

[10]One consequence of this assumption is that we ignore the potential interrelationship between male and female labor supply within families.

Our line of reasoning with respect to female labor supply departs from two features of labor market choices. Firstly, in empirical labor supply modeling it is typically assumed that working time and consumption of goods are the fundamental choice variables of the household. In many occupations it is, however, obvious that other job-related characteristics are at least equally important, such as work activities, job location and other measures related to job satisfaction. Secondly, the assumption that hours of work (and consumption) can be chosen freely subject to the budget constraint is often questionable. In many jobs hours of work are given, which implies that a change in working time either requires or is a consequence of a job change.

Given these features, we assume, as in Aaberge, Dagsvik, and Strøm (1995), that labor supply decisions can be analyzed within a discrete choice approach, where the worker chooses from a finite choice set of jobs, consisting of a fixed (offered) working time, a specific wage rate and non-pecuniary attributes. Generally, these choice sets vary over workers both in composition and number of jobs, since workers have different qualifications and live in different regions of the country. In what follows, however, we assume that the distribution of the choice sets is the same across workers (but the realizations are assumed to be independent), except for wage rates.[11] With the exception of the job actually chosen, the choice set is typically latent to the econometrician. In the empirical specification below it is also assumed that the non-pecuniary attributes of the chosen job are unobserved, but the approach can be modified to take into account that some of these attributes are observed.

Due to the budget constraint the yearly consumption (C_k), which corresponds to a given job k, is

$$C_k = wH_k + I - T(wH_k, I), \quad k \in B, \tag{1}$$

where B is the choice set of jobs, w is pre-tax wage rate, H_k is the fixed annual hours of work associated with job k, I is non-labor income and T is taxes. Wage rates are typically determined in negotiations with trade unions that do not discriminate between full-time and part-time workers in Norway, but there may be considerable gaps between wages in different labor market sectors. We do not have data on job characteristics such as sectors, so these aspects are treated as latent variables. In what follows it is, thus, assumed that wages are individual specific but equal across feasible jobs.

Non-labor incomes (I) include the household's capital incomes, the child benefit and other public transfers as well as labor income of the spouse.[12] Deviations in taxation of various income components are suppressed in the specification of the tax function, as is the taxation of wealth.

The household determines labor supply and consumption of goods by maximizing the preference function subject to the budget constraint. The utility of having a particular job k is assumed to be given by

[11]See Ilmakunnas and Pudney (1990), Dickens and Lundberg (1993) and Aaberge, Dagsvik, and Strøm (1995) for analyses modifying this assumption.

[12]Since we assume that choices with respect to male income have already been made. In contrast to van Soest (1995), we do not include unemployment benefits as income in the non-working state. With low unemployment rates, as was the case in Norway in 1994, it is difficult to stay out of work and receive unemployment benefits.

(2) $$U(H_k,C_k,k) = v(H_k,C_k) + \hat{\varepsilon}(H_k,C_k,k), \quad k \in B,$$

where $v(H_k,C_k)$ is the deterministic part of the utility function and $\hat{\varepsilon}(H_k,C_k,k)$ is a stochastic error term, which is supposed to account for the effects of non-pecuniary variables. We assume that the job-specific error terms, $\hat{\varepsilon}_k \equiv \hat{\varepsilon}(H_k,C_k,k)$, are i.i.d. according to the standard type I extreme value distribution,[13]

(3) $$P(\hat{\varepsilon}_k < \varepsilon) = \exp(-\exp(-\varepsilon)), \quad \varepsilon \in R.$$

The interval from zero hours of work to the upper limit of working time is divided into a number of intervals that can be associated with full-time work, part-time work, etc. Assume there are 0, 1, 2, . . . , j, . . . , J intervals, where interval 0 corresponds to not participating in the labor market and the other intervals are arranged according to increasing hours of work. Let $\underline{h}_j$ and $\bar{h}_j$ denote the lower and upper limit of hours of work in interval j respectively. Given our distributional assumptions about $\hat{\varepsilon}_k$, Dagsvik (2001) has demonstrated that the utility of the job maximizing utility with hours of work in interval j, is given by

(4) $$U_j^* \equiv \max_{\underline{h}_j \le H_k < \bar{h}_j} U(H_k,C_k,k) \stackrel{d}{=} \log\left(\sum_{\underline{h}_j \le H_k < \bar{h}_j} \exp(v(H_k,C_k))\right) + \varepsilon_j,$$

where ε_j has the same distribution as $\hat{\varepsilon}_k$ and $\stackrel{d}{=}$ means equality in distribution.

Moreover, we assume that the following approximation is close

(5) $$\sum_{\underline{h}_j \le H_k < \bar{h}_j} \exp(v(H_k,C_k)) \approx n_j \exp\left(v\left(\tilde{H}_j,\tilde{C}_j\right)\right),$$

where n_j is the number of jobs in interval j, $\tilde{H}_j$ is the average working time in interval j, and $\tilde{C}_j$ is consumption corresponding to working time $\tilde{H}_j$. Substituting this expression into equation (4), leads to

(6) $$U_j^* = \log n_j + v\left(\tilde{H}_j,\tilde{C}_j\right) + \varepsilon_j.$$

The probability of choosing a job within interval j is then given by

(7) $$P_j \equiv P\left(U_j^* = \max_k U_k^*\right) = \frac{\exp\left(v\left(\tilde{H}_j,\tilde{C}_j\right) + \log(n_j)\right)}{\sum_{r=0}^{J} \exp\left(v\left(\tilde{H}_r,\tilde{C}_j\right) + \log(n_r)\right)}.$$

Note that the number of job offers might vary across workers due to differences in education, work experience and place of residence. The n_r's may then be a function of these variables, but this dependency is ignored in what follows. Moreover, the number of jobs might vary across intervals because the densities of offered jobs are particularly high within some intervals, for instance due to economies of scale in production. This is consistent with peaks in the hours of work distribution observed.

[13]There are also errors in the maximization due to imperfect perception and imperfect optimization, also reflected in the error term.

Empirical Specification

We divide working hours into six different intervals: non-participation in the labor market; three types of part-time jobs corresponding to 1–11 hours per week; 12–22 hours per week and 23–33 hours per week; full-time jobs (34–44 hours per week); and overtime jobs (45+ hours per week). In the calculations of $v(\tilde{H}_j, \tilde{C}_j)$, it is assumed that $(\tilde{H}_j/52) \in \{0, 6, 17, 28, 39, 50\}$.

The deterministic part of the household preferences are given by the following "Box-Cox" type utility function,[14]

(8)
$$v(\tilde{H}_j, \tilde{C}_j) \equiv \gamma_0 \frac{\tilde{C}_j^{\alpha_1} - 1}{\alpha_1} + \frac{\left(1 - \frac{\tilde{H}_j}{M}\right)^{\alpha_2} - 1}{\alpha_2} X\beta,$$

where γ_0, α_1, and α_2 are parameters, β is a vector of parameters and $M = 8{,}760$ is the total number of annual hours. The vector of household-specific taste-modifier variables X includes ln(age), ln(age) squared, the number of children under 3 years and the number of children from 3 to 16 years old. The classification according to children's age takes into account the high contribution rate for older pre-schoolers at child-care centers, which makes these children more similar to school children. For the utility function to be quasi-concave, we require $a_1 < 1$, and $a_2 < 1$. Note also that if $a_1 \to 1$ and $a_2 \to 1$, the utility function converges to a log-linear function.

The individuals will find that the number of job opportunities varies with respect to working hours. For instance, there are reasons to assume that there are more full-time opportunities. Thus, in the estimation we allow the full time-parameter n_4 to deviate from the part-time parameters n_1, n_2 and n_3, and the overtime parameter n_5, see equation (7).

Special treatment is also given to the out-of-work alternative since the number of non-working options deviates from the number of jobs in the market. Females not participating in the labor market might also have different preferences for working relative to females participating in the labor market. The value of n_0 is thus allowed to differ from the other values of n.

Equation (7) then implies that the probability of having a job with working time corresponding to interval j is given by

(9)
$$P_j = \frac{\exp(v(\tilde{H}_j, \tilde{C}_j) + \log(n_j))}{\exp(v(\tilde{H}_0, \tilde{C}_0) + \log(n_0)) + \sum_{r=1,2,3,5} \exp(v(\tilde{H}_r, \tilde{C}_r) + \log(n)) + \exp(v(\tilde{H}_4, \tilde{C}_4) + \log(n_4))},$$

where n is the common value of n_1, n_2 , n_3, and n_5, normalized to one in what follows. Thus, the terms log (n_j) cancel out for $j = 1, 2, 3, 5$, while estimation of n_0 and n_4 is carried out by introducing two dummy variables. The dummy variable

[14]See, for instance, Heckman and MaCurdy (1980) and Aaberge, Dagsvik, and Strøm (1995) for empirical analyses applying this specification.

D_4 is 1 for full-time work and 0 otherwise, whereas D_0 is 1 if the person works and 0 otherwise. In the specification of the likelihood function equation (9) is given by

$$P_j = \frac{\exp(v(\tilde{H}_j,\tilde{C}_j)+\delta_0 D_0+\delta_4 D_4)}{\exp(v(\tilde{H}_0,\tilde{C}_0)+\delta_0 D_0)+\sum_{r=1,2,3,5}\exp(v(\tilde{H}_r,\tilde{C}_r))+\exp(v(\tilde{H}_4,\tilde{C}_4)+\delta_4 D_4)}, \tag{10}$$

where $\delta_0 = \log(n_0)$ and $\delta_4 = \log(n_4)$. In the estimation of the model we let

$$Y_{ij} = \begin{cases} 1 & \text{if female } i \text{ has working time in interval } j \\ 0 & \text{otherwise} \end{cases}. \tag{11}$$

Then the logarithm of the likelihood function is given by

$$\log L = \sum_{i=1}^{N}\sum_{j=0}^{5} Y_{ij}\log P_{ij}, \tag{12}$$

where N is the number of females and P_{ij} is the probability that female i works in interval j.

Description of Data and Estimation Results

The IDS does not include information about wage rates or hours of work, but these data are available for a subsample through information from the Standard of Living Survey. The Standard of Living Survey 1995 provides data for about 3,700 individuals, representing the whole population.[15] Because the sample in the Standard of Living Survey 1995 is a subsample of the IDS-94, the two surveys can be linked on the basis of personal identification numbers, in order to connect information on working hours to data on incomes and taxes. Thus, the estimation of the female labor supply model and simulations of the behavioral effects from altering the child benefit scheme are based on a smaller sample of individuals than the calculations of the direct effects. After narrowing down the sample to married female wage-earners/home-workers older than 24 and younger than 65, the sample used in the estimation includes about 500 females. This sample also includes females without children. Descriptive summary statistics are reported in Table 1,[16] while Figure 5 describes the distribution of weekly hours of work among the females in the sample. The figure confirms that Norwegian females have a high participation rate, and that part-time is widespread.

Female pre-tax income corresponding to the discrete choices are derived by multiplying yearly hours in work in each alternative by measures of pre-tax (hourly) wage rates, while male income enters as exogenous income. Measures of female hourly gross wage rates are derived from data by dividing yearly labor income by yearly hours of work. For non-workers the pre-tax wage rates are

[15]The non-response ratio in Standard of Living Survey 1995 is about 25 percent. However, the non-response is most significant among the elderly and less pronounced among middle-aged individuals.

[16]A number of checks have been carried out in order to assess the representativity of this smaller sample of females in comparison with information from the larger sample in the IDS. There is a high degree of correspondence between the distribution of certain characteristics in the two samples, for instance with respect to age, education, and number of children.

TABLE 1

SUMMARY STATISTICS, 504 OBSERVATIONS

Variable	Mean	Standard Deviation
Consumption, household disposable income (NOK)	324,794	102,955
Age of wife	42.7	10.0
Education	11.8	2.35
Work experience	24.9	10.8
Number of children 0–2 years	0.18	0.42
Number of children 3–15 years	0.84	1.00
Full-time dummy	0.46	0.50
Participation dummy (participation rate)	0.86	0.35
Gross wage rate (NOK)	103.1	33.0
Working hours per week	27.0	14.1

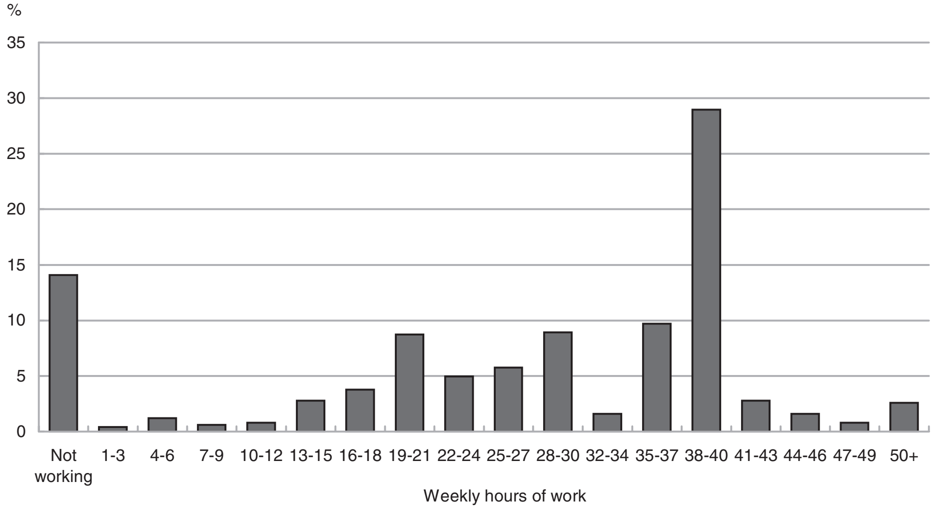

Figure 5. The Distribution of Weekly Hours of Work in Sample

estimated using a wage rate equation, where log(w) is the dependent variable and years of education, experience (measured as age minus years of education and minus pre-school years), and experience squared are the explanatory variables. The estimates are derived by a procedure that permits the possibility of selectivity bias, as suggested by Heckman (1979). However, no such effects are found, as the estimate of the inverse Mills ratio is non-significant. The error terms are ignored in the prediction of the wage rates.

Table 2 reports estimates of the parameters in the utility function and δ_0 and δ_4.[17] Recall that the parameters (α_1, γ_0) and (α_2, β_0, β_1, β_2, β_3, β_4) determine the utility of consumption and leisure respectively. The estimates of α_1 and α_2 are 0.79 and –7.78, respectively, which is consistent with our requirement that preferences be strictly quasi-concave in consumption and leisure. These parameters are determined quite precisely.

[17]A 12-state version has also been estimated with very similar results.

TABLE 2

ESTIMATES OF THE PARAMETERS IN THE UTILITY FUNCTION

Variables	Parameters	Estimates	t-statistic
Consumption	γ_0 (constant)	0.493	2.4
	α_1 (exponent)	0.791	6.6
Leisure	β_0 (intercept)	90.210	2.4
	β_1 (ln age)	–50.918	2.5
	β_2 (ln age squared)	7.248	2.5
	β_3 (# children <3 years)	0.755	2.1
	β_4 (# children 3–16 years)	0.633	2.6
	α_2 (exponent)	–7.782	6.6
Full-time opportunity index	δ_4	0.908	6.3
Participation opportunity index	δ_0	–1.042	5.4

The estimates of β_0, β_1, and β_2 indicate that women's preferences for leisure increase with age, but we cannot rule out that this may be a cohort effect. Preferences for leisure also increase with the presence of children, and the point estimates indicate that preferences for leisure are not much stronger if the children are young (β_3) than if they are older (β_4).

The estimates of the parameters related to the participation dummy (δ_0) and the full-time dummy (δ_4) are both statistically significant. If the estimate of δ_4 only reflects a difference in the number of offered full-time jobs from the number of part-time jobs, there would be about 2.5 times as many full-time jobs as part-time jobs available in the labor market. However, it can be argued that these estimates also reflect particular strong preferences for working full-time. With respect to the application of the parameter estimate, this ambiguity, whether δ_4 reflects preferences or number of job opportunities, does not affect simulation results, since we assume that the policy changes do not influence on estimates. Hence, we are assuming that habit formation plays no role in the determination of preferences; see for instance Neumark and Postlewaite (1998) and Woittiez and Kapteyn (1998).

The estimate of δ_0 is negative. In accordance with our model reasoning this might reflect more options in the non-work choice. There might also be other unobserved characteristics that reduce the utility of participation in the labor market; for instance fixed costs of employment and constraints in the market for child care.

To evaluate the model specifications Figure 6 displays the observed and the model-simulated distribution of weekly hours of work. The model-simulated distribution is found by calculating the average state probabilities across all individuals in the sample. In the calculation of the individual specific state probabilities we have used equation (10), but where the unknown parameters are replaced by their estimates. We see in the figure that the model reproduces the observed distribution quite well.

To put things into perspective, one might compare our labor supply elasticities (for married females) with the findings in other studies. According to our simulation results, the (average over females in sample) uncompensated wage elasticity is 0.47, whereas the income elasticity is –0.37. In another study based on Norwegian data by Aaberge, Colombino, and Strøm (2000), the female wage elasticity is

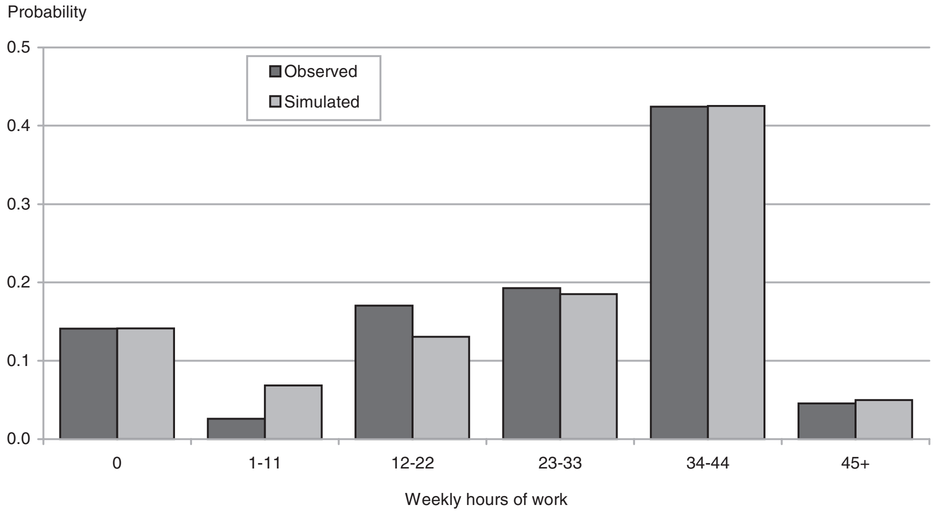

Figure 6. The Distribution of Observed and Simulated Hours of Work

higher (0.9), while the income elasticity is almost equal to zero (–0.06). They used a dataset for another year (1986), and the model specifications differ somewhat from those used in the present study. Blundell, Duncan, and Meghir (1998) find that the wage elasticity for married or cohabiting females in the U.K. is between 0.14 and 0.37, while the income elasticity ranges from 0 to –0.19. Note that differences in tax systems, labor market participation and income levels complicate the comparisons of elasticities across countries.

In the next section the estimated model is used to assess the labor supply responses from means-testing the child benefit.

5. Simulations of Female Labor Supply Responses

The simulations of labor supply responses are carried out only for the child benefit-eligible females, approximately 300 mothers. As noted above, the baseline (pre-reform) probability distribution is derived by calculating individual probability distributions and then calculating sample averages for each state. This pre-reform probability distribution can then be compared with a probability distribution under an altered tax-transfer system. Within our discrete choice approach means-testing the child benefit and redistributing the collected revenue as increased child benefit rates means that disposable income corresponding to a particular working time is increased for gross female wage income below some threshold level, whereas it is reduced for incomes above the threshold level. The probability of working a particular number of hours changes since the probability is a function of both hours of work and consumption (disposable income), see equation (10). The tax-benefit model is employed in order to calculate disposable income for each household, corresponding to the different states of labor supply and the various child benefit schemes.

TABLE 3

AVERAGE CHANGE IN STATE PROBABILITIES WHEN MEANS-TESTING THE CHILD BENEFIT

Weekly Hours of Work	Pre-reform State Probabilities	Change (in %) from Pre-reform Probabilities: Taxation	Income Testing, Alternative A	Income Testing, Alternative B
0	0.14	–0.9	10.9	17.8
1–11	0.07	0.3	8.9	14.2
12–22	0.13	3.5	5.3	5.6
23–33	0.18	2.0	0.6	–2.8
34–44	0.43	–1.5	–6.9	–8.9
45+	0.05	–3.5	–10.2	–10.9

TABLE 4

AVERAGE EXPECTED HOURS OF WORK PER WEEK UNDER VARIOUS CHILD BENEFIT SCHEMES

	Pre-reform	Taxation	Income Testing, Alternative A	Income Testing, Alternative B
Average expected hours of work	25.4	25.3	24.3	23.8
Change in expected hours of work (in %) w.r.t. pre-reform system	–	–0.4	–4.2	–6.1

In the case of taxation, the threshold level is independent of the income of the spouse, except for the dependency through the system of joint filing.[18] In the income testing alternatives the dependency on male income is stronger, as threshold levels depend on income for both spouses. Hence, female threshold levels decrease in male income.

Table 3 shows the simulated average pre-reform probabilities for the six labor market states and changes in the probabilities from taxing and income testing the child benefit and redistributing the collected revenue as increased child benefit rates. In Table 4 the results in Table 3 are summarized into the corresponding change in the average expected hours of work per week.[19]

We see that taxing the child benefit and redistributing the collected revenue reduces the probabilities of working full-time and not participating in the labor market whereas it increases the probability of working part-time, cf. Table 3. Thus, we would expect this reform to entail more part-time work and less full-time work, and a minor increase in the number of women participating in the labor market. Total expected labor supply is reduced by about 0.4 percent.[20]

With respect to the two income testing alternatives, recall that alternative A involves a 50 percent deduction of the child benefit for family incomes between NOK400,000 ($56,700) and NOK600,000 ($85,100) and no child benefit for incomes over NOK600,000. Alternative B means that the child benefit is reduced

[18]However, note that joint filing is only relevant when one spouse has very small income.

[19]The *i*-th female's expected hours of work is derived by multiplying her individual state probabilities by hours of work in each state, and then adding ($P^*_{i0}0 + P^*_{i1}6 + P^*_{i2}17 + P^*_{i3}28 + P^*_{i4}39 + P^*_{i5}50$).

[20]Uncertainty about parameter values and possible selection bias of the sample used in the simulations are ignored.

by 10 percent of a krone for every krone the family earns in excess of NOK250,000 ($35,500).

Both income testing alternatives imply a significant decrease in the probabilities of working full-time and more than full-time, and a significant increase in the probabilities of not working and working short hours, according to our simulations presented in Table 3. These schemes, and, in particular, the linearized deduction scheme, thus lead to a notable reduction in women's labor participation rate. As seen in Table 4, we expect female labor supply to be reduced by approximately 4 and 6 percent for alternatives A and B, respectively. Thus, changing the child benefit scheme to an income tested system reduces female labor supply, which limits resources available for market production.[21] In a perspective of population ageing in the near future, reforms that hamper labor supply might be considered inappropriate.

We notice that the labor supply effects are stronger when targeting the child benefit through income testing than through the taxation alternative, cf. Table 4. Considering our finding in Section 3 too, that income testing is redistributionally more efficient than taxing the child benefit, this result is in line with conventional understanding about the trade-off between distributional gains and labor supply effects.

6. Distributional Effects after Labor Supply Responses

The micro simulation approach applied here makes it feasible to consider the distributional effects of means-testing after labor supply responses. Since the labor supply effects from taxing the child benefit are very small, see Section 5, only the two alternatives of income testing are considered here. The impacts from the behavioral adjustments on income inequality are calculated as follows: for each family we calculate the (mathematical) expected difference between pre-reform equivalent income and equivalent income under the two alternatives of income testing. Similar to the calculation of expected hours of work (cf. description above), expected pre-reform post-tax income for the household is derived by multiplying the pre-reform state probabilities by pre-reform post-tax income belonging to each state, and adding. Conversely, post-reform post-tax income is obtained by replacing the pre-reform state probabilities by the probability distribution following from the income tested transfer system. Post-tax household income is converted into equivalent income as described in Section 3. The deviations between expected pre-reform and post-reform equivalent income are what we characterize as labor supply responses in Figure 7 and Figure 8. In addition to the actual deduction rates of the proposed changes, the second-round distributional effects from income testing depend on the relationship between male and female income across deciles and whether there are differences in the responsiveness of mothers in the high- and low-end of the income distribution scale. Since the ranking of mothers by labor supply effects differs from the ranking by equivalent income, one needs an empirical analysis to establish how the labor supply effects influence the distribution of equivalent income.

[21]Production at home might, however, increase.

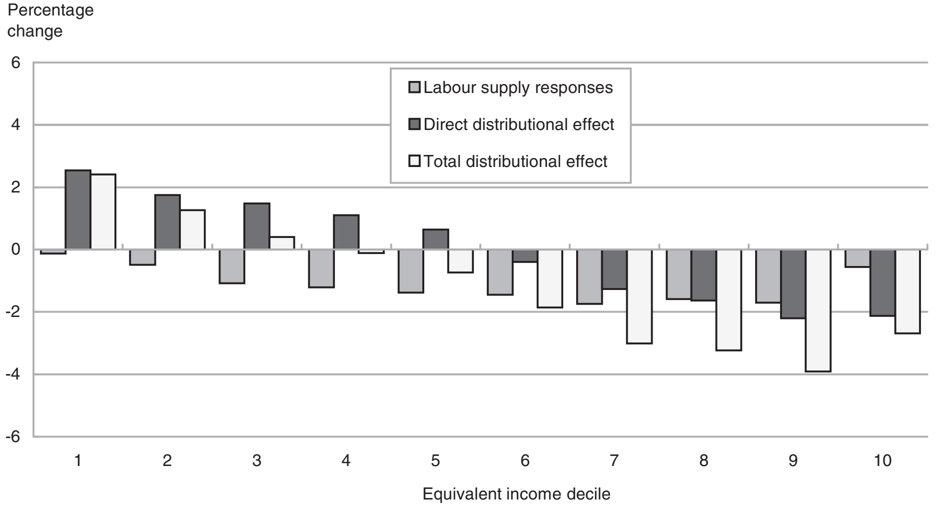

Figure 7. First-Round, Second-Round and Total Distributional Effects from Income Testing, Alternative A. Gain in Percent by Equivalent Income Decile

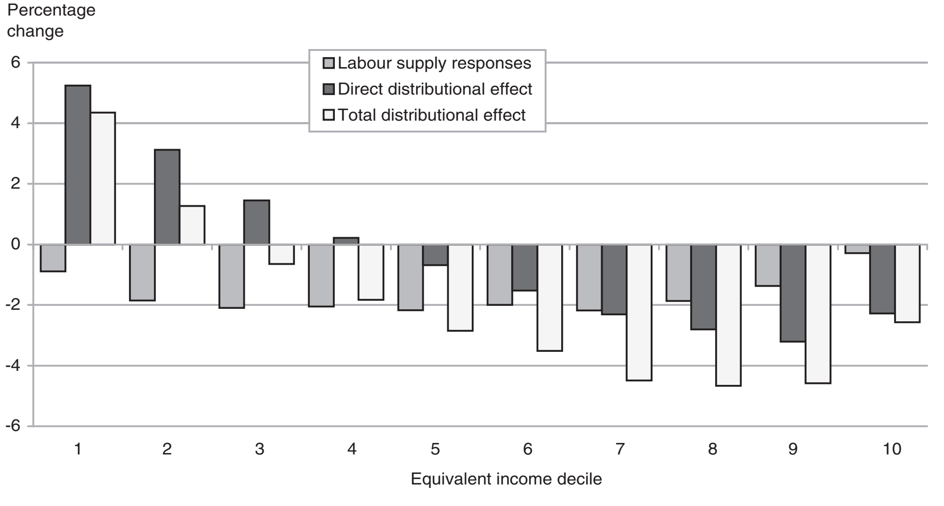

Figure 8. First-Round, Second-Round and Total Distributional Effects from Income Testing, Alternative B. Gain in Percent by Equivalent Income Decile

Figures 7 and 8 show the first-round (direct effects from Section 3), the second-round (labor supply responses) and the total distributional effects (sum of first-round and second-round) of income testing alternatives A and B, respectively.[22] We see that the labor supply responses reduce post-tax income for almost

[22]Remember that the labor supply responses are derived from a smaller sample of families, compared to the sample from which the direct distributional effects are calculated.

all income deciles, but the changes across deciles differ between the two income testing alternatives. In alternative A deciles 1 and 2 are less affected, while the reductions in equivalent income are largest in deciles 3–9. The total distributional effects (effects from labor supply responses + direct distributional effects) are more advantageous than the direct distributional effects only.

Alternative B yields largest income reductions from labor supply responses in the middle of the distribution. Thus, the second-round effects on income distribution are less pronounced under this alternative. Comparing Figures 7 and 8, however, we see that the total redistributional effect is larger under alternative B than under alternative A.

7. Concluding Remarks

The optimal design of a system of transfer to families with children appears to be a major policy issue in many economies, demonstrated for instance by discussions of the Working Families Tax Credit in the U.K. and the Earned Income Tax Credit in the U.S. In Norway a more targeted design of the scheme is frequently called for because the current program is expensive and seen as "inefficient" for poverty-alleviation purposes as it renders equal support to affluent and poor families alike.

Information from non-behavioral and behavioral simulations on micro data should be advantageous for decision-makers considering alterations in the child benefit scheme. With respect to the alternatives of means-testing evaluated here, we want to draw attention to firstly, that expectations of redistributional effects from means-testing might be overstated. For instance, we find that introducing a revenue neutral reform by taxing the child benefit as wage income and simultaneously increasing rates only has negligible effects on income inequality. Secondly, the alternatives of means-testing that imply noticeable redistributional effects (i.e. income testing) have considerable negative consequences for female labor supply. According to our simulation results, the mothers will cut average weekly working hours by approximately 4 percent and 6 percent, respectively, for the two modes of income testing. Thus, our results back the supposition that there is a trade-off between distributional considerations and labor supply concerns in such policy issues. Similar to analysis of the tax credits in the U.K. (Blundell *et al.*, 2000) and the U.S. (Scholz, 1996; Eissa and Hoynes, 1998) we find that transfers that phase-out with respect to income might have unfavorable effects on female labor supply. The results from non-behavioral and behavioral simulations on micro data help spelling out the pros and cons of alternative designs.

The impending ageing of populations and the shortage of labor this will bring about are generating widespread concern. There are reasons to expect an increased demand for female labor in the decades to come. Given this perspective one should be cautious about putting labor supply reducing policies into effect. As seen here, testing the child benefit against income reduces female labor supply. In this sense means-testing is costly, and maintaining a universal transfer system, which is less efficient with respect to poverty alleviation, might be preferred.

Indeed, it might be argued that one should engage in reducing barriers to female employment when considering new policies for support of families with children. As pointed out in this paper, changes in female labor supply have important effects on the distribution of income. This suggests that one should intensify the search for policies that encourage labor supply for the less well-off.

References

Aaberge, R., "Interpretation of Changes in Risk-Dependent Measures of Inequality," *Economics Letters*, 55, 215–19, 1997.

Aaberge, R., J. K. Dagsvik, and S. Strøm, "Labor Supply Responses and Welfare Effects of Tax Reforms," *Scandinavian Journal of Economics*, 97, 635–59, 1995.

Aaberge, R., U. Colombino, and S. Strøm, "Labor 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, 2000.

Atkinson, A. B., "On Targeting Social Security: Theory and Western Experience with Family Benefits," in D. van de Walle and K. Nead (eds), *Public Spending and the Poor: Theory and Evidence*, John Hopkins University and the World Bank, Baltimore and London, 25–68, 1995.

Besley, T., "Means Testing versus Universal Provision in Poverty Alleviation Programmes," *Economica*, 57, 119–29, 1990.

Blundell, R., "UK Taxation and Labour Supply Incentives," in A. B. Atkinson and G. V. Mogensen (eds), *Welfare and Work Incentives: A North-European Perspective*, Clarendon Press, Oxford, 135–60, 1993.

Blundell, R. and T. MaCurdy, "Labor Supply: A Review of Alternative Approaches," in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Volume 3A, Elsevier Science, North-Holland, Amsterdam, New York and Oxford, 1559–695, 1999.

Blundell, R., A. Duncan, and C. Meghir, "Estimating Labor Supply Responses Using Tax Reforms," *Econometrica*, 66, 827–61, 1998.

Blundell, R., A. Duncan, J. McCrae, and C. Meghir, "The Labour Market Impact of the Working Families' Tax Credit," *Fiscal Studies*, 21, 75–104, 2000.

Bradshaw, J., J. Ditch, H. Holmes, and P. Whiteford, "Support for Children. A Comparison of Arrangements in Fifteen Countries," Research Report No. 21, Department of Social Security, HMSO, 1993.

Buhmann, B., L. Rainwater, G. Schmaus, and T. M. Smeeding, "Equivalence-Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates Across Ten Countries using the Luxembourg Income Study (LIS) Database," *Review of Income and Wealth,* 34, 115–42, 1988.

Creedy, J., "Comparing Tax and Transfer Systems: Poverty, Inequality and Target Efficiency," *Economica*, 63, S163–S174, 1996.

———, "Means-Tested versus Universal Transfers: Alternative Models and Value Judgements," *Manchester School of Economic and Social Studies*, 66, 100–17, 1998.

Dagsvik, J., "The Asymptotic Distribution of Maximum of Independent and Nonidentically Distributed Random Variables," unpublished note, Statistics Norway, 2001.

Dickens, W. and S. J. Lundberg, "Hours Restrictions and Labor Supply," *International Economic Review*, 34, 169–91, 1993.

Duncan, A. and C. Giles, "Labour Supply Incentives and Recent Family Credit Reforms," *Economic Journal*, 106, 142–55, 1996.

Eissa, N. and H. M. Hoynes, "The Earned Income Tax Credit and the Labor Supply of Married Couples," NBER Working Paper 6856, National Bureau of Economic Research, Cambridge, MA, 1998.

Garfinkel, I. (ed.), *Income-Tested Transfer Programs: the Case For and Against*, Academic Press, New York, 1982.

Heckman, J. J., "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–61, 1979.

Heckman, J. J. and T. MaCurdy, "A Life Cycle Model of Female Labour Supply," *Review of Economic Studies*, 47, 47–74, 1980.

Ilmakunnas, S. and S. Pudney, "A Model of Female Labour Supply in the Presence of Hours Restrictions," *Journal of Public Economics*, 41, 183–210, 1990.

Immonen, R., R. Kanbur, M. Keen, and M. Tuomala, "Tagging and Taxing: The Optimal Use of Categorical and Income Information in Designing Tax/Transfer Schemes," *Economica*, 65, 179–92, 1998.

Jenkins, S., "The Measurement of Economic Inequality," in L. Osberg (ed.), *Readings on Economic Inequality*, Sharpe, New York, 3–38, 1991.

Lambert, P., "Evaluating Impact Effects of Tax Reforms," *Journal of Economic Surveys*, 7, 205–42, 1993.
Lundberg, S. J., R. A. Pollak, and T. J. Wales, "Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit," *Journal of Human Resources,* 32, 463–80, 1997.
Ministry of Children and Family Affairs, *Offentlige overføringer til barnefamilier*, NOU, Akademika, Oslo, 13, 1996 (in Norwegian).
Neumark, D. and A. Postlewaite, "Relative Income Concerns and the Rise in Married Women's Employment," *Journal of Public Economics*, 70, 157–83, 1998.
Sadka, E., I. Garfinkel, and K. Moreland, "Income Testing and Social Welfare: An Optimal Tax-Transfer Model," in I. Garfinkel (ed.), *Income-Tested Transfer Programs: the Case For and Against*, Academic Press, New York, 291–313, 1982.
Scholz, J. K., "In Work Benefits in the United States: The Earned Income Tax Credit," *Economic Journal*, 106, 156–69, 1996.
van Soest, A., "Structural Models of Family Labor Supply. A Discrete Choice Approach," *The Journal of Human Resources*, 30, 63–88, 1995.
Woittiez, I. and A. Kapteyn, "Social Interactions and Habit Formation in a Model of Female Labour Supply," *Journal of Public Economics*, 70, 185–205, 1998.