Revealed Social Preferences of Dutch Political Parties*

Floris T. Zoutman† Bas Jacobs‡ Egbert L.W. Jongen§

December 31, 2014

Abstract

In a process unique in the world, all major Dutch political parties provide CPB Netherlands Bureau for Economic Policy Analysis with detailed proposals for reforming the tax-benefit system in every national election. This information allows us to uncover the social preferences for income redistribution of each political party by using the inverse optimal-tax method to calculate social welfare weights for each income level. We contribute and amend existing literature by deriving the social welfare weights in the optimal-tax model of Jacquet et al. (2013), which incorporates both an intensive and extensive labor-supply margin. Part of our findings confirm expectations. First, all parties roughly give a higher social weight to the poor than to the rich. Second, left-wing parties give a higher social weight to the poor and a lower social weight to the rich than right-wing parties do. We demonstrate that cross-party differences in social welfare weights are very small and extremely close to the social welfare weights in the existing tax-benefit system. We also uncover two important anomalies for all parties under consideration. First, social welfare weights increase from the working poor to modal income, suggesting that (reverse) redistribution from the poor to middle-income groups raises social welfare. Second, the social welfare weight given to the rich is negative for all political parties, implying that the Dutch government more than 'soaks the rich'. We argue that the high social welfare weights for the middle-income groups can be explained best by political-economy considerations.

Key words: Optimal taxation, revealed social preferences, political parties
JEL-codes: C63, D63, H21

*We thank Nicole Bosch for her assistance in calculating the effective marginal tax rates used in this paper. We have benefited from comments and suggestions by Olivier Bargain, Etienne Lehmann, Erzo Luttmer, Andreas Peichl, Emmanuel Saez, Paul Tang, Danny Yagan and seminar and congress participants at CPB Netherlands Bureau for Economic Policy Analysis, IIPF Michigan, University of California Berkeley and the CPB Workshop Behavioural Responses to Taxation and Optimal Tax Policy. All remaining errors are our own. Zoutman gratefully acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO) under Open Competition grant 400-09-383.

†NHH Norwegian School of Economics, Department of Business and Management Science and Norwegian Center of Taxation. E-mail: floris.zoutman@nhh.no

‡Erasmus University Rotterdam, Tinbergen Institute, Netspar, and CESifo. E-mail: bjacobs@ese.eur.nl. Homepage: people.few.eur.nl/bjacobs.

§CPB Netherlands Bureau for Economic Policy Analysis. E-mail: E.L.W.Jongen@cpb.nl.

"Don't tell me what you value, show me your budget, and I will tell you what you value."
Joe Biden – US Presidential Elections, September 15, 2008

1 Introduction

The quote from Vice-President Joe Biden of the US appeals to many economists who prefer revealed over stated preferences. In this paper we try to go beyond the rhetoric of the political debate, to directly measure the redistributive preferences of political parties. To that end, we use unique data on the proposed tax-benefit system of Dutch political parties in their election campaigns. We use the inverse optimal-tax method to derive the social welfare weights that Dutch political parties attach to different income groups. This allows us to analyze whether parties care more about the poor than the rich and who cares the most about whom. This also allows us to study whether political-economy considerations play a role in the reform proposals.

Revealing the implicit social preferences of tax-benefit systems is an exciting new research area in optimal income taxation. For quite some time, optimal tax theory, which originated from the seminal contribution by Mirrlees (1971), remained rather theoretical and provided little guidance to actual tax policy. However, at the turn of the century Diamond (1998) and Saez (2001, 2002) greatly increased its relevance. In particular, Saez (2001) showed that optimal tax rates can be computed once the elasticity of the tax base, the distribution of gross earnings, and the social preferences for redistribution are known. In principle, both the elasticities of taxable income and the earnings distribution can be determined empirically.[1] However, the social preference for income redistribution is ultimately a political question on which economists have little to say. Indeed, researchers can only determine plausible ranges of optimal marginal tax rates within the boundaries determined by the Rawlsian and utilitarian social welfare functions. Comparing the resulting optimal tax schedules with actual schedules may reveal whether the actual system is optimal, and where there might be room for improvements in social welfare.

A somewhat less ambitious, but equally revealing strategy is to invert the optimal-tax problem and look for the social preferences that render a given tax-benefit system optimal. This so-called inverse optimal-tax method has been developed by Bourguignon and Spadaro (2012). By deriving the social welfare weights in this way, anomalies in current tax-benefit schedules can be detected, and welfare-improving tax reforms can possibly be identified. More importantly, by using this strategy one circumvents the necessity to assume an intrinsically unknown construct such as the social welfare function.

We will use the inverse optimal-tax method to compute the social welfare weights implicit in the actual Dutch tax-benefit system, and to analyze the social welfare weights of Dutch political parties.[2] Since 1986, in a process unique in the world, all major Dutch political parties provide

[1] *E.g.* Brewer et al. (2010) for the UK, Jacquet et al. (2013) for the US and Zoutman et al. (2013) for the Netherlands, all recover the ability distribution using detailed micro data on income, corresponding marginal tax rates and the elasticity of the tax base.

[2] This method has also been applied by Blundell et al. (2009), Bargain and Keane (2010) and Bargain et al. (2011),

CPB Netherlands Bureau for Economic Policy Analysis (CPB) with very detailed policy proposals in every election for the national parliament. CPB then calculates and reports the income, budgetary and behavioral effects of each political party's election program, which then play an important role in the run up to the national elections, but also during the negotiations to form a new government after the elections.[3] These data also contain detailed information on policies to change the Dutch tax-benefit system, which provide us with an opportunity to estimate the redistributive preferences of the political parties.

We invert the optimal-tax model of Jacquet et al. (2013), which allows for both an intensive (hours or effort) and an extensive (participation) decision margin. In doing so, our study is the first in the literature to derive the social welfare weights in a continuous-type model with both intensive and extensive labor-supply responses. Previous studies were only able to analyze social welfare weights in discrete-type models, see e.g. Bourguignon and Spadaro (2012), Bargain and Keane (2010), and Bargain et al. (2013). In addition, we allow for general utility functions allowing for income effects. Finally, our social welfare weights are based on sufficient statistics only: marginal tax rates, participation tax rates, intensive and extensive elasticities, and the labor earnings distribution. Hence, one can employ our model to compute social welfare weights of any tax-benefit system.

We apply this model to the proposed tax-benefit systems of Dutch political parties in the 2002 elections. We use data for 2002 because this is the same year for which Zoutman et al. (2013) recover the ability distribution, using detailed micro data on the income distribution and marginal tax rates. We focus on the proposals by the four main political parties in the Dutch parliament after the 2012 elections that fit into the 'left-wing' and 'right-wing' taxonomy regarding preferences for redistribution. Our main findings are as follows.

In line with prior expectations, all parties attach a larger social weight to the poor than to the rich. Furthermore, again in line with expectations, we find that left-wing parties give a higher social weight to the poor and a lower social weight to the rich than right-wing parties do. What is surprising is how small the differences are in the proposed tax-benefit systems. This implies that revealed social welfare weights are found to be very close across all political parties.

However, we also uncover a number of anomalies. First, we find that social welfare weights are increasing from the working poor to the middle-income groups, rather than decreasing. Parties attach relatively more weight to workers with middle incomes than to workers with low incomes. Indeed, in the Netherlands support schemes are phased out in a relatively dense part of the income distribution, so as to redistribute income towards the middle incomes. Our analysis suggests that this is particularly relevant for the proposals of left-wing parties. Second, all parties, including the most right-wing conservative-liberal party, attach a negative social weight to the rich. Hence, all parties set the top tax rate beyond the 'Laffer rate' at which the government completely 'soaks the rich'. However, it is likely that political parties underestimated the elasticity of the tax base with

see the literature review in Section 2.

[3]See CPB and PBL (2012) for the analysis of the 2012 elections, and the contributions in Graafland and Ros (2003) for the pros and cons of this exercise.

respect to top tax rates, since CPB did as well.

The anomalies we detect are consistent with important political-economy theories. First, the increasing social welfare weights until the middle-income groups can be understood by standard political models of income redistribution, since the support of middle-income voters is crucial to get elected (Meltzer and Richard, 1981; Roberts, 1977; Romer, 1975). Second, the patterns of the social welfare weights – increasing to modal incomes and sharply decreasing thereafter – are in line with Director's law, where the middle-income groups form a successful coalition against the low-income and high-income groups (Stigler, 1970). Third, the high welfare weights for the middle-income groups could be explained by two-dimensional political competition. Even left-wing parties may sacrifice on their redistributive goals if this helps to achieve larger electoral success by attracting more voters on other ideological positions (Roemer, 1998, 1999). Fourth, post-election considerations could explain the strong status-quo bias in announced tax-benefit plans. Political parties may deliberately want to avoid highly pronounced party positions, since they need to form a coalition government with other parties after the elections. Fifth, the strong status-quo bias that we detect, but also the persistence of various anomalies across parties, could also be explained by collective-action problems. Vested interests could be effective in blocking welfare-improving tax-benefit reforms if the benefits of these reforms are dispersed and the costs of the reforms are concentrated at the vested interests (Olson, 1982).

The outline of the paper is as follows. First, in Section 2 we briefly summarize the existing literature on the inverse optimal-tax method. In Section 3 we outline the optimal tax model that is used in the analysis, and then invert the optimality conditions to get an expression for the implicit social welfare weights. In Section 4 we discuss the calibration of the model and illustrate the inverse method by revealing the social welfare weights in the baseline. In Section 5 we then turn to the political parties. We first give a brief overview of the political parties in the 2002 elections, and outline the reform packages they propose for the tax-benefit system. Next, in Section 6 we present the implicit social welfare weights of the proposed systems. Section 7 offers a number of explanations for the anomalies we uncover. Section 8 concludes. An Appendix at the end contains the derivations and some additional graphs.

2 Earlier literature

Pioneering work on the dual approach for tax-benefit systems has been done by Bourguignon and Spadaro (2012).[4] They reveal the social preferences for income redistribution in the French tax-benefit system, using the inverse optimal-tax problem of Saez (2001) with an intensive decision margin (hours or effort), and the inverse optimal-tax problem of Saez (2002) with both an intensive and an extensive decision margin (not only hours or effort, but also participation). For the model

[4] Studying the 'dual' problem of optimal taxation has a longer history, see *e.g.* Stern (1977), Christiansen and Jansen (1978), Ahmad and Stern (1984) and Decoster and Schokkaert (1989). However, only recently have researchers been able to use detailed micro data on incomes and corresponding marginal tax rates to study the social preferences implicit in tax-benefit systems.

with only an intensive margin Bourguignon and Spadaro (2012) find that social welfare weights are always decreasing, but they turn negative at the top-income earners.[5] They obtain these results both when considering only single earners and when considering all income earners and averaging income for couples. When Bourguignon and Spadaro (2012) introduce an extensive margin, they find that social welfare weights are no longer monotonically declining, and can also turn negative for the working poor when participation elasticities are larger than 0.5.

Blundell et al. (2009) consider the social welfare weights of single mothers in the UK and Germany, allowing for both the intensive and extensive decision margin. Their analysis goes a step further than Bourguignon and Spadaro (2012) in that they estimate rather than calibrate the behavioural elasticities, using micro data and a discrete-choice model for labor supply. For both Germany and the UK they find that social welfare weights are not monotonically decreasing with income, as the working poor get a lower weight than middle incomes. For Germany they find a negative social weight for working single mothers with a low income and children younger than school-age.[6]

Bargain and Keane (2010) perform a similar analysis for singles in Ireland. Moreover, they estimate social welfare weights at (four) different points in time (ranging from 1987 to 2005). They find that social welfare weights are remarkably stable over time, despite some significant policy changes. They do not find negative social welfare weights. However, they also find that social welfare weights are not monotonically declining with income as the working poor have a lower social welfare weight than middle-income earners.

Finally, Bargain et al. (2011) conduct a similar analysis for singles in 17 European countries including the Netherlands, and the US. They find that social welfare weights are always positive, although they are not monotonically declining for low-income groups, which is in line with the studies considered above. They further find that there are significant differences in social welfare weights between groups of countries (the US vs. Continental and Nordic Europe vs. Southern Europe), but rather similar social welfare weights for countries within a particular group.

3 Model

We derive the social welfare weights in the optimal-tax model of Jacquet et al. (2013) to calculate the social welfare weights implied by the the tax-benefit schedules proposed by political parties. Jacquet et al. (2013) combine the Mirrlees (1971) model of optimal income taxation with only an intensive labor-supply margin with the Diamond (1980) model of optimal income taxation with only extensive labor-supply margin. This section discusses the key features of the model, and derives

[5] The social weights turn negative even though they do not include indirect taxes (close to 20% of income net of direct taxes), as noted by Bourguignon and Spadaro (2012). Including indirect taxes in marginal tax rates would make the social weights even more negative at the top.

[6] They do not find negative social weights for top incomes. However, behavioral responses in their model are only in hours worked, which are very low at the top. This probably understates the tax base response to changes in the marginal tax rate at the top as suggested by the literature on the elasticity of taxable income (Feldstein, 1999). The same is true for Bargain and Keane (2010) and Bargain et al. (2011).

the formula for the social welfare weights in the optimal tax-benefit system. A full description of the model, as well as derivations of all formulas can be found in the Appendix.

We generalize earlier literature on the inverse optimal-tax method by allowing for continuous skill types and by allowing for income effects on the intensive margin (see e.g. Saez, 2002, Bourguignon and Spadaro, 2012, Bargain and Keane, 2010, and Bargain et al., 2013). Moreover, our formula for the social welfare weights is based on so-called sufficient statistics. To calculate social welfare weights, we only need to know marginal tax rates, participation tax rates, behavioral elasticities for intensive and extensive responses and the earnings distribution. All of these are available from the data.

3.1 Optimal tax-benefit system

We follow Mirrlees (1971) by assuming individuals in the economy differ in their earnings ability $n \in [\underline{n}, \overline{n}]$, $0 < \underline{n} < \overline{n} \leq \infty$. Earnings ability n denotes labor productivity per hour worked. Gross labor earnings are given by $z \equiv nl$, where l is labor supply.[7] Earnings ability is private information. Jacquet et al. (2013) introduce the extensive margin decision through a random participation model. That is, when individuals decide to participate in the labor market they incur an idiosyncratic utility cost (or benefit) $\varphi \in (-\infty, \infty)$, which reflects an individual-specific cost from participation, for example, forgone leisure time or household production, or the cost of commuting to work. However, participation costs can also be negative, for example, because of the value of social contacts at work or by avoiding the sigma of being non-employed. Like ability, participation costs are private information.[8]

Labor earnings and employment status are verifiable to the government. Hence, the government can condition taxes and transfers on gross labor income z and the individual's employment status. The income tax function is non-linear, continuous and denoted by $T(z)$, where $T'(z) \equiv \mathrm{d}T(z)/\mathrm{d}z$ is the marginal tax rate. All net labor income is spend on the consumption good. Consequently, the individual budget constraint is: $c = z - T(z)$. Non-employed workers receive a non-employment benefit b, which generally differs from the net income of employed workers earning zero income, i.e. $-T(0)$. Hence, the non-employed enjoy consumption $c = b$, while they do not provide any labor effort, i.e., $l = 0$.

If individuals decide to participate in the labor market, they maximize utility $u(c, l)$, which increases at a diminishing rate in consumption c, and decreases at an increasing rate in labor supply l. Our analysis assumes that the utility function is separable, i.e., $u(c, l) \equiv v(c) - h(l)$.[9] An individual participates in the labor market if his utility from participation net of participation

[7] Jacquet et al. (2013) demonstrate that the production technology, which transforms earnings ability and labor effort into labor earnings, can be generalized to any general function $z \equiv f(n, l)$.

[8] Both earnings ability n and disutility of participation φ are continuously distributed in the population according to some joint distribution function. Since we will express the optimal tax rules and social welfare weights only in terms of sufficient statistics, we do not need to specify these non-observable distributions. See the Appendix for the formal derivations.

[9] This assumption significantly simplifies the formula for the welfare weight of the non-employed. The analysis can be generalized to settings where we do not assume separability in utility.

costs is larger than the utility obtained while being non-employed, i.e., when $v(c)-h(l)-\varphi \geq v(b)$.

Social welfare is given by a Samuelson-Bergson social welfare function, which is the total sum of a concave transformation of individual utilities $W(u)$, where $W'>0$, $W''\leq 0$. The government is restricted by its budget constraint stating that total revenue from taxing labor income of the employed equals total spending on non-employment benefits and an exogenous revenue requirement. Let λ denote the shadow value of public funds. Then, the social welfare weight given to a worker of earning ability n equals $g_n \equiv W'(u_n)v'(c_n)/\lambda$. The average social welfare weight of all working individuals at income level z is represented by g_z, while the social welfare weight for the non-employed individuals is denoted by g_0. g_z (g_0) measures the monetized gain in social welfare of providing one unit of income to employed individuals with income z (non-employed individuals without labor earnings). Social welfare weights are generally positive and decreasing with income when standard social welfare functions are employed. Hence, social welfare increases if the government redistributes income from rich to poor individuals.

Given that earnings ability is private information, the government cannot redistribute income without distorting labor-supply incentives. Indeed, the government needs to ensure that incentive-compatibility constraints are always respected. Due to the random-participation structure of the model, the incentive-compatibility constraints depend only on earnings ability n, but not on participation costs φ. The government optimally chooses the non-linear tax function $T(z)$, and the non-employment benefits b to minimize resources subject to incentive constraints and a distributional constraint, which specifies an exogenously given level of utility for each individual. This problem can be solved like in Mirrlees (1971) by deriving the optimal second-best allocation using a direct mechanism, and decentralizing this allocation by employing the non-linear tax schedule and non-employment benefits.

We first turn our attention to the formula for the optimal marginal tax schedule, which is expressed in ABC-form as in Diamond (1998) and in terms sufficient statistics as in Saez (2001) – see the Appendix:

$$\frac{T'(z)}{1-T'(z)} = \underbrace{\frac{1}{\varepsilon_z^c}}_{\equiv A_z} \underbrace{\frac{\int_z^{\overline{z}} \left(1-g_{z'}+\eta_{z'}\frac{T'(z')}{1-T'(z)} - \varepsilon_{z'}^P\left(\frac{T(z')+b}{z'-T(z')}\right)\right)\phi(z')\mathrm{d}z'}{1-\Phi(z)}}_{\equiv B_z} \underbrace{\frac{1-\Phi(z)}{\phi(z)z}}_{\equiv C_z}, \tag{1}$$

where $\varepsilon_z^c \equiv -\frac{\partial z^c}{\partial T'}\frac{(1-T')}{z} > 0$ is the compensated earnings-supply elasticity (or the elasticity of taxable income) with respect to the net-of-tax rate $1-T'$ of an employed individual earning income z. $\overline{z}$ is the top earnings level, which can be infinite. $\eta_z \equiv -(1-T')\frac{\partial z}{\partial \rho} \geq 0$ is the income elasticity in earnings supply when a worker with earnings z receives ρ in additional non-labor income. Note that it is defined negatively. $\varepsilon_z^P \equiv \frac{\partial E_z}{\partial(T(z)+b)}\frac{(T(z)+b)}{E_z} > 0$ denotes the participation elasticity of workers at earnings level z, where E_z is the employment rate at earnings level z. Finally, $\Phi(z)$ is the cumulative distribution of earnings of employed workers, where $\phi(z)$ is the corresponding density function of earnings. Equation (1) is a simplification of the original optimal-tax formula in Jacquet et al. (2013), which depends on the joint distribution of unobserved ability and utility costs of work. However, without loss of generality we can express the formula in terms of sufficient statistics, as in Saez (2001), see the Appendix.

The economic function of the marginal tax rate $T'(z)$ at income level z is to raise the tax burden for all individuals with an income above z. The government then uses the tax revenue to either increase the transfers $-T(0)$ workers with income below z or to increase benefits b for the non-employed. However, raising the marginal tax burden $T'(z)$ generates a compensated earnings-supply response for workers with income z. At each point in the earnings distribution the government thus needs to trade off the equity benefits of more income redistribution against the larger distortions in earnings supply. The ABC-formula captures all elements of this trade off.

A_z measures the average distortions of the income tax on the intensive margin. If the compensated elasticity of earnings supply ε_z^c is larger, income taxes distort the intensive earnings-supply margin more. Hence, the optimal tax rate T' at income level z should optimally decline as the elasticity of taxable income increases ε_z^c.

B_z captures the equity gains of higher marginal tax rates. B_z represents the average gain in social welfare of raising one unit of tax revenue from all tax payers above gross income level z. If the government increases the tax burden on everyone with an income above z with one unit, government revenue mechanically increases with one unit, as indicated by the first term inside the integral. However, raising one unit of revenue also inflicts utility losses on all individuals paying one unit higher tax. These utility losses are given by the social welfare weights g_z. Moreover, raising the tax burden also induces an income effect: earnings supply increases when individuals become poorer. This raises tax revenue by $\eta_z T'(z)/(1-T'(z))$. Finally, the additional tax burden raises the participation tax for all incomes above z. As some individuals stop participating, the government loses $\varepsilon_z^P(T(z)+b)/(z-T(z))$ revenue, which is equal to the participation elasticity times the participation tax rate (expressed in terms of net labor earnings). Thus, the distributional benefits of setting a higher marginal tax rate are lower when the extensive margin is more elastic or when participation is more heavily taxed. Finally, the B_z-term averages the direct revenue gains of raising taxes with one unit of income minus associated utility costs and tax-base effects over all individuals with earnings above z.

With standard social welfare functions, the B_z-term typically increases with income, as was demonstrated first by Diamond (1998) in the absence of both income effects and the extensive margin. The prime reason is that the utility losses caused by higher tax payments decline with income z, because richer individuals have lower social welfare weights. Moreover, from a theoretical point of view, income effects in taxable income should be more important for wealthier individuals. Empirically, however, income effects are found to be very small (Saez et al., 2012). Finally,participation elasticities decline with income, both theoretically – since opportunity costs of non-participation rise with earnings – and empirically – see the evidence discussed below. With Rawlsian or 'charitable conservatism' social welfare functions the B_z-term becomes constant when income and participation effects are absent, since social welfare weights become constant. They are equal to zero with the Rawlsian social welfare function.

The C_z-term weighs the efficiency costs A_z per unit of tax base and the average redistributional gains B_z. The numerator in the C_z-term, $1 - \Phi(z)$, captures the number of individuals above z paying a higher marginal tax rate. The larger is the number of individuals are above z, the larger are distributional gains of higher marginal tax rates, and the larger should the optimal tax rate be. As the number of individuals paying higher taxes always declines with income, the total distributional benefits of higher tax rates are continously declining with income – for given average distributional benefits B_z. The denominator in the C_z-term captures the weight on efficiency, since $z\phi(z)$ is the size of the tax base at which the marginal tax rate is levied. The larger is the tax base $z\phi(z)$, the larger are the total efficiency costs of increasing marginal tax rates – for given average costs per unit of tax base A_z –, and the lower should the optimal tax rate be. The efficiency losses of marginal tax rates typically follow the shape of the earnings distribution: low at the bottom, increasing towards the mode, and decreasing thereafter. Consequently, the C_z-term always falls until the mode as its numerator declines and its denominator increases. After the mode, the behavior of C_z becomes theoretically ambiguous as both the numerator and denominator decline with earnings. Empirically, however, the C_z-term increases after the mode in many countries, see for example Saez (2001) and Zoutman et al. (2013). Hence, marginal tax rates should optimally increase after the mode – even with a Rawlsian social welfare function. In the limit, the C_z-term converges to a constant if earnings are Pareto distributed (Diamond, 1998; Saez, 2001). Intuitively, distributional benefits and efficiency costs of higher marginal tax rates decline at equal rates in the Pareto tail of the earnings distribution.

The optimal participation tax implicitly follows from – see Appendix:

$$E \int_{\underline{z}}^{\overline{\overline{z}}} \varepsilon_z^P \left(\frac{T(z) + b}{z - T(z)} \right) \frac{v'(b)}{v'(z - T(z))} \phi(z) \mathrm{d}z = (1 - E)(g_0 - 1), \tag{2}$$

where E is the aggregate employment rate in the economy, and $\underline{z}$ is the lowest earnings level, which is possibly zero. Equation (2) gives the optimality condition for the optimal participation tax. In the optimum, the marginal benefits of redistributing income from the employed to the non-employed (right-hand side) should be equal to the marginal costs of doing so (left-hand side).

The right-hand side gives the total distributional benefits of redistributing resources to the non-employed. Intuitively, suppose that the government raises the participation tax by increasing non-employment benefits b by 1 unit of income. Then, $g_0 - 1$ gives the mechanical welfare gain minus the mechanical cost of this marginal increase in b. The welfare weight of the non-employed g_0, i.e., the individuals who are worst off, is typically larger than 1, since the average welfare weight is approximately one (it is exactly one in the absence of income effects, see below). Hence, redistributing income from the employed to the non-employed raises social welfare. Redistribution towards the non-employed is more valuable, the larger is the number non-employed, i.e., the lower is E.

The left-hand side of equation (2) captures the total participation distortions among working individuals. The participation tax rate is $(T(z)+b)/(z-T(z))$, which equals the total participation tax $T(z)+b$ divided by net income $z-T(z)$. The participation tax consists of two parts. First, when an individual starts working and earns income z he/she faces income taxes $T(z)$. Second, he/she then loses non-employment benefits b. $\varepsilon_z^P\left(\frac{T(z)+b}{z-T(z)}\right)$ captures the social cost of lower participation when the non-employment benefit b is raised as some individuals stop paying taxes and start collecting non-employment benefits. The social cost of the participation tax increases in the participation elasticity ε_z^P and if there are more employed workers, i.e. when the employment rate E is larger.

To further sharpen the intuition for equation (2), suppose that the participation elasticity is constant, $\varepsilon_z^P=\varepsilon^P$, and utility is quasi-linear, such that $v'(b)=v'(z-T(z))$. In that case, we can rewrite equation (2) as:

$$E\int_{\underline{z}}^{\infty}\left(\frac{T(z)+b}{z-T(z)}\right)\phi(z)\mathrm{d}z=(1-E)\frac{(g_0-1)}{\varepsilon^P}. \tag{3}$$

Equation (3) very much resembles the optimal tax expression in the discrete-type model of Saez (2002). The left-hand side gives the participation tax rate for all employed workers, whereas the left-hand side gives the distributional benefits of setting a positive participation tax. The optimal participation tax falls when the participation elasticity is higher or when the non-employed have a lower welfare weight g_0. Finally, the participation tax increases when the non-employed become more important compared to the employed, i.e., when the employment rate E is lower. Equation (2) corrects equation (3), because participation elasticities generally differ with income. In addition, there is a difference because participation costs φ are expressed in utility rather than monetary terms, which disappears with quasi-linear utility.

3.2 Social welfare weights

Our paper aims to uncover the social preferences for income redistribution by Dutch political parties. We do so by using the inverse optimal-tax method. If one is willing to make the assumption that political parties optimally chose the tax-benefit system to satisfy their political desires for income redistribution, then we are able to fully recover each party's social welfare weights for all income groups, including the non-employed. In particular, from the parties' announced election proposals regarding the tax-benefit system, we are able to calculate marginal and participation tax rates by income level. And, by combining this information with data on earnings and estimated earnings and participation elasticities, we are able to distill the social welfare weights that each political party attaches to each income group and to the non-employed.

3.3 Social welfare weights working individuals

We compute social welfare weights g_z for all working individuals by inverting equation (1) – see the Appendix for the derivation:

$$g_z = 1 + (1+\beta_z+\zeta_z)\,\varepsilon_z^c \frac{T'(z)}{1-T'(z)} + \varepsilon_z^c \frac{zT''(z)}{(1-T'(z))^2} + \eta_z \frac{T'(z)}{1-T'(z)} - \varepsilon_z^P \left(\frac{T(z)+b}{z-T(z)}\right), \tag{4}$$

where $\beta_z \equiv \frac{\phi'(z)z}{\phi(z)}$ is the elasticity of the earnings density with respect to gross earnings, and $\zeta_z \equiv \frac{\partial \varepsilon_z^c}{\partial z}\frac{z}{\varepsilon_z^c}$ is the elasticity of the compensated earnings-supply elasticity with respect to gross earnings. Equation (4) shows that the formula for the social welfare weights is based on sufficient statistics only. That is, we can calculate social welfare weights by only using information on observables: marginal and participation tax rates, compensated and income elasticities of earnings supply, participation elasticities, and the earnings distribution.

To gain intuition for the determinants of the social welfare weights, first note that all welfare weights are equal to one ($g_z = 1$) when marginal tax rates are zero. Hence, the government should attach the same welfare weight to all individuals if it does not engage in any income redistribution through distortionary taxes. The most important determinant of the social welfare weights is the *change* of the deadweight loss $DWL_z \equiv \varepsilon_z^c \frac{T'(z)}{1-T'(z)} z\phi(z)$ with earnings z. Here, $\varepsilon_z^c \frac{T'(z)}{1-T'(z)}$ stands for the marginal deadweight loss per unit of tax base at income level z, and $z\phi(z)$ is the size of the tax base at z. To see why, note that the derivative of DWL_z with respect to z is given by:

$$\frac{\partial DWL_z}{\partial z} = \left((1+\beta_z+\zeta_z)\varepsilon_z^c \frac{T'(z)}{1-T'(z)} + \varepsilon_z^c \frac{zT''(z)}{(1-T'(z))^2}\right)\phi(z). \tag{5}$$

By substituting equation (5) into equation (4), the social welfare weights can thus be rewritten as:

$$g_z = 1 + \frac{1}{\phi(z)}\frac{\partial DWL_z}{\partial z} + \eta_z \frac{T'(z)}{1-T'(z)} - \varepsilon_z^P \left(\frac{T(z)+b}{z-T(z)}\right) \approx 1 + \frac{1}{\phi(z)}\frac{\partial DWL_z}{\partial z}, \tag{6}$$

The approximation applies whenever income and participation elasticities are very small.[10] Equation (6) demonstrates that the behavior of deadweight losses DWL_z with income along the optimal tax schedule critically determines the pattern of social welfare weights. Intuitively, if the deadweight losses are found to be increasing at income level z, then the government redistributes from individuals with incomes higher than z to individuals with incomes at z. Hence, if the tax system is optimally set, the government should attach a smaller social welfare weight g_z to individuals with an income higher than z than to individuals with an income at z. Based on this decomposition we can easily understand the behavior of the social welfare weights at the optimal tax system.

In equation (5), $1+\beta_z$ is the elasticity of the tax base $z\phi(z)$ with respect to earnings z, where $\beta_z \equiv \frac{z\phi'(z)}{\phi(z)}$. If $1+\beta_z$ is positive (negative), marginal tax rates generate larger (smaller) distortions

[10] Note that when the welfare weights are expressed in terms of the Diamond (1975)-based social marginal value of income $g_z^* \equiv g_z + \eta_z \frac{T'(z)}{1-T'(z)} - \varepsilon_z^P \left(\frac{T(z)+b}{z-T(z)}\right)$ – which includes the income effects on taxed bases – optimality of the tax system implies that $g_z^* = 1 + \frac{1}{\phi(z)}\frac{\partial DWL_z}{\partial z}$.

above earnings level z than at earnings level z. This indicates that marginal tax rates are optimally lower (higher) above z than they are at z – *ceteris paribus*. This must imply that the government attaches a higher (lower) welfare weight to individuals with earnings above z than to the individuals at earnings level z. Therefore, social welfare weights are increasing in the elasticity of the tax base with earnings $1+\beta_z$.

In equation (5), ζ_z is the elasticity of the compensated earnings-supply elasticity. Taxation becomes more (less) distortionary with earnings if ζ_z increases (decreases) with z, and taxes should optimally be lower – *ceteris paribus*. Hence, if tax systems are optimized, the government attaches a lower (higher) social welfare weight to individuals with earnings above z than to those with income level z – *ceteris paribus*. Social welfare weights thus decline in ζ_z.

In our simulations we impose a constant uncompensated elasticity of taxable income throughout the earnings distribution, which is not unreasonable given the empirical evidence we discuss later. Given that income effects are relatively small and do not vary much with income, the compensated elasticity ε_z^c varies little with income. Hence, ζ_z is close to zero so that the term associated with ζ_z does not matter much in the simulations presented below.

In equation (5), the term $\varepsilon_z^c \frac{zT''(z)}{(1-T'(z))^2}$ captures the non-linearity of the tax schedule on the social welfare weights. If marginal tax rates are increasing (decreasing) with earnings z, so that $T''(z) > 0$ ($T''(z) < 0$), the government attaches a lower (higher) value to people with income above z than to those with income at z, for the simple reason that it taxes them at a higher (lower) rate – *ceteris paribus*. As a result, social welfare weights increase in $T''(z)$.

Social welfare weights display discontinuities if political parties generate spikes in marginal tax rates over small income intervals. For individuals in upward part of the spike, $T''(z)$ is large, and hence the welfare weight is high as well – *ceteris paribus*. For individuals in the downward part of the spike the $T''(z)$ is very low, and hence, the social welfare weight is very low as well. These political parties apparently want to redistribute income towards people just below the spike and away from people just above the spike. This is anomalous, since such a policy generates large differences in social welfare weights for individuals differing only slightly in their earnings.

From equation (6) we see that income income effects on the intensive margin, as represented by $\eta_z \frac{T'(z)}{1-T'(z)}$, raise the social welfare weights – *ceteris paribus*. Intuitively, marginal tax rates result in an income effect in earnings supply, which raises tax revenue. Consequently, stronger income effects raise the distributional benefits of higher marginal tax rates for given deadweight losses. If the tax system is optimized, social welfare weights should therefore be higher if income effects are more important. In our later simulations we assume only small income effects, in line with empirical evidence. Hence, this term should not affect the pattern of social welfare weights much.

Finally, the last term in equation (6) $-\varepsilon_z^P(T(z)+b)/(z-T(z))$ is the participation distortion at income level z. Social welfare weights are lower if participation decisions are more severely distorted at income z – either because of a higher participation tax or larger participation elasticities. Distortions on the extensive margin reduce the redistributional benefits of higher marginal tax rates for given deadweight losses. Consequently, if the participation margin is more heavily distorted, social welfare weights should be lower if the tax-benefit system is optimized.

3.4 Social welfare weights top-income earners

From (4) we can also distill the welfare weights for the top income earners if the top of the earnings distribution is Pareto distributed. This is what we will assume in our simulations, based on our own estimates of the Pareto parameter of the top tail of the Dutch earnings distribution in Zoutman et al. (2013). Moreover, a Pareto tail provides an excellent fit to most top tails of the earnings distribution as Atkinson et al. (2011) have documented.

When the top of the earnings distribution is Pareto with parameter a, we can derive that $1+\beta_z=-a$. If we realistically assume that participation elasticities are negligible for top earners ($\varepsilon_z^P=0$), and that compensated and income elasticities are constant ($\varepsilon_z^c=\varepsilon^c$, $\eta_z=\eta$, $\zeta_z=0$), then optimal top tax rates will be constant as well when social welfare weights for top earners g_∞ are constant. Then, the social welfare weight at the top g_∞ is given by:

$$g_\infty = 1-(a\varepsilon_z^c-\eta)\frac{T'(\infty)}{1-T'(\infty)}. \tag{7}$$

The social welfare weight for the top-income earners g_∞ declines when the government levies a higher marginal tax rate or when elasticities of taxable income are higher. In either case, marginal top tax rates generate larger distortions for given distributional benefits. This is only optimal if the government attaches a lower social welfare weight to top-income earners. The social welfare weight is larger when income effects are more important. The intuition is identical to the one we had above. Income effects in earnings supply raise the distributional benefits of higher marginal tax rates for given deadweight losses. If the tax system is optimized, social welfare weights should therefore be higher if income effects are more important. The social welfare weight g_∞ declines with the Pareto parameter a. Thus, the fatter is the tail of the Pareto distribution (i.e., a lower a), the lower are the deadweight losses, and the higher optimal marginal tax rates. Consequently, welfare weights for top-income earners are lower.

The welfare weights for top income earners are non-negative, i.e. $g_\infty \geq 0$, when the marginal tax rate satisfies:

$$T'(\infty) \leq \frac{1}{1+a\varepsilon_z^c-\eta}. \tag{8}$$

When the inequality is strict, marginal tax rates are set at the top of the 'Laffer rate', beyond which an increase in the top tax rate reduces tax revenue. Setting top rates beyond the Laffer rate is therefore non-Paretian, since a reduction of top rates would both raise utility for top income earners, and raise tax revenue, which can be redistributed to make other individuals better off. Thus, whenever social welfare weights are found to be negative for some individuals, the government is wasting resources by making these individuals worse off (Brendon, 2013; Werning, 2007).

3.5 Social welfare weights non-employed

Finally, we can derive the welfare weight g_0 of the non-employed – see Appendix:

$$g_0 = 1 + \left(\frac{E}{1-E}\right)\int_{\underline{z}}^{\overline{z}} \frac{v'(b)}{v'(z-T(z))}(1-g_z)\phi(z)\mathrm{d}z. \tag{9}$$

Social welfare weights for the non-employed increase when there is more non-employment (E lower), and when the average welfare weights – corrected for the marginal utility of income – of the employed decrease. Note that in the absence of income effects, i.e., $v'=1$, the social welfare weights exactly sum to one at the optimal tax system:

$$(1-E)g_0 + E\int_{\underline{z}}^{\overline{z}} g_z\phi(z)\mathrm{d}z = 1. \tag{10}$$

Equivalently, this equation states that the marginal cost of public funds equals one when the tax system is optimized. Intuitively, the government adjusts the transfers $-T(0)$ and b so that a marginal unit of resources is valued equally in the public and private sector. Thus, the redistributional benefits of taxation should cancel against deadweight losses of taxation at the optimal tax system, see also Jacobs (2013). This result can be generalized to allow for income effects by using the Diamond (1975)-based social marginal value of income to calculate the social welfare weights.[11]

In the analysis that follows we are particularly interested in whether social welfare weights i) are monotonically declining in income, so that political parties always care more about poorer than richer individuals, ii) are always positive, since otherwise Pareto-improving tax reforms exist, and iii) feature discontinuous jumps, so that large differences in social weights exist for individuals differing only marginally in income, which, too, suggests the possibility of welfare-improving tax reforms.

4 Calibration and baseline welfare weights

This section explains in detail the data used in our analysis and the calibration of our model. Further, we construct the baseline welfare weights of the 2002 tax-benefit system, on which the comparisons with the political programs are based in the next section. To calculate the welfare weights we employ data on the income distribution, marginal tax rates, participation tax rates, employment rates, and recent estimates of the elasticity of the tax base for both the extensive and intensive margins.

[11]The Diamond (1975)-based social marginal value of income then equals $g_z^* \equiv g_z + \eta_z \frac{T'(z)}{1-T'(z)} - \varepsilon_z^P \left(\frac{T(z)+b}{z-T(z)}\right)$, so that the marginal cost of public funds equals unity. See also Jacobs (2013).

4.1 Income distribution and marginal tax rates

We define income as gross wage income, excluding employer contributions. We exclude all incomes from capital (interest, dividends and capital gains), self-employment, firm ownership and pensions. The marginal tax rate is defined as the difference between the increase in gross wages and the increase in net disposable income as a fraction of initial gross earnings. We use income data from the *Inkomenspanelonderzoek 2002* (IPO), collected by Statistics Netherlands. IPO is a stratified panel dataset containing adminstrative data on 175,876 individuals in 2002. It covers a little more than 1 percent of the Dutch population. Sampling weights are provided and we use them throughout our analysis. We focus our analysis on working-age individuals, hence we only select individuals from 23 until 65 years of age. Moreover, we exclude all individuals that are enrolled higher education in 2002, since their labor earnings may not be representative of their earnings ability. Our final data set consists of 94,859 individuals. Figure 1 plots a Gaussian kernel density estimate of the earnings distribution using a bandwidth of 5,000 euro.

We have relatively few observations in the top tail of the earnings distribution. Based on the same data, Zoutman et al. (2013) use the method of Clauset et al. (2009) to estimate that the Pareto distribution gives an excellent fit to the top of the Dutch income distribution. The Pareto parameter is estimated to be around 3.0, which is rather high compared to other countries and the estimate indicates that it is lonely at the top in the Netherlands. The estimated Pareto parameter is in line with other studies using Dutch data, see Atkinson and Salverda (2005) and Atkinson et al. (2011). The Pareto tail starts at 45,040 euros, which is approximately the start of the current top tax bracket containing the 8% richest tax payers.

The marginal tax rates are calculated using the tax-benefit calculator MIMOS-2 of CPB Netherlands Bureau for Economic Policy Analysis. MIMOS-2 takes into account all income-dependent subsidies and tax credits to calculate effective marginal tax rates. See Gielen et al. (2009) for more details. Moreover, our measure for the effective marginal tax rates also includes indirect taxes.[12] Figure 2 provides the kernel estimate for the corresponding effective marginal tax rates in the Dutch income distribution for all employed workers. Figure 12 in the Appendix gives a scatterplot of the marginal tax rates. There is large variation in marginal tax rates at each income level, in particular for lower incomes, due to the dependence of the tax-benefit system on other characteristics than individual labor income, such as household income, household composition, and the number of children, but also due to differences in non-labor incomes.[13] The model, however, only allows individuals to differ in their labor income and employment status.Therefore, we use a kernel estimate to smooth out the variation in individual marginal tax rates at each income level, and across individuals at different income levels.[14]

[12]Denote the effective direct marginal tax rate by t_d, the marginal indirect tax rate by t_i and the effective marginal tax rate by t_e. We calculate the effective total marginal tax rate as $t_e = \frac{t_d+t_i}{1+t_i}$.

[13]For example, welfare benefits and various income-support programs (e.g. rent assistance) are typically means-tested and based on household income, whereas most tax credits and the tax system are only based on individual income. The number of (working) family members determines eligibility to and level of various tax credits (e.g. the working family tax credit). Child-care support and child benefits depend on the number of children.

[14]Jacquet and Lehmann (2014) demonstrate that the Mirrlees (1971) framework can be be completely generalized

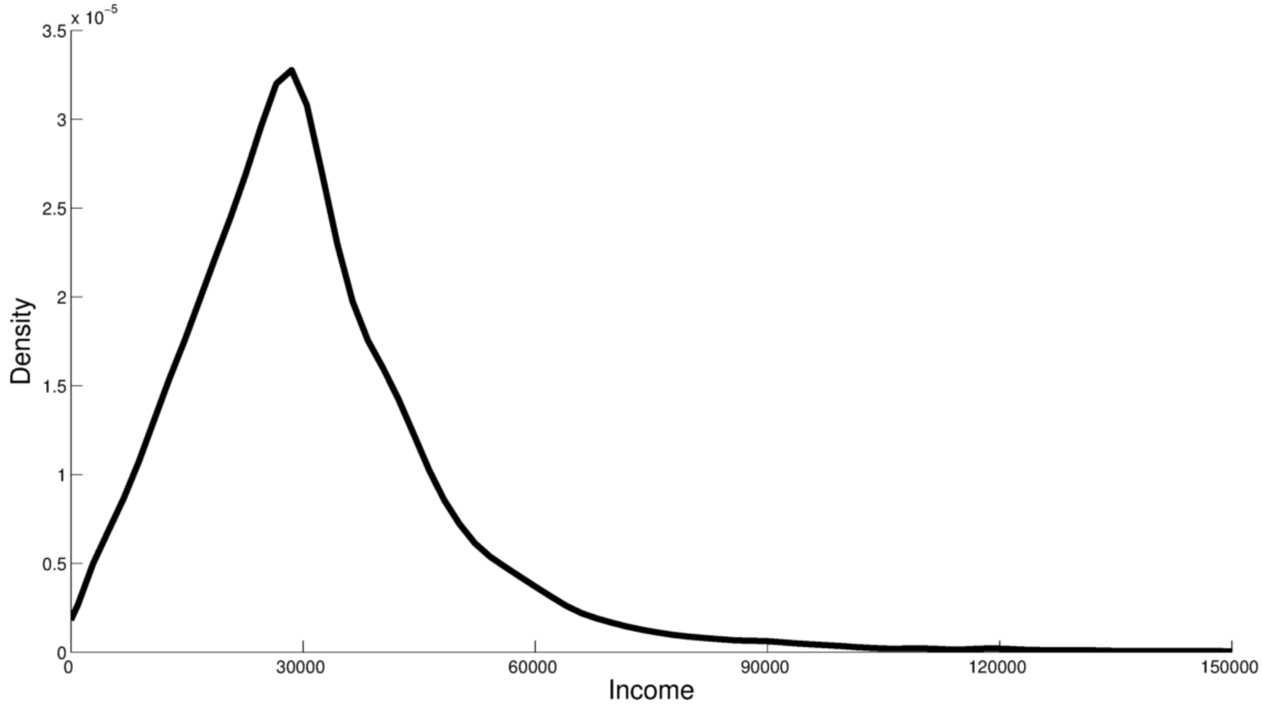

Figure 1: Kernel Density Estimate of Gross Wage Income in the Netherlands, 2002

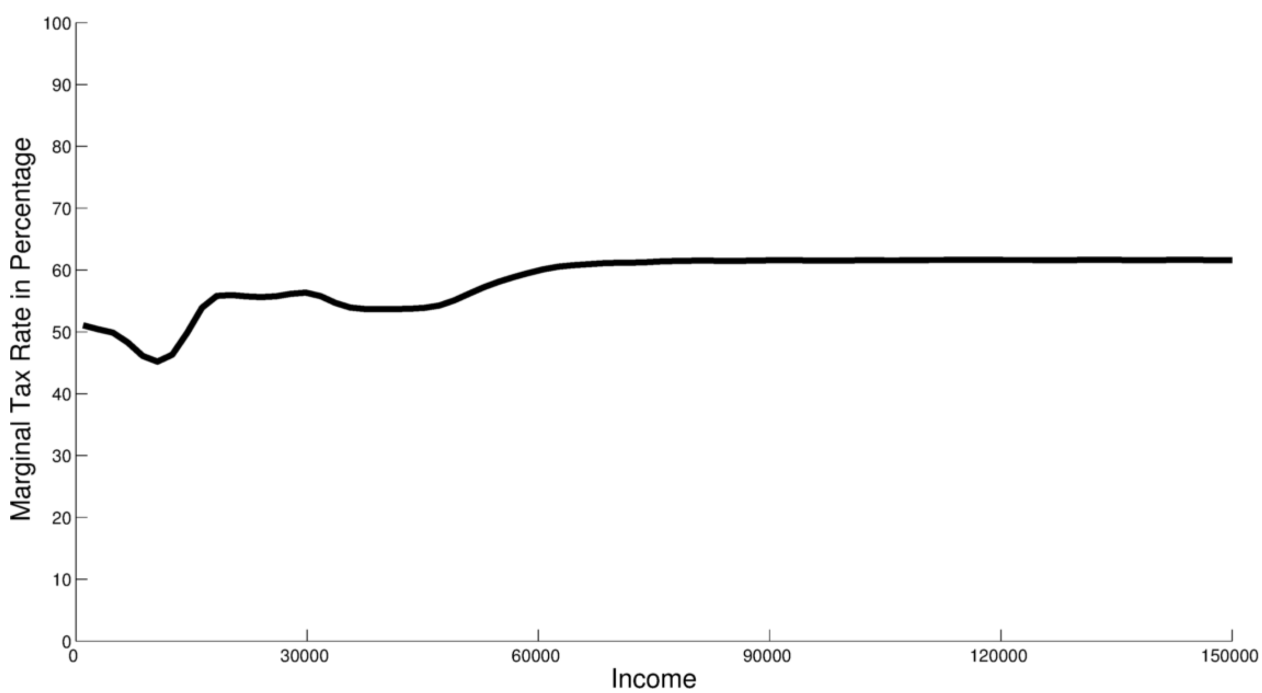

Figure 2: Kernel Density Estimate of Total Effective Marginal Tax Rates in the Netherlands, 2002

Table 1: Tax Brackets and Tax Credits in 2002

	Start	End	Percentage	Maximum amount
Tax brackets				
First tax bracket	0	15,331	32.35	4,960
Second tax bracket	15,331	27,847	37.85	4,737
Third tax bracket	27,847	47,745	42.00	8,357
Fourth tax bracket	47,745	∞	52.00	∞
Tax credits				
General tax credit	0	∞	0	1,647
Earned-income tax credit				
- First part	0	7,692	1.73	133
- Second part	7,692	15,375	10.62	949
Single parent tax credit	0	∞	0	1,301
Earned-income single-parent tax credit	0	30,256	4.30	1,301

To understand the patterns in Figure 2, Table 1 provides some parameters of the Dutch tax system in 2002. In 2002, the Dutch tax system has four tax brackets for labor income, based on individual (not household) income, with rates rising from somewhat below 33% at the bottom to 52% at the top. This explains why marginal tax rates are typically lower for individuals with a low income than for individuals with a high income.

There are also a number of noticeable deviations from statutory tax rates, which result from targeted subsidies and tax credits. The lowest income groups feature marginal tax rates that are higher than the rate in the first tax bracket, because a number of income-support schemes are phased out with income, in particular rent subsidies and a general child tax credit.[15] Thereafter, there is an income segment where marginal tax rates are lower due to the phase-in of the earned income tax credit (EITC). The end of the phase-in range for the EITC nearly coincides with the start of the second tax bracket at around 15,000 euro. Marginal tax rates then rise substantially up to around 40,000 euro.[16] Finally, effectivemarginal tax rates are higher than statutory marginal tax rates due to indirect taxes. Using publicly available input-output tables of Statistics Netherlands we calculate that indirect taxes on private consumption are 11.7% of private consumption in 2002. We assume that these indirect taxes are proportional to net labor income. Bettendorf et al. (2012) show that indirect taxes are close to proportional to consumption in the Netherlands.

to allow for individuals differing in multiple characteristics as long as they make only an earnings-supply choice. Their results should carry over to Jacquet et al. (2013) and thus our paper. This implies that all our derivations remain valid, except that we should take averages of all tax rates and elasticities at each income level.

[15] The exact subsidy levels and taper rates vary with household characteristics other than income, and are therefore not reported in Table 1.

[16] There is an additional jump for individuals earning a gross income close to 40,000 euro. Below a a certain income threshold, individuals can enter the public health insurance scheme with relatively low insurance-premium rates. Beyond this threshold individuals are forced to take private health insurance with relatively high insurance-premium rates. This results in spikes in marginal tax rates around the income threshold. In 2003 this health-care system has been replaced by an obligatory uniform private health insurance scheme, which covers all main health risks. It is financed by a payroll tax and 'lump-sum' premiums paid by individuals. Individuals can voluntarily top up the basic health insurance, with additional private insurance packages.

Table 2: Elasticities Used in the Simulation

	Compensated wage elasticity	Income elasticity	Uncompensated wage elasticity	Participation elasticity
Baseline scenario	0.35	0.10	0.25	0.10
Low-elasticity scenario	0.18	0.05	0.13	0.05
High-elasticity scenario	0.53	0.15	0.38	0.15

Table 3: Employment Rates by Level of Education

Level of education	Net employment rate	Share in population
Elementary school	36.90	11.99
Some high school	53.50	25.79
High school	56.80	10.26
Low-level college	71.20	15.84
Mid-level college	79.10	14.95
Bachelor degree	80.40	13.88
Master degree or higher	84.40	7.28

Table 4: Calibrated Parameters for the Utility Function

Parameter values	Base	Low Elasticity	High Elasticity
α	0.46	0.48	0.45
ε	0.38	0.18	0.60
γ	1981.67	13503.12	1082.11
μ_k	55.95	0.00	82.42
σ_k	271.27	511.00	189.98

4.2 Utility function and elasticity taxable income

Our analysis assumes the following utility function:[17]

$$u = \frac{c^{1-\alpha}}{1-\alpha} - \gamma \frac{l^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}}, \qquad \alpha, \gamma, \varepsilon > 0. \tag{11}$$

α governs the elasticity of the marginal utility of consumption and ε is the Frisch elasticity of earnings supply. α and ε are calibrated so as to match empirically estimated values for the compensated and uncompensated elasticities.[18] Parameter γ is an innocuous scaling parameter that we calibrate to keep the mean of the ability distribution fixed in the different simulations.

The extensive-margin elasticity is assumed to be 0.1 on average, which is based on the weighted average of estimates for this elasticity for different household types by Jongen et al. (2014). More-

[17] This utility function is also used by Mankiw et al. (2009). When $\alpha \to 1$ this utility function converges to the logarithmic Utility Type-II used by Saez (2001). When $\alpha = \frac{1}{\varepsilon}$ this specification is in line with the CES-functions used by Mirrlees (1971) and Tuomala (1984).

[18] All our scenarios have the same ratio $\varepsilon^c/\varepsilon^u$. It turns out that as long as the ratio $\varepsilon^c/\varepsilon^u$ is fixed, the calibrated α is almost similar for different elasticities. This is a useful property, since the demand for redistribution is driven by the concavity in the marginal utility of consumption. When α remains constant across specifications, we can thus isolate the effect of a change in the elasticities from redistributional concerns.

over, we want our model to match the participation rates by skill level that are given in Table 3. To that end, we optimize the parameters of the distribution of participation-costs to minimize the distance between the predicted and observed participation rates and the predicted and observed extensive-margin elasticity. See Zoutman et al. (2013) for more details.

The uncompensated intensive-margin elasticity is assumed to be 0.25, based on recent estimates of the elasticity of taxable income (ETI) in the Netherlands by Jongen and Stoel (2013). We prefer the ETI over the tax elasticity of working hours. The reason is that the ETI not only comprises tax responses in hours worked, but also other margins, such as tax avoidance/evasion, occupational choice, human capital formation, and migration. Based on the few estimates that are available in the literature, we calibrate the income elasticity in earnings supply to be 0.10 on average, see the overviews in Blundell and MaCurdy (1999), Evers et al. (2008), and Meghir and Phillips (2010). The average compensated elasticity of taxable income thus equals 0.35.

Table 2 summarizes the baseline elasticities, and the elasticities we use in a sensitivity analysis where we decrease (low-elasticity scenario) or increase (high-elasticity scenario) the elasticity of the tax base. Table 4 gives the calibrated preference parameters that correspond to these scenarios.

4.3 Government budget constraint

We assume that the government has to collect 9.5% of total labor earnings (i.e., output) to finance government consumption (the benefits of which we ignore in the utility function for simplicity).[19] Government consumption consists of expenditures on public administration, police, justice, defense and infrastructure minus non-tax revenues (from e.g. natural gas) as a percentage of GDP in 2002 (CPB, 2010, Annex 9). With the government revenue requirement set at 9.5% of total output, the government budget balances under the current tax system with a social assistance level of approximately 12,000 euro. This is somewhat higher than the current level of net welfare benefits in 2002 amounting to 9,014 euro for a single-person household. However, we ignored some other forms of social assistance at the local level ('Bijzondere Bijstand'), exemptions from local taxes, and transfers in kind (discounts for arts, public transport, etc.), training, public employment, and labor-market programs, which also act as support schemes for the non-employed.

4.4 Baseline social welfare weights

What are the redistributive preferences of the Dutch government as implied by the current tax-benefit system? By using equations (4) for the employed and (9) for the non-employed, we can calculate the social welfare weights, see figure 3.Dutch social welfare weights only roughly correspond to social welfare weights that are obtained from a standard social welfare function. Indeed, social welfare weights are generally higher for low-income individuals than for high-income individuals.

The Dutch government cares slightly more about non-working poor than the working poor as

[19] Tuomala (2010) uses a similar share.

their social welfare weights are slightly higher. The difference is very small, however. It could be that the non-working poor are considered to be more deserving than the working poor for other reasons than their income. For example, labor handicaps or psychiatric problems migh be more prevalent among the non-working poor. Moreover, we did not control for the fact that benefits for the non-working poor are conditional and subject to monitoring. Zoutman and Jacobs (2014) demonstrate that this implicitly lowers effective marginal tax rates at the bottom. Hence, we could have overestimated the social welfare weights for the non-working poor by ignoring the conditionality of benefits.

However, the Dutch social welfare weights also reveal some important anomalies. First, the social welfare weights do not monotonically decline in income until the mode of the earnings distribution. Specifically, the working poor feature a lower social welfare weight than workers with a median income. This anomaly can be explained by the fact that actual marginal tax rates increase with income towards the mode, but optimal marginal tax rates should always decline with towards the mode – as discussed in the theory section. Hence, there is 'too much' redistribution towards middle-income groups at the expense of low- and high-income groups. Also, higher up the income distribution, close to 60 thousand euro, welfare weights rise again somewhat with income. This is because of the spike in marginal tax rates around 60 thousand euro, which cannot be optimal from an optimal-tax perspective.

Second, for top incomes the welfare weights are negative because the tax rate in the top bracket is set beyond the 'Laffer rate' which maximizes tax revenue at the top. This cannot be an optimal tax policy, see the theory section.

These anomalies are in line with the findings of related studies on other countries. Bourguignon and Spadaro (2012) and Bargain et al. (2011) also find relatively low social welfare weights for the working poor, whereas Bourguignon and Spadaro (2012) also find negative social welfare weights for the top-income earners in France. Below, we calculate the social welfare weights implicit in the tax-benefit proposed system by political parties in the 2002 elections, to determine whether the political parties mitigate or exacerbate these anomalies.

5 Tax-benefit systems proposed by political parties

Having outlined the method, we now take a closer look at the reform proposals for the tax-benefit system of the political parties. We study the proposals for the 2002 elections. We start with a short introduction to the participating political parties in the 2002 elections. Subsequently, we consider the proposals of the four biggest parties (in the 2010 elections) in more detail.

5.1 Political parties in the Netherlands

The Dutch parliament contains 150 seats. Seats are awarded through a system of party-list proportional representation. That is, if a party gets $x\%$ of the votes in the country it is awarded with $1.5x$ seats. Table 5, based on Graafland and Ros (2003), provides an overview of the political parties that received votes in the 2002 elections. For additional perspective we added the most recent, 2012 election outcomes. Political parties are ordered from top to bottom according to their seats in parliament in the period 1998-2002, before the elections in 2002.

Preceding the 2002 elections were two periods of so-called 'purple' governments (Kok-I from 1994-1998, and Kok-II from 1998-2002, named after prime minister Wim Kok). These ruling governments consisted of the 'left' oriented $PvdA$, 'right' oriented VVD, and the smaller liberal democrats $D66$. They had 97 of a total of 150 seats in parliament before the 2002 elections.

However, in a short period of time Pim Fortuyn and his populist party LPF became very popular. Pim Fortuyn himself was murdered in the run up to the 2002 elections, but his party still obtained 26 seats in parliament following the 2002 elections. They formed a coalition together with CDA and VVD, which fell apart less than one year later. The 'traditional' parties CDA, VVD and $D66$ then formed a new coalition.[20] Since the beginning of the century many coalition governments have proven unstable. Since 2012 the ruling coalition consists of VVD and $PvdA$.

In the analysis below we focus on the four largest political parties in the Dutch parliament after the 2012 elections that fit into the 'left-wing' and 'right-wing' taxonomy regarding political preferences for income redistribution. We do not discuss some smaller political parties and the populist party of Pim Fortuyn, because the latter did not submit a tax-benefit plan to CPB in the 2002 elections. We then consider, from 'left' to 'right', the left-wing socialist party SP, the social-democratic party $PvdA$, the christian-democratic party CDA and the conservative-liberal party VVD.

5.2 Description reform proposals

To determine the social welfare weights of Dutch political parties we use the data from the policy packages that these parties submitted for analysis to the CPB Netherland Bureau for Economic Policy Analysis in 2002.[21] Clearly, party proposals during the elections are only not confined exclusively to issues related to income redistribution. Below we outline the policy changes that are most relevant for our analysis: the proposed changes in direct taxes, indirect taxes, corporate taxes and various benefits.

5.2.1 SP

First, consider the proposed changes in direct taxes by the socialist party SP. The SP abolishes health-care premiums. To finance this operation, the tax rate in the first tax bracket is raised by 2.3 percentage points, while the tax rates in the second, third and fourth tax bracket are raised by

[20] Nevertheless, in subsequent years the Balkenende I and II governments, headed by Jan Peter Balkenende of the CDA, converted important elements of the reform proposals of the CDA for the 2002 elections into actual policy. Indeed, there was an increase in income support for low income households, in particular for low income households with children. In line with the CDA election proposal documented in this paper, this led to an increase in marginal tax rates at the lower end of the income distribution, see e.g. (Gielen et al., 2009, Figure 5.5).

[21] CPB (2002b) gives an extensive overview of the proposed policy changes and the resulting effects in Dutch. A brief English summary can be found in CPB (2002a).

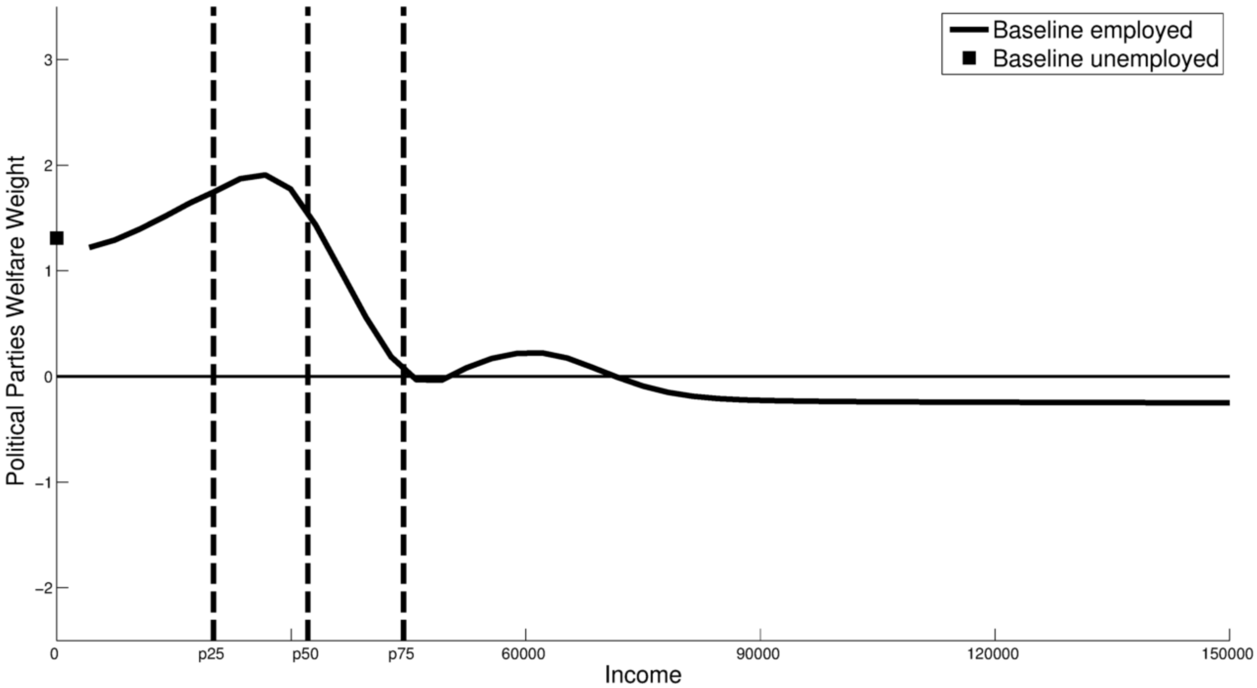

Figure 3: Social Welfare Weights in the Baseline

Table 5: Political Parties in the 2002 National Elections[a]

Name	Acronym	Profile	Seats before 2002 election	Seats after 2002 election	Seats after 2012 election
Partij van de Arbeid	*PvdA*	Social democrat	45	23	38
Volkspartij voor Vrijheid en Democratie	*VVD*	Conservative liberal	38	24	41
Christen Democratisch Appèl	*CDA*	Christian democrat	29	43	13
Democraten 66	*D*66	Social liberal	14	7	20
GroenLinks	*GL*	Environmental progressive	11	10	4
Socialistische Partij	*SP*	Socialist	5	9	15
ChristenUnie	*CU*	Protestant orthodox	5	4	5
Staatkundig Gereformeerde Partij	*SGP*	Protestant orthodox	3	2	3
Lijst Pim Fortuyn	*LPF*	Anti political establishment	-	26	-
Leefbaar Nederland	*LN*	Anti political establishment	-	2	-

[a]Source: Graafland and Ros (2003) and www.tweedekamer.nl.

3 percentage points. This operation increases marginal tax rates, since a part of the initial health-care premiums are lump-sum. The *SP* also introduces a fifth tax bracket of 72% for incomes above 213,000 euro. The *SP* further introduces an additional earned income tax credit (EITC), which is phased in up to the annual minimum wage (15,800 euro in 2002), with a maximum of 1,017 euro, and is phased out between 130% (20,540 euro) and 170% (26,860 euro) of the annual minimum wage. Finally, the *SP* makes the general subsidy per child per year (*Kinderbijslag*) income dependent. Lower incomes receive a higher general subsidy per child per year. At 45,000 euro the subsidy is cut in half. And at 90,000 euro the subsidy is completely abolished. This proposal causes large spikes in marginal tax rates for individuals close to these thresholds (see below).

Next to changes in direct taxes, the *SP* raises environmental levies (2.6 bln euro), which we incorporate in higher indirect taxes. The *SP* also raises corporate income taxes (3.5 bln euro), which is passed on entirely to wages by assuming that capital is perfectly mobile in a small-open economy like the Netherlands. We incorporate this in our analysis a one percentage-point rise in all marginal tax rates at all income levels. Finally, the *SP* wants to raise social-assitance benefits for the unemployed by 5%.

Figure 4 gives the kernel of effective marginal tax rates for the *SP* compared to the baseline marginal tax rates.[22] Higher indirect taxes and corporate taxes raise effective marginal tax rates across the board. The phase-in of the EITC somewhat limits the rise in marginal tax rates at the bottom, but the phase-out range leads to a significant rise in marginal tax rates around 25,000 euro. We clearly see the spike at 90,000 euro where the child subsidy is abolished. This leads to very high marginal tax rates for some households close to the threshold. This effect still shows up in the smoothed kernel estimate.[23] We only plot incomes up to 150,000 euro to focus in detail on what happens to marginal tax rates in the main part of the income distribution. However, this means that the dramatic jump in marginal tax rates beyond 213,000 euro, where the new 72%-bracket kicks in, is not shown.

5.2.2 PvdA

Next, we consider the tax-benefit reforms proposed by the social-democratic party *PvdA*. Regarding direct taxes the *PvdA* integrates public and private health insurance. Health-insurance premiums are raised, but the first tax bracket is reduced with 1.5 percentage points to compensate the lowest incomes. Overall, marginal tax rates hardly change in this operation. The *PvdA* introduces an additional EITC, with a phase-in range between 90 and 100% of the minimum wage, with a maximum of 353 euro (which is much lower than the *SP* above), and a phase-out range between 180 and 240% of the minimum wage. In addition, the *PvdA* phases out the pre-existing EITC between 240 and 400% of the minimum wage. The *PvdA* raises indirect taxes via environmental levies (3.5 bln euro), and lowers corporate taxes (–1.4 bln euro), which we incoporate via a one

[22]Figure 13 in the appendix gives a scatterplot of the individual marginal tax rates before we apply the kernel estimator.

[23]The spike in marginal tax rates at 45 thousand euro is less clear in the smoothed kernel, but is clearly visible in the scatterplot Figure 13 in the appendix.

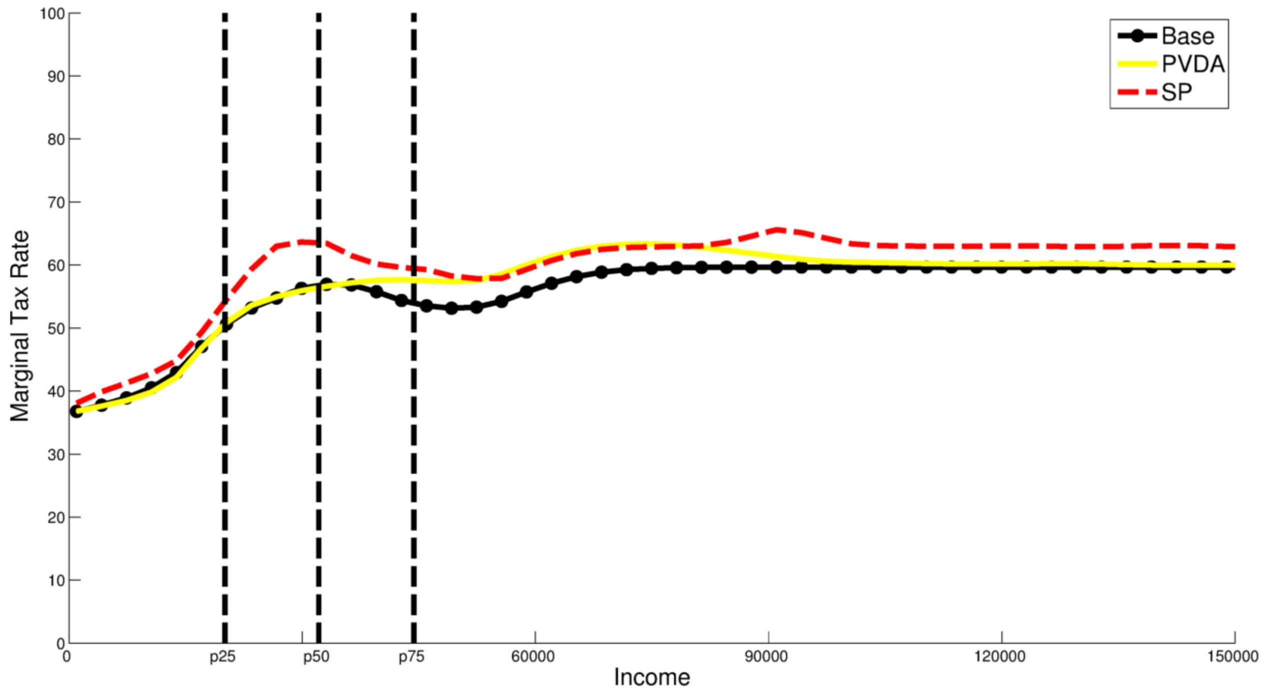

Figure 4: Kernel of Marginal Tax Rates: Social Democratic *PvdA* and (Left-Wing) Socialists *SP*

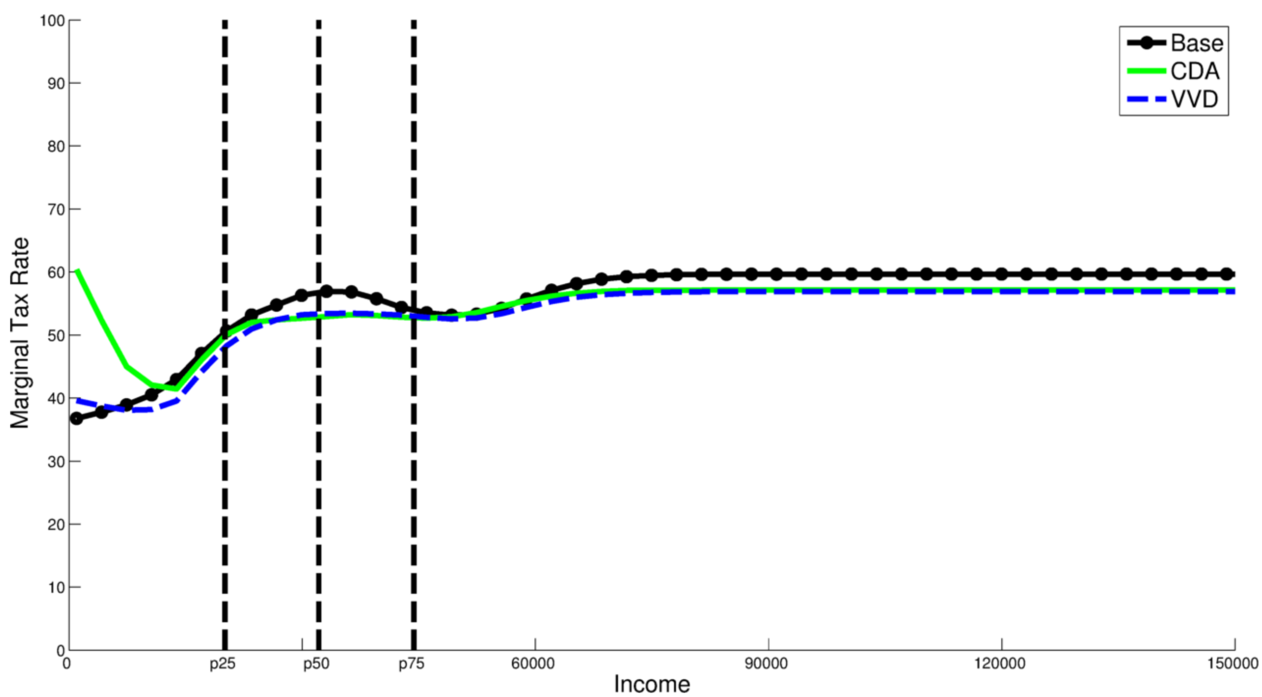

Figure 5: Kernel of Marginal Tax Rates: Christian Democratic *CDA* and Liberal Conservative *VVD*

percentage-point drop in marginal tax rates across the board. The $PvdA$ leaves the level of social assistance benefits basically unchanged.

Figure 4 shows the kernel of resulting effective marginal tax rates for the $PvdA$, along with the baseline marginal tax rates and those proposed by the SP.[24] Higher indirect taxes increase effective marginal tax rates across the board, which is somewhat mitigated by the reduction in corporate tax rates. The phasing-out of the pre-existing EITC and the additional EITC leads to a rise in marginal tax rates over a long income range beyond 30,000 euro.

5.2.3 CDA

The more conservative Christian democratic party CDA also integrates public and private health insurance. The CDA increases the tax rates in the first and second bracket by 1.3 percentage points. Furthermore, they lower the starting point of the fourth tax bracket by 4,440 euro (from 47,745 euro to 43,305 euro), which effectively raises marginal tax rates over this income range. The revenue from increasing the first, second and third tax bracket is used to introduce an income-dependent subsidy for health-care costs and an income-dependent subsidy for children, which are both targeted at low-income families. The CDA reduces the effective top rate by 1.9 percentage points. Furthermore, the CDA introduces a small, additional EITC, with a maximum of 72 euro, which is not phased-out (as opposed to the left-wing parties). The CDA leaves indirect taxes, corporate taxes, and the level of social assistance benefits virtually unchanged.

Figure 5 gives the resulting effective marginal tax rates for the CDA, and compares them with the baseline marginal tax rates and those of the VVD, discussed below.[25] We observe a noticeable increase in marginal tax rates for the lowest incomes, which is the result of the phasing-out of health-care subsidies and subsidies for parents with low incomes. Furthermore, the reduction in the top rate reduces marginal tax rates for top incomes.

5.2.4 VVD

The conservative-liberal party VVD reduces the tax rate in the first tax bracket by 0.4 percentage points. The VVD also introduces an EITC, with a maximum of 232 euro, and they do not phase the EITC out like the CDA. The VVD reduces the top tax rate by 3 percentage points, which is more than the CDA. The VVD slightly increases indirect taxes via environmental levies (0.2 bln euro), reduces corporate taxes (–2.3 bln euro), and leaves the level of social assistance benefits virtually unchanged.

Figure 5 gives the resulting effective marginal tax rates for the VVD, and compares them with the baseline marginal tax rates and those of the CDA.[26] Except for the lowest incomes, the reform

[24] Figure 14 in the appendix gives the scatterplot of individual marginal tax rates before we apply the kernel estimator.

[25] Figure 15 in the appendix gives the scatterplot of the individual marginal tax rates before we apply the kernel estimator.

[26] Figure 16 in the appendix gives a scatterplot of the individual marginal tax rates before we apply the kernel estimator.

package of the liberals reduces marginal tax rates, in particular for individuals with a high income.

6 Social welfare weights political parties

In Section 3 we calibrated the individual utility functions to reproduce the elasticity of the tax base on the extensive and intensive margin. We are now in the position to calculate the social welfare weights of the different political parties by using equations (4) for the employed and (9) for the non-employed, the data on the earnings distribution, the calibrated elasticities and the effective marginal tax rates proposed by the different political parties.

6.1 Left-wing parties

Figure 6 provides the resulting social welfare weights for the left-wing parties SP and $PvdA$. For comparison, we include the social welfare weights of the pre-existing tax-benefit system in 2002. The differences are due to the differences in the proposed marginal tax rates and social assistance benefits.

As expected, the left-wing parties attach a higher welfare weight to the poor than to the rich. However, the SP increases the first anomaly in the baseline by giving more weight to incomes close to middle incomes than to the working poor. This is the result of the phase-out of the EITC just above the minimum wage. Indeed, the SP seems to favor the middle incomes the most.

For high incomes, the proposals of the left-wing parties exacerbate the second anomaly – that social welfare weights are negative at the top – by pushing the top rate further beyond the Laffer rate. As long as the left-wing parties have Paretian social preferences, this finding suggests that they overestimate the thickness of the Pareto tail of the income distribution, or that they underestimate the elasticity of taxable income at the top.

For the $PvdA$ it is also surprising that they attach a higher social weight to the very rich than to the rich, since social welfare weights increase between 80,000 and 110,000 euro. This is the result of the phasing out of the EITC and the higher health-care premiums, both of which lower the welfare weights of the middle- to high-income earners.

For the SP we see a spike in the social welfare weights for incomes close to 90,000 euro. This spike is the result of the spike in the marginal tax rate at that income level as we demonstrate in Figure 17 of the Appendix. It is shown there that the social welfare weights are mainly changing around the spike due to the term $\frac{zT''(z)}{(1-T'(z))^2}$ in equation (4).[27] When a political party suddenly withdraws a subsidy at a given income threshold, it apparently attachs a much higher welfare weight to the individuals just before the threshold than individuals just beyond it. This is inconsistent with with patterns of social welfare weights obtained from a standard social welfare function. Apart from the changes in welfare weights due to spikes in marginal tax rates over some income ranges, the social welfare weights of the left-wing parties still remain very close to the baseline.

[27]The spike from the withdrawal of part of the child subsidy at 45,000 euro is less pronounced, but still visible in Figure 17.

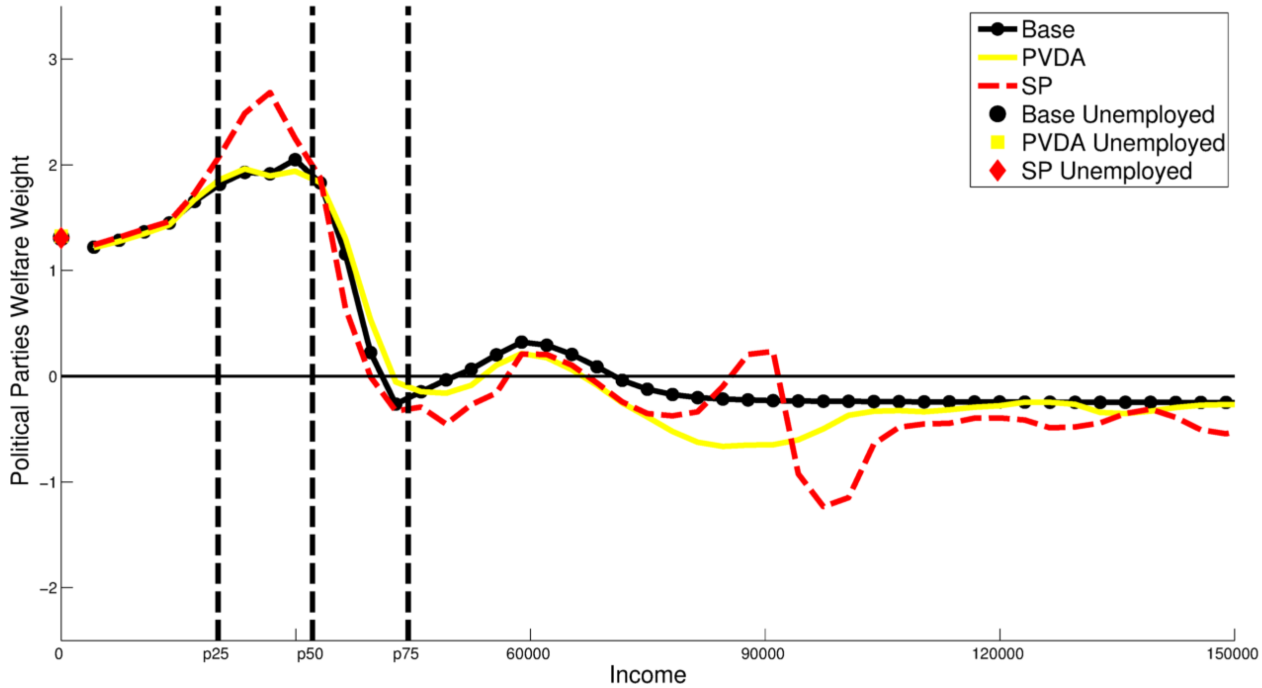

Figure 6: Social Welfare Weights: Social Democratic *PvdA* and (Left-Wing) Socialists *SP*

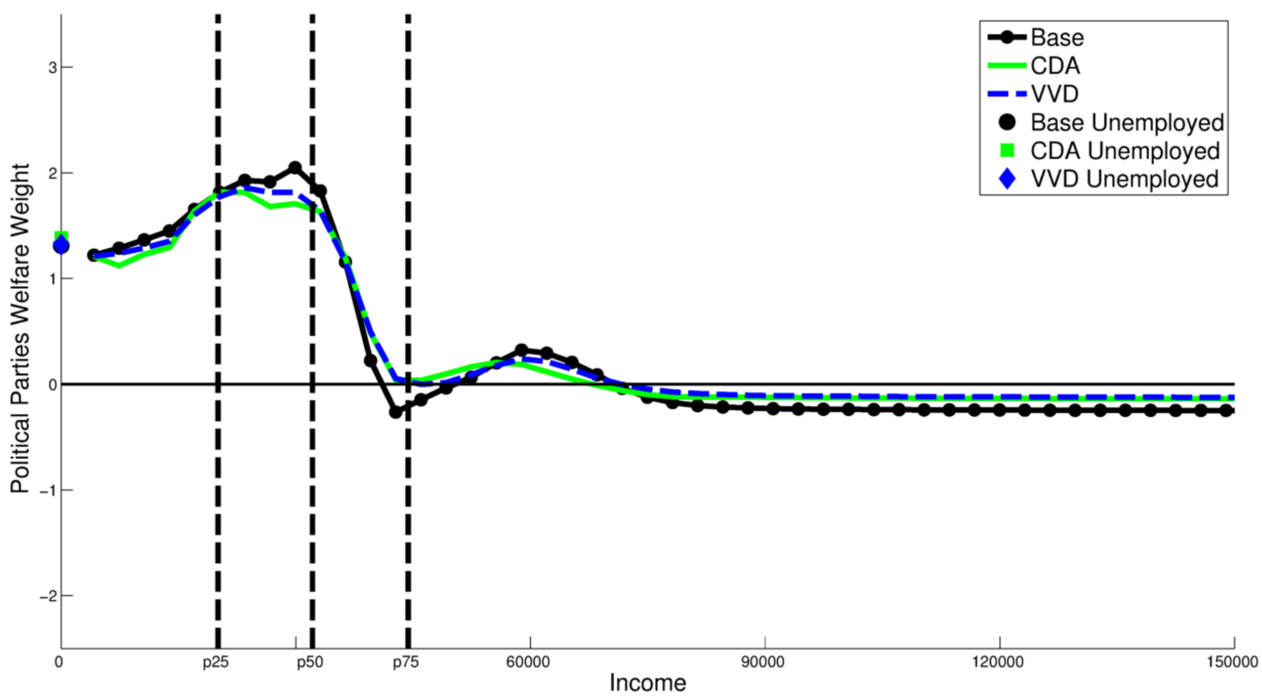

Figure 7: Social Welfare Weights: Christian Democratic *CDA* and Liberal Conservative *VVD*

6.2 Right-wing parties

Figure 7 shows the social welfare weights implied by the proposed tax-benefit systems of the right-wing parties CDA and VVD. The first thing that strikes us is that these parties' redistributive preferences are nearly identical. Indeed, Figure 6 shows that proposed marginal tax rates are quite similar. Only at the lower end we observe that the CDA gives a little bit more weight to the lowest income earners, but otherwise both CDA and VVD have similar patterns of social welfare weights. Hence, the proposals of these parties do not exacerbate, but slightly reduce the existing anomaly of giving a higher more welfare weight to the middle-income groups than to the working poor. Finally, the social welfare weights remain extremely close to the baseline.

We further see that the second anomaly is reduced, though not removed. Although the right-wing parties do lower the top rate, they still set the top rate beyond the Laffer rate. Hence, the social welfare weights remain slightly negative for top-income earners.

Both left- and right-wing political parties maintain the slight difference in social welfare weights between the working poor and the non-working poor.

Perhaps the most striking finding is that the welfare weights remain extremely similar for both the left-wing and the right-wing parties. Indeed, our figures demonstrate a very strong status quo bias, since all political parties, from left to right, hardly deviate from the baseline social welfare weights. This is an important indication that political-economy considerations play a role in the tax-benefit proposals of the political parties.

6.3 Sensitivity with respect to elasticities

How robust are our findings with respect to important parameters of the model? The welfare weights are based on sufficient statistics regarding the marginal tax rates, the income distribution, and the elasticities of taxable income and participation. We think that our results are hardly sensitive to issues in measurement of the marginal tax rates and the income distribution. However, there is some uncertainty regarding the estimates for the elasticity of taxable income and participation. Hence, it is important to know how sensitive the resulting welfare weights are to changes in the assumed elasticities.

Figures 8 and 9 plot the resulting welfare weights for the left-wing parties when the elasticity on both the intensive and extensive margin is 50% lower or 50% higher, respectively. When we assume a lower elasticity of taxable income, the first anomaly is preserved: the social welfare weights still increase from low to middle incomes. However, the second anomaly largely disappears by assuming a much lower elasticity than in the baseline: for the majority of top incomes the social welfare weights are now positive, but very close to zero. The welfare weights can occasionally be negative around some spikes in marginal tax rates. In any case, top marginal tax rates still remain very close to the Laffer rate. Moreover, the working poor now receive a substantially lower welfare weight than the non-working poor.

When we assume a higher elasticity than the base instead, both anomalies in the baseline are exacerbated. The social welfare weights rise much faster in the lower part of the earnings

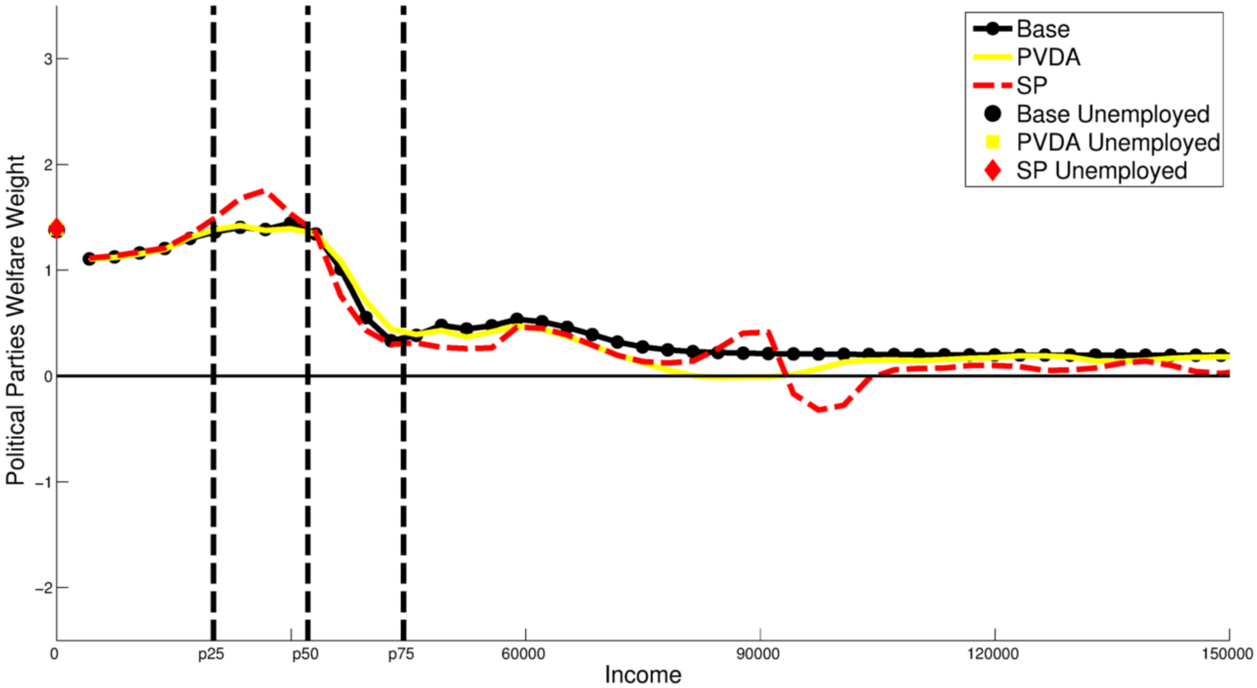

Figure 8: SWW Lower Elasticity of Tax Base: Social Democratic *PvdA* and (Left-Wing) Socialists *SP*

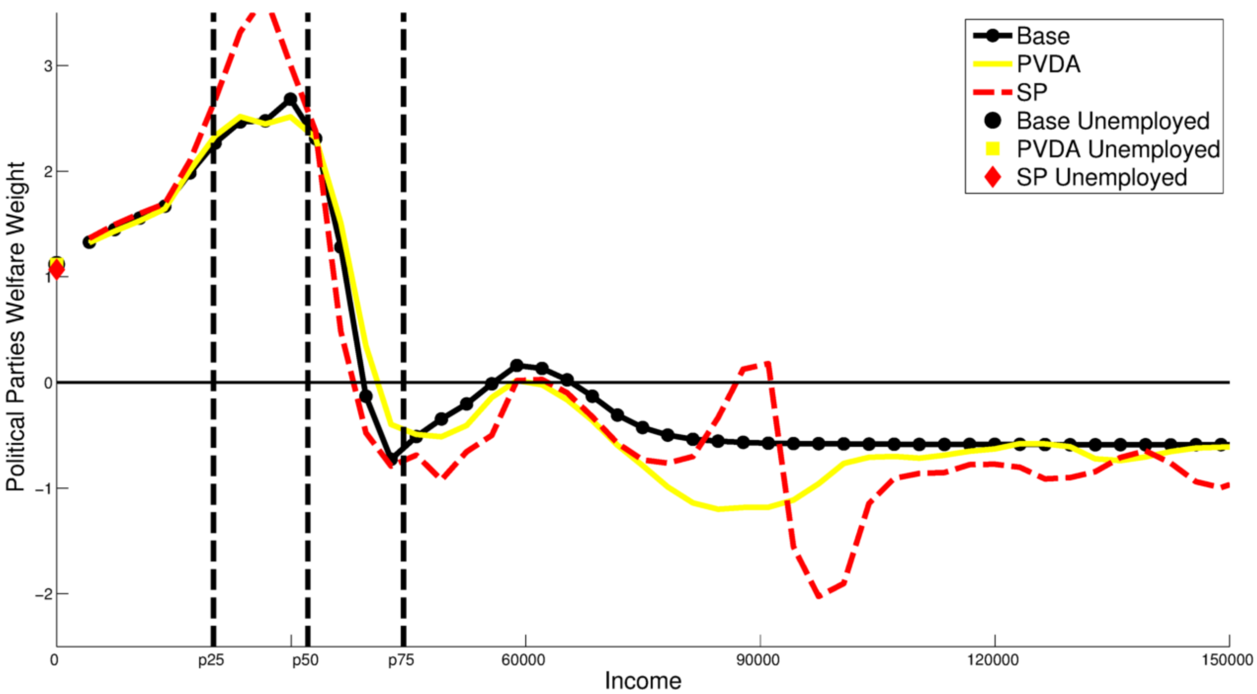

Figure 9: SWW Higher Elast. of Tax Base: Social Democratic *PvdA* and (Left-Wing) Socialists *SP*

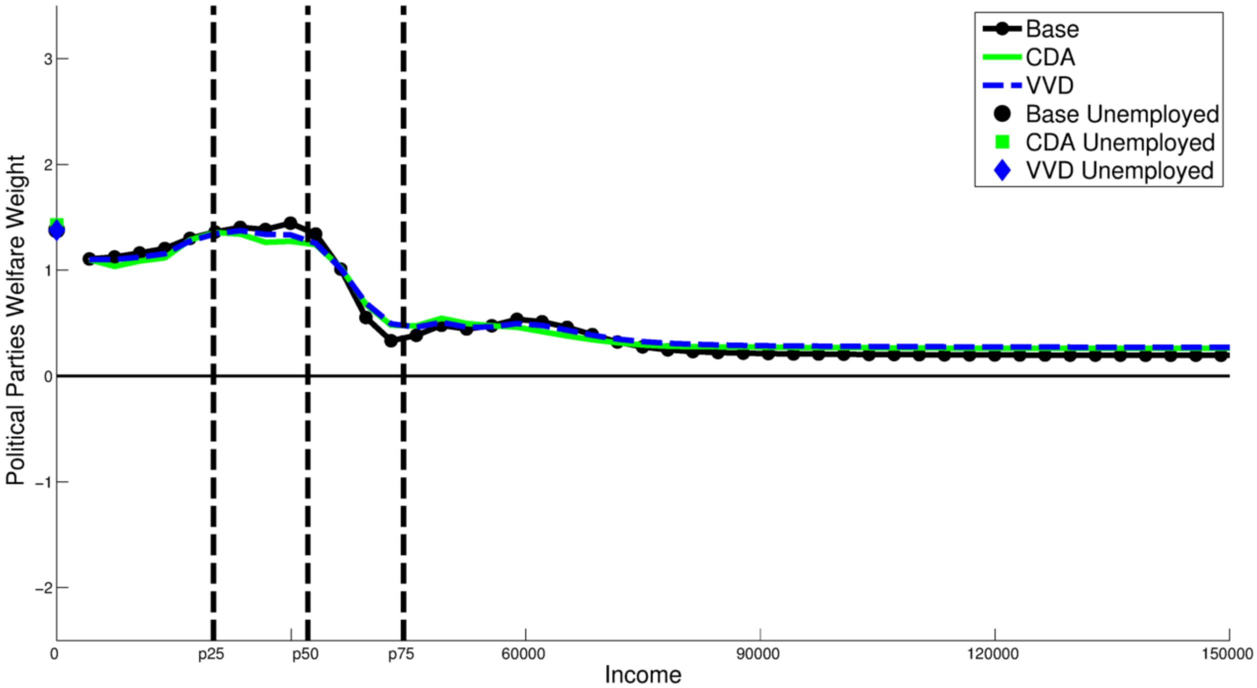

Figure 10: SWW Lower Elast. of Tax Base: Christian Democratic *CDA* and Liberal Conservative *VVD*

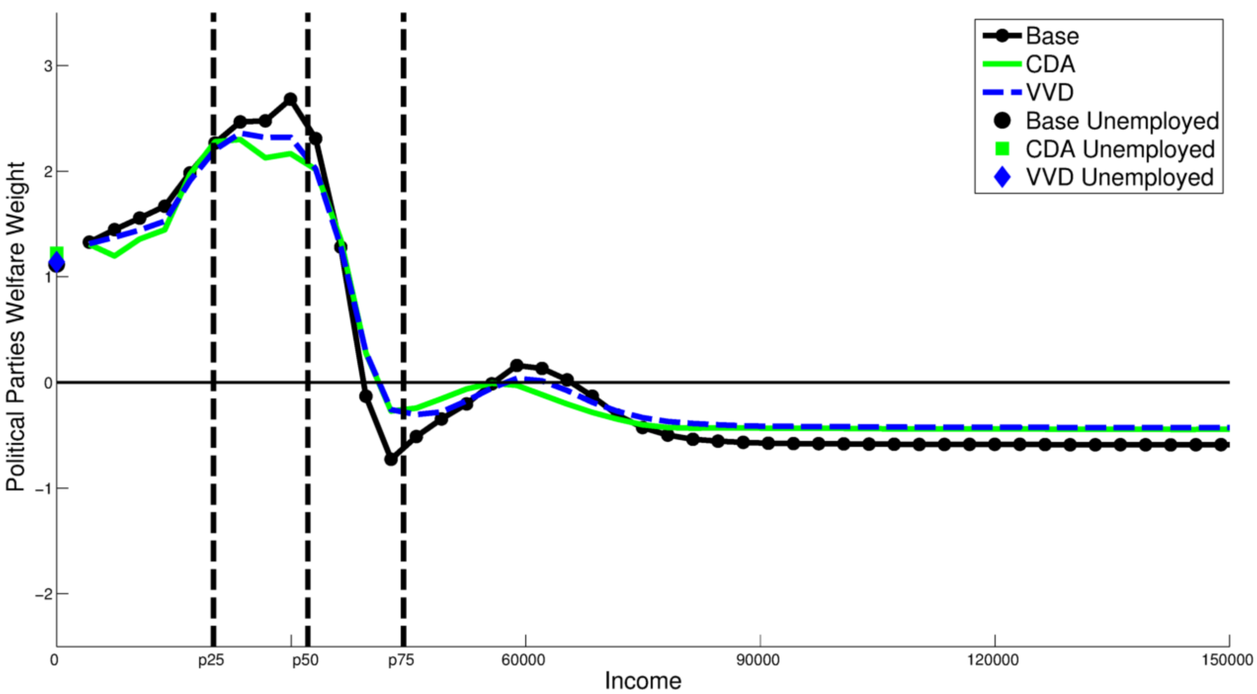

Figure 11: SWW Higher Elast. of Tax Base: Christian Democratic *CDA* and Liberal Conserv. *VVD*

distribution. And weights are much more negative at the top of the earnings distribution. In this case, the non-working poor would receive a lower welfare weight than the non-working poor. The difference, again, is not large. All in all, the results are quantitatively affected by the change in elasticities, but the anomalies we detected largely remain.

Figures 10 and 11 show the social welfare weights for the right-wing parties assuming lower or higher elasticities, respectively. These graphs are completely in line with those of the left-wing parties. The first anomaly is preserved even with low elasticities. Again, the second anomaly disappears with low elasticities: top incomes now obtain a positive welfare weight. However, when the elasticity of the tax base is higher, the anomalies become more pronounced, also for the right-wing parties.

7 Discussion

The analysis above reveals two anomalies in the redistributive preferences of the baseline and of the proposals of the political parties. First, social welfare weights are increasing until modal incomes and they are negative for top-income earners. Furthermore, the social welfare weights look remarkably similar across parties and compared to the baseline. Below we discuss how we could rationalize these findings, both from an economic point of view and from a political point of view.

7.1 Economic interpretations

An explanation for the first anomaly – that the social welfare weights are increasing with income at the bottom of the income distribution – might be that we ignore that many individuals live in multi-person households. For example, secondary earners typically have low income, but high consumption. By ignoring household composition we ignore intra-household redistribution and economies of scale, among other things. There is a small literature looking at family taxation, *e.g.* Boskin and Sheshinski (1983), Apps and Rees (1998), Schroyen (2003), Alesina and Ichino (2011) and Kleven et al. (2009). However, to the best of our knowledge there is not yet an inverse optimal-tax method for families.

Furthermore, we ignore differences in the labor supply elasticity of primary and secondary earners. It is a stylized fact in empirical labor economics that secondary earners have a (much) higher labor-supply elasticity than primary earners (Mastrogiacomo et al., 2013). However, at the lower end of the earnings distribution, one can also find the poor primary income earners, who typically have very low labor-supply elasticities, see also Jongen et al. (2014) and Jongen and Stoel (2013). It is, therefore, unclear whether we might have under or overestimated the elasticity on the intensive margin. By ignoring differences between income and consumption for secondary earners, the higher elasticity of secondary earners might, and the lower elasticity of primary income earners might have biased our results. It is possible, though by no means certain, that we have underestimated the social welfare weights at the bottom of the earnings distribution. More research is needed to verify whether this is indeed the case.

A plausible explanation for the second anomaly – negative social welfare weights at the top, even for the right-wing parties – is that they did not have the right information on the elasticity of taxable income. In the 2002 elections CPB Netherland Bureau for Economic Policy Analysis still conservatively assumed that the elasticity of taxable income at the top was 0.10. In that case, increasing the top marginal tax rate from the baseline value of 52% still generates some tax revenue, while reducing the top tax rate costs revenue. Recent estimates of the elasticity of taxable income in the Netherlands by Jongen and Stoel (2013) show that the elasticity of the tax base is in fact in the order of 0.25 at the top. For this elasticity, and given the high Pareto parameter in the Netherlands, increasing the top tax rate beyond its current level of 52% results in lower tax revenue.

Furthermore, regarding the top rates, we also assumed that utilities are not interdependent. However, it might be that individuals are engaged in 'rat races' (Akerlof, 1976) and 'keeping up with the Joneses' (Layard, 1980). In that case, if one individual decides to supply more labor, negative externalities result as the utilities of other individuals fall due to a loss in relative status or income. Indeed, when there is rivalry in consumption, distortionary income taxes not only have costs, but also have benefits to tame the rat race or to correct status-seeking behavior. The distortions of redistribution are then smaller and optimal tax rates are higher, see also Kanbur et al. (2006).[28]

Admittedly, our simulation model is rather stylized. Still, it remains to be seen how adding more realism to the model will alter the qualitative results of our analysis. In particular, the finding that the social welfare weights of the working poor are lower than the welfare weights of the middle incomes seems to be very robust. The same anomaly can also be found in studies that focus only on singles or single mothers, or use more detailed information on tax base elasticities for subgroups, see also Blundell et al. (2009), Bargain and Keane (2010), and Bargain et al. (2011).

7.2 Political interpretations

There are also other explanations for the anomalies that we detect, which may reflect political constraints in setting the tax-benefit system. Indeed, the anomalies are consistent with a number of political-economics theories that figure prominently in the literature on political economics.

First, the rise in social welfare weights from the working poor to the middle-income groups and the sharp drop in these weights thereafter can be understood by standard political models of income redistribution (Meltzer and Richard, 1981; Roberts, 1977; Romer, 1975). In these models the median voter determines the amount of income redistribution. Consequently, the political system gears income redistribution towards the median voter. The tax-benefit system in the Netherlands is determined by coalition governments. Political-economy models with coalition governments are notoriously hard to solve in theory. However, the underlying mechanism of the basic median voter

[28]However, by the same token, Alesina et al. (2005) argue that for one individual it becomes more attractive to enjoy more leisure if other individuals also enjoy more leisure. This rivalry in leisure thus exacerbates the distortions of income taxation, since not only labor supply choices are distorted, but also a 'leisure multiplier' is put in motion.

model is intuitively appealing. Middle-income groups obtain a high social weight compared to the other groups, because political parties have to attract enough votes from this densily populated group.

Second, the patterns of the social welfare weights – increasing to modal incomes, decreasing thereafter and turning negative for the high-income groups – are consistent with Director's law (Stigler, 1970). According to this theory, the middle-income groups can form a successful, stable political coalition to extract resources from both the low-income and the high-income groups, that cannot align their political interests. This is indeed what we observe in our analysis.

Third, left-wing parties might sacrifice on their ideological preference to redistribute income. Roemer (1998, 1999) shows that the poor – having a larger electorate – may not want to soak the rich through very redistributive tax systems. He develops models of two-dimensional political competition where political parties position themselves on their redistributive preference and some non-economic ideological preference, such as religion. Even left-wing parties may then sacrifice on their redistributive goals if this helps to achieve larger electoral success by attracting more voters on their non-economic, ideological position.

Fourth, post-election considerations could explain the large status-quo bias that we observe in our analysis. Indeed, political parties may not want to deviate too much from the status quo given that they need to form a coalition government with other political parties after elections are held. Coalition agreements are more difficult to achieve if there has been a very polarized political campaign based on sharp ideological differences. See also Persson and Tabellini (2000).

Finally, the status-quo bias and the persistence of various anomalies could also be explained by collective-action problems. The costs of the tax-benefits reforms that remove our anomalies are concentrated in the densely populated middle-income groups, whereas the benefits of such reforms are dispersed among the electorate at large. Vested interests among the middle-income groups could, therefore, be effective in blocking welfare-improving tax-benefit reforms (Olson, 1982).

8 Conclusions

In this paper we have used the inverse optimal-tax method to reveal the redistributive preferences of Dutch political parties in the 2002 elections. We have shown that there a two pre-existing anomalies in the tax-benefit system of 2002. First, social welfare weights are rising from the working poor towards middle-income workers. Second, top-income earners receive a negative social welfare weight.

Both anomalies are preserved by all political parties in their election programs. Indeed, none of the parties features monotonically declining social welfare weights with income. And, all parties attach a negative social welfare weight to top-income earners. The left-wing parties, however, slightly exacerbate these anomalies, and the right-wing parties slightly reduce them. However, our most striking finding is the extreme similarity of the social welfare weights across political parties. We have argued that the anomalies could be explained by economic and political-economy considerations.

Although we put in quite some effort incorporating essential elements – such as the participation decision – into our analysis, our model is admittedly still rather stylized. In future work we hope to tackle some of the difficulties with multi-person households. Furthermore, it would be interesting to study whether the political constraints can be included into the model, and how well they can explain the anomalies detected in this paper.

References

Etisham Ahmad and Nicholas Stern. The theory of reform and Indian indirect taxes. *Journal of Public Economics*, 25:259–298, 1984.

George A. Akerlof. The economics of the caste and of the rat race and other woeful tales. *Quarterly Journal of Economics*, 90(4):599–617, 1976.

Alberto Alesina and A Ichino. L. karabarbounis (2011),gender based taxation and the division of family chores? *American Economic Journal: Economic Policy*, 3(2):1–40, 2011.

Alberto Alesina, Edward Glaeser, and Bruce Sacerdote. Work and leisure in the united states and europe: Why so different? *NBER Macroeconomics Annual*, 20:1–64, 2005.

Patricia Apps and Ray Rees. Taxation and the household. *Journal of Public Economics*, 35(3): 355–369, 1998.

Anthony B. Atkinson and Wiemer Salverda. Top incomes in the netherlands and the united kingdom over the twentieth century. *Journal of the European Economic Association*, 3(4):883–913, 2005.

Anthony B Atkinson, Thomas Piketty, and Emmanuel Saez. Top incomes in the long run of history. *Journal of Economic Literature*, 49(1):3–71, 2011.

Olivier Bargain and Claire Keane. Tax-benefit-revealed redistributive preferences over time: Ireland 1987-2005. *Labour*, 24:141–167, 2010.

Olivier Bargain, Mathias Dolls, Dirk Nuemann, Andreas Peichl, and Sebastian Siegloch. Tax-benefit systems in Europe and the US: between equity and efficiency. IZA Discussion Paper no. 5440, Bonn, January 2011.

Olivier Bargain, Mathias Dolls, Dirk Neumann, Andreas Peichl, and Sebastian Siegloch. Comparing inequality aversion across countries when labor supply responses differ. *International Tax and Public Finance*, 21:845–873, 2013.

Leon Bettendorf, Sijbren Cnossen, and Casper van Ewijk. Btw-verhoging treft hoge en lage inkomens even sterk. *Me Judice*, 25 april, 2012.

Richard Blundell and Thomas MaCurdy. Labor supply: A review of alternative approaches. In Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*, volume 3, chapter 27, pages 1559–1695. Elsevier, Amsterdam, 1999.

Richard Blundell, Mike Brewer, Peter Haan, and Andrew Shephard. Optimal income taxation of lone mothers: an empirical comparison of the UK and Germany. *The Economic Journal*, 119: F101–121, February 2009.

Michael J. Boskin and Eytan Sheshinski. Optimal tax treatment of the family: married couples. *Journal of Public Economics*, 20(3):281–297, 1983.

François Bourguignon and Amedeo Spadaro. Tax–benefit revealed social preferences. *Journal of Economic Inequality*, 10(1):75–108, 2012.

Charles Brendon. Efficiency, equity, and optimal income taxation. EUI MWP; 2013/22, Florence, 2013.

Mike Brewer, Emmanuel Saez, and Andrew Shephard. Means-testing and tax rates on earnings. In James A. Mirrlees, Stuart Adam, Timothy J. Besley, Richard Blundell, Steven Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth D. Myles, and James M. Poterba, editors, *The Mirrlees Review – Dimensions of Tax Design*, chapter 3, pages 202–274. Oxford University Press, Oxford, 2010.

Vidar Christiansen and Eilev S. Jansen. Implicit social preferences in the Norwegian system of social preferences. *Journal of Public Economics*, 10:217–245, 1978.

Aaron Clauset, Cosma R. Shalizi, and Mark E.J. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

CPB. Charting choices 2003-2006: economic effects of eight election platforms. CPB Netherlands Bureau for Economic Policy Analysis, The Hague, 2002a.

CPB. *Keuzes in kaart: 2003-2006.* CPB Netherlands Bureau for Economic Policy Analysis, The Hague, 2002b.

CPB. *Centraal Economisch Plan 2010.* CPB Netherlands Bureau for Economic Policy Analysis, The Hague, 2010.

CPB and PBL. *Charted Choices 2013-2017: Effects of nine election platforms on the economy and the environment.* Number 5 in CPB Book. CPB Netherlands Bureau for Economic Policy Analysis, The Hague, 2012.

Andre Decoster and Erik Schokkaert. Equity and efficiency of a reform of Belgian indirect taxes. *Recherches Economiques de Louvain*, 55(2):155–176, 1989.

Peter Diamond. Income taxation with fixed hours of work. *Journal of Public Economics*, 13(1): 101–110, 1980.

Peter Diamond. Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. *The American Economic Review*, 88(1):83–95, 1998.

Peter A Diamond. A many-person ramsey tax rule. *Journal of Public Economics*, 4(4):335–342, 1975.

Udo Ebert. A reexamination of the optimal nonlinear income tax. *Journal of Public Economics*, 49(1):47–73, 1992.

Michiel Evers, Ruud A. de Mooij, and Daniel J. van Vuuren. The wage elasticity of labour supply: a synthesis of empirical estimates. *De Economist*, 156(1):25–43, 2008.

Martin Feldstein. Tax avoidance and the deadweight loss of the inome tax. *Review of Economics and Statistics*, 81:674–680, 1999.

Miriam Gielen, Joke Goes, Marcel Lever, and Rocus Van Opstal. Ontwikkeling en verdeling van de marginal druk 2001-2011 (development and distribution of marginal tariffs 2001-2011 in dutch). CPB Document 195, The Hague, 2009.

Johan J. Graafland and Arie P. Ros, editors. *Economic assesment of election programmes: does it make sense?* Kluwer Academic Publishers, Dordrecht, 2003.

Bas Jacobs. The Marginal Cost of Public funds is one at the Optimal Tax System. Erasmus University Rotterdam, 2013.

Laurence Jacquet and Etienne Lehmann. Optimal nonlinear income taxation with multidimensional types: The case with heterogeneous behavioral responses. THEMA (THéorie Economique, Modélisation et Applications), Université de Cergy-Pontoise, Paris, 2014.

Laurence Jacquet, Etienne Lehmann, and Bruno Van Der Linden. The optimal marginal tax rates with both extensive and intensive responses. IRES Working Paper 2010-33, Paris, 2010.

Laurence Jacquet, Etienne Lehmann, and Bruno Van Der Linden. Optimal redistributive taxation with both extensive and intensive responses. *Journal of Economic Theory*, 148(5):17701805, 2013.

Egbert L.W. Jongen and Maaike Stoel. Estimating the elasticity of taxable labour income in the netherlands. CPB Background Document, CPB Netherlands Bureau for Economic Research, 2013. May 29.

E.L.W. Jongen, H.-W. De Boer, and P. Dekker. Micsim - a behavioral microsimulation model for the analysis of tax-benefit reform in the netherlands. CPB Background Document, CPB Netherlands Bureau for Economic Research, 2014.

Ravi Kanbur, Matti Tuomala, and Jukka Pirttilae. Non-welfarist optimal taxation and behavioral economics. *Journal of Economic Surveys*, 20(5):1467–6419, 2006.

Henrik J. Kleven, Claus T. Kreiner, and Emmanuel Saez. The optimal income taxation of couples. *Econometrica*, 77(2):537–560, 2009.

Richard Layard. Human satisfactions and public policy. *Economic Journal*, 90(360):737–750, 1980.

N.Gregory Mankiw, Matthew Weinzierl, and Danny Yagan. Optimal taxation in theory and practice. *The Journal of Economic Perspectives*, 23(4):147–174, 2009.

Mauro Mastrogiacomo, Nicole Bosch, Miriam Gielen, and Egbert Jongen. A structural analysis of labour supply elasticities in the Netherlands. CPB Discussion Paper 235, The Hague, March 2013.

Costas Meghir and David Phillips. Labour supply and taxes. In James A. Mirrlees, Stuart Adam, Timothy J. Besley, Richard Blundell, Steven Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth D. Myles, and James M. Poterba, editors, *The Mirrlees Review – Dimensions of Tax Design*, chapter 3, pages 202–274. Oxford University Press, Oxford, 2010.

Allan H. Meltzer and Scott F. Richard. A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–927, 1981.

James A. Mirrlees. An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208, 1971.

James A. Mirrlees. Optimal tax theory:: A synthesis. *Journal of Public Economics*, 6(4):327–358, 1976.

Mancur Olson. *The Rise and Decline of Nations: Economic Growth, Stagflation, And Social Rigidities*. Yale University Press, New Haven and London, 1982.

Torsten Persson and Guido Tabellini. *Political economics: explaining economic policy*. MIT press, 2000.

Kevin W. S. Roberts. Voting over income tax schedules. *Journal of Public Economics*, 8(3):329–340, 1977.

John E. Roemer. Why the poor do not expropriate the rich: an old argument in new garb. *Journal of Public Economics*, 70(3):399 – 424, 1998. ISSN 0047-2727.

John E Roemer. The democratic political economy of progressive income taxation. *Econometrica*, 67(1):1–19, 1999.

Thomas Romer. Individual welfare, majority voting, and the properties of a linear income tax. *Journal of Public Economics*, 4(2):163–185, 1975.

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *The Review of Economic Studies*, 68(1):205–229, 2001.

Emmanuel Saez. Optimal income transfer programs: intensive versus extensive labor supply responses. *Quarterly Journal of Economics*, 117(3):1039–1073, 2002.

Emmanuel Saez, Joel Slemrod, and Seth Giertz. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50:1:350, 2012.

Fred Schroyen. Redistributive taxation and the household: the case of individual filings. *Journal of Public Economics*, 87(11):2527–2547, 2003.

Nicholas Stern. Welfare weights and the elasticity of the marginal valuation of income. In M. Artis and R. Nobay, editors, *Studies in Modern Economic Analysis*. Basil Blackwell Publishers, Oxford, 1977.

George J Stigler. Director's law of public income redistribution. *Journal of Law and Economics*, 13(1):1–10, 1970.

Matti Tuomala. On the optimal income taxation: Some further numerical results. *Journal of Public Economics*, 23:351–366, 1984.

Matti Tuomala. On optimal non-linear income taxation: numerical results revisited. *International Tax and Public Finance*, 17(3):259–270, 2010.

Iván Werning. Pareto efficient income taxation. Mimeo MIT, Cambridge, MA, 2007.

Floris Zoutman, Bas Jacobs, and Egbert Jongen. Optimal redistibutive taxes and redistributive preferences in the Netherlands. mimeo, October 2013.

Floris T. Zoutman and Bas Jacobs. Optimal income redistribution and monitoring of labor effort. mimeo: Erasmus University Rotterdam, 2014.

A Derivation of optimal tax formulae and social welfare weights

A.1 Households

Following Jacquet et al. (2013), individuals differ in their earnings ability n, and utility costs of participation φ.[29] Both characteristics are continuously distributed among individuals. Their joint

[29] Notation throughout this Appendix is taken almost entirely from our companion paper Zoutman et al. (2013). A few exceptions apply. First, in Zoutman et al. (2013) the model is formulated in terms of the unconditional distribution of ability and the conditional distribution of participation costs. Here, the analysis is framed terms of the joint distribution function of ability and participation costs $k(n, \varphi)$. Second, the participation semi-elasticity κ_n, introduced below, is normalized for the marginal utility of workers. Finally, the participation elasticity ε_n^P, introduced below, is defined with respect to a change in the consumption (i.e., net income) of individuals, instead of a change in the participation tax. The first two changes correspond to changes in notation between the working paper version Jacquet et al. (2010) and the published version Jacquet et al. (2013). The latter change allows us to obtain a closer correspondence with the empirical literature on participation elasticities.

density function is given by $k(n, \varphi)$ with support $[\underline{n}, \overline{n}] \times [\underline{\varphi}, \overline{\varphi}]$, where $0 < \underline{n} < \overline{n} \leq \infty$ and $-\infty \leq \underline{\varphi} < \overline{\varphi} \leq \infty$.

Employed individuals with ability n derive utility from consumption c_n, disutility from labor effort l_n, and disutility from participation φ. Utility of an individual with ability n and participation costs φ is assumed to be separable between consumption and labor:

$$U_{n,\varphi} \equiv v(c_n) - h(l_n) - \varphi, \quad v', h', h'' > 0, \quad v'' \leq 0, \quad \forall n, \varphi, \tag{12}$$

$v(c_n)$ is differentiable, increasing, and weakly concave. $h(l_n)$ is differentiable, increasing, and convex. We follow the standard convention in the literature since Mirrlees (1971) that different skill types are perfect substitutes in production. Hence, by normalizing the wage rate for an effective unit of labor supply to unity, we can write gross labor earnings as $z_n = nl_n$.

The tax function $T(z_n)$ is non-linear and continuous and has derivative $T'(z_n) \equiv \mathrm{d}T(z_n)/\mathrm{d}z_n$. All after-tax income is consumed so that the budget-constraint of workers can be written as:

$$c_n = z_n - T(z_n) \quad \forall n. \tag{13}$$

The maximization problem for employed individuals is given by:

$$u_n \equiv \max_{z_n} v(z_n - T(z_n)) - h\left(\frac{z_n}{n}\right), \quad \forall n, \tag{14}$$

where we have substituted the budget constraint and the production function into the utility function. The first-order condition is the same as in the standard Mirrlees (1971) model:

$$\frac{h'(l_n)}{v'(c_n)} = n(1 - T'(z_n)), \quad \forall n. \tag{15}$$

A non-employed individual only derives utility from consuming non-employment benefits: $v(b)$. Consequently, an individual decide to participates in the labor market if his maximized utility when working is larger than the utility obtained from being non-employed: $u_n - \varphi > v(b)$. The employment rate at each level of the income distribution E_z thus depends on the joint distribution $k(n, \varphi)$, the benefit level b, and the tax function $T(z_n)$.

A.2 Incentive compatibility

The allocation is said to be incentive compatible if the following first-order incentive-compatibility constraint holds:

$$\frac{\mathrm{d}u_n}{\mathrm{d}n} = \frac{l_n h'(l_n)}{n}, \quad \forall n. \tag{16}$$

This condition can be derived from totally differentiating utility (12) with respect to ability n and using the first-order condition for labor supply (15).[30]

[30] The incentive-compatibility constraint (16) does not depend on participation costs. Intuitively, a worker with ability n has to incur participation cost φ irrespective of whether the worker self-selects in the consumption-income

The incentive-compatibility constraint (16) is a necessary constraint, but does not ensure that sufficiency conditions for utility maximization (Mirrlees, 1976) are met. This is the case if, in addition, the Spence-Mirrlees and monotonicity constraints are satisfied:

$$\frac{\mathrm{d}\left(\frac{h'(l_n)}{nv'(c_n)}\right)}{\mathrm{d}n} \leq 0, \quad \forall n, \tag{17}$$

$$\frac{\mathrm{d}z_n}{\mathrm{d}n} > 0, \quad \forall n. \tag{18}$$

Intuitively, if the allocation is non-monotonic or does not satisfy the Spence-Mirrlees condition, individuals do not self-select in the consumption-income bundles assigned to them. These assumptions are routinely taken in articles on optimal taxation. See for a discussion e.g Ebert (1992). From the monotonicity condition we can derive that it is never optimal to have higher marginal tax rates than 100%, otherwise the monotonicity condition would be violated, since it implies that $\frac{\mathrm{d}c_n}{\mathrm{d}n} > 0$, see Mirrlees (1976).

In our simulations we will use the first-order approach using (16), assuming that the second-order conditions will be satisfied. After having derived the optimal allocation, we will check ex post whether the sufficiency conditions (17) and (18), are indeed met, which is always the case.

A.3 Government

The objective of the government can be fully summarized by the social welfare weights g_n of employed individuals with ability n and the welfare weight given of the non-employed g_0. The government's redistributive objective is constrained by the fact that it cannot observe earnings ability n and disutility of participation φ. A full description of the informational constraints can be found in Jacquet et al. (2013). The government budget constraint states that total tax revenue from individuals that are employed (i.e., where $\varphi < u_n - v(b)$) equals outlays on transfers b for the non-employed, and exogenous revenue requirement R:

$$\int_{\underline{n}}^{\overline{n}} \int_{\underline{\varphi}}^{u_n - v(b)} T(z_n)k(n,\varphi)\mathrm{d}\varphi\mathrm{d}n = (1-E)b + R, \tag{19}$$

where E is the employment rate.

A.4 Optimal non-linear tax schedule

In Jacquet et al. (2013) the government minimizes total resources subject to obtain a given distribution of utilities and incentive compatibility constraints. Equation (28) of Jacquet et al. (2013)

bundle for type n or decides to mimic a worker of type m to obtain the consumption-income bundle intended for type m.

gives the optimal tax rate for working individuals:

$$\frac{T'(z_n)}{1-T'(z_n)} = \frac{\alpha_n}{\varepsilon_n^c} \frac{\left(\int_n^{\overline{n}} 1 - g_m + \eta_m \left(\frac{T'(z_m)}{1-T'(z_m)}\right) - \kappa_m \left(T(z_m)+b\right)\right) \tilde{k}\,(m)\,\mathrm{d}m}{\tilde{K}(\overline{n}) - \tilde{K}(n)} \frac{\tilde{K}(\overline{n}) - \tilde{K}(n)}{n\tilde{k}(n)}, \quad \forall n. \tag{20}$$

In this expression, $\alpha_n \equiv \frac{\partial z_n}{\partial n}\frac{n}{z_n} > 0$ denotes the elasticity of gross earnings z_n with respect to ability n for workers with ability n. $\varepsilon_n^c \equiv -\frac{\partial z_n}{\partial T'(z_n)}\frac{(1-T'(z_n))}{z_n} > 0$ is the compensated net-of-tax elasticity of labor-supply. $\eta_n = -\left(1-T'(z_n)\right)\frac{\partial z_n}{\partial \rho} \geq 0$ is the income effect on gross earnings of workers with ability n, where ρ is an exogenous change in non-labor income. $\tilde{k}(n) \equiv \int_{\underline{\varphi}}^{u_n - v(b)} k(n,\varphi)\mathrm{d}\varphi$, denotes the density of individuals with ability n that participate in the labor market.[31] $\tilde{K}(n)$ is the fraction of employed workers with ability less than or equal to n in the population, $\tilde{K}(n) \equiv \int_{\underline{n}}^{n} \tilde{k}\,(m)\,\mathrm{d}m$. Finally, $\kappa_n \equiv \frac{k(u_n - v(b), n)}{\tilde{k}(n)} v'\left(c_n\right) > 0$, is the semi-elasticity of employment at ability level n with respect to an exogenous change in non-labor income ρ.

It will be impossible to bring equation (20) to the data, since ability and disutility of participation are unobservable. Hence, we cannot measure $\tilde{K}(n)$. Moreover, while assumption (18) guarantees a one-to-one mapping between the ability of employed workers and their gross earnings, such a mapping does not exist for the non-employed. However, Bayes theorem allows us to decompose $\tilde{K}(n)$ into the distribution of income among the employed, and the employment rate. Since both are observable in the data, we can rewrite equation (20) in terms of sufficient statistics.

In particular, note that by definition we can write $\tilde{K}(n)$ as:

$$\tilde{K}(n) = P\left(n_i \leq n, e_i = 1\right), \quad \forall n. \tag{21}$$

where e_i is a dummy variable, which takes value 1 if individual i is employed, and zero if the individual is not employed. $P\left(n_i \leq n, e_i = 1\right)$ denotes the probability that a random individual i in the population has an ability n_i smaller than or equal to n and is employed. In words the fraction of the population that is employed and has ability smaller than n equals the joint probability that a random individual is employed *and* has ability smaller than or equal to n. By Bayes theorem we can rewrite this probability as:

$$P\left(n_i \leq n, e_i = 1\right) = P\left(n_i \leq n | e_i = 1\right) P\left(e_i = 1\right), \tag{22}$$

$$\tilde{K}(n) = P\left(n_i \leq n | e_i = 1\right) E, \quad \forall n, \tag{23}$$

where $P\left(n_i \leq n | e_i = 1\right)$ denotes the probability that a random individual i in the population has ability smaller than or equal to n_i conditional on being employed, and $P\left(e_i = 1\right)$ is the unconditional probability that a random person i is employed. In the second line, we have used the fact that the unconditional probability that a person in the economy is employed equals the employment rate $E \equiv P\left(e_i = 1\right)$.

[31] Note that Jacquet et al. (2013) also uses $K(u_n - v(b), n)$ to denote the same expression. To avoid confusion we use $h(n)$ throughout this appendix.

By assumption (18) there exists a monotonic mapping between ability n and gross earnings z_n among employed individuals. Therefore, the probability that an individual has ability smaller than n conditional on employment, equals the probability that an individual has an income below z_n conditional on employment:

$$\tilde{K}(n) = P\left(n_i \leq n | e_i = 1\right) E = P\left(z_i \leq z_n | e_i = 1\right) E \equiv \Phi(z_n) E, \quad \forall n, z_n, \tag{24}$$

Hence, we can decompose $\tilde{K}(n)$ entirely into observables $\Phi(z_n)$ and E.

Further, to get an expression for $\tilde{k}(n)$ in terms of observables, take the derivative of expression (24) with respect to ability:

$$\tilde{k}(n) \equiv \frac{\mathrm{d}\tilde{K}(n)}{\mathrm{d}n} = \frac{\mathrm{d}\Phi(z_n)}{\mathrm{d}n} E = \phi(z_n)\frac{\mathrm{d}z_n}{\mathrm{d}n} E = \phi(z_n)\alpha_n \frac{z_n}{n} E. \tag{25}$$

In the second step, we have used the definition of $\tilde{K}(n)$ and α_n, and the fact that the overall employment rate is independent of ability.

We can simplify (20) by substituting expressions (24) and (25) to arrive at:

$$\frac{T'(z_n)}{1-T'(z_n)} = \frac{1}{\varepsilon_n^c} \frac{\left(\int_{z_n}^{\overline{z}} 1 - g_m + \eta_m \left(\frac{T'(z_m)}{1-T'(z_m)}\right) - \kappa_m \left(T(z_m) + b\right)\right) \phi(z_m) \mathrm{d}z_m}{1 - \Phi z_n} \frac{1-\Phi(z_n)}{\phi(z_n) z_n}, \quad \forall z_n, \tag{26}$$

where $\overline{z} \equiv z_{\overline{n}}$, and we have used the fact that $\Phi\left(\overline{z}\right) = 1$. Moreover, for the term inside the integral we have used the fact that (25) can be rewritten as:

$$\tilde{k}(n)\mathrm{d}n = \phi(z_n)\mathrm{d}z_n E, \quad \forall n. \tag{27}$$

Note that by the substitution rule for integration, the bounds of the intergrals also change.

To write equation (26) completely in terms of earnings note again that by equation (18) there exists a monotonic mapping from n to z_n. Hence, there exists (with slight abuse of notation) an x_{z_n} such that $x_{z_n} \equiv x_n$ for all parameters $x = \{g, \eta, \kappa, \varepsilon^c\}$ and for all z_n. In addition, because equation (26) holds for all $z_n \in [\underline{z}, \overline{z}]$ we can rewrite the expression for any particular $z \in [\underline{z}, \overline{z}]$. Finally, to arrive at the expression in the main text, we replace the participation semi-elasticity κ_z by the full participation elasticity ε_z^P with respect to an increase in consumption for workers with income z:

$$\varepsilon_z^P \equiv \kappa_z \left(z - T\left(z\right)\right). \tag{28}$$

This allows us to write expression (26) in the following form which is presented in the main text as (1):

$$\frac{T'(z)}{1-T'(z)} = \frac{1}{\varepsilon_z^c} \frac{\left(\int_{z}^{\overline{z}} 1 - g_{z'} + \eta_{z'} \left(\frac{T'(z')}{1-T'(z')}\right) - \varepsilon_{z'}^P \left(\frac{T(z')+b}{z'-T(z')}\right)\right) \phi(z')\mathrm{d}z'}{1-\Phi(z)} \frac{1-\Phi(z)}{\phi(z) z}, \quad \forall z. \tag{29}$$

A.5 Optimal participation tax

The optimality condition for the participation tax is given by equation (18d) in Jacquet et al. (2013):

$$(g_0 - 1)\int_{\underline{n}}^{\overline{n}}\int_{u_n - v(b)}^{\overline{\varphi}} k(n,\varphi)\mathrm{d}\varphi \mathrm{d}n = v'(b)\int_{\underline{n}}^{\overline{n}} (T(z_n) + b)\, k\,(u_n - v(b), n)\,\mathrm{d}n. \tag{30}$$

Note that this equation is, again, written in terms of the unobservable joint density function $k\,(.)$. To express the equation in terms of the income distribution among workers and the employment rate, note that the double integral on the right-hand side integrates over all individuals that have a participation cost $\varphi > u_n - v(b)$, and hence, decide not to participate in the labor market. Therefore, we can write:

$$\int_{\underline{n}}^{\overline{n}}\int_{u_n - v(b)}^{\overline{\varphi}} k(n,\varphi)\mathrm{d}\varphi \mathrm{d}n = 1 - E. \tag{31}$$

This expression tells us that the number of individuals who decide not to participate in the labor market equals the non-employment rate. Moreover, using the definition of κ_n we can write:

$$k(u_n - v(b), n) = \frac{\kappa_n \tilde{k}(n)}{v'(z_n - T(z_n))}. \tag{32}$$

Hence, we can rewrite equation (30) as:

$$v'(b)\int_{\underline{n}}^{\overline{n}} \frac{(T(z_n) + b)\,\kappa_n}{v'(z_n - T(z_n))}\tilde{k}(n)\mathrm{d}n = (g_0 - 1)(1 - E). \tag{33}$$

Now, use equations (27) and (28) to arrive at expression (2) for the optimal participation tax in the main text which is written in terms of the distribution of income:

$$\int_{\underline{z}}^{\overline{z}} \varepsilon_z^P \left(\frac{T(z) + b}{z - T(z)}\right) \frac{v'(b)}{v'(z - T(z))}\phi(z)\mathrm{d}z E = (g_0 - 1)(1 - E). \tag{34}$$

Expression (30) expresses the welfare weight given to the non-employed to participation tax. Imposing the transversality condition allows us to relate the welfare weight of the non-employed to the welfare weights of workers. Conceptually, the transversality conditions implies an optimality condition that relate the participation tax to the welfare weights given to workers. The differential equation for the co-state variable θ_n is given in equation (18c), in Jacquet et al. (2013):

$$\frac{\mathrm{d}\theta_n}{\mathrm{d}n} = \left[z_n - c_n + b - \frac{(1 - g_n)}{\kappa_n}\right] k\,(u_n - v(b), n)\,\mathrm{d}n, \tag{35}$$

where we have used the assumption that utility is separable between consumption and labor such that the last term in their equation drops out. To introduce the participation tax, as well as the distribution $\tilde{k}(n)$, rewrite this equation using equation (14) and (32):

$$\frac{\mathrm{d}\theta_n}{\mathrm{d}n} = \left[\left(T(z_n) + b\right)\kappa_n - (1 - g_n)\right] \frac{\tilde{k}(n)}{v'(z_n - T(z_n))} \mathrm{d}n. \tag{36}$$

The transversality condition implies that $\theta_{\underline{n}} = \theta_{\overline{n}} = 0$. Integrating expression (36) with respect to n using the fact that $\theta_{\underline{n}} = 0$ yields:

$$\theta_{\overline{n}} = \int_{\underline{n}}^{\overline{n}} \left[\left(T(z_n) + b\right)\kappa_n - (1 - g_n)\right] \frac{h(n)}{v'(z_n - T(z_n))} \mathrm{d}n. \tag{37}$$

Next, impose the transversality condition at the top ($\theta_{\overline{n}} = 0$) to write the optimal participation tax as a function of the welfare weights given to workers:

$$\int_{\underline{n}}^{\overline{n}} \frac{(T(z_n) + b)}{v'(z_n - T(z_n))} \kappa_n h(n) \mathrm{d}n = \int_{\underline{n}}^{\overline{n}} \frac{(1 - g_n)}{v'(z_n - T(z_n))} h(n) \mathrm{d}n. \tag{38}$$

Substituting expression (38) into (30) yields an expression that relates the welfare weight of the non-employed to the welfare weight of workers:

$$\frac{(g_0 - 1)(1 - E)}{v'(b)} = \int_{\underline{n}}^{\overline{n}} \frac{(1 - g_n)}{v'(z_n - T(z_n))} h(n) \mathrm{d}n, \tag{39}$$

Finally, simplify (39) by using equations (27) to arrive at expression (9) in the main text:

$$g_0 = 1 + \frac{E}{1 - E} \int_{\underline{z}}^{\overline{z}} \frac{v'(b)\,(1 - g_z)\,\phi(z)}{v'(z_n - T(z_n))} \mathrm{d}z. \tag{40}$$

A.6 Social welfare weights

To retrieve the welfare weights of a given tax-benefit system we solve equation (1) for the welfare weights g_z at each income level z using Leibniz's rule:

$$g_z = 1 - \eta_z \left(\frac{T'(z')}{1 - T'(z')}\right) - \varepsilon_z^P \left(\frac{T(z) + b}{z - T(z)}\right) + \frac{1}{\phi(z)} \frac{\mathrm{d}\left(\frac{T'(z)}{1-T'(z)} \varepsilon_z^c \phi(z) z\right)}{\mathrm{d}z} \tag{41}$$

We can simplify this expression further by using the product rule on the final term:

$$\begin{aligned} \frac{1}{\phi(z)} \frac{\mathrm{d}\left(\frac{T'(z)}{1-T'(z)} \varepsilon_z^c \phi(z) z\right)}{\mathrm{d}z} &= \varepsilon_z^c z \frac{\mathrm{d}\left(\frac{T'(z)}{1-T'(z)}\right)}{\mathrm{d}z} + \frac{zT'(z)}{1 - T'(z)} \frac{\mathrm{d}\varepsilon_z^c}{\mathrm{d}z} + \frac{\varepsilon_z^c T'(z)}{\phi(z)\,(1 - T'(z))} \frac{\mathrm{d}\left(\phi(z) z\right)}{\mathrm{d}z}, \\ &= \left(1 + \beta_z + \zeta_z\right) \varepsilon_z^c \frac{T'(z)}{1 - T'(z)} + \frac{zT''(z)}{\left(1 - T'(z)\right)^2} \varepsilon_z^c, \end{aligned} \tag{42}$$

where we used the defintions for $\zeta_z \equiv \frac{\partial \varepsilon_z^c}{\partial z} \frac{z}{\varepsilon_z^c}$ and $\beta_z \equiv \frac{\phi'(z) z}{\phi(z)}$ in the second line. Substitute (42) into

(41) to arrive at expression (4) in the main text:

$$g_z = 1 + (1 + \beta_z + \zeta_z)\, \varepsilon_z^c \frac{T'(z)}{1 - T'(z)} + \eta_z \left(\frac{T'(z)}{1 - T'(z)} \right) - \varepsilon_z^P \left(\frac{T(z) + b}{z - T(z)} \right) + \varepsilon_z^c \frac{zT''(z)}{(1 - T'(z))^2}. \qquad (43)$$

Finally, simplify (39) by using equations (27) to arrive at expression (9) in the main text:

$$g_0 = 1 + \frac{E}{1 - E} \int_{\underline{z}}^{\overline{z}} \frac{v'(b)\,(1 - g_z)\,\phi(z)}{v'(z_n - T(z_n))} \mathrm{d}z. \qquad (44)$$

B Scatterplots and kernels of marginal tax rates

Figures 12 to 16 give the scatter plots of effective marginal tax rates in the baseline and for the different political parties, respectively. The figures show that there is considerable variation in marginal tax rates at most income levels, in particular for low incomes.

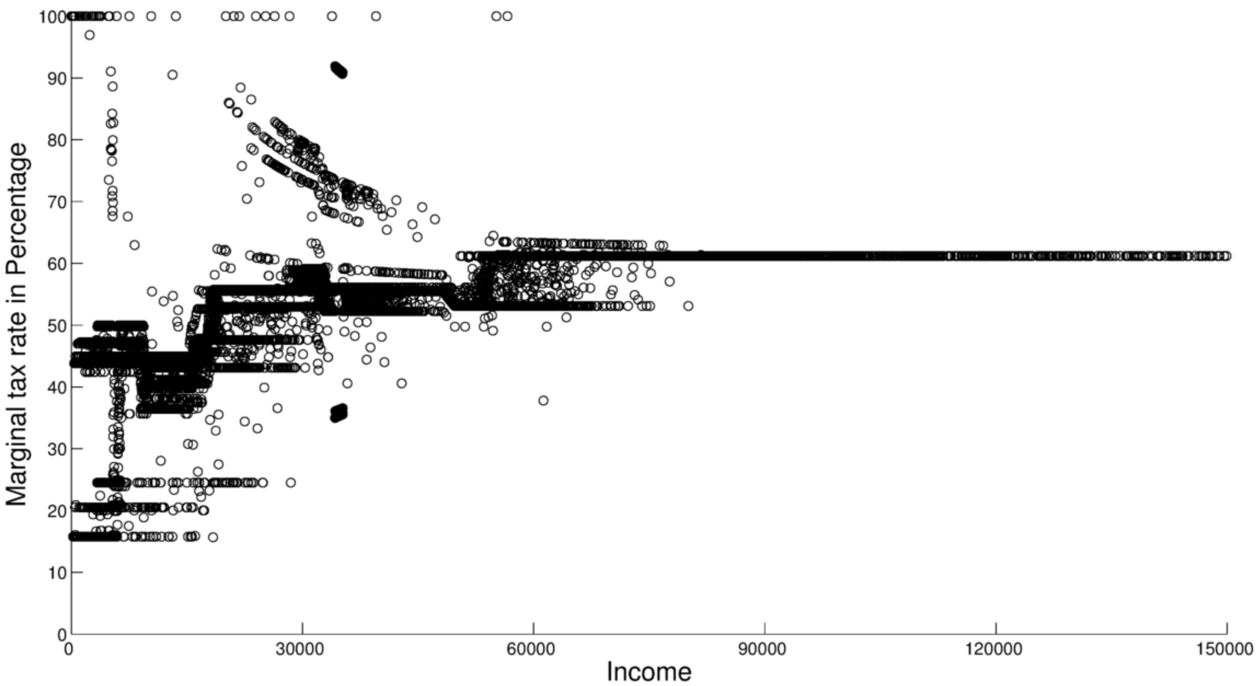

Figure 12: Scatterplot and kernel estimate of effective marginal tax rates: baseline

C Spikes in the welfare weights

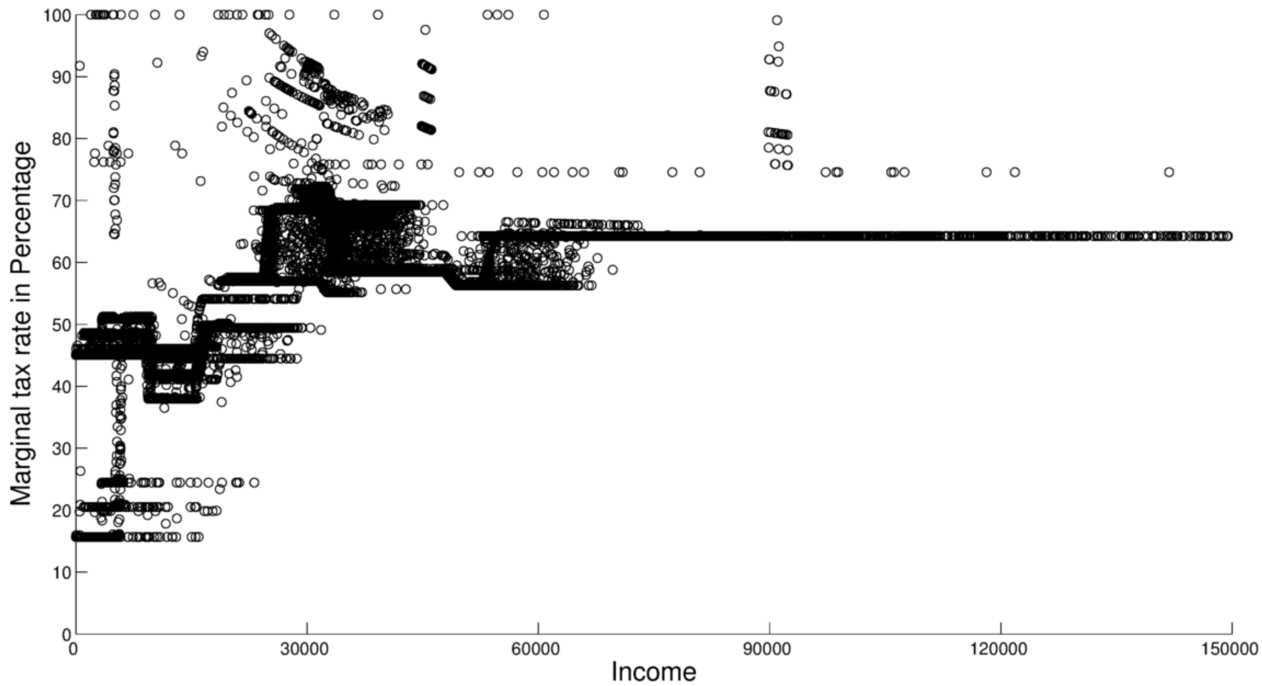

Figure 13: Scatterplot and kernel estimate of effective marginal tax rates: SP

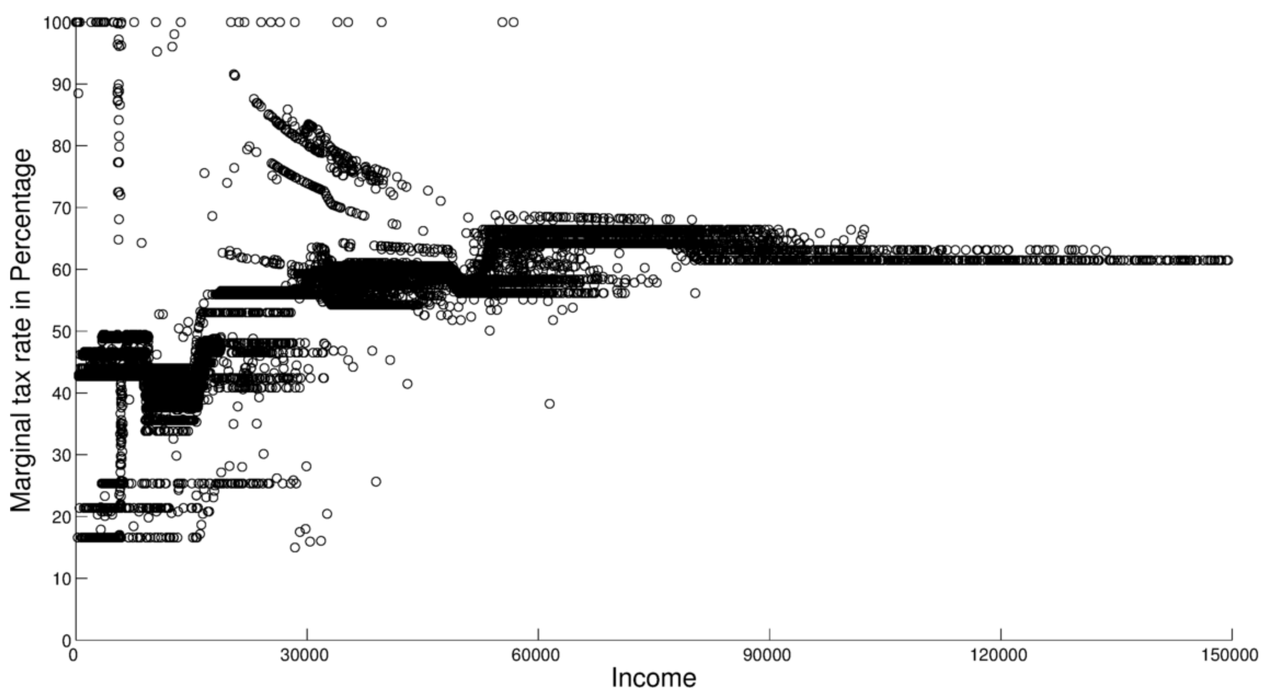

Figure 14: Scatterplot and kernel estimate of effective marginal tax rates: $PvdA$

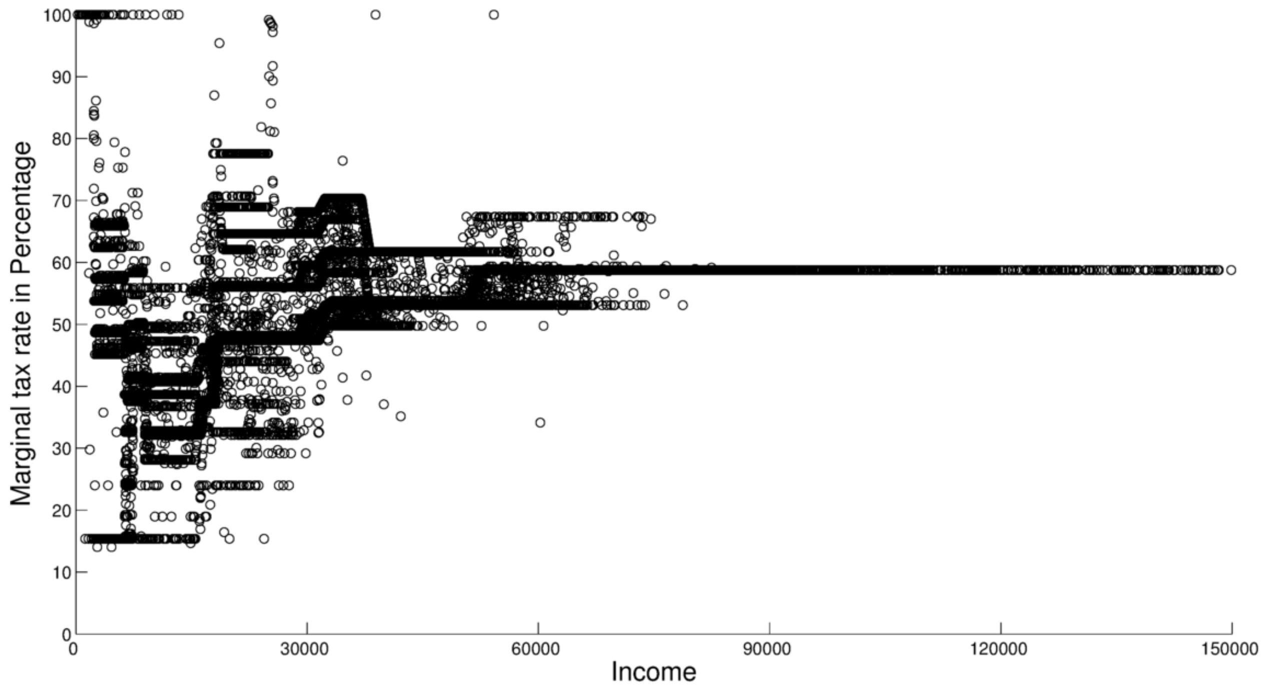

Figure 15: Scatterplot and kernel estimate of effective marginal tax rates: CDA

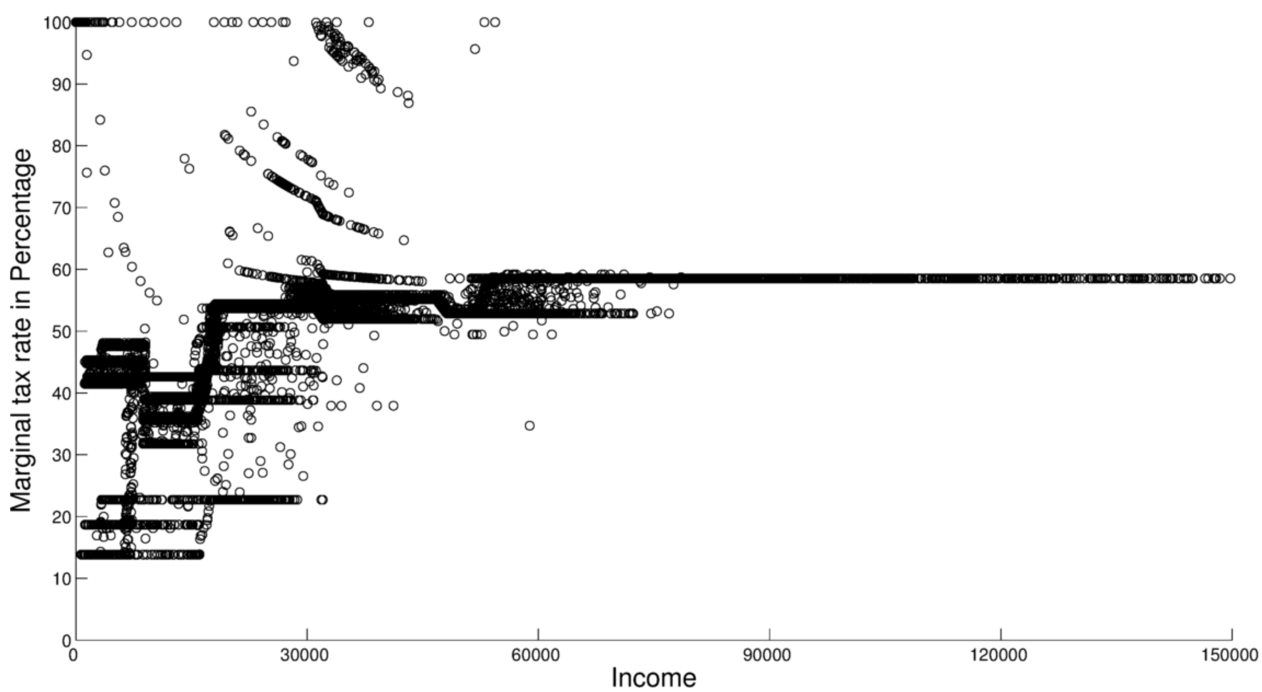

Figure 16: Scatterplot and kernel estimate of effective marginal tax rates: VVD

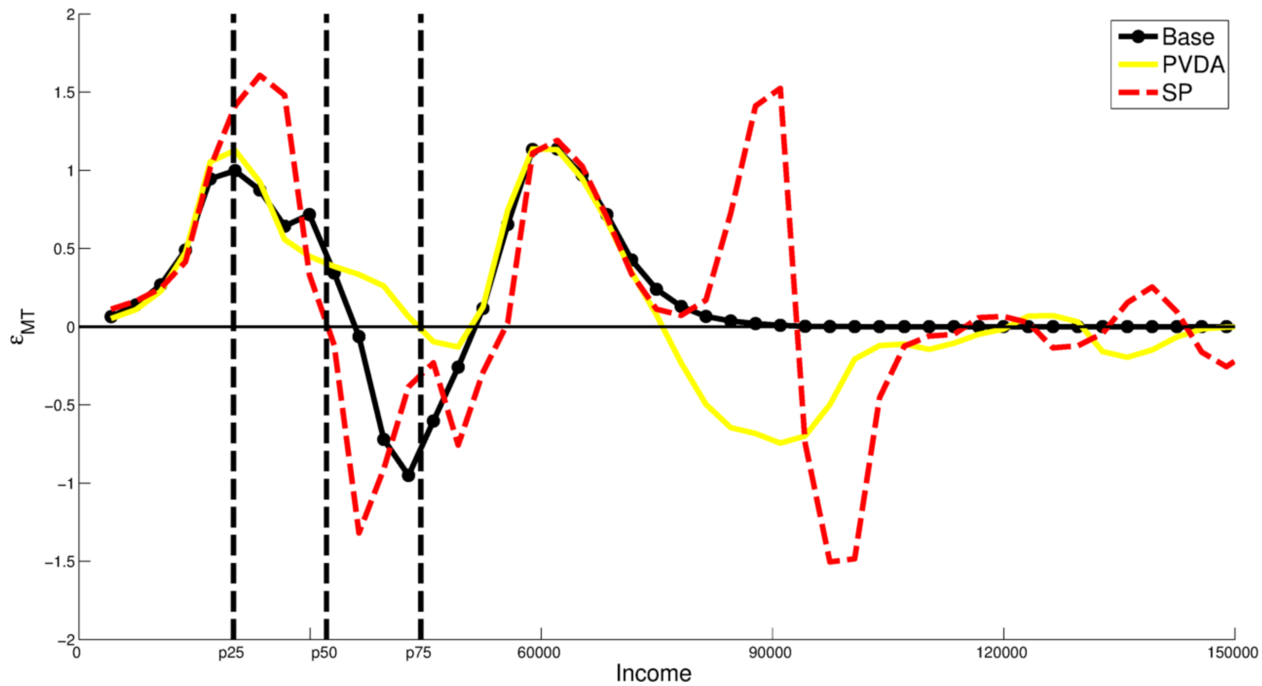

Figure 17: $\frac{zT''(z)}{(1-T'(z))^2}$: Social Democratic *PvdA* and (Left-Wing) Socialists *SP*

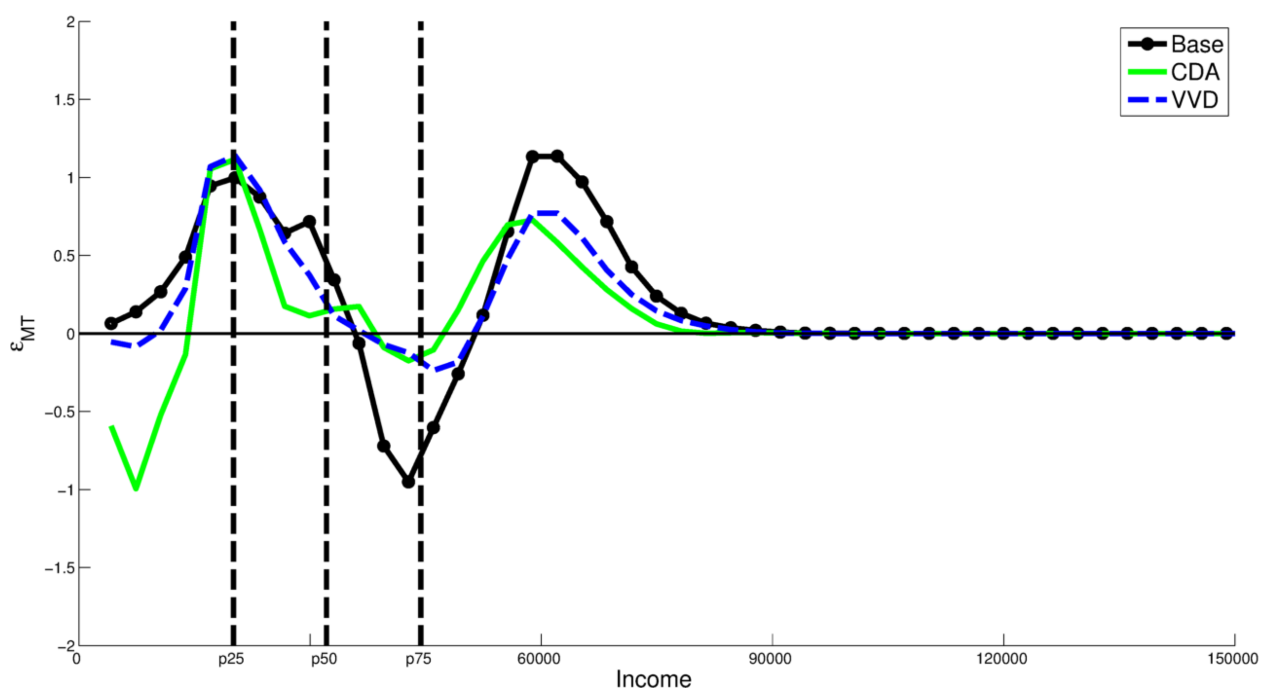

Figure 18: $\frac{zT''(z)}{(1-T'(z))^2}$: Christian Democratic *CDA* and Liberal Conservative *VVD*