I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17566

Normative Judgments Implicit in the Tax System: A Simulation Approach

Niklas Isaak
Robin Jessen

DECEMBER 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17566

Normative Judgments Implicit in the Tax System: A Simulation Approach

Niklas Isaak
RWI and Freie Universität Berlin

Robin Jessen
RWI and IZA

DECEMBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Normative Judgments Implicit in the Tax System: A Simulation Approach*

How much does society value redistribution? The common method to derive inverse-optimum welfare weights is by inverting an optimal-tax model. Our alternative imposes fewer restrictions on labor supply and enables comparisons across household types. We use a structural labor supply model to calculate the marginal value of public funds for various small tax reductions, directly linked to welfare weights. An application to Germany finds: i) The tax-transfer system is optimal if society values one additional Euro for the bottom decile three times as much as for the median. ii) At low-medium incomes, weights for couples exceed those for singles substantially.

JEL Classification: H21, H31, J22

Keywords: inverse optimum, microsimulation, marginal value of public funds, social welfare function, optimal taxation, labor supply, efficiency

Corresponding author:
Robin Jessen
RWI
Zinnowizer Str. 1
10115 Berlin
Germany

E-mail: robin.jessen@rwi-essen.de

* We thank Almut Balleer, Spencer Bastani, Peter Haan, Nathan Hendren, Dominik Sachs, David Seim, and seminar participants at RWI, the 2022 annual congress of the IIPF in Linz, the 2022 EEA congress in Milan, the WATT-Workshop on Inequality and Taxation at DIW Berlin in 2022, the 2023 ECINIEQ conference in Aix-en-Provence, and the 23rd Journées LAGV 2024 in Marseille for valuable comments.

1 Introduction

Modern welfare states redistribute income from higher- to lower-income earners. Labor supply reactions to progressive taxation imply that this redistribution is costly: Redistributing one additional Euro to the poor typically necessitates a reduction in consumption by more than one Euro for those whose taxes are raised. Redistribution is thus only desirable if society values additional consumption by the poor more than additional consumption by the well-off. Tax-transfer systems, if deemed optimal, thus imply normative judgments about this equity-efficiency trade-off. A natural question is what valuation of additional consumption by different income groups is implicit in the status-quo tax-transfer system. The answer to this question can be used to judge future tax reforms that might be necessitated by changes in economic circumstances (see Ayaz et al., 2023). Moreover, if the implicit value judgments appear implausible, for instance, implying negative weights on some income groups, this might be a reason for reform.

The standard approach in economics to quantifying value judgments implicit in tax systems is inverting an optimal tax model and calculating the welfare weights g_i (see, e.g., Mirrlees, 1971; Saez, 2001) under which the status-quo tax-transfer system is optimal (for example, Blundell et al., 2009; Bourguignon and Spadaro, 2012; Jacobs et al., 2017; Lockwood and Weinzierl, 2016; Hendren, 2020). These weights indicate how much society values providing one additional Euro of consumption to an income level i. At the optimal tax system, the value of providing an additional Euro to a household must equal the cost of doing so. Considering labor supply reactions to tax changes, this cost generally differs from one. The literature typically applies workhorse optimal tax models such as Saez (2001) and Saez (2002), which allow for intuitive analytical solutions, while imposing restrictions on labor supply reactions to tax changes.

In this paper we propose an alternative approach: We use microsimulation and a structural labor supply model to simulate responses to hypothetical small tax reforms that increase disposable income at specific parts of the income distributions for specific household types. We then calculate the fiscal cost per Euro redistributed through these marginal reforms. By doing so, we obtain the marginal value of public funds (MVPF; Finkelstein and Hendren, 2020; Hendren and Sprung-Keyser, 2020) of these small reforms. The MVPF 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 relevant impact on the utility of individuals who adjust their labor supply in reaction to it. The calculations of the MVPF are valid even if the current tax-transfer system is not optimal. Additionally assuming that the system is optimal, the MVPF is directly related to the inverse-optimum welfare weight (see Bastani, forthcoming). The marginal social welfare weight of individuals at an income level is then simply given by 1/MVPF of a small reform that redistributes to individuals in that income level. A key advantage of this approach is that do not need to focus on a specific, homogeneous subsample nor do we need to impose restrictions on labor supply, e.g., of couples, in order to obtain analytical solutions. This procedure offers a simple way to identify potential Pareto-improving reforms, namely, if g_i at a specific income level is negative (see also Bierbrauer et al., 2023).[1]

We apply the procedure to Germany and find that the current income tax and transfer system is optimal if society values one additional Euro of consumption for households in the bottom decile of the income distribution about three times as much as one additional Euro for households at the median. At the bottom of the distribution, weights decrease sharply, while they are almost constant for medium to top incomes. A key advantage of our approach is that it allows us to calculate inverse-optimum weights for different household types. To do so, we simulate small reforms that increase net incomes for a specific household type at a specific income level. We find that, for given equivalized net income levels, couple households have higher weights than single households. The reason is that reducing the tax burden of low-to-medium income couples induces a relevant number of higher-income couples to reduce their labor supply. This is not the case for singles and is a consequence of the explicit modeling of couples' labor supply.

Our model is essentially the one in Saez (2002) with the key difference that we allow for more flexible labor supply reactions to tax changes. In the model, there are discrete income groups and the efficiency cost of a change in the tax schedule depends on how many households move from one income group to another in reaction to a tax change. In Saez (2002), there are no income effects and labor supply reactions are restricted to 'neighboring' groups and into or out of unemployment. In contrast, we do not impose these restrictions and allow for income effects. In our application, we use the microsimulation model EMSIM (Einkommensteuer-Mikrosimulationsmodell, see Bechara et al., 2015), which is based on the the Socio-Economic Panel (SOEP), a representative survey of German households, and a discrete choice structural labor supply model following Aaberge

[1] Note that the procedure only tests for local Pareto optimality as, in principle, tax-transfer systems that are very different from the status quo could exist that would make everyone better off.

et al. (1995) and van Soest (1995) to simulate labor supply reactions. This workhorse labor supply model allows flexible modeling of labor supply reactions and does not impose restrictions of the budget constraint. Importantly, in future research, alternative labor supply models could be used for our approach. We compare and contrast the welfare weights implicit in the German tax system that are attained by our simulation approach to those resulting from the Saez (2002) model using a common sample of households in the SOEP. Our approach yields higher weights for non-working households and, importantly, higher weights for couple households relative to single households at low-medium incomes compared to the calculated weights following Saez (2002). The reason is that the Saez (2002) model imposes restrictions on labor supply, which we show to be particularly important for couple households. This highlights the importance of carefully modeling labor supply of couples. In the Saez (2002) model with no income effects, the weights for a specific household type add up to one. With no income effects, taxing away a specific amount of money uniformly from households of one type and giving it uniformly to households of a different type, leads to no labor supply reactions. Therefore, assuming optimality, the average weights of different household types are the same. In contrast, we allow for income effects, which enables meaningful comparisons between household types.

A number of papers quantify marginal social welfare weights implicit in various countries' income tax schedules. Typically, these papers apply well-known optimal tax models (see Atkinson and Stiglitz, 2015) such as Saez (2001), which rules out participation effects of tax changes (Ayaz et al., 2023; Bourguignon and Spadaro, 2012; Lockwood and Weinzierl, 2016), or the model in Saez (2002), which allows for participation responses, but restricts labor supply reactions to neighboring income groups and rules out income effects (Bargain et al., 2014; Bourguignon and Spadaro, 2012; Blundell et al., 2009; Jessen et al., 2022). In several of these papers (Blundell et al., 2009; Jessen et al., 2022; Bargain et al., 2014), 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. Haan and Navarro (2008) extend the approach in Saez (2002) and allow for changes into non-neighboring income groups, but also assume away income effects. Moreover, they hold labor supply of primary earners in couple households fixed. Jacobs et al. (2017), Bastani and Lundberg (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, i.e., large increases or decreases in hours worked in reaction to small tax changes. 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 (Keane, 2011). For a given utility function, the labor supply elasticity then depends on the budget set and an elasticity obtained in one context cannot readily be applied to another. Non-convex budget sets are an important feature of European welfare states with high marginal transfer withdrawal rates.[2] 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.[3]

The inverse optimal taxation literature typically finds optimal welfare weights that do not decrease strictly with income as would be the case under an inequality averse social planner (e.g., Bargain et al., 2014; Blundell et al., 2009; Jessen et al., 2022). 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, particularly for singles, their weights are only slightly smaller than those at the median. 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. Stronger decreases in labor supply due to tax reductions imply a higher fiscal cost and hence higher inverse-optimal weight. Jumps in labor supply 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 this paper, we show what types of labor supply reactions determine the fiscal cost of marginal additional redistribution to specific income groups. For couples, intensive "jumps" of total income play a more important role than for singles, which leads to higher weights at low-medium incomes.[4] In contrast, in the Saez (2002) model, a local tax break will only induce individuals with slightly higher income

[2]As an illustration, Figure A.1 in the appendix shows the budget constraint of a typical single household in Germany.

[3]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.

[4]Few optimal tax papers model labor supply of couples. One is Kleven et al. (2009), who restrict the labor supply decision of the secondary earner to the participation decision. The labor supply reactions of singles and couples are identical if couples are taxed jointly and labor supply is characterized by a single intensive labor supply elasticity that is the same for both partners (Bach et al., 2012).

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.

Assuming optimality of the tax-transfer schedule, our results imply that society values one additional Euro for couples more than one additional Euro for singles. Economic theory and empirical work starting with Lazear and Michael (1980) provide a rationale for this. Some goods and services consumed within a household provide non-rival benefits to all its members. Therefore, one additional Euro for couples allows more additional consumption than one Euro for singles, implying a higher welfare weight for individuals in couple households at a given individual productivity (Golosov and Krasikov, 2023). A second reason is that it can be optimal for an inequality averse social planner to incentivize people to form couple households. Golosov and Krasikov (2023) study the optimal taxation of singles and couples in a model of couple formation. With imperfect assortative matching in the marriage market, couples provide some socially valuable redistribution since spouses share resources within the household. Therefore, the optimal tax schedule includes a "subsidy" for couples. The resulting tax schedule then appears to imply higher welfare weights for individuals in couple households in a simple model like ours with exogenous household formation.

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 the resulting marginal social welfare weights implied by the status-quo tax-transfer schedule. Section 4 documents the labor supply reactions to marginal, local tax cuts that lead to these results. A comparison of social welfare weights obtained by our model to those following the Saez (2002) model is provided in section 5. Section 6 concludes.

2 Model and Methodology

2.1 Inverse Optimal Taxation

2.1.1 The Optimal Tax Model

The general model — We build on the model in Saez (2002), but with two key differences. First, we impose fewer restrictions on labor supply, and second, we distinguish between different household types. There are K household types, denoted by $k = 1, \ldots .K$. The number of households of each type is n_k These types can be thought of as singles and couples with or without children, but additional distinctions are possible. There is no need to impose restrictions on labor supply reactions of the household members to tax reforms. Households do not change type. For a given gross income, the tax burden might differ between household types. Households are indexed by m and decide on their labor supply, which determines their gross income group within their household type, $i = 0, ..., I$, in order to maximize utility $u_{m,k}(c_{i,k}, i)$. Consumption equals net income, $c_{i,k} = y_{i,k} - T_{i,k}$. $T_{i,k}$ can be positive or negative, representing both taxes and transfers. $h_{i,k}$ denotes the fraction of households in income group i within household type k. Disutility from work is well distributed such that $h_{i,k}$ is differentiable.

M_k with $m \in M_k$ is a set normalized to measure one, divided into subsets $M_{0,k}, .., M_{I,k}$. The measure of subset $M_{i,k}$ is simply $h_{i,k}$. The measure of households on M_k is $d\nu_k(m)$. The social planner decides on the tax-transfer system and maximizes social welfare given by

$$W = \sum_{k=1}^{K} n_k \int_{M_k} \mu_{m,k} u_{m,k}(y_{i^*,k} - T_{i^*,k}, i^*_m) \mathrm{d}\nu_k(m) \tag{1}$$

where $\mu_{m,k}$ are weights and the government budget constraint is

$$G = \sum_{k=1}^{K} n_k \sum_{i=0}^{I} T_{i,k} h_{i,k}, \tag{2}$$

where G is an exogenous spending requirement. Thus, government can redistribute both within and between household types. The first-order condition with respect to $T_{i,k}$ is

$$-\int_{M_{i,k}} \mu_{m,k} \frac{\partial u_{m,k}(c^*_{i,k}, i*)}{\partial c_{i,k}} \mathrm{d}\nu_k(m) + \lambda \left(h_{i,k} + \sum_{j=0}^{I} T_{j,k} \frac{\partial h_{j,k}}{\partial T_{i,k}} \right) = 0, \tag{3}$$

essentially equation (9) in Saez (2002) with added subscript k, where λ is the Lagrange multiplier of the government budget constraint. Due to the envelope theorem, the change in the measure of the set $M_{i,k}$ has no first-order impact on welfare. Saez (2002) imposes specific restrictions on labor supply, see below. Under these restrictions, the second term on the left hand side is an elegant formula containing two elasticities. Instead, we allow for flexible labor supply reactions and directly obtain the change in government revenue through simulation.

As in Saez (2002), we define marginal social welfare weights as

$$g_{i,k} = \frac{1}{\lambda h_{i,k}} \int_{M_{i,k}} \mu_{m,k} \frac{\partial u_{m,k}(c^*_{i,k}, i^*)}{\partial c_{i,k}} \mathrm{d}\nu_k(m). \tag{4}$$

Substituting equation (4) into equation (3) and rearranging, we obtain

$$g_{i,k} h_{i,k} = h_{i,k} + \sum_{j=0}^{I} T_{j,k} \frac{\partial h_{j,k}}{\partial T_{i,k}}. \tag{5}$$

It is worth noting that the sum of equations (5) for all income groups within a household type does not need to equal one as we allow for income effects. Consequently, redistributing money uniformly from one household type to another might lead to behavioral reactions and hence additional fiscal cost. This is not the case in a scenario with no income effects, where it is impossible to quantify the social planner's taste for redistribution between household types.

Consider the government wants to redistribute 1 Euro to each individual in income group i. This leads to an increase in social welfare by $g_{i,k}h_{i,k}$. The direct cost of this tax reduction is $h_{i,k}$ and an additional cost (or a cost reduction) occurs because of labor supply reactions, the second term on the right-hand side. At the optimum, the benefit and cost of a small tax change are equal. Dividing equation (5) through $h_{i,k}$, we obtain cost and benefit per additional Euro redistributed:

$$g_{i,k} = \underbrace{1}_{\text{mechanical cost per add. 1 Euro redistributed}} + \underbrace{\frac{1}{h_{i,k}} \sum_{j=0}^{I} T_{j,k} \frac{\partial h_{j,k}}{\partial T_{i,k}}}_{\text{behavioral cost per add. 1 Euro redistributed}} \tag{6}$$

This interpretation is also discussed in Hendren (2020)—at the optimum, society's valuation of an additional Euro of consumption for an individual in household type k at income level i, $g_{i,k}$, must equal the marginal cost of a taxation reduction.[5] The behavioral cost is often referred to as fiscal externality and the concept is closely related to the marginal value of public funds (MVPF), which is calculated as MVPF $= 1/(1+$ behavioral cost per add. 1 Euro redistributed) (Hendren and Sprung-Keyser, 2020). Thus,

$$g_{i,k} = 1/\mathrm{MVPF}_{i,k}, \tag{7}$$

[5] Similarly, Chetty (2008), Schmieder et al. (2012), and Schmieder and Von Wachter (2016) express the value and cost of additional unemployment insurance per additional $1 redistributed.

where $\text{MVPF}_{i,k}$ is the MVPF of a small increase in disposable incomes at income level i for household type k.

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 $g_{i,k}$ is negative.

Special case: Saez (2002) — We nest the model in Saez (2002), which allows for only one household type, assumes no income effects, and that labor supply reactions to tax changes are defined through two types of elasticities. These elasticities are allowed to differ between income groups. As there is only one household type, we now omit the subscript k. The participation elasticity is defined as

$$\eta_i = \frac{c_i - c_o}{g_i} \frac{\partial h_i}{\partial (c_i - c_0)}. \tag{8}$$

The participation reactions of group i depend only on the difference in net incomes between group 0 and group i. The intensive elasticity is defined as

$$\zeta_i = \frac{c_i - c_{i-1}}{g_i} \frac{\partial h_i}{\partial (c_i - c_{i-1})}. \tag{9}$$

Again, as there are no income effects, movements between groups i and $i-1$ depend only on the difference in disposable income between these two groups.

Under these restrictions, it is straightforward to show that, at the optimal tax schedule, the marginal social welfare weight for the highest income group is

$$g_I = 1 - \eta_i \frac{T_I - T_0}{C_I - C_0} - \eta_I \frac{T_I - T_{i-1}}{C_I - C_{i-1}}. \tag{10}$$

For income groups $i = 1, ...I - 1$, the marginal social welfare weights are[6]

$$g_i = 1 - \eta_i \frac{T_i - T_0}{C_i - C_0} - \zeta_i \frac{T_i - T_{i-1}}{C_i - C_{i-1}} + \zeta_{i+1} \frac{T_{i+1} - T_i}{C_{i+1} - C_i} \frac{h_{i+1}}{h_i}. \tag{11}$$

The behavioral cost of a tax reduction for group i, and hence its marginal social welfare weight (see equation (6)) if the current tax-transfer schedule is optimal, depends only on behavioral reactions from non-participation and the two neighboring groups. With no

[6] Bourguignon and Spadaro (2012) also invert the optimal tax formulae by Saez (2002) with the canonical restrictions to labor supply, their equations (20) and (21). In the formulae derived in that paper, g_i is a function of other g_j, elasticities, tax levels, and income levels. By substituting one can obtain formulae for g_i that only depend on elasticities, income levels, and tax levels.

income effects, the weighted sum of marginal social welfare weights is one (Saez, 2002). Therefore, the weight for the lowest income group, $i = 0$, can simply be calculated as

$$g_0 = \left(1 - \sum_{j=1}^{I} h_j g_j\right) / h_0. \tag{12}$$

2.1.2 Simulating the Behavioral Cost of a Marginal Tax Change

Typically, optimal taxation papers impose restrictions and derive analytical expressions for the fiscal externality based on labor supply elasticities. In contrast, we simulate the behavioral cost 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 given their hourly wage. For couple households, we allow for hours adjustments of both partners. For instance, a married heterosexual couple can move to a higher income group by having the wife, the husband, or both work more hours.

For each household type, we divide the sample into deciles of the distribution of net equivalized incomes and, decile by decile, increase disposable income for all households at a specific decile of the distribution by a marginal amount. We choose 100 Euro per year. Using smaller or slightly larger amounts has no significant impact on the results. If for a specific household, income level i is within reach, the probability that it chooses i 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 with specific characteristics in the population who work that number of hours. With the sample weights and choice probabilities in the status quo and a hypothetical tax cut scenario in hand, we calculate the total (behavioral and mechanical) decrease in government revenue due to a small tax cut at income level i divided through the mechanical reduction in revenues (100 Euro times the number of households at income level i in the status quo) for every small reform for all I income levels and K household types. In our case, we partition the sample into 4 household types with 10 income groups each. Hence, we perform 40 separate simulations. Thereby, we directly obtain $g_{i,k}$ as in equation (4).

In Germany, increasing disposable incomes for given income levels can be achieved by increasing transfers or reducing taxes. Married couples are taxed jointly, i.e. $T_m(y_1, y_2) = 2 * T((y_1 + y_2)/2)$, where y_1 and y_2 are gross incomes of the spouses and $T()$ is the tax schedule applicable to singles. Thus, the tax rate only depends on the total gross income of the couple. Instead, for unmarried couples, the tax liability is given by $T_u = T(y_1)+T(y_2)$ and depends on the within-household distribution. In our simulations, we increase disposable incomes at a specific level of household (net or gross) income. In practice, such reforms could easily be conducted for singles and married couples by changing the income tax schedule or means-tested benefits. For unmarried couples, increasing disposable incomes at specific income levels independent of the within-household income distribution represents a marginal move toward joint taxation, which might be considered unrealistic. In section 3, we therefore also show results for married couples only and find that they are essentially equivalent to those for all couples.

2.2 Microsimulation and Labor Supply Model

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

We estimate a utility function for every household depending on consumption and leisure (of both partners in the case of couples). For 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 the estimated utility function, we can calculate probabilities for each hours category (combination) that a certain household chooses.

The discrete hours categories are 0, 10, 20, 30, 40, and 45 hours per week for women and men. Hence, a household in which both partners are flexible in their labor supply can choose from 25 combinations of hours per week. For every labor supply alternative, the EMSIM simulates the disposable income according to the tax-transfer system. The observed gross income and hours worked in the SOEP are used to construct hourly wage rates. Whenever the wage is unobserved because an individual is not working in the status quo, potential wage rates are imputed via a Heckman-type (Heckman, 1979) selectivity-corrected wage regression based on human capital-related variables.

Labor supply responses are modelled following the approach by Aaberge et al. (1995) and van Soest (1995). We suppress individual subscripts to ease notation. Each household values consumption and leisure as displayed in the utility function in equation (13), which depend on the selected amount of hours worked z.

$$V_z = U_z(Lf_z, Lm_z, C_z) + \varepsilon_z \tag{13}$$

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 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

$$U_z = \beta_1 ln(C_z) + \beta_2 ln(C_z)^2 + \beta_3 ln(Lf_z) + \beta_4 ln(Lm_z) + \beta_5 ln(Lf_z) ln(Lm_z) + I_z, \tag{14}$$

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 capture heterogeneity. This is crucial since, for example, typically female leisure time is valued higher in couples with children.

The error term ε in equation (13) is i.i.d. across the hours categories and households and assumed to follow an extreme value type I 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,

$$P_z = Pr(V_z > V_j, \forall j \neq z) = \frac{exp(U_z)}{\sum_{j=1}^{J} exp(U_j)}. \tag{15}$$

If a reform—in our case a small tax cut—changes the disposable income associated with certain choices of labor supply, it also changes the deterministic part of the utility function and hence the household's probability to choose that hours category. This, in turn, results in changes in the optimal amount of labor supplied and the behavioral responses that we are interested in in this study.

2.3 The Data

We use data from the German Socio-Economic Panel (SOEP), a representative annual survey of about 20,000 households. For more information see Goebel et al. (2018) and Schröder et al. (2020). The survey has been conducted since 1984. We use wave 38 and employ retrospective information for the year 2019 only. The data contain detailed information on employment (including hours worked per week and earnings), and characteristics of all household members, which allows us to predict hypothetical wages for the unemployed and model labor supply of different household types.

We restrict the sample to those households with at least one member with flexible labor supply[7]. 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 a sample of 16,698 households, which accounts for 36,255,535 German households with 66,309,145 persons (80 percent of the German population) using the SOEP's representative weighting factors. For the main results, households of each household type are split into deciles along the distribution of post-government income equivalized using the modified OECD equivalence scale. [8]

2.4 Implied Elasticities

Table 1 shows the own-wage labor supply elasticities for different household types resulting from a 1 percent increase in gross wages. The elasticities implied by the model are in line with the literature (see Keane, 2011, for an overview). In particular, labor supply elasticity of women in couples is substantially higher than that of men. For instance, a one-percent increase in female wages leads to a 0.36-percent increase in hours worked by women in couples where both partners can adjust their work hours, while their partners only increase their hours worked by 0.12 percent in response to a one-percent wage rise.

[7]Fixed labor supply is assumed for civil servants, students, members of the military, etc.

[8]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.

Table 1: Own-Wage Labor Supply Elasticities

	women	men
single women	0.23	.
single men	.	0.23
couples, both flexible	0.36	0.12
couples, woman flexible	0.34	.
couples, man flexible	.	0.11
all	0.30	0.14

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

Table 2 reports responses at the extensive margin, also simulated with a 1 percent increase in gross wages. Again, women tend to react stronger to wage changes than men. Overall, a one-percent increase in female wages leads to a 0.09-percentage-point increase in their participation rate, while the equivalent semi-elasticity of men is only 0.04.

Table 2: Own-Wage Participation Semi-Elasticities

	women	men
single women	0.11	.
single men	.	0.13
couples, both flexible	0.13	0.06
couples, woman flexible	0.14	.
couples, man flexible	.	0.06
all	0.09	0.04

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

3 Marginal Social Welfare Weights Implied by the Tax Schedule

Figure 1 reports the marginal social welfare weights—or one divided through the marginal value of public funds (MVPF) of small tax reductions—for the deciles of the distribution of net equivalent income implied by the 2019 German tax-transfer schedule. We show weights along the distribution of equivalized net incomes because this makes it possible to make comparisons across households with a different number of household members. As a starting point, we do not distinguish between household types. The resulting weights are essentially weighted means of the income-specific weights for different household types, which we show below. The welfare weights were calculated by simulating the labor supply responses to small tax cuts by 100 Euro per household in each decile of the income distribution. The implied welfare weights directly reflect the cost of the tax cut. At the optimum, the social planner is indifferent to a marginal tax change. Thus, at the optimum the social value of a tax reduction must equal its cost.

The figure implies that the current tax-transfer schedule is optimal if society values one additional Euro of consumption for the bottom decile almost three times as much as for the median. Throughout most of the distribution weights are decreasing with income, but we observe a small local minimum for the third decile. As we show below, this is driven by single households.

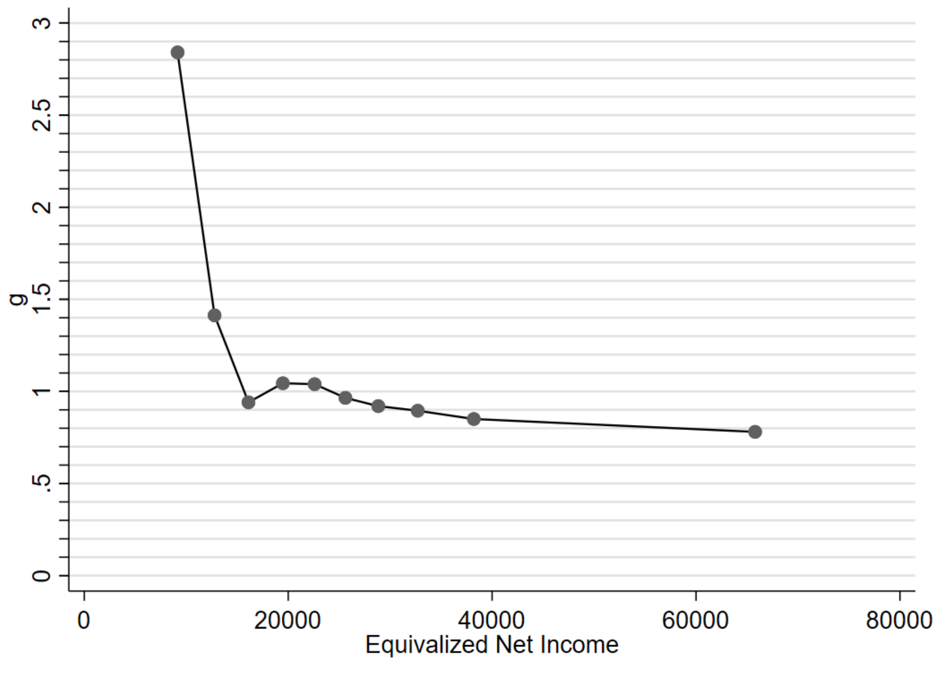

Figure 1: Marginal social welfare weights along the income distribution for all households

Note: The figure displays marginal social welfare weights along the distribution of net incomes, equivalized using the modified OECD scale. Each marker represents a decile of the income distribution. The position of the marker on the horizontal axis indicates average equivalized net income at the decile.

Figure 2 shows the inverse-optimum welfare weights for four different household types. We consider singles and couples with and without children. A first observation is that

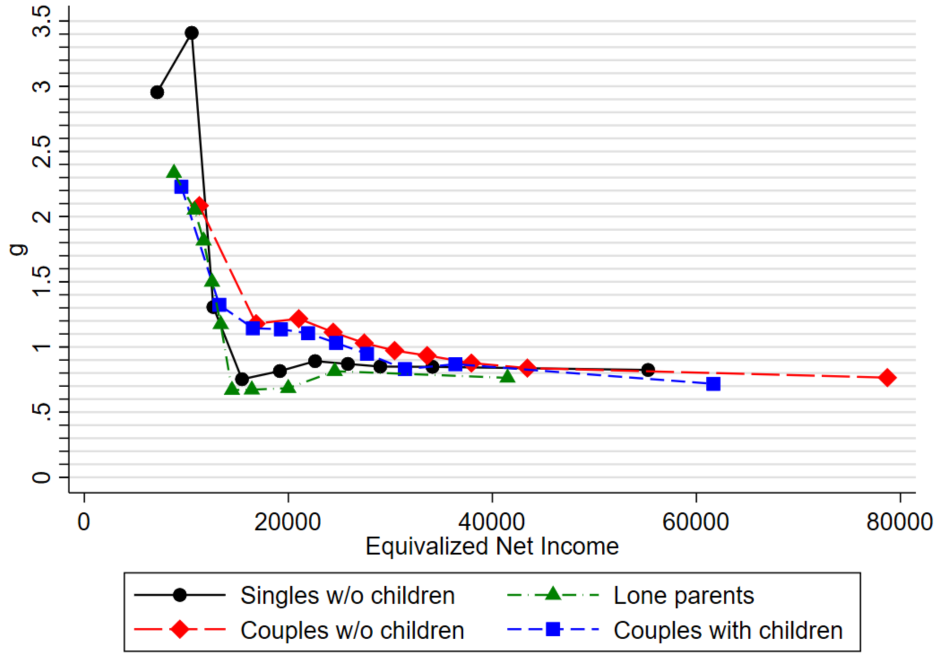

Figure 2: Marginal social welfare weights along the income distribution for various household types

Note: The figure displays marginal social welfare weights for different household types along the distribution of net incomes, equivalized using the modified OECD scale. As we are using equivalized incomes, the deciles are based on persons. Each marker represents a decile of the income distribution of each household type. The position of the marker on the horizontal axis indicates average equivalized net income at the decile.

the income distributions differ between the household types. For instance, average net equivalized earnings in the tenth decile of childless couples are around 80,000 Euro, while for couples with children and childless singles they are around 60,000 Euro. For lone parents, they are just slightly above 40,000 Euro.

For all household types, the cost of marginal tax reductions or, equivalently, marginal increases in transfers—and thus the implied marginal social welfare weight—for low income earners is higher than for the rest of the income spectrum. For instance, for singles without children, the computed welfare weight for the very bottom decile is 3, then even increases slightly for the 2nd decile and subsequently drops to 0.8 for the 4th decile. Hence, for childless singles, the current tax-transfer system is optimal if the social planner values one additional Euro for the households in the 1st decile almost four times as much as an additional Euro for the 4th decile. Strikingly, the welfare weights increase slightly from the 4th to the 6th decile before declining again for higher income earners. A similar pattern is observed for lone parents. A local minimum of the welfare function for low-medium incomes is in line with previous findings in the literature (e.g., Bargain et al., 2014; Blundell et al., 2009; Jacobs et al., 2017; Jessen et al., 2022).

For the other household types, the weight for very low incomes is lower than is the case for singles without children. The reason is that this group reacts more to increases in net income at the bottom. Note also that the second decile also includes individuals that do not work, but have relevant non-labor income. An advantage of our method is that we do not need to assume that all individuals who stop working end up in the lowest income group.

Comparing marginal social welfare weights between household types at various income levels, a striking finding is that couples have higher weights at low-medium income levels than singles. At an income of 20,000 Euro, the welfare weight is 1.1 for couples with children and 0.7 for lone parents. Thus, at this income level, society values one additional Euro for the former group about 60 percent more than one additional Euro for the latter group. The reason for the result is that the labor supply reactions of couples with higher income to increases in disposable income at low-medium income levels are larger than those of singles with higher income. Thus, reducing the tax burden (or increasing transfers) for singles at this income level is relatively inexpensive. Not doing so is only optimal if the welfare weight is quite low. The result for couples is obtained because our modeling approach allows for flexible adjustments of hours of work of secondary earners. If net incomes at, say, the 3rd decile are increased for childless couples, some couples might jump to that income group by having the secondary earner leave the labor force. This type of jump is not allowed for in comparable approaches.

In the appendix, we also show the equivalent of Figure 2, but along the distributions of gross household incomes of the four household types in Figure A.2. These household incomes are not equivalized and may include non-labor income. Naturally, the x-axis is less 'compressed' for the gross income distribution. The patterns of welfare weights are very similar to those observed in Figure 2. Importantly, at low-medium incomes we still observe substantially higher marginal social welfare weights for couples than for singles.[9]

[9] While the results are similar, the social marginal welfare weights are not exactly the same for the various deciles of each household type. This can only occur if the deciles contain different households. One reason for this is that, as is common, for the net equivalized income we construct deciles for persons, while for household gross income we construct deciles for households. A second reason is that in some cases the ranks in the household income distribution differ from those of the distribution of equivalized net incomes for a given household type. For households with different numbers of children, equalizing incomes can change ranks. A second reason is that in some cases the welfare state leads to a rank reversal, for instance because the magnitude of transfers depends on the place of residence or because incomes from different sources are taxed at different rates.

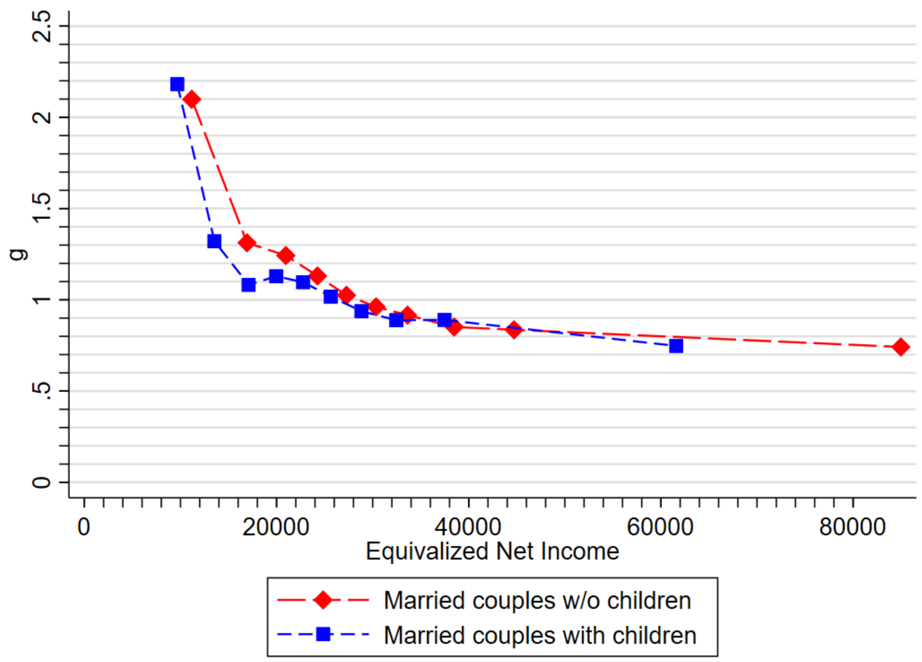

Figure 3: Marginal social welfare weights along the income distribution for married couples with and without children

Note: The figure displays marginal social welfare weights for different household types along the distribution of net incomes, equivalized using the modified OECD scale. As we are using equivalized incomes, the deciles are based on persons. Each marker represents a decile of the income distribution of each household type. The position of the marker on the horizontal axis indicates average equivalized net income for the household type at the decile.

In the following section, we show in detail, how different labor supply reactions of couples and singles to net income increases for the working poor result in the different patterns of welfare weights. Before, we present marginal social welfare weights implied by the German tax-transfer schedule separately for married and unmarried couples in Figure 3. As described in subsubsection 2.1.2, our modeling approach is arguably more suited for singles and married couples with joint taxation than for unmarried couples. It turns out that the welfare weights for married couples are essentially the same as those for all couples shown in Figure 2.

4 Decomposition of the Cost of Marginal Tax Reductions

In this section we demonstrate how labor supply reactions at different deciles explain our results for the inverse-optimum welfare weights. To this end, we decompose the cost per Euro distributed to households in a given decile into the mechanical cost and behavioral costs due to various labor supply reactions. The sum of mechanical cost and behavioral cost equals the inverse-optimum marginal social welfare weight or the inverse of the marginal value of public funds (MVPF). We focus on childless singles and couples with children to demonstrate the main difference in labor supply reactions. For completeness, we provide the equivalent graphs for the other household types in the appendix in Figures A.3, A.4, A.5, A.6, and A.7.

For childless singles, the bars in Figure 4 show for a tax reduction at every decile the labor supply reactions that lead to behavioral costs. The red bar is the mechanical cost per Euro redistributed to a given decile and is, by definition, one. The gray bars indicate the additional cost (or cost reduction for negative values) due to households adjusting their labor supply. In the case of no behavioral adjustment, all bars except for the red bar equal zero. The line shows the inverse-optimum marginal social welfare weight at each decile. The 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, hence the marker for the marginal social welfare weight lies above the red bar. Households at that income level receive the highest net transfers. Labor supply reactions to an increase in disposable income for the lowest decile, i.e. households "moving" into that decile, therefore unambiguously increase the cost of additional redistribution. The height of the bars representing all other income deciles depicts the size of the fiscal externality caused by the behavioral response of households moving from that pre-reform decile into the targeted decile. It is calculated as the product of the number of households that reduce labor supply times the difference in net taxes paid (or transfers received) by these households in their pre-reform decile relative to their post-reform decile divided through the mechanical cost of the tax reduction. In this case, for a small tax cut targeted at the first decile, the fiscal externality is the largest in deciles three and four, although even the ninth decile contributes slightly to the fiscal externality. These labor supply reactions to income increases at the very bottom of the distribution indicate that it is important to capture the extensive labor supply margin.

For decile 2, the inverse-optimum weight is slightly higher than for decile 1. The reason is that an increase in disposable income at the income level of decile 2 attracts more households from higher up in the distribution than is the case for decile 1. One reason might be that some households with substantial non-labor market income do not move to decile 1 even if they stop working. Instead, dropping out of the work force would bring them to decile 2.

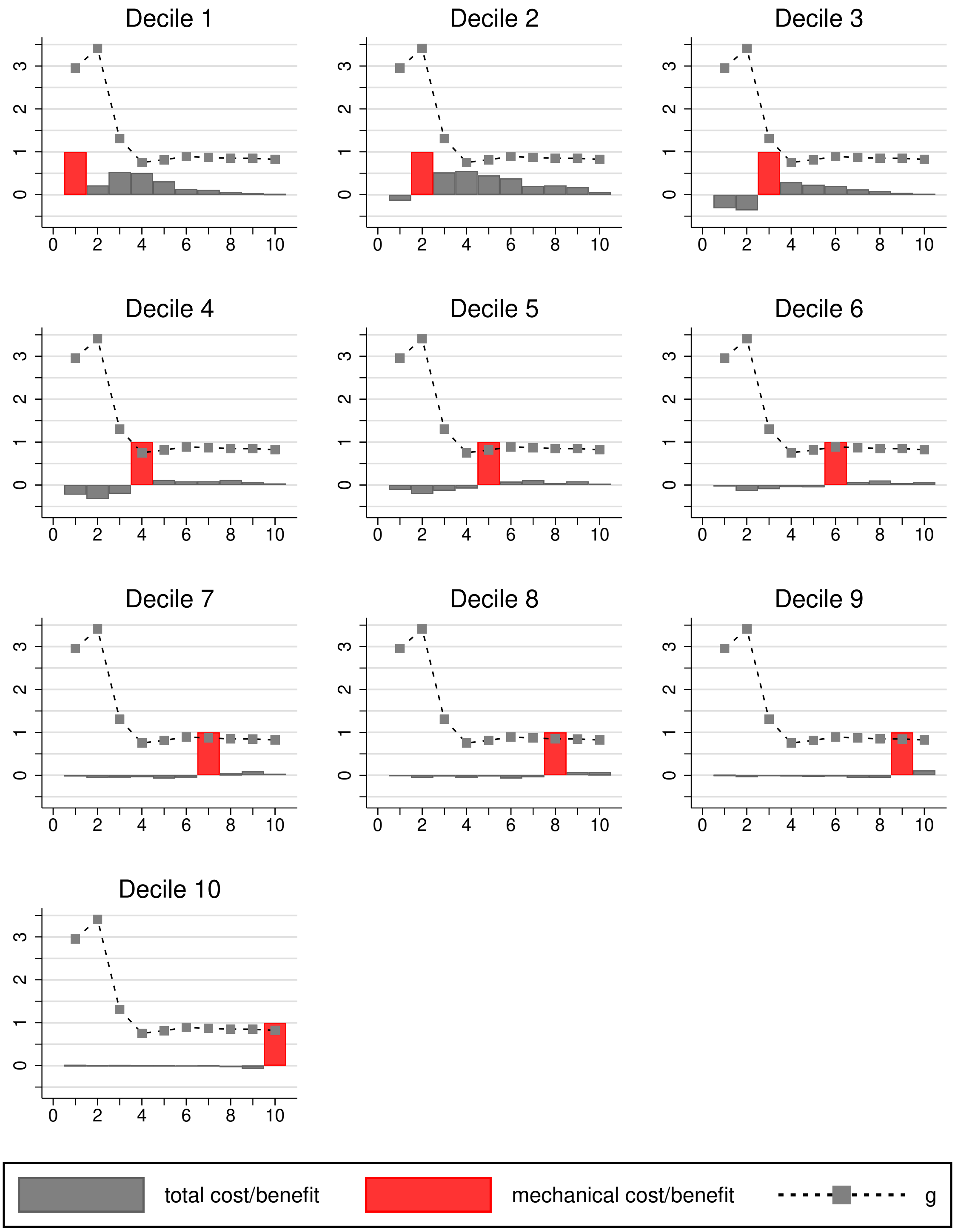

Figure 4: Decomposition of Marginal Cost of Redistribution: Childless singles

Note: The red bars show the mechanical cost of distributing one Euro to households in the relevant decile. The gray bars show the fiscal costs due to labor supply changes from each decile to one specific decile, where a marginal tax reduction occurs, divided through the mechanical cost of the tax reduction. In other words, it shows the fiscal cost per Euro redistributed to individuals in the respective decile. In the case of no behavioral adjustment, all bars except for the red bar equal zero. The line shows the inverse-optimum marginal social welfare weight at each decile. It equals the inverse of the marginal value of public funds (MVPF). For each decile, the inverse-optimum marginal social welfare weight equals the sum of the bars.

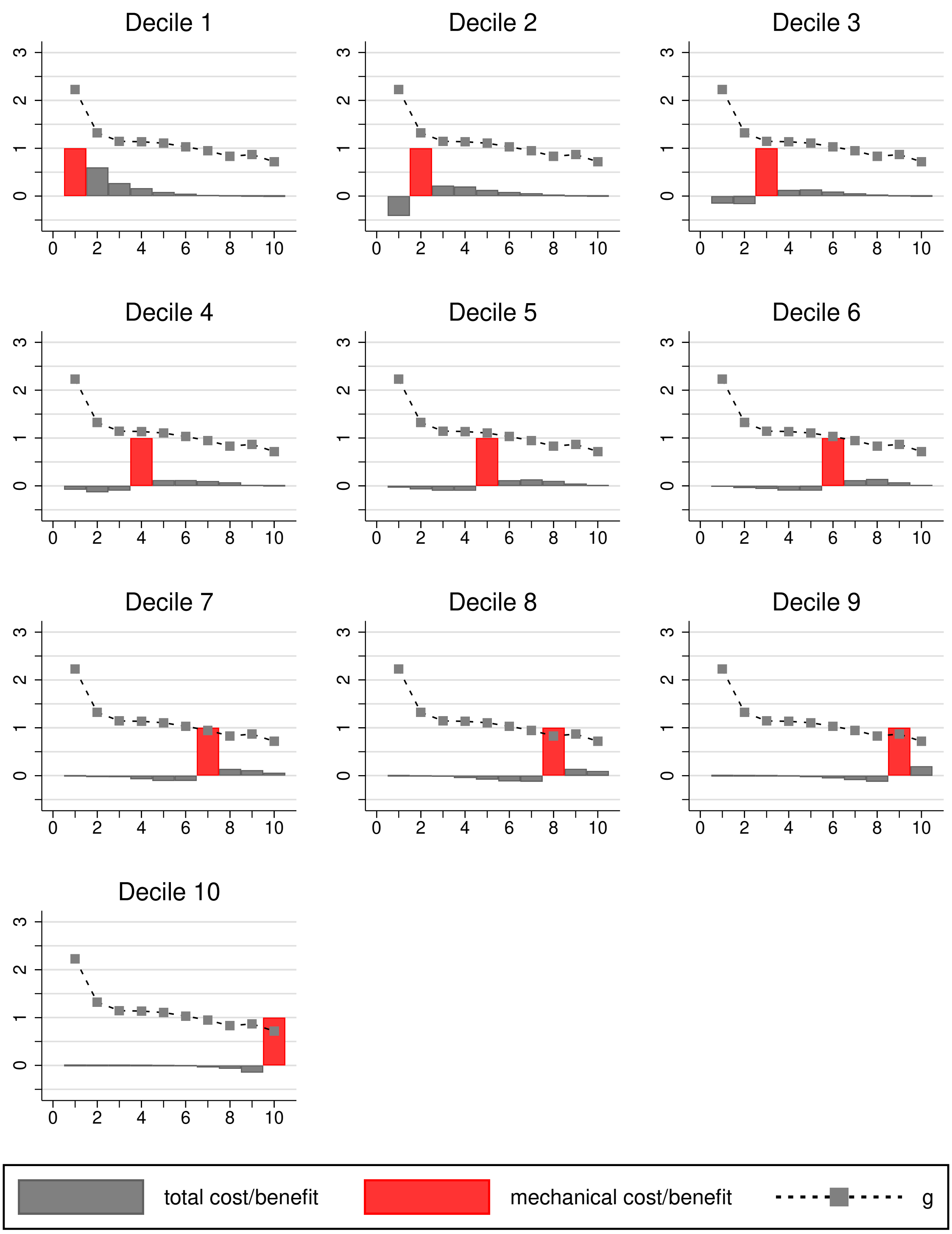

Figure 5: Decomposition of Marginal Cost of Redistribution: Couples with Children

Note: The red bars show the mechanical cost of distributing one Euro to households in the relevant decile. The gray bars show the fiscal costs due to labor supply changes from each decile to one specific decile, where a marginal tax reduction occurs, divided through the mechanical cost of the tax reduction. In other words, it shows the fiscal cost per Euro redistributed to individuals in the respective decile. In the case of no behavioral adjustment, all bars except for the red bar equal zero. The line shows the inverse-optimum marginal social welfare weight at each decile. It equals the inverse of the marginal value of public funds (MVPF). For each decile, the inverse-optimum marginal social welfare weight equals the sum of the bars.

Turning to decile 4, this 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. This positive fiscal externality outweighs the negative fiscal effects of labor supply reductions from deciles 5-10.

For small tax reductions in deciles 2-6, labor supply reactions from all deciles play a relevant role. These labor supply reactions are not explicitly accounted for in optimal tax models like Saez (2001) and Saez (2002). As we move to tax reductions for high income earners, only labor supply reactions by workers higher up in the income distribution matter.

Figure 5 shows the equivalent decomposition of the cost of marginal redistribution and hence the inverse-optimum social welfare weight for couples with children. Consider decile 5. Here the average income is just slightly higher than at decile 5 of childless singles, about 20,000 Euro. The marginal welfare weight is above one for couples with children, while it is below one for childless singles. The reason is that the bars at the non-directly affected deciles 6-8 are larger for couples, implying larger labor supply reductions. These reactions, which are mostly due to secondary earners, explain why the fiscal cost of a tax reduction at that decile is larger than it is for singles, which, in turn, implies a higher marginal social welfare weight for couples at that income level.

5 Comparison to the Saez (2002) Approach

For a comparison of the calculated implied marginal social welfare weights based on our approach and in the model proposed by Saez (2002), we apply both methods on the same sample. In contrast to our simulation approach, the Saez (2002) approach calls for relatively homogeneous households. For this reason, it is typically necessary to exclude many 'non-typical' households from the analysis. For instance, Bargain et al. (2014) exclude households where capital incomes represent more than 25 percent of market income. Moreover, the lowest group, which is relevant for the participation decision, is typically defined as non-working and net incomes in this group must be lower than in the next-lowest group. In contrast, our simulation approach allows for non-working households to be in higher income groups. For comparability, we restrict the sample, making sure that gross incomes for non-working households do not exceed 1,000 Euro per year. Moreover, to avoid edge cases of high-income households in the out-of-work group we exclude households with incomes from self-employment, from capital, or from rent of more than 100 Euro per year as well as those couples with one working partner with fixed labor supply.

In contrast to the previous analysis, we divide the sample into a group 0 of households who work zero hours and nine quantiles along the household gross income distribution of households with positive labor earnings.[10] As is common in papers applying the Saez (2002) model, we show results along the gross income distribution.

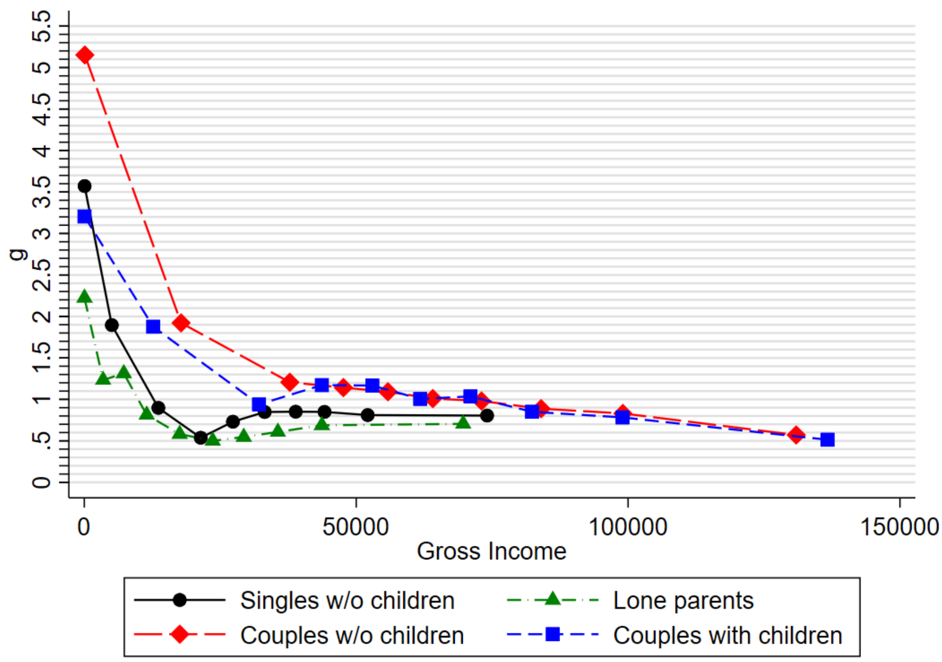

Figure 6: Simulated marginal social welfare weights along the income distribution for various household types using Saez-style groups

Note: The figure displays marginal social welfare weights for different household types along the distribution of gross incomes. The deciles are based on households. Each marker represents a group of the income distribution of each household type. The position of the marker on the horizontal axis indicates average gross income within the group.

We first report the inverse-optimum welfare weights following our simulation approach, but for the restricted sample, in Figure 6. We display these weights the distribution of gross incomes. Compared to the weights simulated in section 3 and Figure A.2, couples have drastically higher marginal social welfare weights at lower incomes. The reason is the different sample restriction. While in the main analysis, some couple households could not "move" to the lowest income group because either labor supply of one partner is assumed fixed or because they have substantial non-labor market income, this is not the case in

[10]Households in group 0 may still have other sources of income, hence average household gross income in this group can be larger than zero.

the restricted sample. Therefore, labor supply reactions to an increase in net income at the bottom are stronger, and hence the inverse-optimum wage is higher.

The local minimum of welfare weights for the working poor is still more prevalent for single households than for couples and except for group 0, the key result that welfare weights at any given gross income are lower for singles than for couples still holds.

In contrast, Figure 7 documents the calculated marginal social welfare weights using the Saez (2002) model on our data. In particular, we use the labor supply model combined with the microsimulation model to simulate labor supply reactions to increases in income groups $i = 1, ..., I$, but then disregard labor supply reactions except for that from 'neighboring' groups and from group 0. We also assume no income effects, implying that an increase in net income in group i by a small amount leads to the same movement from group $i-1$ to i as a small decrease in net income in group $i-1$. We then calculate the implied elasticities and plug them into equations (10) and (11) to calculate welfare weights. Blundell et al. (2009) and Bargain et al. (2014) take the same approach. The components used for the calculation of the social welfare weights following Saez (2002), i.e., shares, gross and net incomes as well as extensive and intensive elasticities of the various income groups, are documented in detail for the four household types in tables A.1 to A.4 in the appendix. In addition, the tables include a column showing the respective welfare weights we attained through microsimulation based on the same sample.

Here, couples' welfare weights are lower than singles' at gross incomes between 25,000 and 50,000 Euro. In group 0, the weights are slightly lower than in the microsimulation approach across all household types. Out of all income groups with non-zero hours worked, the second highest income quantile displays the highest implied social welfare weight according to this model, with couples' welfare weight clearly spiking in that income group. Interestingly, for singles without children and couples, we observe clear minima for the third-lowest income group, the first group with a positive net tax (see Tables A.1 to A.4 in the appendix). This observation is in line with results for childless singles in various countries Bargain et al. (2014). For lone parents, all but the four groups with the highest income are net receivers and the lowest weight is observed for the group with the fifth-lowest income, the highest-income group with a very high implicit backward-looking marginal tax rate ($(T_1 - T_{i-1})/(Y_i - Y_{i-1})$, see Blundell et al., 2009) of of 80 percent. The intuition is simple: An increase in net income for a given group leads to households from other groups moving to this group. When the marginal tax rate is high, the positive revenue effect of households moving to this group from lower groups is substantial. When the backward-looking marginal tax rate of the next higher group is comparatively low, the negative fiscal effect of households moving down is lower. Therefore, a tax cut for this group pays to a large degree for itself. Not reducing taxes/increasing transfers for this group is only optimal if its marginal social welfare weight is quite low.

Importantly, the Saez (2002) approach does not result in the relatively high welfare weights for low-medium income couples. The reason is that some labor supply reactions to increases in net incomes for this income group, which we showed to be important in the previous section, are neglected. More generally, average welfare weights cannot differ between household types in a model without income effects.

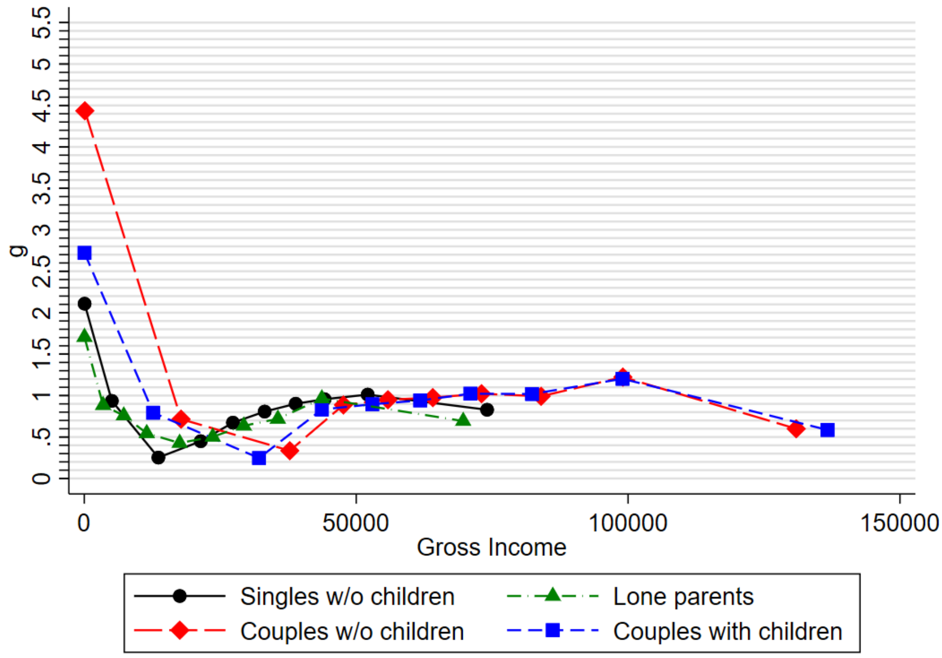

Figure 7: Saez (2002)-style marginal social welfare weights along the income distribution for various household types using Saez-style groups

Note: The figure displays marginal social welfare weights for different household types along the distribution of gross incomes. The deciles are based on households. Each marker represents a group of the income distribution of each household type. The position of the marker on the horizontal axis indicates average gross income within the group.

The comparison shows that, while both approaches yield higher inverse-optimum weights at the very bottom and almost flat weights for higher income, loosening the restrictions of the Saez (2002) model allows for a more meaningful comparison between household types.

6 Conclusions

We have simulated labor supply reactions to small tax cuts for various household types throughout the German income distribution and calculated their marginal value of public funds. This allowed us to calculate the marginal social welfare weights implied by the German tax-transfer schedule. Our approach has two key advantages. First, it is not restricted to homogeneous single households and allows for the comparison of social welfare weights across different household types. Second, it does not impose strong restrictions on labor supply behavior and allows for flexible labor supply adjustments of secondary earners and nonlinear and non-convex budget sets.

We have found that the current tax-transfer system is optimal if the social planner values one Euro for households at the 1st decile about three times as much as one Euro for households at the median of the distribution. At low-medium incomes, implied weights for couples are substantially higher than for singles. The reason is that tax reductions for couples at this income level lead to substantial labor supply reductions. This result highlights the importance of explicitly modeling labor supply of couples. For comparison, we calculated inverse-optimum weights implied by the optimal tax approach by Saez (2002), which puts more restrictions on labor supply behavior. We find that this approach does not recover the substantial difference in weights between different household types with low-medium income.

Our result that, at low-to-medium incomes, the marginal social welfare weight of couples is lower than that of singles has various possible interpretations. First, since the result was obtained under joint taxation for couples, a potential policy implication is that the relatively high efficiency loss of tax reductions for low-medium income couples calls for a different type of taxation of married couples. Second, society might actually value one additional Euro for couples more than one additional Euro for singles, perhaps because one Euro for couples offers more "bang for the buck" due to the public good character of some of the goods consumed within the household. Third, our model might be too simple to capture all relevant behavioral reactions to changes in the tax schedule. For instance, Golosov and Krasikov (2023) show in a model with endogenous household formation that it is optimal for an inequality-averse social planner to "subsidize" couples since couples provide socially valuable within-household redistribution. Neglecting this margin might bias the implied marginal social welfare weights of individuals in couple households. In principle, our approach of calculating inverse optimum weights can accommodate more complex models, which could be a fruitful avenue of future research.

References

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

Atkinson, A. B. and J. E. Stiglitz (2015): *Lectures on Public Economics: Updated Edition*, Princeton University Press.

Ayaz, M., L. Fricke, C. Fuest, and D. Sachs (2023): "Who Should Bear the Burden of COVID-19 Related Fiscal Pressure? An Optimal Income Taxation Perspective," *European Economic Review*, 104381.

Bach, S., G. Corneo, and V. Steiner (2012): "Optimal Top Marginal Tax Rates under Income Splitting for Couples," *European Economic Review*, 56, 1055–1069.

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

Bastani, S. (forthcoming): "The Marginal Value of Public Funds: a Brief Guide and Application to Tax Policy," *International Tax and Public Finance*.

Bastani, S. and J. Lundberg (2017): "Political Preferences for Redistribution in Sweden," *The Journal of Economic Inequality*, 15, 345–367.

Bechara, P., T. Kasten, and S. Schaffner (2015): "Dokumentation des RWI-Einkommensteuer-Mikrosimulationsmodells (EMSIM)," RWI Materialien 86, RWI—Leibniz-Institut für Wirtschaftsforschung.

Bierbrauer, F. J., P. C. Boyer, and E. Hansen (2023): "Pareto-Improving Tax Reforms and the Earned Income Tax Credit," *Econometrica*, 91, 1077–1103.

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

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

Burtless, G. and J. A. Hausman (1978): “The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment,” *Journal of political Economy*, 86, 1103–1130.

Chetty, R. (2008): “Moral Hazard versus Liquidity and Optimal Unemployment Insurance,” *Journal of Political Economy*, 116, 173–234.

Finkelstein, A. and N. Hendren (2020): “Welfare Analysis Meets Causal Inference,” *Journal of Economic Perspectives*, 34, 146–67.

Fullerton, D. and L. Gan (2004): “A Simulation-Based Welfare Loss Calculation for Labor Taxes with Piecewise-Linear Budgets,” *Journal of Public Economics*, 88, 2339–2359.

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

Golosov, M. and I. Krasikov (2023): “The Optimal Taxation of Couples,” Tech. rep., National Bureau of Economic Research.

Haan, P. and D. Navarro (2008): “Optimal Income Taxation of Married Couples: An Empirical Analysis for Germany,” Tech. rep., IZA working paper.

Heckman, J. J. (1979): “Sample Selection Bias as a Specification Error,” *Econometrica*, 153–161.

Hendren, N. (2020): “Measuring Economic Efficiency Using Inverse-Optimum Weights,” *Journal of Public Economics*, 187.

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

Jacobs, B., E. L. Jongen, and F. T. Zoutman (2017): “Revealed Social Preferences of Dutch Political Parties,” *Journal of Public Economics*, 156, 81–100.

Jessen, R., M. Metzing, and D. Rostam-Afschar (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, 961–1075.

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

Lazear, E. P. and R. T. Michael (1980): "Real Income Equivalence Among One-Earner and Two-Earner Families," *The American Economic Review*, 70, 203–208.

Lockwood, B. B. and M. Weinzierl (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," .

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

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

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

Schmieder, J. F. and T. Von Wachter (2016): "The Effects of Unemployment Insurance Benefits: New Evidence and Interpretation," *Annual Review of Economics*, 8, 547–581.

Schmieder, J. F., T. Von Wachter, and S. Bender (2012): "The Effects of Extended Unemployment Insurance over the Business Cycle: Evidence from Regression Discontinuity Estimates over 20 Years," *The Quarterly Journal of Economics*, 127, 701–752.

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

van Soest, A. (1995): "Structural Models of Family Labor Supply: A Discrete Choice Approach," *Journal of Human Resources*, 30, 63–88.

Appendix

A Additional Figures

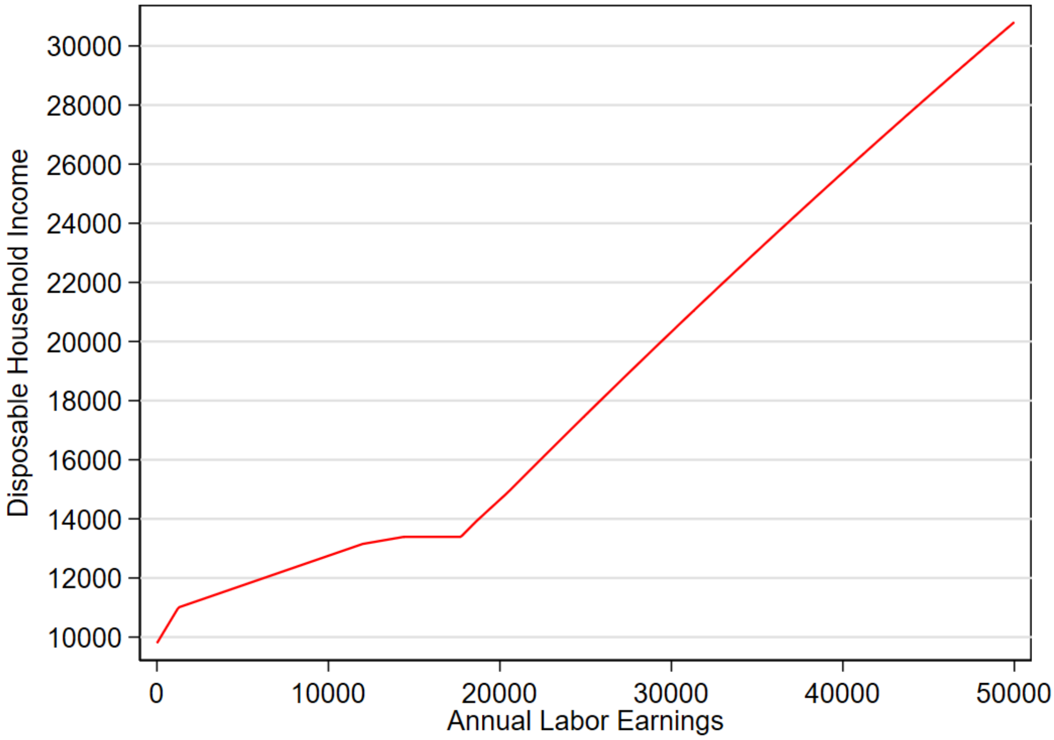

Figure A.1: Budget constraint of a childless single

Note: Disposable incomes for various gross income levels; no market income except for labor earnings. The graph is constructed using the microsimulation model EMSIM.

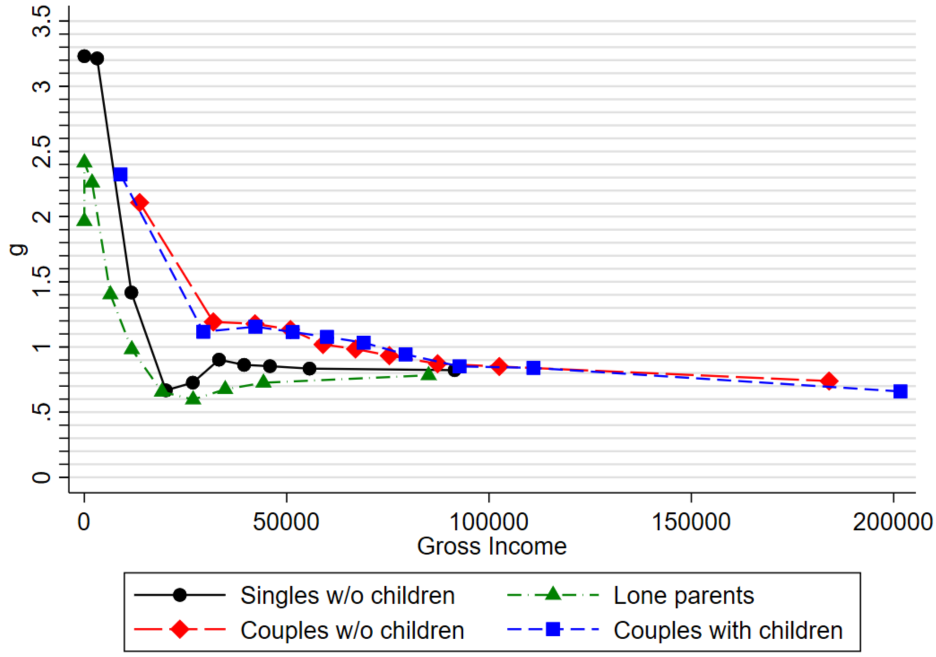

Figure A.2: Marginal social welfare weights along the distribution of gross household incomes for various household types

Note: The figure displays marginal social welfare weights for different household types along the distribution of gross incomes. Each marker represents a decile of the income distribution of each household type. The position of the marker on the horizontal axis indicates average gross income at the decile.

Table A.1: Singles without children

Group	Gross Income	Net Income	Net Tax	Intensive Elasticity	Extensive Elasticity	Share	Saez (2002) Weights	Simulated Weights
0	108	8294	-8186	.	.	0.18	2.11	3.57
1	5197	10771	-5573	0.20	0.20	0.09	0.94	1.90
2	14141	13303	837	0.14	0.33	0.09	0.25	0.90
3	21627	16302	5325	0.14	0.26	0.09	0.45	0.54
4	27429	19107	8321	0.08	0.20	0.09	0.67	0.73
5	33217	22294	10923	0.07	0.14	0.09	0.81	0.85
6	38841	25232	13608	0.06	0.08	0.09	0.90	0.85
7	44082	27999	16083	0.06	0.06	0.09	0.96	0.85
8	52093	32000	20093	0.08	0.04	0.09	1.01	0.81
9	73132	42719	30412	0.15	0.02	0.09	0.83	0.80

Note: Groups are defined as one group with no labor income (group 0) and nine equal-sized groups along the gross household income distribution. All monetary values are expressed in Euro. All measures of income and taxes are the group-specific means. Saez weights refer to the calculation method expressed in equations (10) and (11), simulated weights refer to the microsimulation approach.

Table A.2: Lone parents

Group	Gross Income	Net Income	Net Tax	Intensive Elasticity	Extensive Elasticity	Share	Saez (2002) Weights	Simulated Weights
0	49	16620	-16571	.	.	0.31	1.71	2.22
1	3407	18339	-14931	0.09	0.09	0.08	0.88	1.23
2	7153	19878	-12724	0.04	0.23	0.08	0.76	1.31
3	11647	20608	-8961	0.01	0.24	0.08	0.55	0.81
4	17639	21794	-4154	0.02	0.23	0.08	0.43	0.58
5	23887	24589	-701	0.06	0.24	0.08	0.50	0.50
6	29888	27264	2624	0.04	0.20	0.08	0.64	0.55
7	35478	30118	5360	0.05	0.17	0.08	0.72	0.61
8	44396	35814	8581	0.07	0.14	0.07	0.96	0.69
9	68529	46817	21711	0.15	0.10	0.08	0.69	0.71

Note: Groups are defined as one group with no labor income (group 0) and nine equal-sized groups along the gross household income distribution. All monetary values are expressed in Euro. All measures of income and taxes are the group-specific means. Saez weights refer to the calculation method expressed in equations (10) and (11), simulated weights refer to the microsimulation approach.

Table A.3: Couples without children

Group	Gross Income	Net Income	Net Tax	Intensive Elasticity	Extensive Elasticity	Share	Saez (2002) Weights	Simulated Weights
0	75	13849	-13773	.	.	0.04	4.44	5.15
1	18075	19599	-1523	0.22	0.22	0.11	0.72	1.92
2	38490	28232	10257	0.47	0.14	0.11	0.33	1.21
3	47922	34262	13660	0.38	0.05	0.11	0.89	1.14
4	56164	38792	17371	0.21	0.03	0.11	0.95	1.09
5	64308	43367	20941	0.20	0.01	0.11	0.98	1.01
6	73251	48278	24972	0.18	0.01	0.11	1.02	0.98
7	83983	53925	30057	0.20	0.00	0.11	0.99	0.89
8	98949	62910	36038	0.26	0.00	0.11	1.22	0.83
9	132549	79394	53155	0.39	0.00	0.11	0.60	0.57

Note: Groups are defined as one group with no labor income (group 0) and nine equal-sized groups along the gross household income distribution. All monetary values are expressed in Euro. All measures of income and taxes are the group-specific means. Saez weights refer to the calculation method expressed in equations (10) and (11), simulated weights refer to the microsimulation approach.

Table A.4: Couples with children

Group	Gross Income	Net Income	Net Tax	Intensive Elasticity	Extensive Elasticity	Share	Saez (2002) Weights	Simulated Weights
0	30	21017	-20986	.	.	0.09	2.72	3.20
1	12978	24394	-11416	0.15	0.15	0.10	0.79	1.88
2	33385	30889	2495	0.30	0.14	0.10	0.25	0.94
3	43710	36456	7254	0.25	0.07	0.10	0.83	1.17
4	53060	41595	11465	0.22	0.05	0.10	0.90	1.17
5	61905	46722	15183	0.20	0.03	0.10	0.94	1.01
6	71051	51871	19180	0.17	0.02	0.10	1.02	1.04
7	82297	57883	24414	0.20	0.01	0.10	1.02	0.85
8	99080	67357	31722	0.27	0.00	0.10	1.20	0.78
9	138176	88069	50107	0.47	0.00	0.10	0.59	0.51

Note: Groups are defined as one group with no labor income (group 0) and nine equal-sized groups along the gross household income distribution. All monetary values are expressed in Euro. All measures of income and taxes are the group-specific means. Saez weights refer to the calculation method expressed in equations (10) and (11), simulated weights refer to the microsimulation approach.

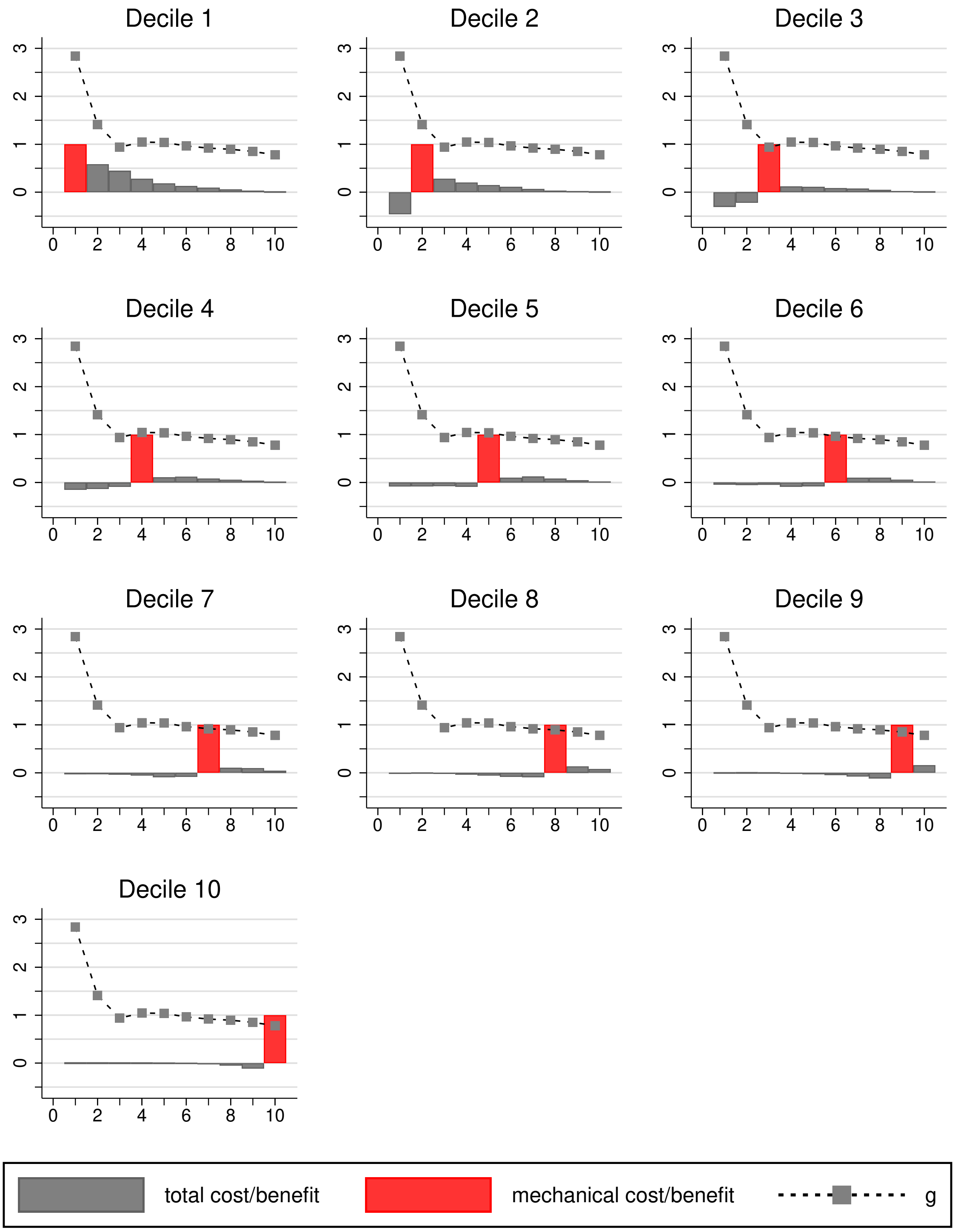

Figure A.3: Decomposition of Marginal Cost of Redistribution: Full sample

Note: The red bars show the mechanical cost of distributing one Euro to households in the relevant decile. The gray bars show the fiscal costs due to labor supply changes from each decile to one specific decile, where a marginal tax reduction occurs, divided through the mechanical cost of the tax reduction. In other words, it shows the fiscal cost per Euro redistributed to individuals in the respective decile. In the case of no behavioral adjustment, all bars except for the red bar equal zero. The line shows the inverse-optimum marginal social welfare weight at each decile. It equals the inverse of the marginal value of public funds (MVPF). For each decile, the inverse-optimum marginal social welfare weight equals the sum of the bars.

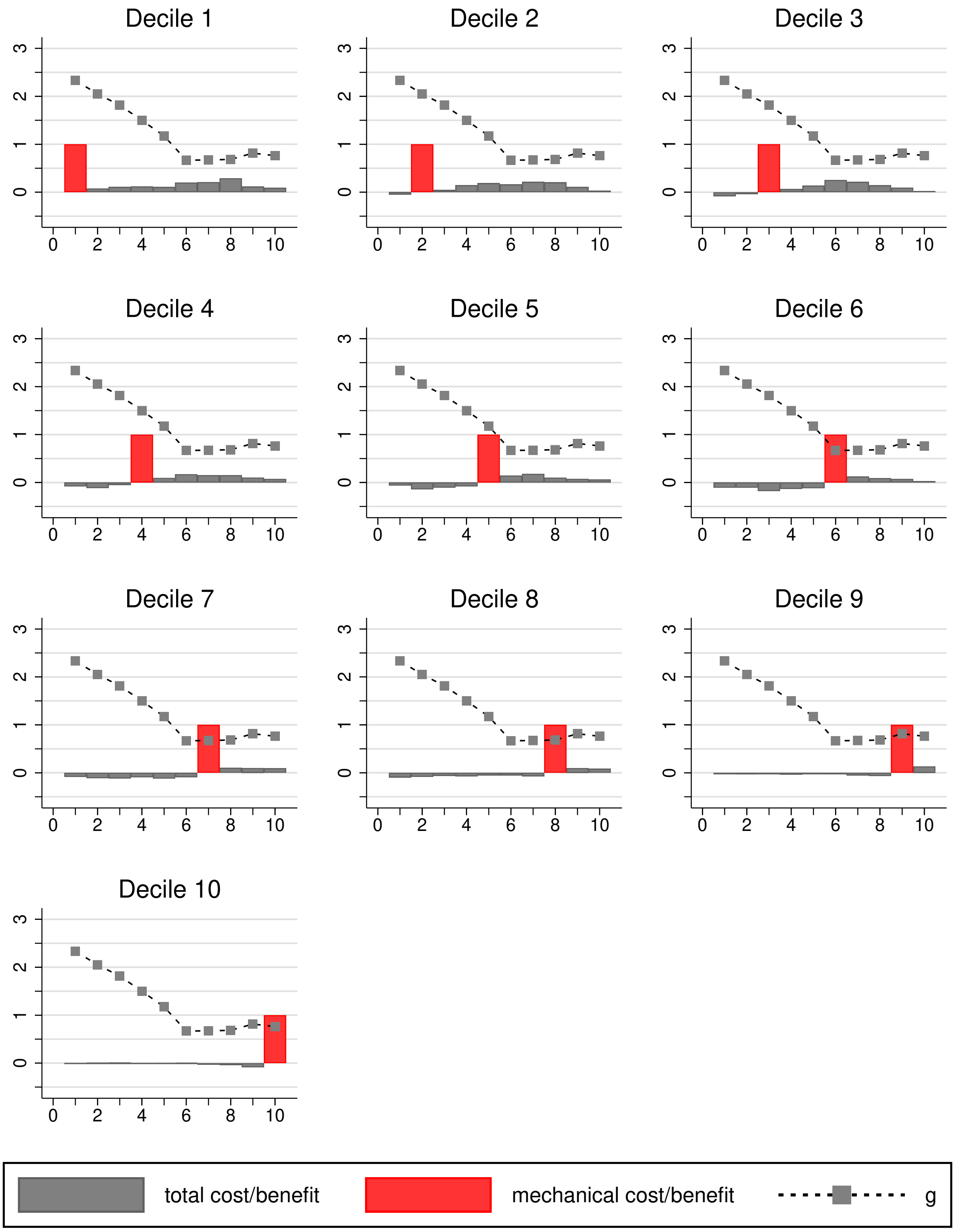

Figure A.4: Decomposition of Marginal Cost of Redistribution: Lone parents

Note: The red bars show the mechanical cost of distributing one Euro to households in the relevant decile. The gray bars show the fiscal costs due to labor supply changes from each decile to one specific decile, where a marginal tax reduction occurs, divided through the mechanical cost of the tax reduction. In other words, it shows the fiscal cost per Euro redistributed to individuals in the respective decile. In the case of no behavioral adjustment, all bars except for the red bar equal zero. The line shows the inverse-optimum marginal social welfare weight at each decile. It equals the inverse of the marginal value of public funds (MVPF). For each decile, the inverse-optimum marginal social welfare weight equals the sum of the bars.

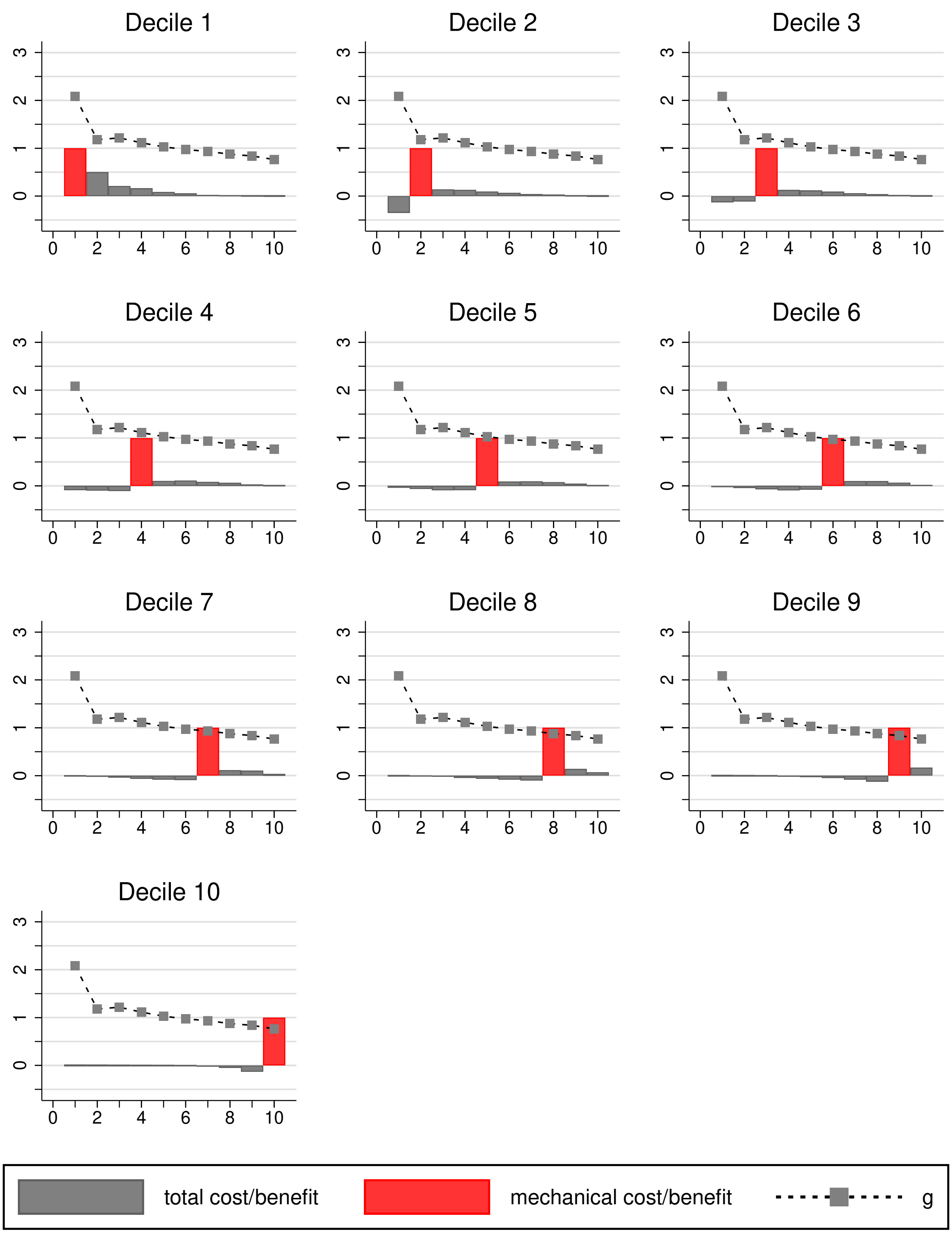

Figure A.5: Decomposition of Marginal Cost of Redistribution: Couples without Children

Note: The red bars show the mechanical cost of distributing one Euro to households in the relevant decile. The gray bars show the fiscal costs due to labor supply changes from each decile to one specific decile, where a marginal tax reduction occurs, divided through the mechanical cost of the tax reduction. In other words, it shows the fiscal cost per Euro redistributed to individuals in the respective decile. In the case of no behavioral adjustment, all bars except for the red bar equal zero. The line shows the inverse-optimum marginal social welfare weight at each decile. It equals the inverse of the marginal value of public funds (MVPF). For each decile, the inverse-optimum marginal social welfare weight equals the sum of the bars.

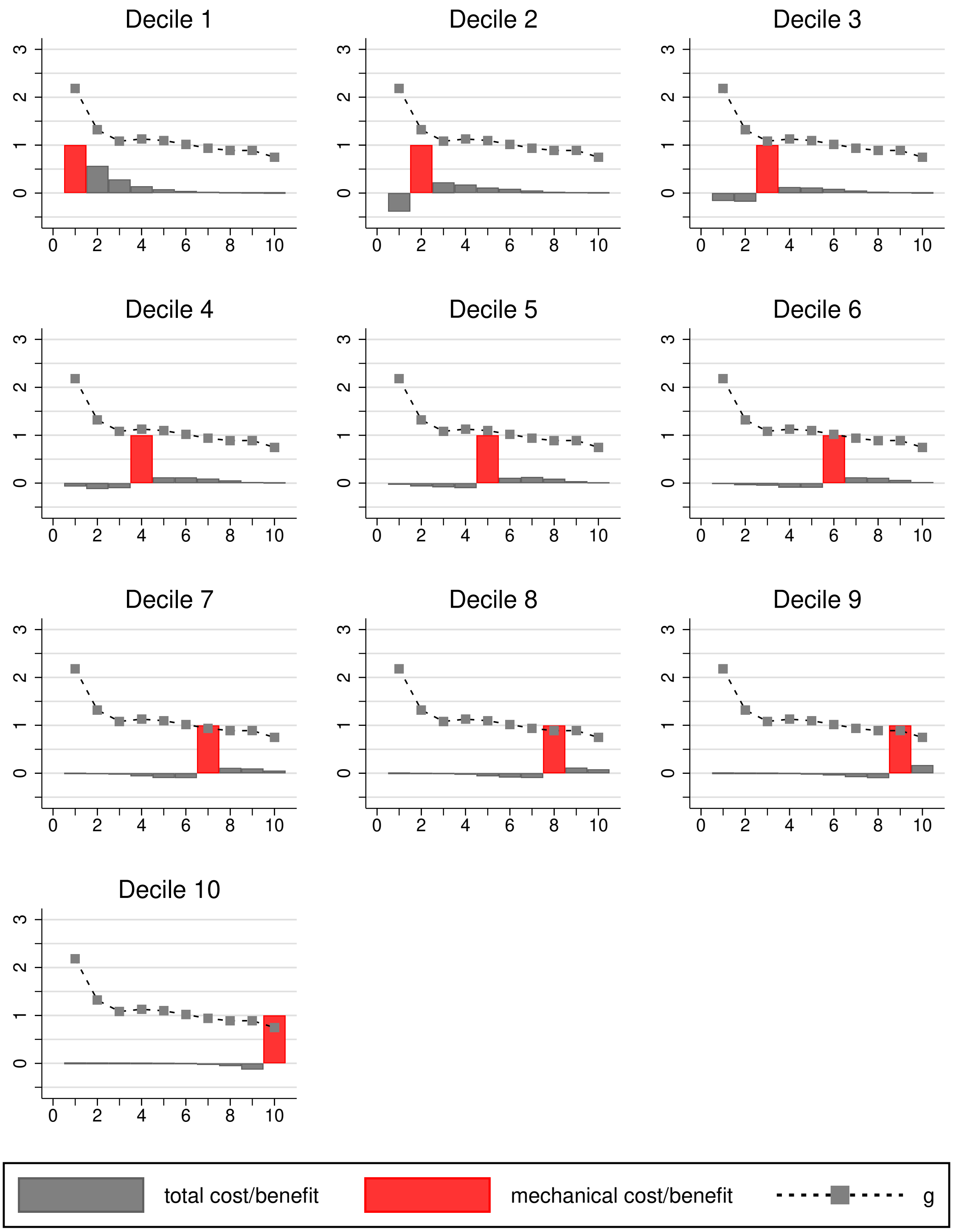

Figure A.6: Decomposition of Marginal Cost of Redistribution: Married couples with Children

Note: The red bars show the mechanical cost of distributing one Euro to households in the relevant decile. The gray bars show the fiscal costs due to labor supply changes from each decile to one specific decile, where a marginal tax reduction occurs, divided through the mechanical cost of the tax reduction. In other words, it shows the fiscal cost per Euro redistributed to individuals in the respective decile. In the case of no behavioral adjustment, all bars except for the red bar equal zero. The line shows the inverse-optimum marginal social welfare weight at each decile. It equals the inverse of the marginal value of public funds (MVPF). For each decile, the inverse-optimum marginal social welfare weight equals the sum of the bars.

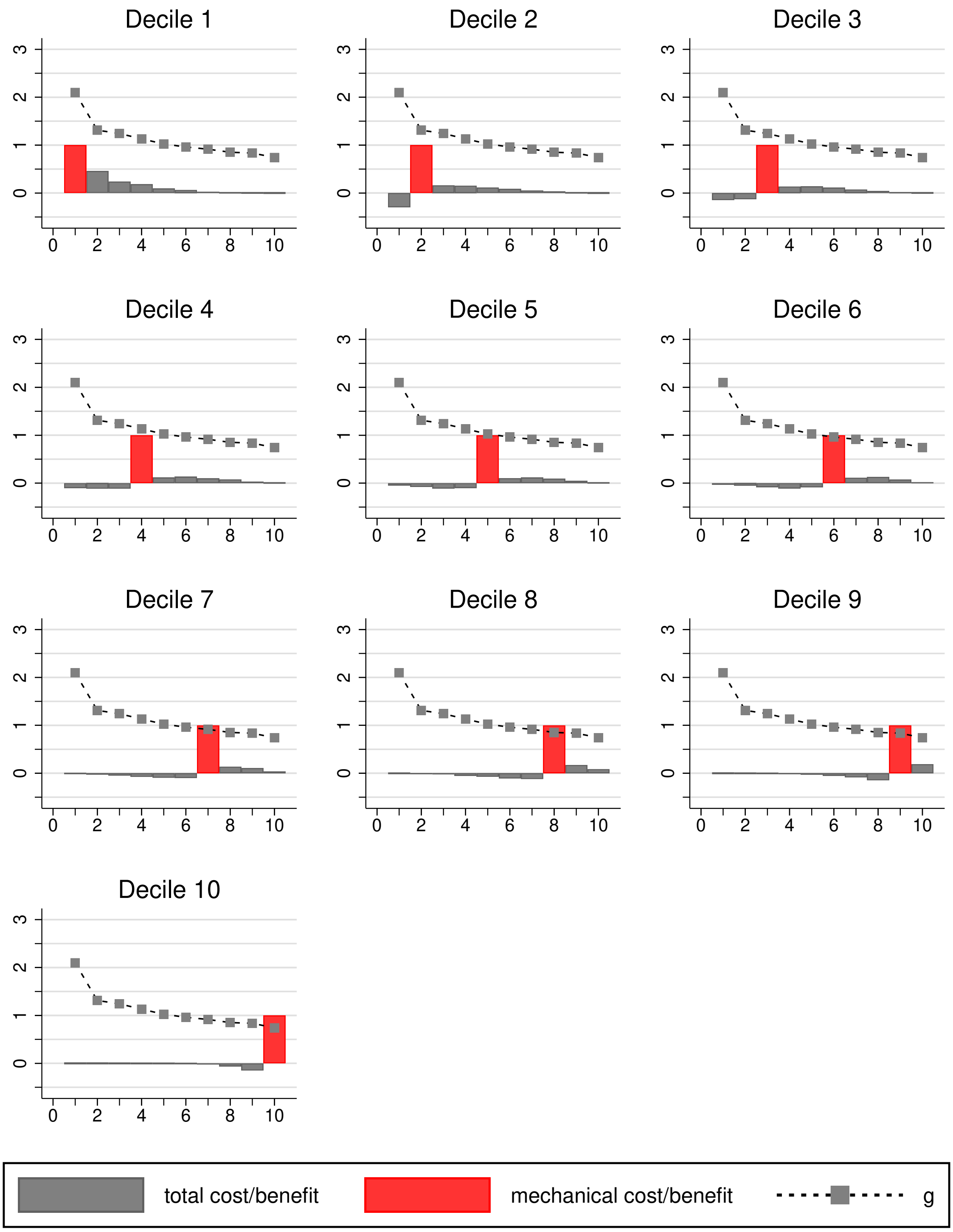

Figure A.7: Decomposition of Marginal Cost of Redistribution: Married couples without Children

Note: The red bars show the mechanical cost of distributing one Euro to households in the relevant decile. The gray bars show the fiscal costs due to labor supply changes from each decile to one specific decile, where a marginal tax reduction occurs, divided through the mechanical cost of the tax reduction. In other words, it shows the fiscal cost per Euro redistributed to individuals in the respective decile. In the case of no behavioral adjustment, all bars except for the red bar equal zero. The line shows the inverse-optimum marginal social welfare weight at each decile. It equals the inverse of the marginal value of public funds (MVPF). For each decile, the inverse-optimum marginal social welfare weight equals the sum of the bars.