THE AUSTRALIAN ECONOMIC REVIEW

The Australian Economic Review, vol. 56, no. 3, pp. 328–354 DOI: 10.1111/1467-8462.12511

Alternatives to Paying Child Benefit to the Rich: Means-Testing or Higher Tax?

Ray Rees, Thor O. Thoresen and Trine E. Vattø*

Abstract

There appears to be a general movement away from universal child benefits and towards means-testing. In the present article we argue that instead of suppressing the labour supply of middle-income parents by withdrawing the transfer as a function of income, one should consider the alternative of financing a generous universal child benefit by increasing taxation of income. The implications of means-testing compared with a tax-financed universal alternative are discussed analytically in a piecewise linear schedule and by combining information from behavioural and non-behavioural micro-simulation models. Our results provide support for making child benefit universal instead of means-tested.

JEL CLASSIFICATION

J13; J22; C25

1. Introduction

As it is often argued that scarce resources should go primarily to those in need, providing child benefit support to the very rich can be seen as a waste. Views along these lines have been expressed by institutions such as the World Bank, the OECD and the European Commission. For example, the OECD (2011, p. 58) argues that in times of constraints on public budgets, it should be ensured that those most at risk do not emerge as the losers.

Several countries have adopted or are considering means-testing child benefit schemes. With the introduction of the American Rescue Plan Act of 2021 (Congressional Research Service 2021) the United States is effectively providing child benefit to families with children 0–17 years old,[1] but this is phased out for high-income earners. Similarly, the United Kingdom introduced the High Income Child Benefit Charge in 2013 where child benefit is tapered off for high-income levels. The child benefit of several other countries is means-tested too, for example in the schemes of Canada and Australia. The Australian Family Tax Benefit

* Rees: LMU, Munich, Germany; Thoresen: Research Department, Statistics Norway and the Department of Economics, University of Oslo; Vattø: Research Department, Statistics Norway. Corresponding author: Thoresen, email <tot@ssb.no>. Ray Rees sadly passed away in December 2022, before the revised version of this manuscript was compiled for resubmission. We are deeply grateful to Ray for sharing his profound knowledge of household economics with us—we feel privileged to have had Ray as our co-author. Financial support from the Norwegian Ministry of Children and Equality is gratefully acknowledged. We thank Sigvart Bretteville-Jensen for research assistance and Bard Lian and Zhiyang Jia for assisting in applying simulation models. We have benefited from comments to earlier versions of the article by two anonymous referees, Patricia Apps, Erling Holmøy, Axel West Pedersen, Terje Skjerpen and participants at the modelling tax policy and compliance workshop (Exeter, March 2019), the Norwegian research forum on taxation (Holmen, June 2019), the Norwegian–German seminar on public sector economics (Munich, December 2019) and the IIPF Annual Congress (August 2020).

is income-tested and linked to the Australian income tax system.[2] In the same vein the Norwegian discussion has emphasised that the universal child benefit, which has been nominally frozen for several years, has lost its impact, and it has proposed re-designing it as means-tested support directed at low-income families (Ministry of Children and Equality 2017).

The main message of the present article is that before establishing a means-tested scheme, one should evaluate the alternative of a universal child benefit scheme financed by increasing the general income tax. We compare these two alternatives by discussing two revenue-neutral schemes: either increasing the child benefit transfer, but cutting down on recipients through means-testing, or increasing the (universal) transfer, and financing the additional costs by increasing the income tax rate. The two alternatives are described analytically in a piecewise linear tax system and through the results of micro-simulation models. For the latter, we employ a behavioural micro-simulation tool for Norway to analyse how the two alternative designs for the child benefit scheme affect labour supply and the distribution of income and welfare.

In practice, we define the means-testing alternative by doubling the 2019 Norwegian child benefit rate, bringing it up to an annual total of NOK25,300 per child, which is equivalent to US$2,900.[3] Then we let the entire child benefit be phased out for household income above the average, to maintain revenue neutrality. The alternative to means-testing, which we refer to as a 'tax-financed universal scheme', uses the same start rate as for means-testing (NOK25,300), but it is offered as a universal rate to all families regardless of parental income. We let the increased expenditure of this universal scheme be financed by a 1.2 percentage-point increase in the labour income tax for all individuals.

Our results illustrate the drawbacks of letting child benefit be means-tested according to household income, finding that parents in the middle of the income distribution in particular face the highest labour supply disincentives. This illustrates a rather general result of means-testing: excluding the well-off from obtaining support is costly in terms of reduced labour supply incentives for households in the middle of the income distribution. When the labour supply of females is relatively more elastic than that of males, a large part of this response is due to mothers reducing their working hours. We find that the tax-financed universal child benefit scheme is less costly in terms of reduced labour supply than the means-testing scheme, and it also redistributes income to a greater extent.

In addition to providing analytical and empirical evidence on the consequences of the child benefit design, our article also contributes to the broader literature on the design of transfer policies. First, one strand of the literature on optimal taxation emphasises that the existence of children or household size can be used as a tagging device for earning ability, along the lines of Akerlof (1978), addressing the screening problem of governments.[4] For example, Immonen et al. (1998) and Blumkin, Margalioth and Sadka (2015) discuss the optimal design for tax/transfer schemes which involves elements of both tagging and means-testing. However, it is acknowledged, not surprisingly, that the optimal design of transfer programs depends on the relevant empirical evidence and on how society trades off gains to some individuals against losses to others (Acs and Toder 2007; Kaplow 2007). Our study complements these studies by describing the trade-offs between redistribution and labour supply distortions for various child benefit scheme designs.

Second, the present study adds to the literature on the effects of family transfers and their design on the labour supply of parents (Kornstad and Thoresen 2004; Milligan and Stabile 2009; Gonzalez 2013; Schirle 2015; Hener 2016; Bibler, Guettabi and Reimer 2019). For example, a substantial number of studies address labour supply effects of the Earned Income Tax Credit of the United States and various versions of the same type of support in the United Kingdom (Hotz and Scholz 2003; Blundell 2006;

Brewer, Saez and Shephard 2010; Brewer and Hoynes 2019). As these transfers are phased out according to income, the analysis of the present study parallels the considerations of this literature.

Third, although we do not evaluate child outcomes, our study connects to the literature that discusses the effect of transfers on the well-being of children and families (Baker, Gruber and Milligan 2008; Atkinson 2015; Hener 2016; Burton and Phipps 2017; Hoynes and Patel 2018; Hendren and Sprung-Keyser 2020; Aizer, Hoynes and Lleras-Muney 2022). Our behavioural modelling approach does not include a childbirth choice component. There is a strand of literature on how family policies and child benefits affect fertility (see, for example, Doepke et al. 2022) but such effects are beyond the scope of the present analysis.

The rest of the article is organised as follows. Section 2 describes the institutional background and the debate on child benefit policy in Norway. In Section 3 we consider the means-testing problem from an analytical perspective, placing the universal and the means-tested child benefit in a piecewise linear tax scheme. In Section 4 we present the micro-simulation models that are applied to describe the effects on income and labour supply of the two alternative child benefit schemes: a detailed (non-behavioural) tax-benefit model and a discrete choice labour supply model. Then in Section 5 we first describe the benchmark scheme and explain how we have designed a means-tested scheme and an equally generous universal scheme financed by increased taxation of income. Second, we describe the labour supply responses of the two alternative schemes, leading up to a discussion of distributional effects, presenting both direct (non-behavioural) and total distributional effects after taking account of labour supply responses. Finally, we show robustness checks of the specific design of the means-tested scheme and the choice of equivalence scales for comparing income across individuals and households. Section 6 provides a conclusion for the article.

2. Institutional Background

Norway is a country with an active family policy, orientated towards families being able to combine having children and maintaining a strong connection to the labour market for both genders, the so-called dual-earner model. Encouraging both genders to participate in the labour market is also reflected in the personal income tax system, in which taxation is based on individual income rather than household income.[5]

However, Norwegian family policy has been an arena for substantial political controversy over the last couple of decades, and certainly not all new directions can be seen as promoting the dual-earner family model. In particular, the cash-for-care reform, which was introduced in 1998, generated a heated debate on the rationalisations and directions of family policies. The reform introduced monetary compensation for not using subsidised care at child-care centres, for parents of children aged one and two. The three main aims of the reform were to give parents more time to care for their own children, to give families freedom of choice with regard to care provider and to equalise public support to families, irrespective of the care option chosen (Ellingsæter 2003). The support equalisation argument was strengthened by the fact that access to subsidised care in centres at that time (late 1990s) was severely constrained.

Since then, there has been massive expansion in the child-care centre participation rate in Norway, particularly for children under three years of age. Policy-makers formalised their efforts to increase the supply of centre-based care through the so-called ‘childcare compromise’, approved by the Storting (Norwegian parliament) in 2003. The agreement included a plan for eliminating queues for care at child-care centres, and introduced a substantial reduction in child-care fees, regulated by a maximum monthly parental fee. By 2009, the policy initiative had resulted in a market for centre-based care which enabled the government to guarantee all families of children older than one access to a slot at a centre. The parental fee covers approximately

14 per cent of the actual costs for children under three, and approximately 25 per cent for children aged three to five (Lunder 2015). In 2019, the maximum monthly fee was set at NOK3,040 (US$356).

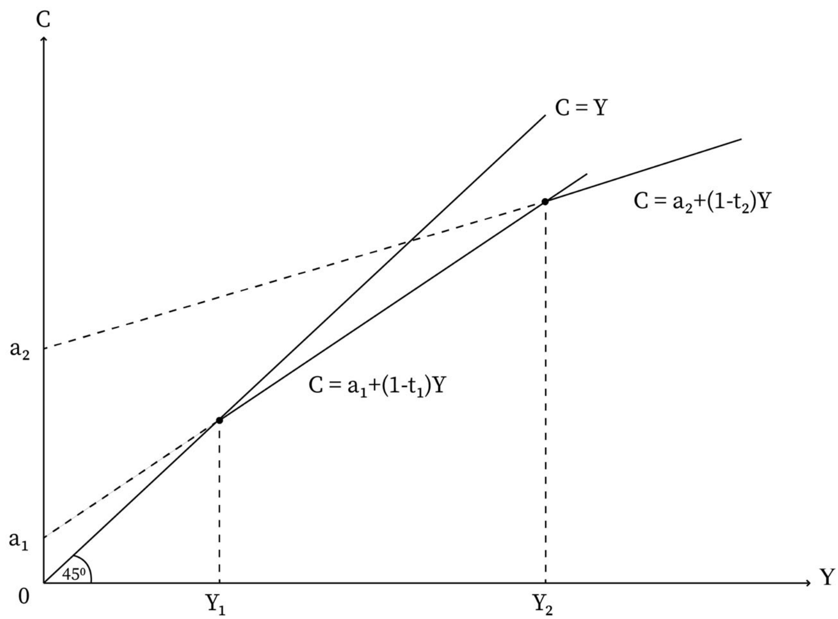
Norway has a relatively modest child benefit support scheme with a benefit rate of NOK1,054 (US$120) per month for each child under 18 years old in 2019. The relatively small support in recent years is a consequence of the policy-makers keeping child benefit nominally frozen in the period 1996–2018 in order to finance extensive development of subsidised child-care services. As a result, the benefit for one child amounted to almost 4 per cent of the median household income for families with children in 1996, but only 1.5 per cent in 2019. Partly because of this reduced impact over time, means-testing has been discussed as a way of using scarce resources more efficiently. As an example, in a recent White Paper (Ministry of Children and Equality 2017), a group of experts was split on whether to means-test or not, with the majority arguing that child benefit should be redesigned as means-tested support directed at low-income families.

The debate on means-testing of child benefit in Norway and elsewhere has motivated us to evaluate the effects of means-testing in comparison to other designs for support schemes. Specifically, we discuss two main alternatives: either increasing child benefit support while reducing the number of recipients through means-testing based on household income, or increasing the (universal) transfer and financing the additional costs by increasing the general labour income tax rate. Our main focus is on the effects of the two schemes on labour supply and their implications for the distribution of income. We provide further details of the two alternative child benefit designs in Section 5.1.

3. Means-Tested Child Benefit in a Piecewise Linear Tax System

In the following, we describe means-tested child benefits in the context of a piecewise linear tax system.[6] The system assigns each tax unit—in the Norwegian context a single individual—to one of a number of specified tax brackets on the basis of the level of their taxable income. Figure 1 illustrates this by plotting disposable

Figure 1 A Piecewise Linear Tax System



income or, in the absence of saving, consumption, C, as a piecewise linear function of the tax base, taxable income Y, which gives all taxpayers a convex budget set in the (Y, C) plane.[7]

In Figure 1, tax units with income less than Y_1 pay no tax, $T_0 = t_0 Y$ where $t_0 = 0$, those with income in the interval (Y_1, Y_2) pay $T_1 = t_1(Y - Y_1)$, and those in the interval (Y_2, ∞) pay $T_2 = t_2(Y - Y_2) + t_1(Y_2 - Y_1)$, which defines the tax system in terms of a set of pairs of marginal tax rates and upper bracket limits $\{(t_0, Y_1), (t_1, Y_2), (t_2, \infty)\}$ where $t_0 = 0$. An equivalent alternative interpretation is to define the tax system as a set of pairs of bracket-specific lump sum transfers and marginal tax rates that are applied to the tax unit's total income, the set of pairs $\{(t_0, a_0), (t_1, a_1), (t_2, a_2)\}$, with $a_0 = 0$, $a_1 = t_1 Y_1$ and $a_2 = (t_2 - t_1)Y_2 + a_1$. The budget constraint is then simply defined as

$$C = Y - T_j = a_j + (1 - t_j)Y, \; j = 0, 1, 2 \qquad (1)$$

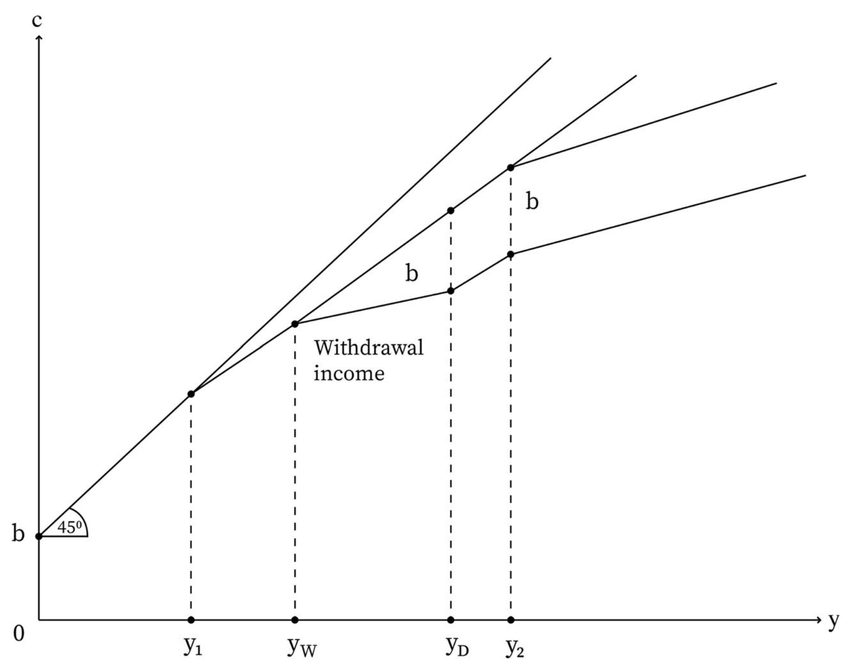
as shown in Figure 1, with a_0, $t_0 = 0$.[8] It follows from this that any universal transfer can be clawed back by an appropriately designed piecewise linear tax system.

The effect of the child benefit on households receiving it can be represented by simply shifting the bracket-specific lump sum transfers upward by b units as illustrated in the upper function of Figure 2. Since the presence and age of children can be observed, this is a 'tagged' transfer, available only to households with children under 18 years old.

On the horizontal axis of Figure 2, we measure an individual earner's income,[9] and use y_1, y_2 to denote the initial bracket limits on individual incomes in the absence of means-testing.

Suppose that it is now decided to means-test the child benefit and phase it out over the income range $y_W - y_D$ by choosing an appropriate withdrawal rate, $t_w > 0$, which is just sufficient to reduce the lump sum benefit to zero at y_D. Adding this to the marginal tax rate over this range gives the new effective tax rate $t_1 + t_W$. In the case shown this adds two additional tax brackets, which replace the portion of the previous scheme over the range

Figure 2 A Piecewise Linear Tax System with Transfers and Means-Testing



(y_W, y_2). In the figure as drawn, this implies that the marginal tax rate for taxpayers receiving child benefit is not only higher than that for non-recipients of child benefit in the (y_W, y_D) bracket, but also exceeds the top marginal tax rate t_2.[10] Obviously the narrower the bracket (y_W, y_D) and the higher the child benefit b, the higher t_W will be.

Means-tested child benefit thus creates a non-convexity in the budget set, and involves changes in the effective marginal tax rates which are expected to affect labour supply incentives and the distribution of disposable income. In the following, we use non-behavioural and behavioural micro-simulation models to provide empirical evidence of labour supply effects of means-testing and the accompanying distributional effects.

4. Using Micro-Simulation Tools

4.1 The Tax-Benefit Model

Statistics Norway's tax-benefit model LOTTE-Skatt (Aasness, Dagsvik and Thoresen 2007) is a detailed tax calculator used to simulate the direct (non-behavioural) effects of tax and benefit changes on tax revenues and distributional outcomes. The version of the model used in the present study is based on extensive information on individuals and households based on administrative register data, including detailed information from income tax returns, for the Norwegian population in 2019. To reduce computational time we use a representative sample with accompanying population weights to represent the entire population.

4.2 Behavioural Effects of a Discrete Choice Labour Supply Model

We use a behavioural micro-simulation model based on a discrete choice framework to simulate labour supply responses of alternative tax and transfer schemes. In the category of structural labour supply modelling approaches, the discrete choice labour supply model based on the random utility modelling approach (van Soest 1995) has gained widespread popularity among public finance practitioners (Creedy and Kalb 2005). This type of model can handle non-convex budget sets and two-earner households.

The labour supply model applied here is based on a particular type of discrete choice model denoted as the 'job choice model' (Dagsvik et al. 2014; Dagsvik and Jia 2016). According to this framework, labour supply decisions are viewed as the outcomes of individuals choosing among jobs, with additional constraints on the set of available jobs. The job choice model is specified as follows: each individual is assumed to have preferences within a set of 'jobs', where each market job (indexed by $z = 1, 2, ...,$) is characterised by disposable income $C(z)$, hours of work $h(z)$ and other non-pecuniary job attributes such as job-specific tasks to be performed, workplace locations and working environment quality. Disposable income for a given job is defined as $C(z) = f(h(z)w(z), I)$, where $w(z)$ is the offered wage rate for the given job z, I is non-labour income and $f(\cdot)$ is the net-of-tax function. The offered wage rate w is assumed to be constant across jobs for a given individual.[11] The individual's utility of choosing job z is represented as $U(C, h, z)$, where the utility function is assumed to be additively separable, that is $U(C, h, z) = v(C, h) + \varepsilon(z)$.

The sets of available jobs from which the individuals choose are individual-specific. Dagsvik and Jia (2016) show that it is sufficient to identify the model by introducing a latent measure of job opportunities representing the number of available jobs for a given working time option h, $m(h)$, where the number of non-working opportunities is normalised to one, that is, $m(0) = 1$.

Hours of work for each job take a value within a given set H. It can be shown that applying the assumption of i.i.d. extreme value distributed error terms, $\varepsilon(z)$, the probability of a worker choosing one of the jobs with working time $h \in H$ can be written as

$$\varphi(h) = \frac{m(h)\exp(v(C(h), h))}{\sum_{x \in H} m(x)\exp(v(C(x), x))} = \frac{\exp V(C(h), h)}{\sum_{x \in H} \exp V(C(x), x)} \quad (2)$$

This expression is analogous to a multinomial logit model with payoff $V(C(h), h) = \nu(C(h), h) + \log m(h)$.

The preference parameters and parameters for individuals' wage rates are estimated by the method of maximum likelihood, combining (cross-sectional) information from the Income and Wealth Statistics for Households (Statistics Norway 2018a) and the Labour Force Survey (Statistics Norway 2018b). The model is estimated for prime aged (25–62 years old) wage earners, separately for couples, single females and single males. For the present study it should be noted that the wage earner couple is seen as a single decision-making unit.[12] More details of the empirical specification, the data used and estimation of the model are provided in Section A in the Appendix.

Given the probabilistic nature of the labour supply model (follows from the random utility framework of discrete choice), we obtain a predicted probability distribution for the discrete set of working hours alternatives for each individual which is altered when the budget constraint changes. The labour supply responses we report are based on changes in the expected hours of work for each individual.[13]

Table 1 presents our estimated labour supply elasticities with respect to the wage rate for females and males in couples and for male and female singles, at both the extensive and the intensive margin. In particular, the model predicts that females in couples are more responsive than others to changes in both the wage rate and the (effective) marginal tax rate.[14] This is a general finding in the labour supply literature (Whalen and Reichling 2017; Eckstein, Keane and Lifshitz 2019; Keane 2022),[15] and the elasticity estimates are in line with other comparable microsimulation studies; see for example, Thoresen and Vattø (2015), Bargain and Peichl (2016) and de Boer and Jongen (2023). Estimation results (not reported here) also show standard income response regularities, with negative labour supply response to increased non-labour income (child benefit), but the effects are small.

As the labour supply model is designed to provide simulation results for prime aged (25–62 years old) wage earners, labour supply effects in other groups are set equal to zero. In effect, this means that the responses of the self-employed are neglected. A tax simulation model for the self-employed requires a different decision model. It is worth noting, however, that the share of self-employed in the total workforce is low in Norway, around 7 per cent (Berg and Thoresen 2020).

5. Simulation Results—Means-Testing or Higher Tax?

5.1 Description of the Benchmark and the Two Alternative Schemes

We use the income tax and child benefit system of Norway in 2019 as a benchmark for

Table 1 Simulated Labour Supply Elasticities with Respect to the Wage Rate for Individuals in Couples and Singles

	Own wage		*Cross-wage*	
	Female	*Male*	*Female*	*Male*
Individuals in couple				
Participation (ext. margin)	0.135	–	–0.048	–
Hours cond. on working (int. margin)	0.197	0.095	–0.043	–0.009
Total elasticity	0.332	0.095	–0.091	–0.009
Single individuals				
Participation (ext. margin)	0.012	–		
Hours cond. on working (int. margin)	0.057	0.009		
Total elasticity	0.069	0.009		

Note: The elasticities reflect the simulated percentage change (average across individuals) in the probability of participation (extensive margin) and working hours conditional on working (intensive margin) when the hourly wage rate is increased by 1 per cent for all wage earners.

our empirical results. In the benchmark scheme, parents (usually the mother) are entitled to child benefit for each child under the age of 18. In 2019 the recipients received NOK1,054 (US$120) per child per month, which means that the annual support was NOK12,650 (US$1,440). In addition, single parents are entitled to extended child benefit and infant supplement. Extended child benefit means receiving benefit for one child more than the parent actually has, while an extra infant supplement, which was NOK660 (US $75) in 2019, is paid for children zero to three years of age.

Norway has a dual income tax system, with a flat tax rate on all taxable income (net of deductions), as well as a graduated four-tier tax rate structure (so-called bracket tax) on labour income.[16] In 2019, the individual's labour income was taxed at a flat rate of 22 per cent, as well as a social security contribution of 8.2 per cent, in addition to the bracket tax; with the following additions corresponding to four different income ranges (NOK174,500–245, 650, NOK245,650–617,500, NOK617,500–964,800 and above NOK964,800): 1.9 per cent, 4.2 per cent, 13.2 per cent, and 16.2 per cent. Thus, the top marginal tax rate on labour income for wage earners was 46.4 per cent in 2019. Histograms describing the distributions of observed individual labour income and household income, for households with and without children, in 2019, are provided in Figure 3.

Table 2 summarises the characteristics of the benchmark scheme (2019) and the two alternative child benefit designs—the means-tested and the tax-financed universal schemes. The total cost of the child benefit scheme is NOK15.5 billion (US$1.8 billion) in the benchmark, which corresponded to approximately 3 per cent of the revenue from income taxes in that year. The means-testing alternative is defined by doubling the Norwegian child benefit rate of 2019, bringing it up to an annual total of NOK25,296 per child, which corresponds to US$2,900. Then we let the entire child benefit be phased out at a rate of 10 per cent for household income above NOK756,000 (US$86,700), to maintain revenue neutrality.[17] It follows that the length of the phase-out interval is determined by the size of the support, which depends, on how many children there are in the family. Given the benchmark income distribution, 28 per cent of the households with children have household income below the phase-out range, 12 per cent are in the phase-out range and 60 per cent have income above the phase-out range. See more details on the distribution of income for households with children in Panel C of Figure 3. The alternative to means-testing, which we refer to as the tax-financed universal scheme, uses the same initial child benefit rate as for means-testing (NOK25,296), but now it is offered at a universal rate to all families regardless of parental income. In order to make this alternative revenue-neutral, increased expenditure is financed by increasing the rates in the four-tier tax scheme for individual labour income by 1.22 percentage points. As the increased taxation affects all individuals with individual labour income above the first bracket of NOK174,500, around 95 per cent of the adult Norwegian population is involved in financing this alternative.

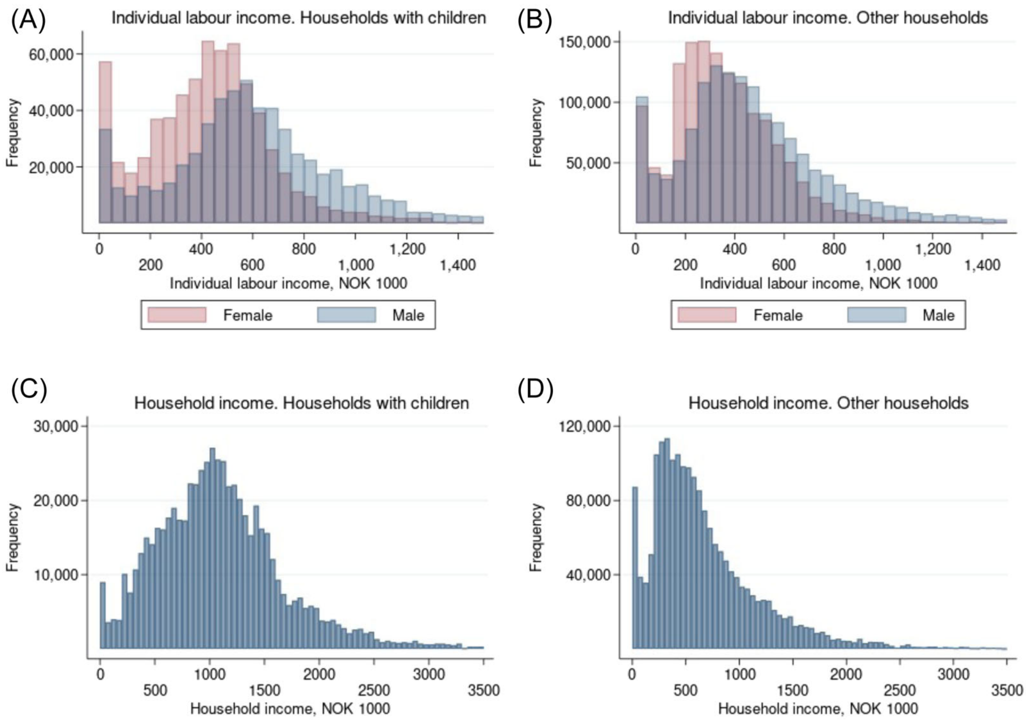
5.2 Labour Supply Responses

5.2.1 Aggregated Labour Supply Responses and Effects on Revenue Balance

The upper part of Table 3 summarises the labour supply effects of the two alternatives (compared to the benchmark), for wage-earner households.[18] We find that the aggregated reduction in working hours is considerably larger for the means-testing alternative, 0.32 hours on average, compared to 0.20 hours on average for the tax-financed universal scheme. Converted into reductions in man-years, these effects correspond to withdrawals from market work of approximately 13,000 man-years and 8,000 man-years, respectively.

In the lower part of Table 3 we draw attention to how the estimated labour supply changes affect tax revenues and child benefit

Figure 3 Distribution of Individual Labour Income and Household Income in the Benchmark.



Note: Panel A and Panel B show histograms of individual labour income (basis for the four-tier tax rate) for adult males and females (at least 18 years old) belonging to households with children (under 18 years old), and other households, respectively. Individual labour income here also includes income that replaces earnings, such as pension, sick pay, disability and unemployment benefits. Panel C and Panel D show histograms of household income (basis for means-tested child benefit) for households with children (under 18) and other households, respectively. Household income includes individual labour income for all household members, as well as positive capital income. The frequencies in Panel A and Panel B refer to the number of individuals, whereas the frequencies in Panel C and Panel D refer to the number of households.

expenditures. We see that the behavioural effects of means-testing weaken the budget more than the universal tax-financed alternative. Under means-testing, income tax revenues are reduced by NOK2.4 billion (US$0.27 billion) because of reduced work, and this, in turn, results in increased child benefit expenses of NOK0.4 billion (US$45 million). The latter is due to more households falling below the means-testing threshold when households reduce their labour supply. The total effect on the budget is then NOK−2.8 billion (US$−0.32 billion). In comparison, the reduction in tax payments due to the labour supply responses under the tax-financed universal child benefit alternative is lower, NOK1.5 billion (US$0.17 billion).

5.2.2 *Heterogeneity in Labour Supply Responses*

Table 4 demonstrates that the labour supply effects differ with respect to gender. The table provides simulation results for subgroups—females and males in couples, single households with children, and other households. It shows that females in couples with children, in particular, reduce their labour supply under means-testing: on average they reduce their working hours by 1.2 hours per week (extensive margin+intensive margin). In

Table 2 Descriptions of Benchmark and Alternative Child Benefit Schemes

	Benchmark	*Means-testing*	*Tax-financed universal*
Child benefit scheme			
Benefit rate (per child)	12,650	25,296	25,296
Inc. threshold (household income)	–	756,000	–
Phase-out rate of benefit	–	0.10	–
Tax rate change (all brackets)	–	–	+1.22 pp.
Revenue balance			
Income tax revenue (NOK bn)	525.5	525.5	541.0
Child benefit expense (NOK bn)	15.5	15.5	31.0
Revenue balance (NOK bn)	510.0	510.0	510.0

Note: The child benefit rate and the household income threshold are measured in NOK (per annum). The benchmark corresponds to the 2019 tax-benefit system of Norway, with a four-tier scheme for personal income tax and a (low-rate) universal child benefit. Both the means-testing and the tax-financed universal scheme are revenue-neutral compared to the benchmark (before labour supply adjustments).

Table 3 Aggregated Labour Supply Responses and Effects on Revenue Balance of the Means-Tested and Tax-Financed Universal Child Benefit Schemes

	Benchmark		*Change*			
			Means-testing		*Tax-financed universal*	
	Hours per week	*Man-years (in 1000)*	*Hours per week*	*Man-years (in 1000)*	*Hours per week*	*Man-years (in 1000)*
All wage earners	34.56	1,406	−0.32	−13	−0.20	−8
Households with children	34.61	625	−0.73	−13	−0.33	−6
Mothers	31.90	319	−1.08	−11	−0.48	−5
Fathers	37.99	306	−0.30	−2	−0.14	−1
Other households	34.52	781	–	–	−0.10	−2

		Change			
		Means-testing		*Tax-financed universal*	
	Benchmark	*Direct*	*Total*	*Direct*	*Total*
Revenue effects					
Income tax revenues (NOK bn)	525.5	0.0	−2.4	15.5	14.0
Child benefit expenses (NOK bn)	15.5	0.0	0.4	15.5	15.5
Revenue balance (NOK bn)	510.0	0.0	−2.8	0.0	−1.5

Note: Hours refer to mean working hours per week. A man-year is defined by calculating annual hours of work corresponding to working 37.5 hours per week. Wage earners are divided into two categories, depending on whether they belong to a household where there are children (<18) or not. The total revenue effect includes predicted labour supply responses.

contrast, the effect on fathers' labour supply (for the same policy change) is much lower, a reduction of 0.32 hours per week.

Table 5 further describes female response heterogeneity for mothers for combinations of wage levels for themselves and their spouses:[19] nine combinations of low (L), medium (M) and high (H) wages. While the labour supply effects of the tax-financed universal schemes are relatively equally

Table 4 Labour Supply Effects of Changes in Child Benefit by Wage Earner Groups

				Households with children				
				Couple		*Single*		
			All	*Mothers*	*Fathers*	*Mothers*	*Fathers*	*Other households*
	Benchmark	Participation	0.970	0.953		0.940		0.969
		Work. hours, int. marg.	35.49	33.36		33.83		35.49
		Uncond. working hours	34.56	31.87	38.04	31.98	37.53	34.52
Change	Means-testing	Participation	–0.003	–0.015		–0.010		–
		Work. hours, int. marg.	–0.23	–0.76		–0.47		–
		Uncond. working hours	–0.32	–1.20	–0.32	–0.77	–0.13	–
	Tax-financed universal	Participation	–0.002	–0.004		–0.010		–0.001
		Work. hours, int. marg.	–0.15	–0.26		–0.44		–0.08
		Uncond. working hours	–0.20	–0.38	–0.15	–0.75	–0.10	–0.10
	Number of wage earners (in 1,000)		1,526	270	270	105	32	849

Note: Changes measured in absolute values of participation or working hours per week.

Table 5 Effects on Female Labour Supply of Changes in the Child Benefit Schedule for Combinations of Female–Male Wage Levels in Couples with Children

			Wage combinations: Female/male wage rate levels								
			L/L	*L/M*	*L/H*	*M/L*	*M/M*	*M/H*	*H/L*	*H/M*	*H/H*
	Benchmark	Participation	0.927	0.924	0.919	0.965	0.958	0.951	0.981	0.974	0.967
		Work. hours, int. marg.	32.21	32.23	32.05	33.79	33.49	33.24	34.71	34.34	34.05
		Uncond. working hours	29.95	29.83	29.50	32.66	32.14	31.67	34.07	33.48	32.97
Change	Means-testing	Participation	–0.023	–0.022	–0.019	–0.013	–0.015	–0.016	–0.008	–0.010	–0.009
		Work. hours, int. marg.	–1.00	–0.85	–0.62	–0.88	–0.79	–0.66	–0.65	–0.67	–0.46
		Uncond. working hours	–1.60	–1.46	–1.15	–1.27	–1.23	–1.13	–0.90	–1.00	–0.74
	Tax-financed universal	Participation	–0.007	–0.006	–0.005	–0.003	–0.004	–0.004	–0.002	–0.003	–0.003
		Work. hours, int. marg.	–0.31	–0.28	–0.26	–0.26	–0.26	–0.26	–0.23	–0.24	–0.23
		Uncond. working hours	–0.49	–0.44	–0.40	–0.36	–0.36	–0.37	–0.29	–0.32	–0.32
Number of wage earners (in 1,000s)			29.3	31.1	7.1	31.5	72.6	30.9	6.5	31.3	29.5

Note: Wage rate levels (per hour) defined by percentiles: L, low wage, 0–25; M, medium wage, 25–75; H, high wage 75–100. Changes measured in absolute values of participation rate or working hours per week.

distributed across wage combinations, the table shows that means-testing has the strongest labour supply-reducing effects for combinations involving low wage levels; see, for example, the effects in the case where both spouses have low-income (L/L). These results therefore suggest that means-testing is particularly harmful for the labour supply of low-wage households, and in that sense can be seen as tending towards a poverty trap, inducing families to move to positions with less market work.

5.3 Distributional Effects

5.3.1 Direct and Total Distributional Effects on Income

In the description of distributional effects we first use income as our measure of well-being, before showing results when money metric utility is used as a welfare metric in Section 5.3.2. We measure distributional effects over the entire population.

The measure of income is based on equivalent disposable income, derived by aggregating income over all individuals in the household and adjusting to account for household size and economies of scale. In the following we use the square root equivalence scale, which is commonly used, as for example in recent OECD publications (OECD 2023). Results for other choices of equivalence scale are shown in Section 5.4.

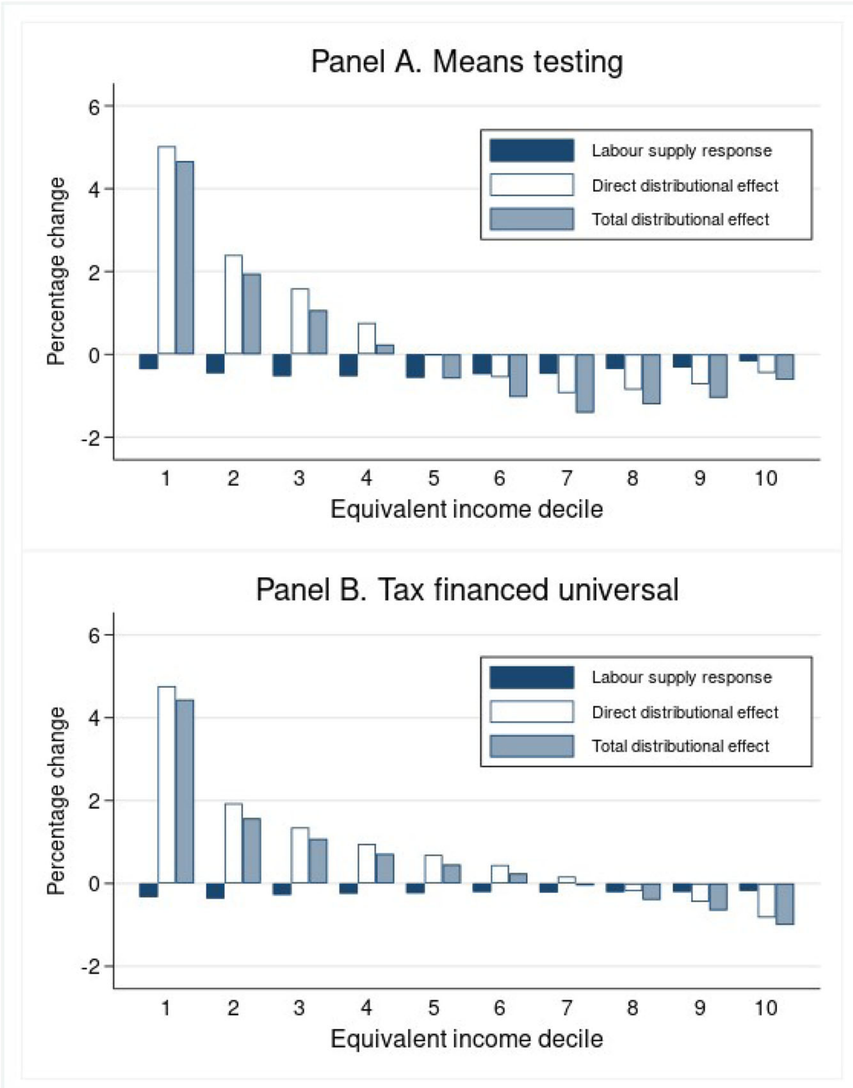
The individual is the unit of analysis. This means that the equivalent income of each household is represented *n* times, where *n* is the number of individuals in the household, see Ebert (1997).

In Figure 4 we divide the population into deciles, ranging from lowest to highest equivalence scale-adjusted disposable household income in the 2019 benchmark. Panel A describes the percentage change in (equivalence scale-adjusted) disposable income for the means-tested alternative, compared to the 2019 benchmark. Similarly, Panel B illustrates the effect of the tax-financed universal scheme, compared to the 2019 benchmark. We describe the distributional effects of changes in the child benefit scheme both before (direct distributional effects) and after behavioural effects (total distributional effects). Thus, we see the extent to which the labour supply effects modify the initial direct (non-behavioural) effects of changing the child benefit scheme. Given the probabilistic nature of the labour supply model (follows from the random utility framework of discrete choice), we arrive at income responses by taking expectations across the discrete choices of each individual.

The figure illustrates the main findings of this study. First, we see that the negative effect on income resulting from reduced labour supply is larger for the means-tested scheme than for the tax-financed universal scheme. However, as expected, the direct distributional effects have stronger effects on incomes than the effects of labour supply adjustments. Second, the figure reveals that there is a clear difference in the direct distributional effects (in the absence of any behavioural effects) between the means-tested and the tax-financed universal alternatives, compared to the 2019 benchmark.[20] Negative effects on income (both direct and total) are observed for individuals in deciles 5–10 under means-testing, whereas for the tax-financed universal alternative the negative effects start at higher income levels, in deciles 8–10. This shows that the means-tested alternative, in contrast to the tax-financed universal scheme, not only reduces the income of the rich, but negatively affects incomes of the middle-income deciles. For the tax-financed universal scheme, the redistribution is more evenly spread across the income distribution and the losses are concentrated in the top three deciles.

Table 6 presents the direct and total distributional effects for a range of aggregate measures of inequality: the Gini coefficent,

Figure 4 Direct Distributional Effect, Labour Supply Response and Total Distributional Effect of Changes in Child Benefit, All Households



Note: Incomes are measured and households are ranked in terms of equivalent (square root scale) household income in 2019, with the individual as the unit of analysis. The effects of the alternative designs are measured as differences from the 2019 benchmark. Given the probabilistic nature of the labour supply model (follows from the random utility framework of discrete choice), we arrive at income by taking expectations across the discrete choices of each individual. In the ranking of individuals we use observed (equivalence scale-adjusted) income, such that the comparison of effects is carried out for a fixed decile ranking. The direct distributional effect refers to the (non-behavioural) mechanical effect, whereas the total distributional effect includes both non-behavioural and behavioural effects.

child poverty and a number of percentile ratios: P90/P10, P90/P50 and P50/P10 ratios. The well-known Gini coefficient ranges between 0 per cent in the case of perfect equality and 100 per cent in the case of perfect inequality. P90/P10 is the ratio of the upper bound of the ninth decile to that of the first decile; P90/P50 of the upper

Table 6 Direct and Total Distributional Effects on Inequality and Child Poverty

		Change			
		Means-testing		*Tax-financed universal*	
Distributional effects	*Benchmark*	*Direct*	*Total*	*Direct*	*Total*
Gini (%)	26.40	−0.51	−0.42	−0.51	−0.51
P90/P10	3.15	−0.10	−0.10	−0.09	−0.09
P50/P10	1.89	−0.06	−0.07	−0.03	−0.03
P90/P50	1.66	0.00	0.01	−0.02	−0.02
Child poverty (share in %)	14.1	−2.6	−2.1	−2.4	−2.0

Notes: The benchmark refers to the 2019 income distribution as obtained from the tax-benefit model (see Section 4.1), whereas the changes are reported as absolute differences to the benchmark. P90/P10 is the ratio of the upper bound of the ninth decile to that of the first decile; P90/P50 of the upper bound of the ninth decile to the median income; and P50/P10 of median income to the upper bound of the first decile. Child poverty is defined as the share of children living in households, where the household's equivalent disposable income is less than 60 per cent of the median equivalent disposable income in the population. The total effect includes predicted labour supply responses.

bound of the ninth decile to the median income; and P50/P10 of median income to the upper bound of the first decile. We define child poverty as the share of children living in households with an equivalent disposable income that is less than 60 per cent of the median equivalent disposable income in the population.

We find that the inequality reduction measured by the Gini coefficient is identical for the direct effects (non-behavioural) of the two schemes. However, given that means-testing, as we have already seen, implies reduced working hours in the middle of the income distribution, the total inequality reduction is somewhat smaller for the means-testing scheme: −0.42 compared to −0.51 for the tax-financed universal scheme.

The difference in distributional effects between the two schemes is also reflected by the percentile ratios. When focusing on total effects, the reductions in the P90/P10 ratio are fairly similar, whereas the reductions in the P50/P10 and the P90/P50 ratios are highest for means-testing and tax-financed universal, respectively. In other words, means-testing reduces inequality between low-income households (P10) and median-income households (P50), but it slightly increases inequality between high-income households (P90) and median-income households (P50). Overall, the effects are quite small.

Finally, Table 6 shows that there is a small difference between the two alternatives with respect to the effects on child poverty: means-testing leads to a direct effect reduction in the share of children living in poor families of 2.6 percentage points, compared to a reduction of 2.4 percentage points under the tax-financed universal scheme. When we take account of labour supply responses, we find that the effect on child poverty is close to identical in the two alternatives: −2.1 percentage points (means-testing) and −2.0 percentage points (tax-financed universal).

5.3.2 Results for Alternative Measure of Well-Being—Using Money Metric Utility

As an alternative to describing distributional effects in terms of disposable income, we here use changes in money metric utility as a welfare metric. We employ the equivalent variation (EV) measure, which is the maximum amount of money the individual is willing to pay to avoid the policy change, using pre-reform prices as reference prices. The random utility framework of the labour supply model implies that compensating variation (*CV*) and *EV* become random variables. Instead of using an analytic formula for *EV* (Dagsvik and Karlström 2005) we employ a practical simulation approach for obtaining average estimates of *EV* based on McFadden (1999), seen as $\max[\nu(C_{pre}(z) + EV, h(z)) + \varepsilon(z)] = \max[\nu(C_{post}(z), h(z)) + \varepsilon(z)]$, where $C_{pre}(z)$ and $C_{post}(z)$ reflect different tax treatments before (subscript pre) and after (subscript post) the policy change.

In practice this means that we derive measures of *EV* by simulating the choices of the economic agents, before and after the policy change, obtained from the labour supply model.[21] We find it convenient to measure *EV* in terms of negative values, which means that Figure 5 shows measures of how much the agents are willing to pay to let the policy change happen. For individuals others than wage earners, the *EV* is set equal to the amount of the transfer in NOK.

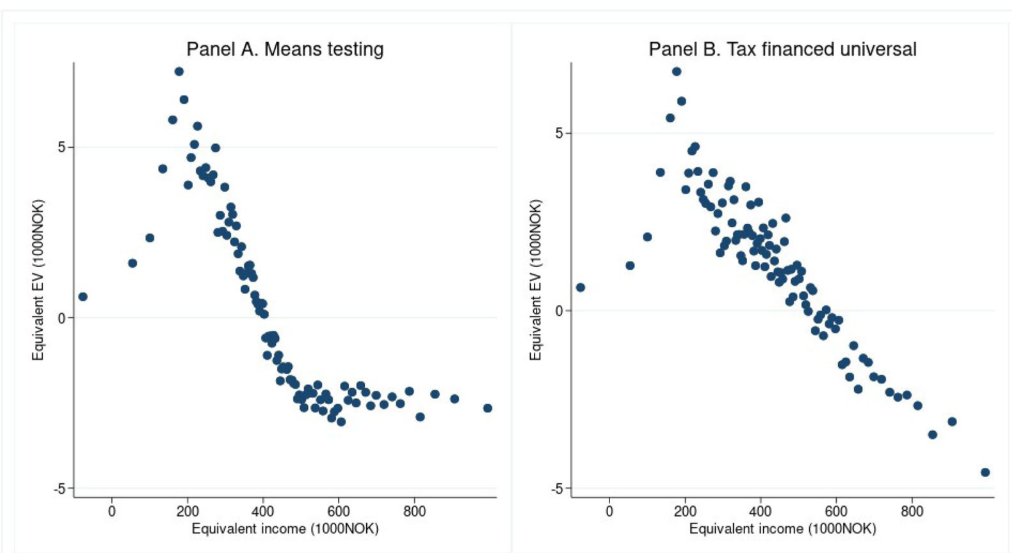
We basically see the same pattern as for income in Figure 4: the two alternative schemes are positively valued by the poor. The difference between the two schemes is found for households with middle and high equivalent income, where the gradient for the means-testing alternative is steeper, making middle-income households worse off under means-testing than under the tax-financed universal scheme. Thus, we again see that the main difference between the two alternatives is manifested in the treatment of the middle part of the income distribution.

5.4 Robustness Checks

In this section we present some robustness checks of the design of the means-tested scheme and methodological choices.

First, in Table 7 we assess the extent to which results are robust with respect to how the means-testing alternative is established. More specifically, we present results for two alternatives to the baseline phase-out rate of 10 per cent. We find that a flatter phase-out rate of 5 per cent reduces the labour supply effects of

Figure 5 Effects of Changes in Child Benefit Measured as Equivalent Variation (*EV*) Against Equivalent Household Income (All Households)



Note: Measures of *EV* are derived using the optimal choices of the economic agents, pre-reform and post-reform, obtained from the labour supply model simulations. We measure *EV* in negative values, which means that the figures are measures of how much the agents are willing to pay to let the policy change happen. Equivalent income is obtained by dividing by equivalence scale (square root scale) and letting the individual be the unit of analysis.

means-testing somewhat, but distributional effects are not very different from the main alternative. Tapering off with a steeper phase-out rate of 20 per cent increases revenue costs because of stronger labour supply effects, and the revenue balance is weakened by NOK1 billion (US$0.11 billion) compared to the main means-testing procedure. But again, the redistributional effects are very close to those found for the main alternative.

Table D1 in Section D in the Appendix provides results for a different means-testing procedure, where the testing is based on the income of the person with the highest income (primary earner) of the household. This is similar to the means-testing of the UK's High Income Child Benefit Charge. As expected we find that this alternative leads to smaller negative labour supply effects, following from primary earners of the household being less responsive to changes in economic incentives. However, means-testing based on the primary earner's income also results in less redistribution compared to the two child benefit designs discussed in the main text.

We also show how the distributional effects of the alternatives depend on the choice of equivalence scales. Recall that in our main analysis we used the square root of the household size to construct consumption units. In Table 8 we present results for other equivalence scales, referred to as the OECD Oxford equivalence scale and the modified OECD scale (also referred to as the EU-scale). The OECD Oxford scale assigns one to the first household member, 0.7 to each additional adult and 0.5 to each child (under 18 years old). The OECD modified equivalence scale

Table 7 Results of Alternative Child Benefit Designs for Different Versions of Means-Testing

	Benchmark	*Means-testing*	*Alternative means-testing I*	*Alternative means-testing II*	*Tax-financed universal*
Child benefit scheme					
Benefit rate (per child)	12,650	25,296	25,296	25,296	25,296
Inc. threshold	–	756,000	503,000	890,000	–
Phase-out rate	–	0.10	0.05	0.20	–
Tax rate change (all brackets)	–	–	–	–	+1.22 pp.
		Change to benchmark			
Labour supply effects					
All wage earners	34.56	–0.32	–0.25	–0.36	–0.20
Households with children	34.61	–0.73	–0.57	–0.82	–0.33
Mothers	31.90	–1.08	–0.84	–1.21	–0.48
Fathers	37.99	–0.30	–0.24	–0.33	–0.14
Other households	34.52	–	–	–	–0.10
Revenue balance (NOK bn)	510.0	–2.8	–2.0	–3.8	–1.5
Distributional effects					
Gini %	26.40	–0.42	–0.37	–0.43	–0.51
P90/P10	3.15	–0.10	–0.09	–0.09	–0.09
P50/P10	1.89	–0.07	–0.06	–0.07	–0.03
P90/P50	1.66	0.01	0.00	0.01	–0.02
Child poverty (share in %)	14.1	–2.1	–1.9	–2.1	–2.0

Note: The child benefit rate and the household income threshold are measured in NOK. The benchmark refers to Norway's 2019 tax-benefit system (as obtained from the tax-benefit model, see Section 4.1), with a four-tier scheme for personal income tax and a (low-rate) universal child benefit. All the schemes are revenue-neutral compared to the benchmark (before behavioural adjustments). Labour supply effects are measured in working hours per week. P90/P10 is the ratio of the upper bound of the ninth decile to that of the first decile; P90/P50 of the upper bound of the ninth decile to the median income; and P50/P10 of median income to the upper bound of the first decile. Child poverty is defined as the share of children living in households whose equivalent disposable income is less than 60 per cent of the median equivalent disposable income in the population.

Table 8 Distributional Effects with Alternative Equivalence Scales

	Benchmark	*Means-testing*	*Tax-financed universal*
Baseline: Square root scale			
Gini %	26.40	−0.42	−0.51
P90/P10	3.15	−0.10	−0.09
P50/P10	1.89	−0.07	−0.03
P90/P50	1.66	0.01	−0.02
Child poverty (share in %)	14.1	−2.1	−2.0
OECD Oxford scale			
Gini %	25.41	−0.31	−0.60
P90/P10	2.88	−0.10	−0.12
P50/P10	1.74	−0.07	−0.05
P90/P50	1.65	0.01	−0.02
Child poverty (share in %)	15.8	−2.2	−2.2
OECD modified scale			
Gini %	25.45	−0.41	−0.51
P90/P10	2.95	−0.09	−0.09
P50/P10	1.81	−0.06	−0.03
P90/P50	1.63	0.00	−0.02
Child poverty (share in %)	12.4	−2.1	−2.0

Note: In our main results we use the square root of the household size to construct consumption units. The Oxford scale (old OECD equivalence scale) assigns one to the first household member, 0.7 to each additional adult and 0.5 to each child (under 18 years old). The OECD modified equivalence scale (also called the EU scale) assigns one to the first household member, 0.5 to each additional adult and 0.3 to each child. The Gini coefficient is based on the comparison of cumulative proportions of the population against cumulative proportions of income they receive, and it ranges between 0 per cent in the case of perfect equality and 100 per cent in the case of perfect inequality. P90/P10 is the ratio of the upper bound of the ninth decile to that of the first decile; P90/P50 of the upper bound of the ninth decile to the median income; and P50/P10 of median income to the upper bound of the first decile. Child poverty is defined as the share of children living in households whose equivalent disposable income is less than 60 per cent of the median equivalent disposable income in the population.

assigns one to the first household member, 0.5 to each additional adult and 0.3 to each child. Although the choice of equivalence scale clearly affects the income inequality in the benchmark, the effects of the means-testing and tax-financed universal scheme compared to the benchmark are quite robust to this methodological choice.

6. Conclusion

Universal transfers, in the form of child benefit to all, are often described in the public debate as a waste of money on the rich. But is the answer to this to direct child benefits towards lower-income families through means-testing? Several child benefit schemes around the world, including those in the United Kingdom, Canada and Australia, and the new one in the United States, involve means-testing, in the sense that the transfer is phased out above an income threshold. The main contribution of the present article is to demonstrate that a tax-financed universal child benefit scheme is a viable alternative to means-testing.

A key to our findings is that the withdrawal of child benefit implies that the effective marginal tax rates become very high in the middle of the parental income distribution. Given that mothers have significantly higher labour supply elasticities at both the intensive and the extensive margin than men, aggregate labour supply losses resulting from a means-testing policy will be higher than those associated with increasing the tax rate on earnings in general.

We use a detailed tax-benefit model for Norway, and estimate a structural labour supply model, to compare a generous tax-financed universal benefit scheme and a means-testing alternative. The two schemes reduce child poverty by approximately the same. Our simulation results show that the means-testing scheme reduces overall labour supply more than the tax-financed universal child benefit scheme and weakens the government budget to a larger degree. We find that means-testing is harmful to mothers' labour supply, especially in households with low and medium income. Overall, from the alternatives considered here, we find that a universal scheme is (moderately) preferable to a means-tested scheme with respect to both economic efficiency and redistribution.

Lastly, it should be noted that our study is based on a static cross-sectional framework. Whether the universal and means-tested schemes affect fertility and long-term inequality differently is beyond the scope of the present article.

Endnotes

1. The increased tax credit was only in place in 2021. However, several states are enacting a state-level child tax credit.

2. We use the term ‘means-testing’ as synonymous with income-testing. Of course, many child benefit schemes would include other modes of means-testing, such as giving preferential treatment to single parents.

3. Here and in the following we use average exchange rates for 2019 to obtain values in US dollars, US\$1 = NOK8.80.

4. Then, children as an indicator of earning capacity may come from spesialisation on the quantity of children from low-ability parents.

5. Joint taxation of couples (‘taxclass 2’) was completely abolished in 2018, but also before this couples were rarely taxed jointly.

6. In this discussion we draw extensively on Apps, Van Long and Rees (2014), Andrienko, Apps and Rees (2015) and Apps and Rees (2018).

7. The convexity of the budget set is satisfied by virtually all formal tax systems, but is often not maintained when additional aspects of the tax/transfer system are taken into account, as we see below in the case of means-tested child benefits. See Apps et al. (2014) and Slemrod et al. (1994) for discussion of the non-convex case.

8. This is well known in econometric labour supply analysis, see, for example, Pudney (1989), van Soest (1995), Creedy and Kalb (2005) and Dagsvik et al. (2014), and the labour supply model of the present study, presented in Section 4.2. Apps et al. (2014) use this approach in the optimal taxation context.

9. Strictly speaking, it represents a situation where only one earner receives the transfer and pays the withdrawal rate out of their income as the formal tax base but with benefit withdrawal based on joint income, as in the Norwegian context, one could add a further dimension to the figure to represent the income of the other earner.

10. Apps and Rees (2022) show that this has been the case in the Australian tax system.

11. Ideally, one would prefer to allow for unobserved heterogeneity across both jobs and workers, but identification of such a model is not guaranteed (Dagsvik and Jia 2016).

12. However, given this assumption, it is worth noting that empirical evidence suggests that children may benefit from the child benefit being transferred to mothers (and not fathers), see Lundberg, Pollak and Wales (1997), which signifies that parents do not always pool their income. There is also evidence suggesting that parents use the child benefit money differently from other types of income; see, for example, Kooreman (2000). We abstract from such effects here.

13. We have, however, derived results for adding random draws of the error terms too, basically obtaining the same results as without drawn error terms.

14. Note that leisure here is simply defined as time used not working in the market; thus it may include parental child care and other types of household work.

15. Some studies report higher wage elasticity estimates than we find here; see, for example, Keane (2022) for micro evidence and Mertens and Montiel Olea (2018) and Zidar (2019) for macro studies. Widely differing estimates of responses to changes in tax or wage rates, in particular between the micro and macro evidence, have resulted in efforts to explain and reconcile the gaps; see, for example, Chetty et al. (2011) and Keane and Rogerson (2015). Although the present study is static, and agents do not face a dynamic optimisation problem, which margins should be included amongst the behavioural responses is open to discussion (Creedy and Duncan 2005; Blundell, Brewer and Francesconi 2008; Peichl and Siegloch 2012). It can be argued that the job choice model goes further in taking the demand side into account than the standard static labour supply model, as it brings demand-side job opportunities into the choice problem. However, in simulations of policy changes, the feedback effects on job choice sets and on individual wages are ignored.

16. Labour income here also includes all income which replaces earnings, such as pensions, sick pay, disability and unemployment benefits.

17. We abstract from behavioural effects when defining revenue neutrality. See Appendix, Section C, for results of an alternative approach to defining revenue neutrality, where the revenue effects are established after taking labour supply responses into account. Results are less influenced by employing this alternative procedure to define revenue neutrality.

18. Household members are defined as wage earners or potential wage earners, that is, households with self-employed persons, pensioners and unemployed persons are excluded.

19. Table B2 in Section B in the Appendix presents results of similar calculations for males.

20. Table B1 in Section B in the Appendix describes the direct distributional effects in more detail. It also shows that the present child benefit system has a (modest)

redistributional effect, which is due to families with children (or many children) being overrepresented in the lower- and middle-income deciles. Extended child benefit for single parents, who on average have relatively low income, also contributes to enhanced redistribution.

21. Dagsvik, Locatelli and Strøm (2009) and Creedy, Herault and Kalb (2011) provide further details of how measures of change in money metric utility can be obtained given that a discrete choice labour supply model is employed. It should be noted that there is some controversy concerning interpersonal comparison of measures of utility in distributional analyses (see the review in Slesnick 1998).

ORCID

Thor O. Thoresen http://orcid.org/0000-0002-5764-5825

References

Aasness, J., Dagsvik, J. K. and Thoresen, T. O. 2007, 'The Norwegian tax-benefit model system LOTTE', In *Modelling our Future: Population Ageing, Health and Aged Care, International Symposia in Economic Theory and Econometrics*, eds A. Gupta and A. Harding, Elsevier Science, 513–8.

Acs, G. and Toder, E. 2007, 'Should we subsidize work? Welfare reform, the earned income tax credit and optimal transfers', *International Tax and Public Finance*, vol. 14, no. 3, pp. 327–43.

Aizer, A., Hoynes, H. and Lleras-Muney, A. 2022, 'Children and the US social safety net: Balancing disincentives for adults and benefits for children', *Journal of Economic Perspectives*, vol. 36, no. 2, pp. 149–74.

Akerlof, G. A. 1978, 'The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning', *American Economic Review*, vol. 68, no. 1, pp. 8–19.

Andrienko, Y., Apps, P. and Rees, R. 2015, 'Gender bias in tax systems based on household income', *Annals of Economics and Statistics/Annales d'Economie et de Statistique'*, vol. 117/118, pp. 141–55.

Apps, P. and Rees, R. 2018, 'Optimal family taxation and income inequality', *International Tax and Public Finance*, vol. 25, no. 5, pp. 1093–128.

Apps, P. and Rees, R. 2022, 'Inequality measurement and tax/transfer policy', *International Tax and Public Finance*, vol. 29, no. 4, pp. 953–84.

Apps, P., Van Long, N. and Rees, R. 2014, 'Optimal piecewise linear income taxation', *Journal of Public Economic Theory*, vol. 16, no. 4, pp. 523–45.

Atkinson, A. B. 2015, *Inequality: What Can Be Done*, Cambridge, MA.

Baker, M., Gruber, J. and Milligan, K. 2008, 'Universal child care, maternal labor supply, and family well-being', *Journal of Political Economy*, vol. 116, no. 4, pp. 709–45.

Bargain, O. and Peichl, A. 2016, 'Own-wage labor supply elasticities: Variation across time and estimation methods', *IZA Journal of Labor Economics*, vol. 5, no. 10, pp. 10.

Berg, K. and Thoresen, T. O. 2020, 'Problematic response margins in the estimation of the elasticity of taxable income', *International Tax and Public Finance*, vol. 27, no. 3, pp. 721–52.

Bibler, A., Guettabi, M. and Reimer, M. 2019, 'Universal cash transfers and labor market outcomes', manuscript, University of Alaska, Anchorage.

Blumkin, T., Margalioth, Y. and Sadka, E. 2015, 'The re-distributive role of child benefits revisited', *International Tax and Public Finance*, vol. 22, no. 3, pp. 476–501.

Blundell, R. 2006, 'Earned income tax credit policies: Impact and optimality. The Adam Smith Lecture, 2005', *Labour Economics*, vol. 13, no. 4, pp. 423–43.

Blundell, R., Brewer, M. and Francesconi, M. 2008, 'Job changes and hours changes: Understanding the path of labor supply adjustment', *Journal of Labor Economics*, vol. 26, no. 3, pp. 421–53.

Brewer, M. and Hoynes, H. 2019, 'In-work credits in the UK and the US', *Fiscal Studies*, vol. 40, no. 4, pp. 519–60.

Brewer, M., Sae, E. and Shephard, A. 2010, 'Means-testing and tax rates on earnings', In *Dimensions of Tax Design: The Mirrlees Review*, eds J. A. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, G. M., P. Johnson, G. Myles and J. Poterba, Oxford University Press, Oxford, pp. 90–174.

Burton, P. and Phipps, S. 2017, 'Economic well-being of Canadian children'. *Canadian Public Policy*, vol. 43, no. 4, pp. 299–30.

Chetty, R., Guren, A., Manoli, D. and Weber, A. 2011, 'Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins' *American Economic Review: Papers and Proceedings*, pp. 471–75.

Congressional Research Service. 2021, *The Child Tax Credit: Temporary Expansion for 2021 Under the American Rescue Plan Act of 2021 (ARPA; P.L. 117-2)*, Congressional Research Service, Washington, DC.

Creedy, J. and Duncan, A. 2005, 'Aggregating labour supply and feedback effects in microsimulation'. *Australian Journal of Labour Economics*, vol. 8, no. 3, pp. 277–90.

Creedy, J., Herault, N. and Kalb, G. 2011, 'Measuring welfare changes in behavioural microsimulation modelling: Accounting for the random utility component', *Journal of Applied Economics*, vol. 14, no. 1, pp. 5–34.

Creedy, J. and Kalb, G. 2005, 'Discrete hours labour supply modelling: Specification, estimation and simulation', *Journal of Economic Surveys*, vol. 19, no. 5, pp. 697–734.

Dagsvik, J. and Jia, Z. 2016, 'Labor supply as a choice among latent jobs: Unobserved heterogeneity and identification', *Journal of Applied Econometrics*, vol. 31, no. 3, pp. 487–506.

Dagsvik, J., Jia, Z., Kornstad, T. and Thoresen, T. O. 2014, 'Theoretical and practical arguments for modeling labor supply as a choice among latent jobs', *Journal of Economic Surveys*, vol. 28, no. 1, pp. 134–51.

Dagsvik, J. K. and Karlstrom, A. 2005, 'Compensating variation and Hicksian choice probabilities in random utility models that are nonlinear in income', *Review of Economic Studies*, vol. 72, no. 1, pp. 57–76.

Dagsvik, J. K., Locatelli, M. and Strøm, S. 2009, 'Tax reform, sector-specific labor supply and welfare effects', *Scandinavian Journal of Economics*, vol. 111, no. 2, pp. 299–321.

de Boer, H.-W. and Jongen, E. L. 2023, 'Analysing tax-benefit reforms in the Netherlands using structural models and natural experiments', *Journal of Population Economics*, vol. 36, no. 1, pp. 179–209.

Doepke, M., Hannusch, A., Kindermann, F. and Tertilt, M. 2022, 'The economics of fertility: A new era', NBER Working Paper No. 29948, National Bureau of Economic Research.

Ebert, U. 1997, 'Social welfare when needs differ: An axiomatic approach', *Economica*, vol. 64, no. 254, pp. 233–44.

Eckstein, Z., Keane, M. and Lifshitz, O. 2019, 'Career and family decisions: Cohorts born 1935–1975', *Econometrica*, vol. 87, no. 1, pp. 217–53.

Ellingsæter, A. L. 2003, 'The complexity of family policy reform. The case of Norway', *European Societies*, vol. 5, pp. 419–43.

González, L. 2013, 'The effect of a universal child benefit on conceptions, abortions, and early maternal labor supply', *American Economic Journal: Economic Policy*, vol. 5, no. 3, pp. 160–88.

Hendren, N. and Sprung-Keyser, B. 2020, 'A unified welfare analysis of government policies', *The Quarterly Journal of Economics*, vol. 135, no. 3, pp. 1209–318.

Hener, T. 2016, 'Unconditional child benefits, mothers' labor supply, and family well-being: Evidence from a policy reform', *CESifo Economic Studies*, vol. 62, no. 4, pp. 624–49.

Hotz, V. J. and Scholz, J. K. 2003, 'The Earned Income Tax Credit', In *Means-Tested Transfer Programs in the United States*, ed R. A. Moffitt, University of Chicago Press, Chicago, pp. 141–97.

Hoynes, H. W. and Patel, A. J. 2018, 'Effective policy for reducing poverty and inequality? The Earned Income Tax Credit and the distribution of income', *Journal of Human Resources*, vol. 53, no. 4, pp. 859–90.

Immonen, R., Kanbur, R., Keen, M. and Tuomala, M. 1998, 'Tagging and taxing: The optimal use of categorical and income information in designing tax/transfer schemes', *Economica*, vol. 65, no. 258, pp. 179–92.

Kaplow, L. 2007, 'Optimal income transfers', *International Tax and Public Finance*, vol. 14, no. 3, pp. 295–325.

Keane, M. and Rogerson, R. 2015, 'Reconciling micro and macro labor supply elasticities: A structural perspective', *Annual Review of Economics*, vol. 7, no. 1, pp. 89–117.

Keane, M. P. 2022, 'Recent research on labor supply: Implications for tax and transfer policy', *Labour Economics*, vol. 77, pp. 102026.

Kooreman, P. 2000, 'The labeling effect of a child benefit system', *American Economic Review*, vol. 90, no. 3, pp. 571–83.

Kornstad, T. and Thoresen, T. O. 2004, 'Means-testing the child benefit', *Review of Income and Wealth*, vol. 50, no. 1, pp. 29–49.

Lundberg, S. J., Pollak, R. A. and Wales, T. J. 1997, 'Do husbands and wives pool their resources?', *Journal of Human Resources*, vol. 32, no. 3, pp. 463–80.

Lunder, T. 2015, 'Nasjonale satser til private barnehager', TF-notat nr 97/ 2015, Telemarksforsking, Bø i Telemark (in Norwegian).

McFadden, D. 1999, 'Computing willingness–to–pay in random utility models', In *Trade, Theory and Econometrics: Essays in Honour of John S. Chipman, Chapter 15*, J. R. Melvin, J. C. Moore and R. G. Riezman eds Routledge, London, pp. 253–74.

Mertens, K. and Montiel Olea, J. L. 2018, 'Marginal tax rates and income: New time series evidence', *The Quarterly Journal of Economics*, vol. 133, no. 4, pp. 1803–84.

Milligan, K. and Stabile, M. 2009, 'Child benefits, maternal employment, and children's health: Evidence from Canadian child benefit expansions', *American Economic Review: Papers Proceedings*, vol. 99, no. 2, pp. 128–32.

Ministry of Children and Equality. 2017, 'Overføringer til barnefamiliene', Norges offentlige utredninger (NOU) 2017: 6, Akademika (in Norwegian).

OECD. 2011, *Doing Better for Families*. Organisation for Economic Cooperation and Development Publishing, Paris.

OECD. 2023, 'What are equivalence scales?', viewed 15 March 2023. <www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf>.

Peichl, A. and Siegloch, S. 2012, 'Accounting for labor demand effects in structural labor supply models', *Labour Economics*, vol. 19, no. 1, pp. 129–138.

Pudney, S. 1989, *Modelling individual choice*. Basil Blackwell, Oxford.

Schirle, T. 2015, 'The effect of universal child benefits on labour supply', *Canadian Journal of Economics*, vol. 48, pp. 437–463.

Slemrod, J., Yitzhaki, S., Mayshar, J. and Lundholm, M. 1994, 'The optimal two-bracket linear income tax', *Journal of Public Economics*, vol. 53, no. 2, pp. 269–290.

Slesnick, D. T. 1998, 'Empirical approaches to the measurement of welfare', *Journal of Economic Literature*, vol. 36, no. 4, pp. 2108–65.

Statistics Norway. 2018a, 'Income and wealth statistics for households', viewed 15 March 2023. <https://www.ssb.no/en/inntekt-og-forbruk/statistikker/ifhus>.

Statistics Norway. 2018b, 'Labour force survey 2018', viewed 15 March 2023. <https://www.ssb.no/en/arbeid-oglonn/statistikker/aku>.

Thoresen, T. O. and Vattø, T. E. 2015, 'Validation of the discrete choice labor supply model by methods of the new tax responsiveness literature', *Labour Economics*, vol. 37, pp. 38–53.

van Soest, A. 1995, 'Structural models of family labor supply: A discrete choice approach', *Journal of Human Resources*, vol. 30, no. 1, pp. 63–88.

Whalen, C. and Reichling, F. 2017, 'Estimates of the Frisch elasticity of labor supply: A review', *Eastern Economic Journal*, vol. 43, no. 1, pp. 37–42.

Zidar, O. 2019, 'Tax cuts for whom? Heterogeneous effects of income tax changes on growth and employment', *Journal of Political Economy*, vol. 127, no. 3, 1437–72.

Appendix

A. More Details on the Behavioural Micro-Simulation Model

Empirical specification and estimation of model parameters

In the following, we present further details of the empirical specification and estimation of the discrete choice labour supply model. Following Dagsvik and Jia (2016), we assume that the deterministic part of the utility function can be represented by a Box-Cox function,

$$\log \ \nu(C,h) = \alpha_0 \frac{(C - C_0)^{\alpha_1} - 1}{\alpha_1} + \beta_0 \frac{(\bar{h} - h)^{\beta 1} - 1}{\beta_1} \qquad \text{(A1)}$$

where C_0 represents the minimum or subsistence household-adjusted consumption level, set here to NOK80,000. $\bar{h}$ is defined as 80 hours per week, such that ($\bar{h}$-h measures leisure time. To allow for preference heterogeneity, we let β_0 be a function of individuals' age and number of children. If $\alpha_0, \beta_0 > 0$ and $\alpha_1, \beta_1 \leq 1$, the utility function is increasing and concave in consumption and leisure.

The job opportunity measure (see Section 4.1), $m(h) = \theta g(h)$, is considered to be a sufficient statistic for the choice sets of available jobs and represents labour market restrictions. We interpret θ as the normalised total number of jobs (relative to non-participation) and $g(h)$ as the fraction of jobs available to the agent with offered hours of work equal to h. We allow for θ depending on the individual's education level. $g(h)$ on the other hand, is considered to be independent of individual characteristics, as restrictions on hours are intended to reflect institutional labour market regulations and negotiations between employers' associations and workers unions. In particular, we let $g(h)$ be uniformly distributed among working-time options, except for two possible peaks (estimated within the model) for full-time and part-time jobs. For all individuals, we specify eight weekly hours of work alternatives, that is, $h \in \{0, 4, 12, 20, 28, 37.5, 45, 50\}$. The full-time peak corresponds to 37.5 hours a week, while the part-time peak corresponds to 20 hours a week. It follows that in practice the job opportunity measure is estimated as alternative-specific constants for non-participation and part-time/full-time work, reflecting $\log \theta^{-1}$ and $\log g(h)$, respectively.

The model parameters are estimated using cross-sectional data from the 2014 Norwegian Labour Force Survey, linked to administrative tax-return information. The estimation sample is restricted to wage earners, singles and couples, in the age group 26–62 years. Hourly wage rates are measured as the sum of labour income from the main and the second job (if the individual has a second job) divided by reported hours of work (for main and second job). The wage parameters of the model, reported in Table A1, are estimated by OLS regressions (separately for men and women) where the log of the wage rate is explained by individual characteristics on education, work experience (age minus years of education) and an indicator variable for being married.

Further, the tax-benefit model (see Section 4.1) is used to compute disposable income for each working hours alternative. Individuals are assigned to the closest discrete working hours alternative, and parameter estimates are obtained by maximum likelihood estimation. The estimated parameters are reported in Table A2. We find that the utility function is concave and increasing with respect to both consumption and leisure, and the job opportunity measure also has the expected sign.

Table A1 Wage Rate Equation Estimates

Dependent variable: Log wage rate	*Women*	*Men*
Years of education (scale 10^{-1})	0.4331	0.4661
Experience (scale 10^{-1})	0.1713	0.2152
Experience, squared (scale 100^{-1})	−0.0255	−0.0348
Married	−0.0030	0.0457
Constant	4.7103	4.7931
Measure of standard deviation	0.2782	0.3395

Note: The parameters reflect OLS estimates obtained from the 2014 Labour Force Survey.

Simulations

In the simulations each individual is assigned a wage rate based on the estimates of the wage regression, see Table A1, inflated to 2019 level. We use the assigned wage rate together with information on household composition, non-labour income and the income tax scheme to compute disposable income (consumption) for each alternative of working hours. For a given tax-benefit schedule, the probability distribution for each discrete alternative follows from the estimated model parameters in Table A2. Thus the model can

Table A2 Parameter Estimates of the Labour Supply Model

Variables		*Couples*	*Single females*	*Single males*
Preferences, $v(C,h)$				
Consumption				
Constant	α_0	0.8254	0.3828	0.2981
Exponent	α_1	0.6998	0.6095	0.5816
Leisure, female ln age (scale 10^{-1})	β_{01f}	−6.5803	−3.1794	
ln age, squared (scale 10^{-1})	$\beta_{02f}\beta$	2.7112	1.1054	
Children under age 6	$\beta 03f$	0.4098		
Children above age 6	$\beta 04f$	0.1977		
Constant	$\beta 00f$	7.4317	4.1719	
Exponent	$\beta 1f$	−1.4459	−1.0618	
Leisure, male ln age (scale 10^{-1})	$\beta 01m$	−8.2440		−39.450
ln age, squared (scale 10^{-1})	$\beta 02m$	3.4316		13.929
Children under age 6	$\beta 03m$	−0.0576		
Children above age 6	$\beta 04m$	−0.3489		
Constant	$\beta 00m$	13.3842		28.261
Exponent	$\beta\ 1m$	0.1410		0.0000
Interaction	γ	1.4393		
Female leisure x male leisure				
Job opportunity measure, $m(h)$	$f00,f$	2.3830	3.0673	
Non-participation, female				
Years of education (scale 10^{-1}), female	$f01,f$	−0.0084	−0.9986	
Part-time peak, female	fpt,f	0.1569	−0.0987	
Full-time peak, female	fft,f	1.4627	2.0724	
Part-time peak, male	fpt,m	0.0207		−0.3478
Full-time peak, male	fft,m	2.7803		2.9524

Note: Parameter estimates are obtained by maximum likelihood estimation, with data from the 2014 Labour Force Survey.

be used to obtain aggregate labour supply responses to alternative child benefit and tax schemes by obtaining probability distributions for each individual.

B. Additional Table Results

(see Tables B1 and B2)

C. Results When Schemes are Adjusted to Ex-post Revenue Neutrality

In the main text of the article we defined revenue neutrality before any behavioural change. Thus, in Table C1 we show results for the means-tested and tax-financed alternatives when the schemes are adjusted to achieve revenue neutrality after taking behavioural effects into account. Results are very little influenced by this alternative way of defining revenue neutrality—compare the results of Table C1 to the second and fifth columns of Table 7.

D. When Child Benefit is Tapered Off by Income of the Household's Primary Earner

Our main means-testing alternative is one where child benefit is tapered off by household income. Another possibility is to means-test on the basis of the household's primary earner's income. This is similar to the means-

Table B1 Direct (Non-behavioural) Distributional Effects of the Alternative Child Benefit Schedules (All Households)

	Benchmark			*Means-testing*	*Tax-financed universal*
Decile	*Disp. income*	*Income tax*	*Child benefit*	*Δ Disp. income*	*Δ Disp. income*
1	136.9	27.1	5.9	5.8	5.5
2	257.4	49.4	6.6	6.6	5.1
3	316.2	69.8	6.2	5.3	4.0
4	362.6	91.5	6.3	2.8	3.3
5	404.7	111.8	6.4	−0.3	2.7
6	447.0	132.7	6.4	−3.5	2.0
7	494.1	158.4	6.0	−5.2	0.8
8	555.7	192.6	5.4	−5.2	−0.8
9	643.7	246.4	4.5	−4.5	−3.1
10	1000.9	492.6	3.9	−3.8	−8.1

Note: Incomes are measured in terms of equivalence-adjusted (square root scale) household income in NOK1,000s in 2019, with the individual as the unit of analysis. The effects of means-testing and the tax-financed universal scheme are measured as average differences in an individual's equivalence-adjusted household income under the two child benefit schemes compared to the 2019 benchmark.

Table B2 Effects on Male Labour Supply of Changes in the Child Benefit Schedule for Combinations of Female/Male Wage Levels in Couples with Children

	Wage combinations: Female/male wage rate levels								
	L/L	*L/M*	*L/H*	*M/L*	*M/M*	*M/H*	*H/L*	*H/M*	*H/H*
Benchmark: uncond. work. hours	37.53	38.03	38.61	37.48	38.06	38.64	37.26	38.00	38.58
Means-testing: change in uncond. work. hours	−0.42	−0.33	−0.23	−0.41	−0.33	−0.26	−0.35	−0.30	−0.19
Tax-fin. univ.: change in uncond. work. hours	−0.15	−0.14	−0.15	−0.14	−0.14	−0.15	−0.14	−0.14	−0.15
Population (1,000 individuals)	29.3	31.1	7.1	31.5	72.6	30.9	6.5	31.3	29.5

Note: Wage-rate levels (per hour) defined by percentiles: L, low wage, 0–25; M, medium wage, 25–75; H, high wage 75–100. Changes are measured in absolute values of unconditional working hours per week (extensive + intensive margin).

Table C1 Results of Alternative Child Benefit Designs When Revenue Neutrality is Determined after Taking Behavioural Effects into Account

	Benchmark	*Means-testing*	*Tax-financed universal*
Child benefit scheme			
Benefit rate (per child)	12,650	25,296	25,296
Income threshold	–	620,000	–
Phase-out rate	–	0.10	–
Tax rate change (all brackets)	–	–	+1.33 pp.
		Change to benchmark	
Labour supply effects			
All wage earners	34.56	–0.30	–0.21
Households with children	34.61	–0.68	–0.34
Mothers	31.90	–1.01	–0.50
Fathers	37.99	–0.26	–0.15
Other households	34.52	–	–0.11
Revenue balance (NOK bn)	510.0	–	–
Distributional effects			
Gini %	26.40	–0.30	–0.51
P90/P10	3.15	–0.09	–0.09
P50/P10	1.89	–0.08	–0.03
P90/P50	1.66	0.02	–0.02
Child poverty (share in %)	14.1	–2.0	–1.9

Note: The child benefit rate and the household income threshold are measured in NOK. The benchmark corresponds to Norway's 2019 tax-benefit system, with a four-tier scheme for personal income tax and a (low-rate) universal child benefit. In this robustness check all the schemes are revenue-neutral compared to the benchmark after the predicted behavioural adjustments (ex-post). Labour supply effects are measured in working hours per week. The Gini coefficient is based on the comparison of cumulative proportions of the population against cumulative proportions of income they receive, and it ranges between 0 per cent in the case of perfect equality and 100 per cent in the case of perfect inequality. P90/P10 is the ratio of the upper bound of the ninth decile to that of the first decile; P90/P50 of the upper bound of the ninth decile to the median income; and P50/P10 of median income to the upper bound of the first decile. Child poverty is defined as the share of children living in households whose equivalent disposable income is less than 60 per cent of the median equivalent disposable income in the population.

Table D1 Alternative Means-Tested Scheme: Based on the Income of Households' Primary Earner

	Benchmark	*Means-testing: household income*	*Means-testing: primary earner's income*	*Tax-financed universal*
Child benefit scheme				
Benefit rate (per child)	12,650	25,296	25,296	25,296
Income threshold	–	756,000	389,000	–
Phase-out rate	–	0.10	0.10	–
Tax rate change (all brackets)	–	–	–	+1.22 pp.
		Change to benchmark		
Labour supply effects				
All wage earners	34.56	–0.32	–0.17	–0.20
Households with children	34.61	–0.73	–0.39	–0.33
Mothers	31.90	–1.08	–0.52	–0.48
Fathers	37.99	–0.30	–0.23	–0.14
Other households	34.52	–	–	–0.10
Revenue balance (NOK bn)	510.0	–2.8	–1.8	–1.5
Distributional effects				
Gini %	26.40	–0.42	–0.34	–0.51
P90/P10	3.15	–0.10	–0.08	–0.09
P50/P10	1.89	–0.07	–0.04	–0.03
P90/P50	1.66	0.01	–0.00	–0.02
Child poverty (share in %)	14.1	–2.1	–1.5	–2.0

Note: The child benefit rate and the income threshold of the household income and the primary earner's income are measured in NOK. The benchmark corresponds to Norway's 2019 tax and benefit system, with a four-tier scheme for personal income tax and a (low-rate) universal child benefit. All the schemes are revenue-neutral compared to the benchmark (before any behavioural adjustments). Labour supply effects are measured in working hours per week. Distributional effects are measured by the equivalence-adjusted household income of individuals, using the square root scale. The Gini coefficient is based on the comparison of cumulative proportions of the population against cumulative proportions of income they receive, and it ranges between 0 per cent in the case of perfect equality and 100 per cent in the case of perfect inequality. P90/P10 is the ratio of the upper bound of the ninth decile to that of the first decile; P90/P50 of the upper bound of the ninth decile to the median income; and P50/P10 of median income to the upper bound of the first decile. Child poverty is defined as the share of children living in households whose equivalent disposable income is less than 60 per cent of the median equivalent disposable income in the population.

testing of the UK's High Income Child Benefit Charge. Results of this alternative design are reported in Table D1. As expected we find that this alternative leads to smaller negative labour supply effects, following from primary earners of the household, predominantly males, being less responsive to changed economic incentives. However, this type of means-testing results in smaller redistributive effects, both compared to the means-testing alternative based on household income and the tax-financed universal scheme.