A Simulation Method to Calculate Inverse-Optimum Social Marginal Welfare Weights*

Niklas Isaak†

Robin Jessen‡

26th January 2023

Abstract

We propose a method to calculate implied marginal social welfare weights if the current tax-transfer system is optimal. At the optimum, the cost of providing one Euro to a specific group must equal its weighted benefit. For every percentile of the income distribution, we calculate the cost of transferring 100 Euro to individuals in that percentile. To calculate behavioral costs, we estimate a structural labor supply model and simulate labor supply reactions to these small local tax reductions. The advantage of this simulation approach is that we do not need to restrict labor supply in any way to obtain analytical solutions. For instance, the method allows for non-convex budget sets and labor supply reactions of secondary earner. We apply the approach to Germany and find that the tax-transfer system is optimal if society is inequality averse and values one Euro for households at the 10th percentile twice as much as one Euro for households at the median. At medium incomes, implied weights for couples are higher than for singles as tax reductions for middle-income couples lead to substantial labor supply reductions by secondary earners.

JEL: H21; H31; J22

Keywords: Optimal taxation, Microsimulation, Marginal Value of Public Funds, Labor Supply, Efficiency, Inequality Aversion

*We thank seminar participants at RWI, the 2022 annual congress of the IIPF in Linz, the 2022 EEA congress in Milan and the WATT-Workshop on Inequality and Taxation at DIW Berlin in 2022 for valuable comments.

†RWI, Invalidenstraße 112, 10115 Berlin, Germany (e-mail: nisaak@rwi-essen.de); Department of Economics, Freie Universität Berlin.

‡RWI, Invalidenstraße 112, 10115 Berlin, Germany (e-mail: rjessen@rwi-essen.de)

1 Introduction

In this paper we estimate social marginal welfare weights $g(y)$ (see, e.g., Saez and Stantcheva, 2016) along the distribution of net incomes in Germany implied by the current tax-transfer system. These weights indicate how much society values providing one additional Euro of consumption to a specific individual earning gross income y. As our first main contribution, we propose a new method to obtain implied welfare weights: We simulate small policy reforms, which increase disposable net income for households in one percentile of the income distribution. We use a structural labor supply model to simulate behavioral reactions and thereby the total fiscal cost of each of these reforms. The advantage of our approach is that it is very flexible and simple to apply. We do not need to impose any restrictions on labor supply reactions or focus on specific, homogeneous subsamples in order to obtain analytical solutions. The approach allows for flexible labor supply adjustments of couples. Our second main contribution is the quantification of normative judgements implicit in the German tax-transfer system. We find that the current tax-transfer system is optimal if the social planner values one Euro for households at the 10th percentile twice as much as one Euro for households at the median of the distribution. At medium incomes, implied weights for couples are higher than for singles as tax reductions for couples lead to substantial labor supply reductions.

At the optimal tax schedule, the ratios of weights $g(c)$ between two individuals equal the cost of redistributing one Euro from one individual to the other. In the presence of behavioral reactions, this cost typically differs from one: an increase in transfers for the unemployed by one Euro induces additional people to become unemployed leading to an additional increase in transfer payments. In this case, the cost of redistributing one Euro exceeds one. Conversely, a tax break might induce some individuals to increase their hours of work. In this case the tax break partially pays for itself and providing an additional Euro of consumption costs less than one Euro. The literature on the marginal value of public funds (Finkelstein and Hendren, 2020; Hendren and Sprung-Keyser, 2020) measures the "bang for the buck" of policy reforms by quantifying the revenue effects of behavioral reactions to policy changes. Due to the envelope theorem, a "small" reform has no impact on the utility of individuals who adjust their labor supply in reaction to it. Thus, only the effect of an increase in income for those who do not adjust labor supply is relevant for welfare analysis, while behavioral reactions matter due to their effect on government revenue.

To simulate labor supply responses to tax breaks, we the microsimulation model EMSIM (Einkommensteuer-Mikrosimulationsmodell, see Bechara et al., 2015) and a discrete choice structural labor supply model following Aaberge et al. (1995) and van Soest (1995). We make use of the Socio-Econmic Panel (SOEP), a representative survey data set of German households. We analyze the tax year 2017 using retrospective data for that year.

A number of papers quantify marginal social welfare weights implicit in different countries tax schedules. Typically, these papers apply "standard" optimal tax models (see Atkinson and Stiglitz, 2015) such as Saez (2001), which rule out participation effects of tax changes (Ayaz et al., 2021; Bourguignon and Spadaro, 2012; Lockwood and Weinzierl, 2016), or the model in Saez (2002), which allows for participation effects, but rules out income effects and restricts intensive labor supply reactions to discrete changes to "neighbor" income groups (Bargain et al., 2014; Bourguignon and Spadaro, 2012; Blundell et al., 2009; Jessen et al., 2022). Jacobs et al. (2017) and Hendren (2020) obtain inverse optimal weights in model economies that allow for both participation effects and intensive labor supply reactions while ruling out intensive "jumps" in labor supply. In contrast, standard labor supply models imply that with non-convex budget sets small changes in marginal tax rates can lead to "jumps" in labor supply. Importantly, in this case labor supply elasticities are not directly linked to parameters of the utility function and are not constant. Non-convex budget sets are an important feature of European welfare states with high marginal transfer withdrawal rates. In this case it is necessary to specify the utility function in order to simulate labor supply responses to tax changes. Considering this, many papers starting with Burtless and Hausman (1978) estimate structural labor supply models, see Keane (2011) for a discussion. In an approach that allows for non-convex budget sets, Fullerton and Gan (2004) calibrate stochastic utility functions to calculate the welfare effects of specific tax reforms measured through the equivalent variation. In contrast to that paper, we estimate utility functions and calculate the welfare effects of marginal reforms, which allows us to quantify implicit welfare weights.

Moreover, optimal taxation models models typically do not explicitly account for labor supply decisions of couple household. We do not impose any of these restrictions. In many applied papers, labor supply elasticities used to calibrate optimal taxation models are obtained from structural labor supply models such as the one we employ in this paper. These labor supply models do not impose the same restrictions as the optimal taxation models, implying a discrepancy between the models used.

The inverse optimal taxation literature typically finds optimal welfare weights that do not decrease strictly with income. Often, low-income working households are found to have the lowest weights, in some cases even zero or negative, implying a tax system that is not Pareto optimal. While we too find increasing welfare weights for low-income working households, their weights are only slightly smaller than those at the median and larger than weights of top-income earners. This finding results from the fact that in our simulations a slight increase in disposable income for low-income earners can lead to a decrease in labor supply even for individuals with relatively high income. Such jumps can occur i) due to non-convex budget sets, ii) if individuals have a preference not to work a particular number of hours, or iii) due to labor supply responses of couples. In contrast, in the Saez (2002) model, such a tax break will only induce individuals with slightly higher income to reduce their labor supply, while the participation rate will increase. This reduces the potential fiscal cost of a tax break for the working poor, implying a very low marginal welfare weight for this group.

The next section describes how we obtain implicit welfare weights and presents the data as well as the microsimulation model and structural labor supply model. Section 3 reports our results, section 4 shows the labor supply reactions that lead to these results and section 5 concludes.

2 Method

2.1 Inverse Optimal Taxation

As discussed in Hendren (2020), at the optimum, society's valuation of an additional Euro of consumption for an individual with gross income y, $g(y)$, must equal the marginal cost of taxation,

$$g(y) = 1 + FE(y), \tag{1}$$

where the mechanical cost of a small tax cut of one Euro equals one and $FE(y)$ denotes the "fiscal externality" of a small tax cut for individuals earning y, i.e., the effect of labor supply reactions to the tax cut on government revenue. $T(y)$ denotes the tax liability of households earning y (which can be negative for transfer recipients), $f(y)$ the density of households earning y, and x all possible gross income levels.

The fiscal externality of a small tax cut by $d[T(y)]$ at gross income level y with density

$f(y)$ is given by

$$FE(y) = -\sum_{x \neq y} \frac{d_x[f(y)]}{d[T(y)]} \frac{T(y) - T(x)}{f(y)} = \frac{\text{behavioral effect}}{\text{mechanical effect}}, \quad (2)$$

the revenue effect due to behavioral reactions divided through the mechanical revenue effect due to the tax reduction for inframarginal individuals. $d_x[f(y)]$ is the number of households that move from earnings level x to earnings level y. $T(y)$ and $T(x)$ denote the tax liabilities at earnings levels y and x. When the fiscal externality is negative, tax cuts partially pay for themselves. When it is smaller than one, a tax cut increases revenues, the tax schedule is on the right hand side of the Laffer curve and $g(y)$ is negative. Typically, optimal taxation papers impose restrictions and (often implicitly) derive analytical expressions for the fiscal externality based on labor supply elasticities. For instance, Hendren (2020) reports a formulation for the fiscal externality in terms of three elasticities.

In contrast, we simulate $FE(y)$ using the structural labor supply model described more in detail below. In this probabilistic model, households maximize utility by choosing a particular number of work hours per week. We split the sample into percentiles of the income distribution, each associated with an income level y, and, one by one, increase disposable income at each income level of the distribution by 100 Euro per year. If the income level y is available for a specific household, the probability that it chooses y increases in reaction to a tax cut at this income level. As each household in the sample represents a larger number of households (indicated by the sample weight), a change in the probability to work a particular number of hours can be thought of as a change in the share of households of a specific type in the population who work that number of hours. With the sample weights and choice probabilities in the status quo and a reform scenario in hand, we calculate the total (behavioral and mechanical) decrease in government revenue due to a small tax cut at income level y divided through the mechanical reduction in revenues (100 Euro times the number of people at income level y in the status quo) for every small reform for all 100 income levels. Thereby, we directly obtain $g(y)$ as in equation (1).

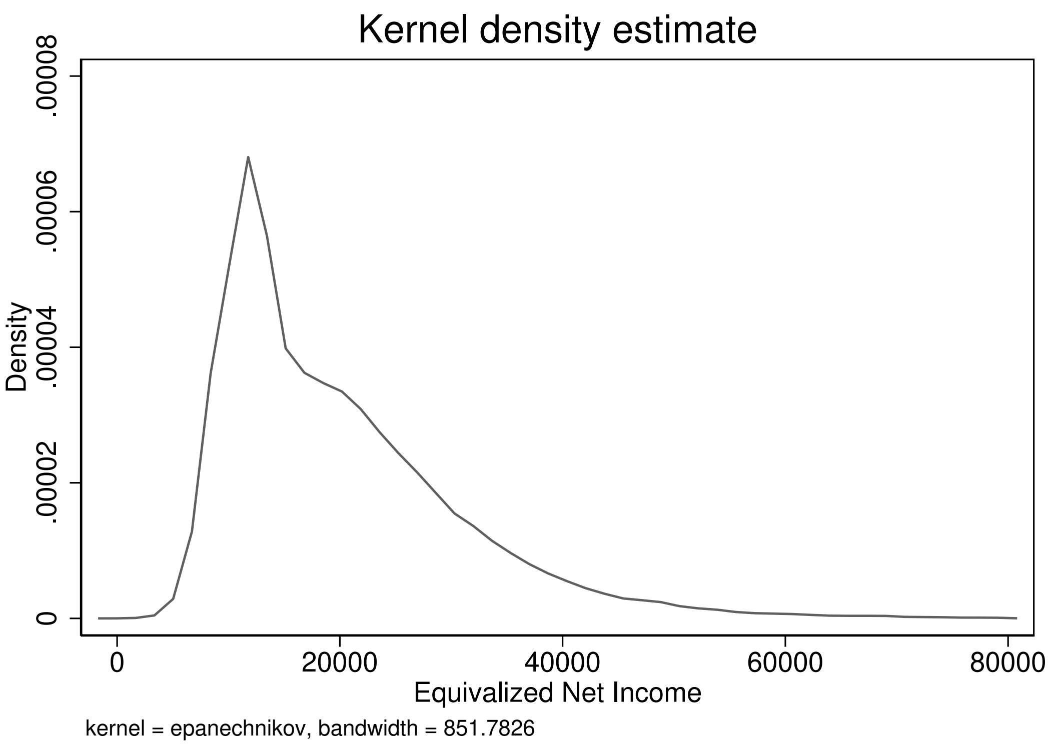
2.2 The Data

For the empirical part of this analysis, we use data from the German Socio-Econmic Panel (SOEP). It is a representative annual survey of about 15,000 households. For more information see Goebel et al. (2018) and Schröder et al. (2020). The survey has been conducted since 1984. We use wave 36. We employ retrospective information on the year 2017 only. The data contain detailed information on employment (including hours worked per week and wage rates), personal characteristics and the household context, which allows us to model labor supply and hypothetical wages in counterfactual scenarios.

We restrict the sample to those households with at least one member with flexible labor supply[1]. Furthermore, pensioners and people in parental leave are excluded as well as the self-employed since labor supply is difficult to model in self-employment.

These sample restrictions leave us with 12,911 households that account for 23,383,105 German household with 53,944,893 persons using the SOEP's representative weighting factors. The households are split into percentiles along the distribution of post-government income — equivalized using the new OECD equivalence scale.[2] Within this sample, we observe a typical positively skewed (equivalized net) income distribution in Figure 1.

Figure 1: Income distribution



[1]Inflexible labor supply is assumed for civil servants, students, members of the military, etc.
[2]The OECD-modified scale puts a weight of 1 to the first adult in a household, 0.5 to every other person aged 14 and older and 0.3 to every child under 14.

2.3 Microsimulation and Labor Supply Model

We use household data to simulate labor supply under the altered conditions for every household and individual in our sample. The general framework for this simulation is the microsimulation model EMSIM (Einkommensteuer-Mikrosimulationsmodell, see Bechara et al., 2015). It replicates all major aspects of the German system for income taxes, social security contributions and transfers. Hence, the EMSIM is able to use the SOEP data on observed (or imputed) market income and other relevant characteristics to compute the disposable income for every household in the sample. We combine the microsimulation model with a discrete labor supply model.

In principle, the labor supply model operates as follows. We estimate a utility function for every household depending on consumption and leisure (for both partners separately in couples). We simulate labor supply behavior of both singles and couples. In the case of couples, it is assumed that both partners choose the labor supply that maximizes their combined utility. The budget constraint is determined by non-labor income, both partners' wage rates and the tax and transfer system. Using this utility function, we can estimate probabilities for each hours category (combination) that a certain household chooses.

The discrete possible hours choices are 0, 10, 20, 39, and 45 hours per week for women and 0, 20, 40, and 48 for men. Hence, a household in which both partners are flexible in their labor supply can choose from 20 combinations of hours per week. For every labor supply alternative, the EMSIM simulates according to the tax and transfer system the disposable income which is ultimately equivalent to the household consumption in this one-period model. The observed gross income and hours worked in the SOEP are used to construct wage rates. Whenever this is not possible because an individual is not working in the status quo, potential wage rates are imputed via a Heckman-type (Heckman, 1979) wage regression based on human capital-related variables.

The labor supply responses are modelled following the approach by Aaberge et al. (1995) and van Soest (1995). Each household values consumption and leisure as displayed in the utility function in equation (3) depending on the selected amount of hours worked z.

$$V_z = U(Lf_z, Lm_z, C_z) + \varepsilon_z \tag{3}$$

Here, Lf and Lm are the hours of leisure of the female and the male household member, respectively, and C is consumption which is equivalent to the disposable income in this one-period model, while ε denotes a category-specific error term. The deterministic part of the utility function, U, is a translog utility function of the form

$$U = \beta_1 ln(C_z) + \beta_2 ln(C_z)^2 + \beta_3 ln(Lf_z) + \beta_4 ln(Lm_z) + \beta_5 ln(Lf_z) ln(Lm_z) + I_z, \quad (4)$$

where I_z is an hours-category specific fixed effect. Additionally, the coefficients in the utility function depend on socio-demographic characteristics like age, education and number of children in the household to allow for heterogeneity. This is crucial since, for example, typically female leisure time is valued higher in couples with children.

The error term ε in equation 3 is i.i.d. across the hour categories and households and assumed to follow an extreme-value type I (EVI) distribution. This allows a closed form solution for the probability P_z that the household chooses hour category z (McFadden, 1974). The resulting probability that alternative z is preferred by the household is given by the conditional logit model,

$$P_z = Pr(V_z > V_j, \forall j \neq z) = \frac{exp(U_z)}{\sum_{j=1}^{J} exp(U_j)}. \quad (5)$$

If a reform—in our case for a tax cut—changes the disposable income associated with certain choices of labor supply, it also changes the deterministic part of the labor supply function and hence the household's probability to choose that hour category. This, in turn, results in changes in the optimal amount of labor supply and the behavioral responses that we are interested in in this study. Table 1 shows the own-wage labor supply elasticities for different household types resulting from a 1 percent increase in gross wages. Similarly, Table 2 responses at the extensive margin, also simulated with a 1 percent increase in gross wages. These labor supply elasticities with larger elasticities for women in couples than for the rest of the population are in line with the literature (Keane, 2011).

Table 1: Own-Wage Labor Supply Elasticities

	women	men
single women	0.23	.
single men	.	0.29
couples, both flexible	0.39	0.15
couples, woman flexible	0.38	.
couples, man flexible		0.20
all	0.33	0.20

Note: Simulated with a 1 percent increase in gross wages.

Table 2: Own-Wage Participation Semi-Elasticities

	women	men
single women	0.11	.
single men	.	0.14
couples, both flexible	0.15	0.08
couples, woman flexible	0.14	.
couples, man flexible	.	0.10
all	0.09	0.06

Note: Simulated with a 1 percent increase in gross wages.

3 Results

We simulate a simple reform that increases the net incomes for all household in a specific quantile by 100 Euro, say, via an income tax cut that does not lead to further transaction cost. Hence the cost amounts to the mechanical cost of 100 Euro per receiving household plus the fiscal externality that originates from the labor supply reactions simulated in the model described in section 2.3. We then calculate the welfare weights $g(y)$ for each percentile in the full sample and display the results in Figure 2. Evidently, the marginal

Figure 2: All households

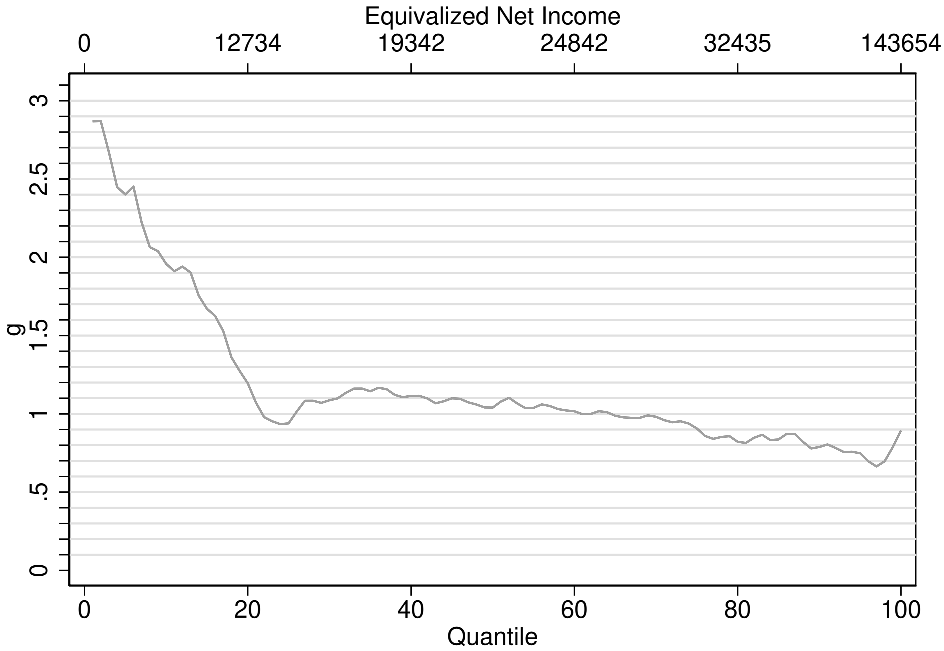
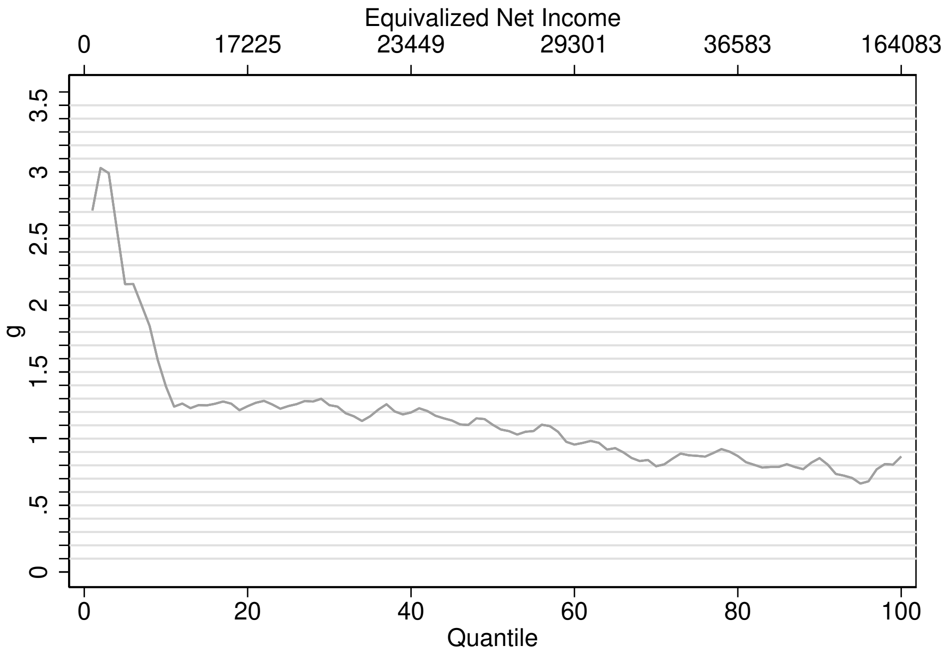


Figure 3: Childless couples

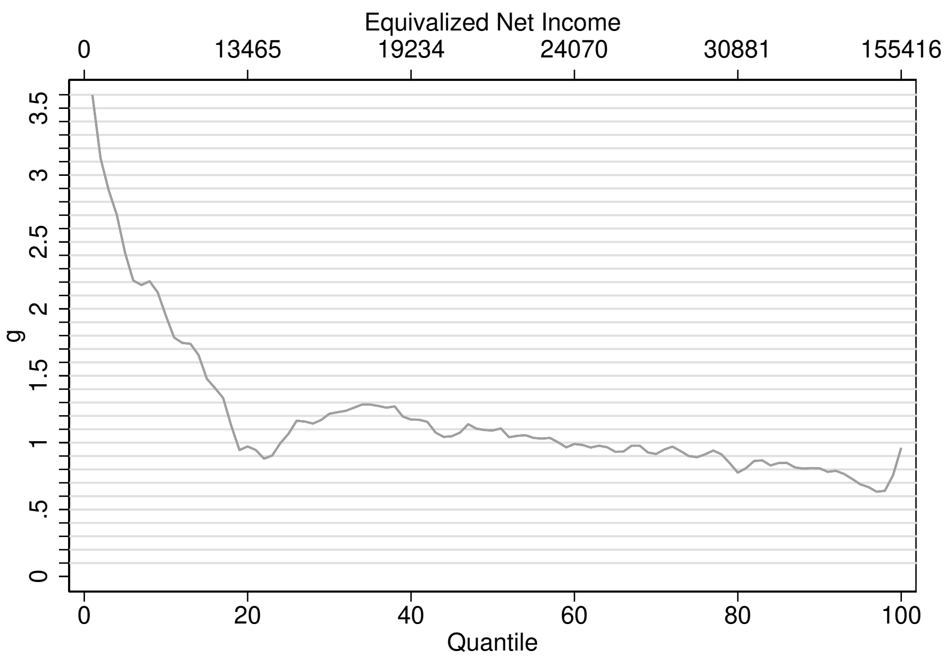


cost of transfers for low income earners is higher than for the rest of the income spectrum. The computed welfare weight for the very bottom percentile is close to 2.9 and subsequently drops to 2.0 around the 10th percentile and further to just below 1 at the 20th percentile. Hence, the current tax-transfer system is optimal if the social planner indeed values one additional Euro for the households in the 10th percentile twice as much as an additional Euro for the 20th percentile. Welfare weights decrease rather steeply at low incomes, while they only decrease moderately at higher income levels.

Strikingly, the welfare weights increase slightly between the 25th and 35th percentile before dropping again towards the high income earners with a notable increase in cost of transfers only at the very top of the income distribution. The local minimum at roughly one quarter along the distribution is likely due to the substantial participation elasticities and hence positive behavioral revenue effects of transfers to the working poor. This is in line with results from the existing literature (Bargain et al., 2014; Blundell et al., 2009; Jessen et al., 2018).

Figure 3 shows the result for the simulations in a restricted sample of only childless couples. At the lower end of the income distribution, the welfare weights are equally high with a similarly steep decline as in the full sample. However, the weights between the 10th and 40th percentile remain stable and well above 1 as opposed to the dip around the 20th and 25th percentile including all households. Beyond the 40th percentile, the curve of social welfare weights is slightly downward sloping, reaching its lowest point at the 95th percentile.

Figure 4: Couples with children



In contrast to this, couples with children have even higher weights at the lower end of the income distribution (Figure 4). Generally, couples with children have a notably lower equivalized net income (upper axis) than childless couples. The low marginal cost of a tax cut is observable in this subsample around the 20th percentile. Increasing weights for the working poor in the full sample appear to be driven by couples with children. This is likely due to two factors: i) women in couples with children have particularly high participation elasticities leading to a relatively high revenue increases due to participation responses. ii) Families might be transfer recipients even with relatively higher gross incomes than singles such that marginal withdrawal rates are quite high even at the 20th percentile.

The computed welfare weights of singles—without and with children—are reported in Figures 5 and 6. Singles' equivalized net income in the sample at the given percentiles is even lower than the equivalent for couples. The welfare weights are more volatile along the income distribution due to the smaller sample sizes. Nonetheless, we observe the same steep decrease in marginal social welfare weights at low income levels. At higher income levels, weights are essentially flat.

Figure 5: Childless Singles

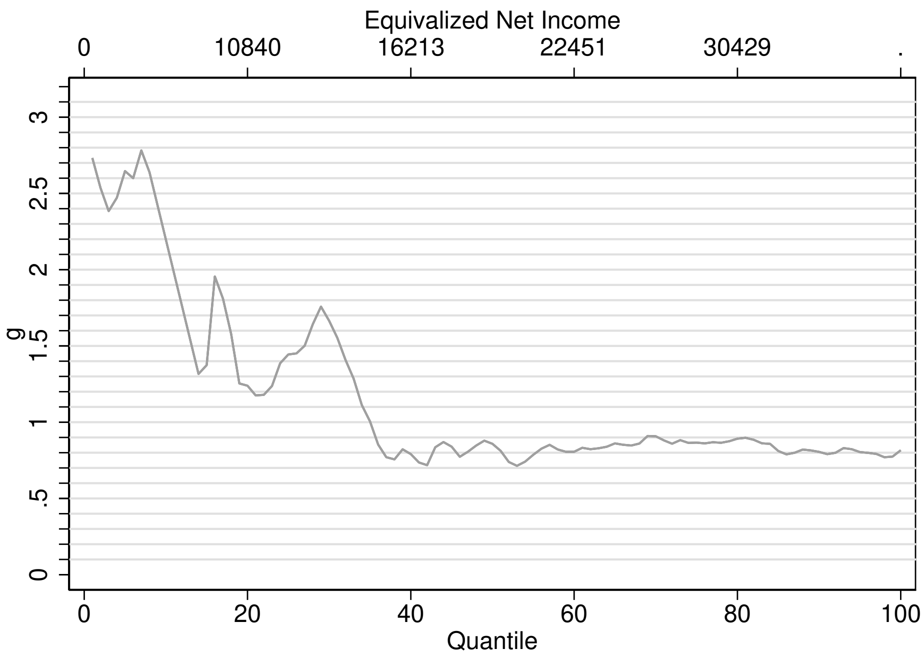


Figure 7 shows marginal welfare weights for different household types, simulated with local tax reforms for every decile, which leads to substantially smoother welfare weights, and with equivalized net incomes instead of deciles on the horizontal axis. At medium income levels implied welfare weights for couples are lower than for children. In other words, tax cuts for couples at these income levels are more expensive than tax cuts for singles. This is likely due to labor supply reactions of secondary earners.

Figure 6: Lone parents

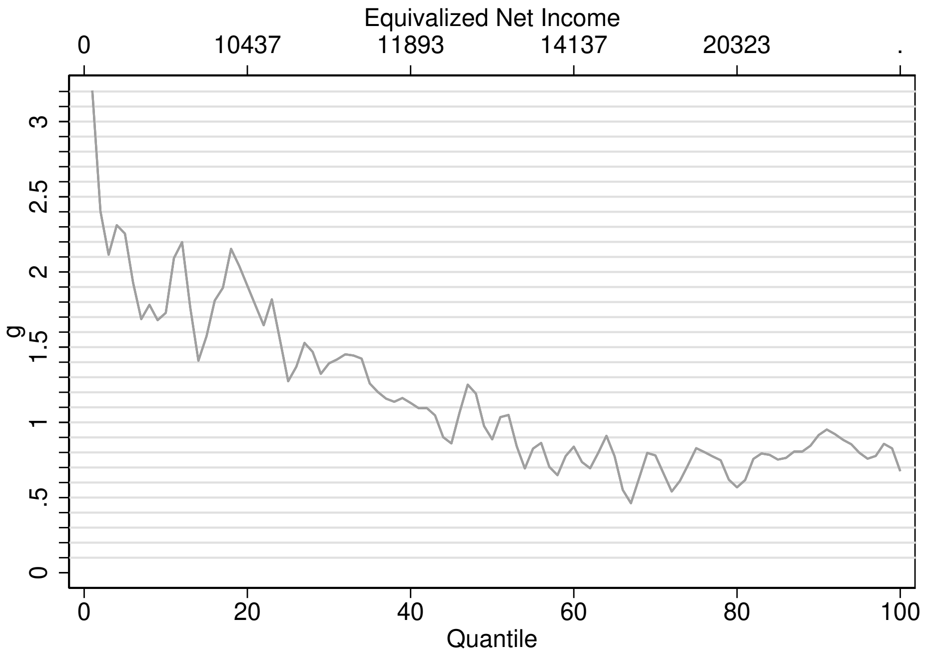
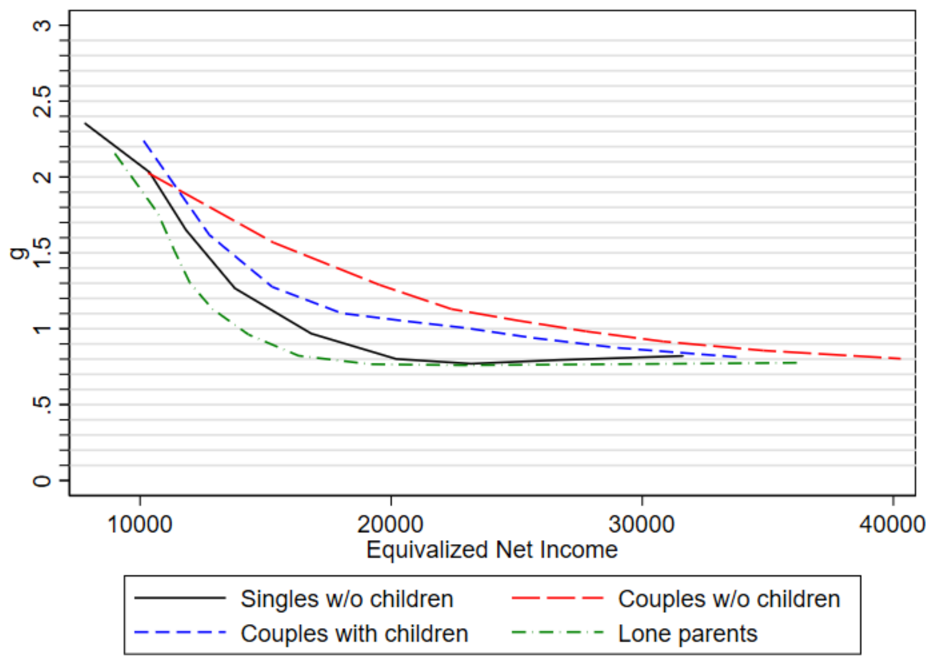


Figure 7: Different household types



4 Decomposition of Fiscal Externalities

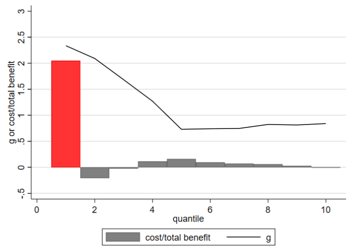
In this section we demonstrate how labor supply reactions at different deciles explain our results. For childless singles, the bars in Figure 8 show the fiscal costs at the deciles of the income distribution induced by small tax reductions at the decile in red divided through the mechanical benefit. Each bar represents the fiscal cost at the income level that responded to the relevant decile before the small tax reform. In the case of no behavioral adjustment, the red bar equals one and all other bars equal zero. The line shows the inverse-optimum social marginal welfare weight at each decile. All bars add up to the value of the marginal welfare weight of the decile in red (the value of the line at the relevant decile).

Consider the first decile. The mechanical cost of a small tax reduction divided through the benefit is one. Apparently, labor supply reactions increase the fiscal cost observed at income levels corresponding to the first decile. The reason is that households at that income level are net transfer recipients. Thus, households "moving" into that decile increase costs. Households at the second and third decile are still on average net recipients, therefore households moving away from that decile by itself reduces the fiscal cost, which corresponds to negative values for the bars at the second and third decile. The reduction of labor supply by households at the upper seven deciles leads to a reduction in tax revenue (represented by positive bars at these deciles) in addition to the increase in transfer payment due to households moving to the first decile. At the first decile, the size of the bar is the product of the number of households who move to this income level times the average transfer received by this households divided through the benefit of the tax reduction. At deciles 2-9, the size of each bar is the product of the number of households that reduce labor supply times the net taxes paid (or transfers received) on average by these households divided through the benefit of the tax reduction. Apparently, the extensive labor supply margin plays a role for all deciles except for the highest. Thus, it is important to capture the extensive margin.

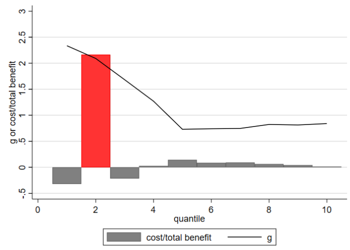
Decile 4 is the first, where labor supply reactions overall reduce the fiscal cost of a small local tax reduction, implying a marginal welfare weight below unity. The reason is that former transfer recipients increase their labor supply and move to decile 4, where they become net tax payers.

For small tax reductions in deciles 2-5, labor supply reactions from all deciles play a role. Importantly, tax reductions for decile 1 lead to labor supply reactions from all but the top decile. This increases the fiscal cost of this small tax reduction, leading to a higher weight than is often found for the working poor in positive optimal tax papers. These labor supply reactions are ignored in optimal tax models like Saez (2001) and Saez (2002). As we move to tax reductions for high income earners, only labor supply increases by workers higher up in the income distribution matter.

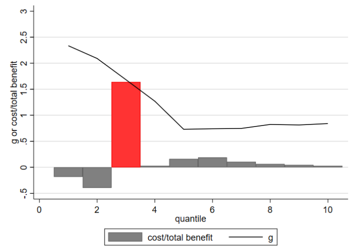
Figure 8: Decomposition of Fiscal Externalities: Childless singles



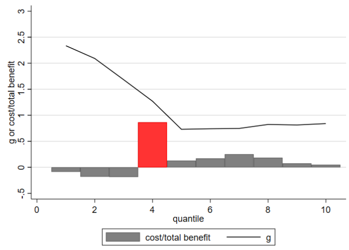
(a) Decile 1



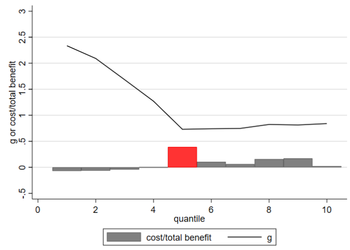
(b) Decile 2



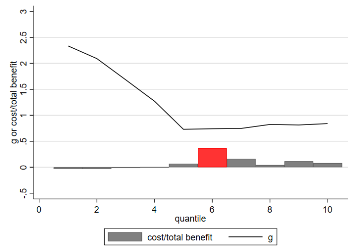
(c) Decile 3



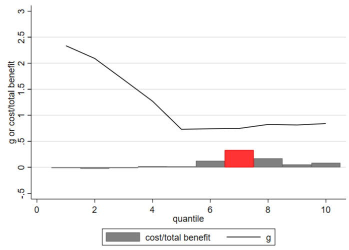
(d) Decile 4



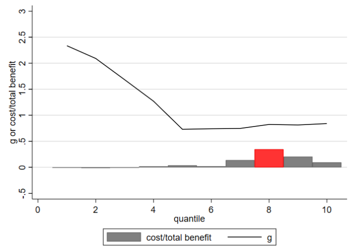
(e) Decile 5



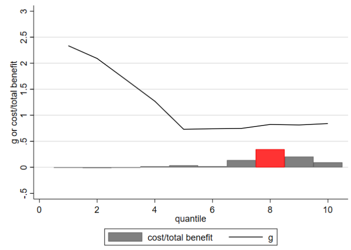
(f) Decile 6



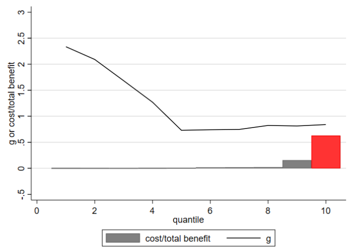
(g) Decile 7



(h) Decile 8



(i) Decile 9

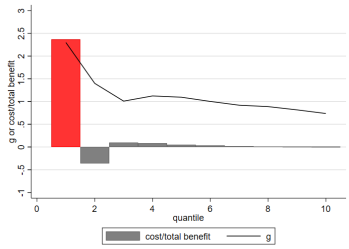


(j) Decile 10

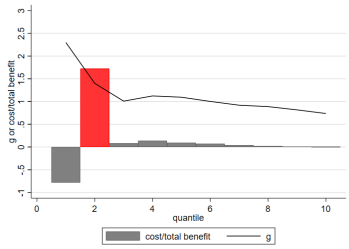
Note: The bars show the fiscal costs at the deciles of the income distribution induced by small tax reductions at the decile in red divided through the mechanical benefit. In the case of no behavioral adjustment, the red bar equals one and all other bars equal zero. The line shows the inverse-optimum social marginal welfare weight at each decile.

The equivalent decomposition of fiscal externalities for couples with children is reported in Figure 9. Consider decile 5. In comparison to singles, the bars at non-directly affected deciles are larger, implying larger labor supply reactions. In particular, households ni the sixth and seventh decile substantially reduce their labor supply. These reactions, which are due to secondary earners, explain why the fiscal cost of a tax reduction at that decile is larger than it is for singles.

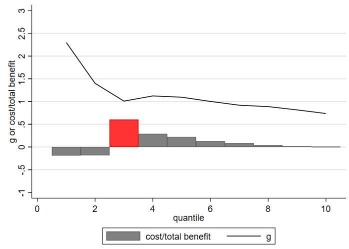
Figure 9: Decomposition of Fiscal Externalities: Couples with Children



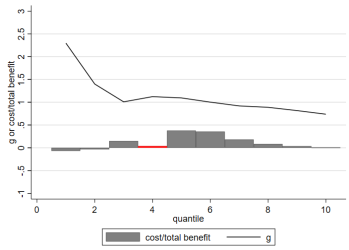
(a) Decile 1



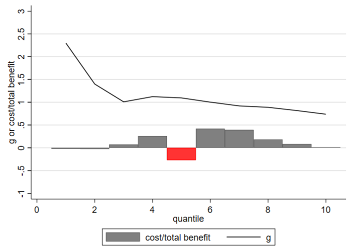
(b) Decile 2



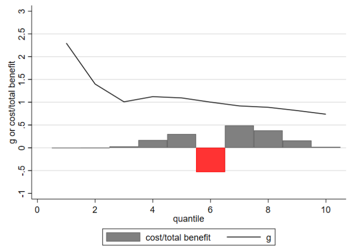
(c) Decile 3



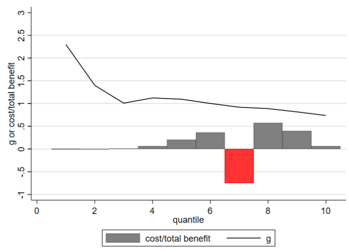
(d) Decile 4



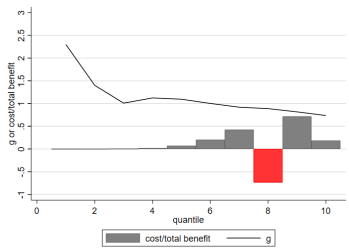
(e) Decile 5



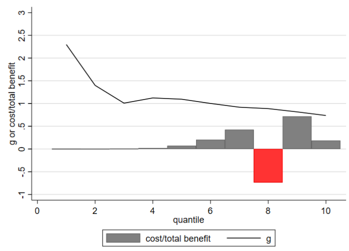
(f) Decile 6



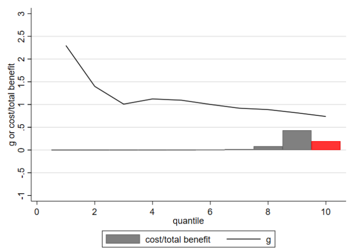
(g) Decile 7



(h) Decile 8



(i) Decile 9



(j) Decile 10

Note: The bars show the fiscal costs at the deciles of the income distribution induced by small tax reductions at the decile in red divided through the mechanical benefit. In the case of no behavioral adjustment, the red bar equals one and all other bars equal zero. The line shows the inverse-optimum social marginal welfare weight at each decile.

5 Conclusion

By simulating labor supply reactions to small tax cuts throughout the German income distribution, we have estimated social marginal welfare weights implied by the German tax-transfer schedule. The main advantage of our approach is that it does not impose strong restrictions on labor supply behavior and is applicable to both single and couple households. This approach offers a simple method to identify feasible Pareto improving reforms, namely, if $g(y)$ at a specific income level is negative. It should be noted that our approach only tests for local Pareto optimality as, in principle, tax-transfer systems that are very different from the status quo could exist which would make everyone better off.

We have found that the current tax-transfer system is optimal if the social planner values one Euro for households at the 10th percentile twice as much as one Euro for households at the median of the distribution. At medium incomes, implied weights for couples are higher than for singles as tax reductions for couples lead to substantial labor supply reductions.

Avenues for future research include the application to other countries or time-periods to analyze implicit value judgements in other settings. Moreover, it would be fruitful to compare the results to those obtained from widely used optimal taxation models for the same sample and labor supply elasticities as we have used. Finally, the proposed approach could be used for normative optimal taxation purposes. To this end, net incomes associated with specific gross income levels would have to be adjusted iteratively until the implied marginal social welfare weights were in line with the normative judgement imposed by the researcher.

References

Aaberge, R., Dagsvik, J. K., and Strøm, S. (1995). Labor supply responses and welfare effects of tax reforms. *Scandinavian Journal of Economics*, 97(4):635–659.

Atkinson, A. B. and Stiglitz, J. E. (2015). *Lectures on public economics: Updated edition*. Princeton University Press.

Ayaz, M., Fricke, L., Fuest, C., and Sachs, D. (2021). Who Should Bear the Burden of Covid-19 Related Fiscal Pressure? An Optimal Income Taxation Perspective. CESifo Working Paper Series 9420, CESifo.

Bargain, O., Dolls, M., Neumann, D., Peichl, A., and Siegloch, S. (2014). Comparing Inequality Aversion across Countries when Labor Supply Responses Differ. *International Tax and Public Finance*, 21(5):845–873.

Bechara, P., Kasten, T., and Schaffner, S. (2015). Dokumentation des rwi-einkommensteuer-mikrosimulationsmodells (emsim). RWI Materialien 86, RWI - Leibniz-Institut für Wirtschaftsforschung.

Blundell, R., Brewer, M., Haan, P., and Shephard, A. (2009). Optimal Income Taxation of Lone Mothers: An Empirical Comparison of the UK and Germany. *Economic Journal*, 119(535):F101–F121.

Bourguignon, F. and Spadaro, A. (2012). Tax-Benefit Revealed Social Preferences. *Journal of Economic Inequality*, 10(1):75–108.

Burtless, G. and Hausman, J. A. (1978). The effect of taxation on labor supply: Evaluating the gary negative income tax experiment. *Journal of political Economy*, 86(6):1103–1130.

Finkelstein, A. and Hendren, N. (2020). Welfare analysis meets causal inference. *Journal of Economic Perspectives*, 34(4):146–67.

Fullerton, D. and Gan, L. (2004). A simulation-based welfare loss calculation for labor taxes with piecewise-linear budgets. *Journal of Public Economics*, 88(11):2339–2359.

Goebel, J., Grabka, M., Liebig, S., Kroh, M., Richter, D., Schröder, C., and Schupp, J. (2018). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161.

Hendren, N. (2020). Measuring economic efficiency using inverse-optimum weights. *Journal of Public Economics*, 187(C).

Hendren, N. and Sprung-Keyser, B. (2020). A Unified Welfare Analysis of Government Policies. *The Quarterly Journal of Economics*, 135(3):1209–1318.

Jacobs, B., Jongen, E. L., and Zoutman, F. T. (2017). Revealed social preferences of dutch political parties. *Journal of Public Economics*, 156:81–100.

Jessen, R., Metzing, M., and Rostam-Afschar, D. (2018). Optimal taxation under different concepts of justness. Ruhr Economic Papers 762, RWI - Leibniz-Institut für Wirtschaftsforschung, Ruhr-University Bochum, TU Dortmund University, University of Duisburg-Essen.

Jessen, R., Metzing, M., and Rostam-Afschar, D. (2022). Optimal taxation when the tax burden matters. *Finanzarchiv/Public Finance Analysis*, forthcoming.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature*, 49(4):961–1075.

Lockwood, B. B. and Weinzierl, M. (2016). Positive and normative judgments implicit in u.s. tax policy, and the costs of unequal growth and recessions. *Journal of Monetary Economics*, 77:30–47.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68:205–229.

Saez, E. (2002). Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses. *The Quarterly Journal of Economics*, 117(3):1039–1073.

Saez, E. and Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, 106(1):24–45.

Schröder, C., König, J., Fedorets, A., Goebel, J., Grabka, M. M., Lüthen, H., Metzing, M., Schikora, F., and Liebig, S. (2020). The economic research potentials of the German Socio-Economic Panel study. *German Economic Review*, 21(3):335–371.

van Soest, A. (1995). Structural models of family labor supply: A discrete choice approach. *Journal of Human Resources*, 30(1):63–88.