Welfare Effects of a Shift of Joint to Individual Taxation in the German Personal Income Tax

André Decoster and Peter Haan*

Received 5 December 2011; in revised form 25 November 2013; accepted 27 November 2013

We empirically derive the welfare effects of a shift from joint to individual taxation of married households in Germany. For the welfare evaluation we estimate the preference heterogeneity and use normative welfare concepts proposed by Fleurbaey (2006) to address the difficulties of comparison between and aggregation of heterogeneous agents. Our results suggest that the normative choice of the welfare metric can be crucial, in particular for the simple question who are the gainers and who are the losers in a reform. Further, we find that it is possible to design a system of individual taxation that would be welfare-improving for different welfare metrics.

Keywords: taxation of couples, welfare measures, labor supply, preference heterogeneity

JEL classification: C 35, D 63, D 78, H 24, H 31

1. Introduction

For couple households, joint taxation with full income splitting (i.e., in equal parts), as implemented in several countries such as France or Germany, imposes much higher marginal tax rates on the secondary earner – in general the wife – than does individual taxation; see, e.g., Apps and Rees (1999). It is not surprising, therefore, that numerous empirical studies provide evidence of strong disincentive effects on the labor supply behavior of secondary earners under joint taxation.[1] This disincentive effect is not compensated by the positive incentives for the first earner, in general the husband; see, among others, LaLumia (2008) for the U.S. and Steiner and Wrohlich (2004) for Germany. This empirical finding is the central reason why in theoretical models of optimal taxation joint taxation of couple households is in general not the optimal tax schedule. In contrast, given the higher elasticities of

* Decoster: KULeuven, Naamsestraat 69, B-3000 Leuven, Belgium (andre.decoster@kuleuven.be); Haan: DIW Berlin and FU Berlin, Mohrenstraße 58, 10117 Berlin, Germany (phaan@diw.de). The authors would like to thank Katharina Wrohlich, the editor, and three anonymous referees. Financial support by the Thyssen Foundation in Project AZ.10.11.2.085 is gratefully acknowledged.

1 In this analysis we do not analyze the effect of joint taxation on the marriage decision.

FinanzArchiv 70 (2014), 599–624 doi: 10.1628/001522114X685500
ISSN 0015-2218

the wife, Alesina et al. (2011) and Boskin and Sheshinski (1983) show that gender-based taxation with lower marginal tax rates for the secondary earner would be optimal, since it is more efficient in terms of labor supply.

There is a large body of literature on joint taxation based on either empirical models of labor supply or theoretical models of optimal taxation. Therefore, it is all the more surprising how little is known about the empirical *welfare* effects of a policy move from joint to individual taxation, either at the individual or at the aggregate level. The reason is that most papers confine their analysis to aggregate labor supply changes or, when taking account of individual heterogeneity, present changes in labor supply and/or changes in disposable income for deciles of the gross wage distribution.[2] However, leisure being an argument in the individual utility function, it is clear that the variable of interest at the individual (or household) level should be the change in *welfare* as measured by household preferences over consumption (or income) and leisure. Also, the ordering of individuals or households from worst to best off should preferably be based on the same welfare concept. It is true that money-metric utilities and, as far as welfare changes are concerned, equivalent or compensating variations derived from these money metrics have been used widely for this purpose; see, e.g., Creedy and Kalb (2005) or Eissa et al. (2008). But a normative analysis in a framework of heterogeneous preferences is well known to pose difficult problems of comparability and aggregation of these individual money metrics or changes therein (see Boadway and Bruce, 1984, chapter 9; Auerbach, 1985). One solution to these problems rests on the willingness to mold the normative analysis in the framework of some representative or reference preference ordering (see, e.g., King, 1983; Aaberge et al., 2004; Aarberge and Colombino, 2013 for empirical applications). Since the major aim of this paper is to provide empirical evidence of the impact on individual welfare of a switch to individual taxation, we could have followed this track.

Yet, the choice of a reference preference ordering has its own drawbacks. First, it obviously raises the question how to choose the representative agent, and how sensitive the welfare evaluation is to the choice.[3] Secondly, the choice of a representative agent in the normative step of the analysis is perpendicular to the important role of preference heterogeneity in the positive part of the analysis. Indeed, it has now become standard to analyze research

2 Beblo et al. (2007) look at the welfare effect of introducing individual taxation; however, they only consider within-household changes, and their method does not allow for any comparison between households or aggregation of individual welfare effects into social welfare.

3 See Aarberge and Colombino (2013) for a proposal to estimate the representative preference ordering on the basis of the heterogeneous preferences in the structural labor supply model.

questions concerning changes in the tax and transfer system by means of static structural models of labor supply in which preference heterogeneity is one of the main building blocks (besides a detailed representation of the budget constraint). Removing this preference heterogeneity in the normative step of the analysis might be considered unfortunate by some. It might, e.g., lead to the conclusion that a tax reform is welfare-improving, though it does not make any individual better off in terms of its own preference ordering.[4]

In this paper we also use a static structural model of labor supply to derive the potential labor supply effects of replacing joint taxation by individual taxation. But, contrary to the standard practice of relying on a representative-agent framework for the normative analysis, when trying to perform distributional, or more general, normative analyses, we make use of a recently developed strand in the literature that proposes individual welfare measures that fully respect preference heterogeneity. We explain the measures developed by Fleurbaey and coauthors (see, e.g., Fleurbaey, 2006 for an overview) and more extensively in Decoster and Haan (2010), and explain in section 5.2 how we have extended the scope of these measures to also be applicable to couple households.

As far as our results are concerned, we find that, in line with the previous studies for Germany, e.g., Steiner and Wrohlich (2004), a shift from joint taxation with full income splitting to a system of strict individual taxation would substantially increase the labor supply of married women and increase the disposable net household income. The labor supply effects are lower if we allow for a joint exemption level for the spouses when introducing individual taxation (quasi-individual taxation). Moreover, we can extend these findings and show the distributional and welfare effects for two different welfare measures that start from very different normative foundations. Our results suggest that the choice of the welfare metric can be crucial, in particular for the simple question who are the gainers and who are the losers. In our specific application we find that for the case of quasi-individual taxation *any* inequality-averse social planner who expresses social welfare by means of a social preference ordering defined over these individual welfare metrics would come to the same conclusion: a policy change that replaces joint taxation with full splitting by individual taxation with a joint exemption level would be welfare-improving for both individual welfare metrics. The results for strict individual taxation are less clear cut.

The paper proceeds as follows. In the next section, we translate the tax systems of joint and individual taxation into the budget constraints faced

4 This is known in the literature as the incompatibility between "respecting preferences" and the "Pareto" principle; see Decoster and Haan (2010) for an extensive discussion of the incompatibility.

by couple households. Section 3 briefly summarizes the micro data and the microsimulation model used to estimate the labor supply model. The estimated model is presented in section 4, together with the simulated labor supply effects of the switch from joint to individual taxation. Section 5 discusses the individual welfare effects of this reform, and also answers the question whether social welfare, defined as some weighted aggregate of individual welfare, is higher under individual than under joint taxation. Section 6 concludes.

2. Joint Taxation with Full Splitting Versus Individual Taxation of Couples

In this section we compare joint taxation with full splitting and individual taxation in terms of a stylized household budget constraint and the working incentives for both spouses. We also provide a brief overview of the current system of joint taxation with full income splitting in Germany and describe the income effects of two hypothetical reforms replacing joint taxation with different scenarios of individual taxation: (i) strict individual taxation with two individual exemption levels, and (ii) quasi-individual taxation with a joint tax exemption level for both spouses.

2.1. Budget Constraint

We use a stylized and simplified setting to discuss the effect of joint and individual taxation on the budget constraint of a household. We assume that labor earnings are the only source of income and that households do not save and are credit-constrained. Furthermore, we abstract from any governmental transfers and social security contributions.

When married couples are taxed jointly with full income splitting, the budget constraint for a household h can be written as

$$C^h_{\text{joint}} = \omega^h_f l^h_f + \omega^h_m l^h_m - 2T\left(\frac{\omega^h_f l^h_f + \omega^h_m l^h_m}{2}\right), \tag{1}$$

where ω^h_m and ω^h_f are the gross wages of husband and wife respectively, and l^h_m and l^h_f are the hours of work of each spouse. $T(\cdot)$ denotes the progressive income tax schedule, including a tax exemption level. The schedule is applied on half the joint taxable income (whence the term "splitting"), and the resulting tax liability is then multiplied again by a factor 2.

Under strict individual taxation with two individual exemption levels the same household h would face the following budget constraint:

$$C^h_{\text{individual}} = \omega^h_f l^h_f + \omega^h_m l^h_m - T(\omega^h_f l^h_f) - T(\omega^h_m l^h_m). \tag{2}$$

The household net income is now determined by the gross labor income of the two spouses minus the sum of the individual tax payments, which are themselves determined independently from each other by means of the same progressive schedule and the same "individual" exemption level. When the tax schedule is progressive, in terms of disposable income a household benefits from joint taxation with full income splitting.[5] This gain, the *splitting advantage*, depends on the inequality between the partners' earnings and the level of overall household taxable income. But more importantly, under individual taxation marginal tax rates of one spouse only depend on that spouses' individual labor income. In contrast, under joint taxation the labor income of the partner matters for the marginal tax rate of the other spouse. In a setting where we define the first earner to be the partner with the higher labor income and the secondary earner the one with the lower labor income, the first earner will always benefit from joint taxation, while the secondary earner faces higher marginal tax rates under joint taxation. Also, these effects increase with the difference between the partners' income.

2.2. Income Taxation of Couples in Germany

In theory, the German income tax is based on the principle of comprehensive income taxation. That is, the sum of a household's income from all sources is taxed at a single rate after several deductions have been applied. The tax schedule is piecewise linear and progressive. Married couples are taxed jointly. As described above, the income tax of a married couple is calculated by applying the tax function to half of the sum of the spouses' incomes; this amount is then doubled to determine the tax amount of the couple. Steiner and Wrohlich (2004) provide a detailed discussion and analysis of the German system of joint taxation with full splitting. For different household types they calculate the splitting advantage, that is, the tax reduction of joint taxation with full splitting relative to individual taxation.

In this analysis we consider two different scenarios of individual taxation. First we study the effect of strict individual taxation, in which both individuals enjoy the individual tax exemption level. In the tax legislation of the year 2008 the individual tax exemption amounted to 7664 euros. In the current German system this strict individual taxation applies to individuals living in unmarried couples. Secondly we focus on a quasi-individual tax system for married couples. In that case, the system consists of individual-specific marginal tax rates, but uses a tax exemption at the household level that is twice the individual exemption of 7664 euros. For a married household this household exemption level of quasi-individual taxation will be beneficial

5 Not surprisingly, therefore, one measure of progressivity of a tax schedule is based on the gain from splitting; see Lambert (2001).

compared to strict individual taxation when under the latter the individual exemption levels, of 7664 euros each, cannot be fully exhausted. In general, however, the pattern of changes (compared to the current joint system) in disposable income induced by quasi-individual taxation is very similar to the one following from strict individual taxation. As we will illustrate in the empirical analysis, the only difference is that the changes are smaller.

The introduction of strict individual taxation without changing the tax schedule would significantly increase the government's tax revenues, since it would absorb the splitting advantages of the current system; see Steiner and Wrohlich (2004). At the same time, the disposable net income of households would decrease. Each household would suffer a loss equal to the splitting advantage it enjoys under the current system of joint taxation with full splitting. The same holds for the budgetary effects of quasi-individual taxation, though at a lower level. For most of the analyses, we decided to focus on non-revenue-neutral tax reforms, since all empirical results strongly depend on the design of redistributing the additional revenue. In this way it is possible to present clear distributional effects of joint versus individual taxation in terms of income or of welfare according to the different welfare measures. In the final section, where we will discuss overall welfare effects, we will change the reform setting and make the two hypothetical individual tax scenarios revenue-neutral by means of a lump sum transfer.

3. Data, Microsimulation, and Descriptive Evidence

Our empirical analysis is based on the German Socio-Economic Panel (SOEP), which is a representative household survey for Germany with detailed information about the socio-economic situation of households (Wagner et al., 2007). Moreover, the data includes information about the employment behavior of all household members. For this analysis we make use of data collected in 2008 and 2009, which yields the information for the tax years 2007 and 2008. We restrict the sample to married households in which both partners are aged between 20 and 60 years and neither partner is self-employed, retired, or in full-time education. This sample restriction is motivated by the behavioral analysis in which we focus on the labor supply of both partners.

The SOEP is used as the input data set for the microsimulation model STSM (Steiner et al., 2008), which describes in detail the German tax and transfer system. Given observed gross earnings of households, it is possible to determine the tax liabilities of the current tax system with joint taxation and the resulting net household income for all observed households. Moreover it is possible to derive the tax payments and the net income as-

Table 1
Changes in Net Household Income from a Shift to Individual Taxation

	Income in Euro Joint taxation	Change in income Quasi Individual taxation	Change in income Individual taxation
	Quintiles of gross household earnings		
I	1997	−12	−74
II	2621	−36	−155
III	3185	−50	−153
IV	3733	−70	−164
V	5469	−110	−201
	Quintiles of inequality of earnings between spouses		
I	2898	0	−2
II	3016	−1	−23
III	3051	−19	−106
IV	3228	−61	−216
V	4811	−198	−400
	West versus East Germany		
West Germany	3504	−65	−171
East Germany	3038	−21	−73

Note: Own calculations based on SOEP and STSM. Changes in income are calculated as the difference between pre- and post-reform net household income. All income information is per month, household income is not equivalised. Inequality of earnings between spouses is calculated as the absolute distance of the wife's labour earning to the average labour earnings of the household.

suming hypothetical scenarios of taxation, viz., strict and quasi-individual taxation.

Table 1 presents disposable net household income for different subgroups under joint taxation and the resulting income differences relative to strict and quasi-individual taxation. In this descriptive table we do not consider any behavioral reaction of the households.

We compare the income differences by the level of gross household earnings, by the inequality of earnings between the two spouses, and by region. Inequality of earnings between spouses is calculated as the absolute difference of the wife's labor earning from the average labor earnings of the household.

The distributional analysis in table 1 reveals that introducing individual taxation affects different household types differently. Overall the distribu-

tional effects of introducing strict individual taxation and quasi-individual taxation are similar, yet the effects of the latter reform are smaller.

Consistent with the fact that the splitting gain is increasing in earnings, we find that households in the lowest quintile of gross household earnings face the lowest income losses when individual taxation is introduced. In particular, in the scenario of quasi-individual taxation the income effects of about 12 euros per month are relatively low. The income effects become more substantial in the higher quintiles. In the top quintile households lose on average slightly more than 200 euros per month in the case of strict individual taxation and 110 in the case of quasi-individual taxation.

Also according to expectations, the second panel in the table shows the highest losses for households with high earnings inequality between the spouses, and only moderate effects for households with the least unequal intrafamily distribution of earnings. The latter households, on average, hardly benefited from the splitting advantage under joint taxation. The households with more unequal earnings between spouses, on the other hand, lose the splitting advantage, which in extreme cases can be substantial. As is well documented in the literature (Steiner and Wrohlich, 2004), households in west Germany would markedly lose from the introduction of individual taxation, whereas the effect on households in the east would be smaller. This is explained both by the lower household incomes of east Germans and by the greater equality of earnings between the spouses.

4. The Labor Supply Model

We estimate the labor supply effects of introducing individual taxation by means of a static structural discrete-choice model of labor supply, similar to Aaberge et al. (1995) or van Soest (1995). The model is structural in that it starts from a specification of a utility function of the household that depends on household consumption and the leisure time of the husband and the wife.[6] The model is a discrete-choice model and it reduces the choices of the individual (in this case the number of hours worked) to a finite number of discrete alternatives. Identification of the model is based on the assumption that households choose the alternative that yields the highest utility. The main advantage of this discrete specification over the continuous framework is the possibility to allow for nonlinearities in the budget set and to cope with the endogeneity of net household income in a relatively straightforward way.

6 Given the available information in the SOEP data, it is not possible to differentiate empirically between home production and leisure time. Based on time use data, van Soest and Stancanelli (2010) are able to distinguish between market and nonmarket work.

4.1. Description of Participation and Hours of Work in a Discrete Setting

In the discrete-choice model the utility level of household i at a finite number of discretely chosen levels of labor supply is specified. We index the discrete points by means of the subscript $j = 1, ..., J$. For the female labor supply, we define six discrete working alternatives: nonparticipation, three part-time alternatives, full-time work, and overtime; since very few men work part-time, we specify only four discrete alternatives. Thus for each household we consider $J = 4 \times 6 = 24$ combinations of working alternatives.[7]

To derive net household income according to the tax legislation in Germany at each discrete alternative of working hours, we use the microsimulation model STSM, mentioned above.[8] More precisely, for each discrete hours point we calculate gross household earnings as the sum of the employment-state-specific earnings of the two spouses. The gross earnings of each spouse are simply the state-specific hours multiplied by the expected market wage. For working individuals we take the observed wage information as their market wage, while for the nonworking we impute their expected market wage using an estimated wage equation with selection correction (see Appendix for a detailed discussion of the wage estimation and estimation results). The information on gross earnings is the key input for the microsimulation model, which describes, in detail, all relevant transfer programs, social security contributions, and income taxation and which delivers the state-specific net household income. In addition we take explicit account of child-care costs.

The income tax is computed by applying the progressive income tax function to the taxable income of the household. Taxable income includes, in addition to gross labor earnings, income from capital and renting. Income tax and employee's social security contributions are deducted from gross income, and social transfers that depend on the employment state are added to derive net household income. Social transfers include child benefits, child-rearing benefits, unemployment compensation, housing benefits, and social assistance.

As is well documented in the literature, the presence of young children is an important determinant of female labor supply behavior (Mroz, 1987). This is particularly the case in Germany, where child-care costs are very high. This is due to the limited number of subsidized child-care facilities (Wrohlich, 2006). To allow for the effect of child-care cost on labor supply

7 The median of the empirical distribution of working hours in the following intervals define the discrete points: for women, 0, [0, 12], [13, 20], [21, 34], [35, 40], > 40; for men, 0, [0, 20], [21, 40], > 40. The estimation results are robust to changes in the approximation of the distribution of working hours.

8 For a similar procedure and a more detailed discussion, see, e.g., Steiner and Wrohlich (2004) or Steiner et al. (2008).

behavior we follow Wrohlich (2006) and estimate expected child-care costs for part-time and full-time care, separately for children younger than three years and between three and six years.[9] Households with children in these age groups need to pay for child care if both parents work full time (full-time child care), or if both work at least part time (part-time care).

4.2. The Utility Function

The state-specific level of utility of household i, denoted v_{ij}, in the discrete states $j = 1, ..., J$ described in the previous subsection, consists of a deterministic and a stochastic part:

$$v_{ij} = u(c_{ij}, 1 - lf_{ij}, 1 - lm_{ij}; \mathbf{z}_i, \alpha_i) + \epsilon_{ij}\,, \tag{3}$$

where c_{ij} denotes the state-specific net household income c_{ij} at choice j, and lf_{ij} and lm_{ij} are measures of the female and the male labor supply in household i at choice j. The leisure time at each hours point is the time endowment $T = 80$ minus the working time of the spouse. In (3) we normalize the time endowment at 1, so that the leisure time of the spouses can be simply expressed as $1 - lf_{ij}$ and $1 - lm_{ij}$.

The deterministic part of the utility function is represented by $u(c_{ij}, 1 - lf_{ij}, 1 - lm_{ij}; \mathbf{z}_i, \alpha_i)$, and ϵ_{ij} is a stochastic random error term that varies independently between the households and the discrete points. Observed preference heterogeneity is captured by the vector $\mathbf{z}_i$; unobserved heterogeneity is captured by α_i.

For the main specification, we choose a separable functional form for the deterministic part of the utility function. The main motivation for this functional form is that it is general enough to provide an accurate description of the behavior of individuals and that it is transparent so that calculation of the welfare metrics is possible without complications. More flexible and complex functional forms might lead to multiple solutions or equilibria when calculating the welfare metrics. To check if our functional form is general enough, we present, in addition to our main specification, estimation results for different functional form specifications, and we calculate the related labor supply responses. Labor supply elasticities hardly differ between the different specifications, and therefore we opt for the following functional form:

$$\begin{aligned} u(c_{ij}, 1 - lf_{ij}, 1 - lm_{ij}; \mathbf{z}_i, \alpha_i) = \beta_c \ln c_{ij}(\alpha_i) + \beta_{Lf} \ln Lf_{ij}(\mathbf{zf}_i) \\ + \beta_{Lm} \ln Lm_{ij}(\mathbf{zm}_i)\,, \end{aligned} \tag{4}$$

9 The expected costs take account of the fact that subsidized child care is heavily rationed, particularly for children younger than three years.

where $\mathbf{zf}_i$ and $\mathbf{zm}_i$ are subvectors of $\mathbf{z}_i$. Preference heterogeneity is introduced by means of taste shifters in the following form:

$$\beta_c(\alpha_i) = \beta_{c0} + \beta_{c1}\alpha_i\,,$$
$$\beta_{Lf}(\mathbf{z}_i) = \beta_{Lf0} + \beta'_{Lf1}\mathbf{zf}_i\,,$$
$$\beta_{Lm}(\mathbf{z}_i) = \beta_{Lm0} + \beta'_{Lm1}\mathbf{zm}_i\,,$$

and the vectors $\mathbf{zf}_i$ and $\mathbf{zm}_i$ include individual measures of age and education, as well as household variables such as the number and age of children and a regional dummy indicating if the household resides in west Germany. Following Heckmann and Singer (1984), we introduce unobserved heterogeneity in a nonparametric way with discrete mass points. In the main specification we introduce unobserved heterogeneity in the preferences for consumption, and we describe α_i with two mass points. For a similar specification, see Haan (2006). More complex specifications, with either a higher number of discrete mass points or additional unobserved terms in preferences for leisure time of the spouses, either did not affect the results or had convergence problems.

The estimation procedure is based on the assumption that the error terms ϵ_{ij} are i.i.d. and follow an extreme-value distribution. This gives an expression for the probability of each discrete working alternative, and thus the model can be estimated by maximum likelihood where the household specific contribution is determined by the working alternative observed in the data.

4.3. Estimated Preferences for Income and Leisure

Table 2 presents the estimated parameters of the different specifications of the utility function. In addition to the results from the preferred specification (Model III) in columns 5 and 6, we also show results of a model (i) without unobserved heterogeneity (in columns 1 and 2) and (ii) without unobserved heterogeneity but with nonseparabilities in the leisure terms of the spouses (in columns 3 and 4). Overall the estimation results of the different specifications are quite similar. The coefficient on income is positive and significant, which implies that the households respond to financial incentives. As discussed above, in Model III we introduce unobserved heterogeneity in the preferences for consumption. We can identify two different groups for which the taste for consumption is significantly different.

Turning to the estimated coefficients on the leisure terms, we find substantial preference heterogeneity, particularly for women.[10] In line with previous

10 See Decoster and Haan (2010) for a more detailed discussion of preference heterogeneity using this empirical specification.

studies, we find that the taste for female leisure increases with the presence of children, particularly for children younger than 3 years (the coefficient of female leisure, equal to 1.46, increases by 4.00 for each additional child younger than three). The preference for leisure is negatively correlated with the level of education: the coefficient for higher education is −1.07 for women. This implies that *ceteris paribus* women with low and medium education have a higher preference for leisure than women with the highest educational attainment. Finally, we find important preference differences between women in east and west Germany. Women in west Germany have a significantly higher leisure preference than women in east Germany. The different pattern in female employment behavior has often been analyzed and is mainly explained by the different history and socialization of the two parts of Germany before the reunification. For male leisure time, the heterogeneity is considerably smaller or not significant.

In addition to the coefficients, we display at the bottom of the table the labor supply elasticities derived from the different specifications. The elasticities measure the relative change in the average participation rate and the average number of working hours induced by a 1% increase in gross wages (see below for a more detailed discussion about the calculation of labor supply responses). The elasticities provide important information with a clear economic interpretation, namely, how strongly individuals respond to changes in financial incentives. Importantly, the labor supply elasticities are very similar across the different specifications. Thus without loss of generality we can use (4) as our main specification. In the following we will only focus on the results obtained from this model.

Moreover, the derived labor supply elasticities are very similar to elasticities derived in previous studies for Germany, e.g., Steiner and Wrohlich (2004).

5. Effects of the Tax Reform

5.1. Labor Supply Effects

Based on our preferred specification of the estimated structural model, we simulate the labor supply responses of men and women when replacing joint with strict and quasi-individual taxation. We calculate the labor supply responses of the spouses with regard to participation effects and the effect on expected working hours. Calculating the participation effect in the discrete-choice model is straightforward. This is directly related to the individual probability of not working. In order to calculate the expected working hours, we use the individual probabilities at each discrete point and multiply these with the working hours that are assigned to this point. This measure thus

Table 2
Labour Supply Estimation

	Model I		Model II		Model III	
	Coef.	Std.	Coef.	Std.	Coef.	Std.
Income	3.74	0.18	3.57	0.19	9.09	1.02
Female leisure × Age of wife/100	5.79	1.86	6.37	1.94	6.69	2.07
Female leisure time × Age of husband/100	−3.05	1.88	−3.07	1.96	−3.58	2.10
Female leisure time × Child03	4.00	0.34	4.38	0.35	4.19	0.37
Female leisure time × Child36	1.02	0.26	1.11	0.27	0.92	0.28
Female leisure time × East Germany	−1.87	0.15	−2.02	0.16	−2.00	0.17
Female leisure time × high education	−1.07	0.16	−1.26	0.17	−1.06	0.18
Female leisure time	1.46	0.46	3.51	0.51	1.78	0.51
Male leisure time × Age of wife/100	2.50	1.81	3.29	1.90	2.83	1.92
Male leisure time × Age of husband/100	2.26	1.82	1.72	1.91	2.20	1.93
Male leisure time × Child03	−0.14	0.23	0.51	0.25	−0.19	0.24
Male leisure time × Child36	−0.52	0.23	−0.29	0.25	−0.60	0.24
Male leisure time × East Germany	−0.58	0.15	−0.94	0.16	−0.59	0.16
Male leisure time × high education	−0.42	0.18	−0.73	0.18	−0.40	0.19
Male leisure time	−1.87	0.46	0.82	0.53	−1.73	0.49
Female leisure time × Male leisure time			−9.87	0.85		
Mass point 1					−5.74	0.88
log(Probability)					1.13	0.35
Log likelihood	−9795.0948		−9717.2553		−9760.987	
	Labour Supply Elasticities					
Participation elasticity of wife	0.08		0.08		0.07	
Hours elasticity of wife	0.3		0.35		0.29	
Participation elasticity of husband	0.08		0.08		0.07	
Hours elasticity of husband	0.2		0.23		0.19	

includes changes on both the intensive and the extensive margins. For a more technical discussion of this procedure, see van Soest (1995).

Tables 3 and 4 summarize the labor supply effects induced by strict and quasi-individual taxation.

Broadly speaking, the results for strict individual taxation are in line with earlier studies, such as Steiner and Wrohlich (2004) or Bach et al. (2011). The labor supply responses for women are positive and substantially larger than the negative effects for men. In the case of strict individual taxation, the participation rate of married women (extensive margin) increases across the board by 3.01%. The total labour supply in working hours (extensive and intensive margin combined) even goes up by 8.72%. For men the comparable numbers are −0.3% (participation) and −1.5% (working hours). The variation of the labour supply effects across the different subgroups can be explained by the differences in preferences (table 2) and in the working incentives, i.e., changes in the disposable net household income due to the tax reform (table 1).

Female labor supply effects, both in terms of participation and in terms of working hours, increase monotonically with the gross household earnings. For men, the effect is opposite. Interestingly, in the top quintile we find small positive participation effects for men. Obviously, in some households

Table 3
Labour Supply Effects of Strict Individual Taxation

	Women		*Men*	
	Change in part. (in %)	Change in hours (in %)	Change in part. (in %)	Change in hours (in %)
All households	3.01	8.72	−0.3	−1.5
Quintiles of gross household earnings				
I	1.09	4.43	−0.68	−2.23
II	2.47	8.29	−0.69	−2.44
III	3.18	9.11	−0.30	−1.56
IV	3.88	10.3	−0.05	−0.99
V	4.48	11.5	0.23	−0.17
Quintiles of inequality of earnings between spouses				
I	1.57	4.49	0.25	−0.04
II	2.05	5.77	0.01	−0.76
III	2.90	8.60	−0.52	−2.17
IV	3.69	10.73	−0.84	−2.97
V	4.94	14.25	−0.40	−1.53
Household characteristics				
West Germany	3.51	9.95	−0.48	−2.01
East Germany	1.31	4.47	0.31	0.24
Without children younger 3	2.95	8.51	−0.23	−1.32
With children younger 3	3.52	10.52	−0.94	−3.0

Note: Own calculations based on SOEP and STSM. All income information is per month. Inequality of earnings between spouses is calculated as the absolute distance of the wife's labour earning to the average labour earnings of the household.

husbands are the secondary earners and therefore their working incentives would increase with individual taxation. On average this effect dominates the negative responses of the male first earners in the top quintile. Similarly, the positive effects for women tend to increase with inequality of earnings between the two spouses, while the pattern for men is not so clear cut. For men the positive effects are found in the lowest quintile, which shows that if the husband is the secondary earner, the earnings of the wife are not substantially higher.

By demographic characteristics the most striking difference is for women in east and west Germany. Whereas west German wives increase their labor supply by 3.5% and their working hours by 9.9%, married women in the

Table 4
Labour Supply Effects of Quasi-Individual Taxation

	Women		*Men*	
	Change in part. (in %)	Change in hours (in %)	Change in part. (in %)	Change in hours (in %)
All households	1.13	3.18	−0.06	−0.51
Quintiles of gross household earnings				
I	0.25	0.98	−0.13	−0.48
II	0.67	2.22	−0.17	−0.71
III	1.00	2.85	−0.09	−0.60
IV	1.50	3.99	−0.03	−0.53
V	2.30	5.95	0.12	−0.16
Quintiles of inequality of earnings between spouses				
I	0.48	1.11	0.11	0.02
II	0.68	1.69	0.06	−0.13
III	0.96	2.68	−0.07	−0.54
IV	1.31	3.80	−0.25	−1.01
V	2.3	6.76	−0.16	−0.83
Household characteristics				
West Germany	1.34	3.71	−0.12	−0.68
East Germany	0.43	1.32	0.15	0.11
Without children younger 3	1.12	3.12	−0.04	−0.45
With children younger 3	1.29	3.63	−0.25	−0.91

Notes: Own calculations based on SOEP and STSM. All income information is per month. Inequality of earnings between spouses is calculated as the absolute distance of the wife's labour earning to the average labour earnings of the household.

eastern part react by 1.3% (participation) and 4.5% (hours): substantially less. This is indeed in line with the different changes in the disposable net household income discussed above. In addition, preferences are important in explaining this difference. As mentioned above, women in west Germany have a significantly lower inclination to work, leading to lower participation rates and lower working hours. This implies *ceteris paribus* that they can change their behavior more than the east German women, who to a large extent worked already before the reform.

In the case of quasi-individual taxation the positive and negative labor supply effects are, as expected, considerably smaller than in the case of strict individual taxation. However, the pattern of the changes across the heterogeneous population is very similar.

5.2. Welfare Effects of the Tax Reform: Gainers and Losers

In this section of the paper we go beyond the previous literature on the effects of replacing joint taxation with individual taxation in Germany. In fact that literature limited itself to describing the results as in the previous subsection, that is, in terms of effects on labor supply.[11] Yet, for the evaluation of this kind of reform, which triggers a considerable effect on labor supply, one needs a more comprehensive individual- or household-specific welfare measure than simply net household income. Decreased leisure being the counterpart of the increase in participation or working hours, one has to take account of this in a genuine welfare analysis of the reform.

In classical applied welfare analysis, equivalent or compensating variations have been used intensively as metrics of individual welfare changes. But as noted in the introduction, interpersonal comparability of these changes necessitates applying them in the context of a reference preference ordering. This drives a wedge between the role of heterogeneous preferences in the positive model (where heterogeneity plays a crucial role in explaining and simulating behavior) and in the normative model (where we are forced to remove heterogeneity from the scene).

In this paper we avoid this price, by pursuing a different strategy that is based on the explicitly normative individual welfare concepts proposed in Fleurbaey (2006). We refer to Decoster and Haan (2010) for a more detailed exposition of these concepts. Here it suffices to say that the individual preference information, in the form of indifference curves, is used to translate the chosen bundle into an "equivalent" income concept (see chapter 4 in Fleurbaey and Blanchet, 2013). [12] In practice, e.g., since one has identified preferences in the structural labor supply model, one can easily calculate the "nonlabor income" that, if one were to supply no labor, would bring the individual to the same indifference curve as the currently chosen bundle. In the standard graphical representation of the labor supply decision, with labor supply on the horizontal and net income on the vertical axis, the equivalent income is determined by the intersection of the vertical axis, at zero hours of work, with the indifference curve passing through the chosen bundle. This metric was proposed earlier as the so-called intercept income by Preston and Walker (1999) and is called the *rent criterion* in Fleurbaey (2006). But the contribution of the recent literature is that it makes the normative choices

11 As mentioned in the introduction, Beblo et al. (2007) look, in their welfare analysis, only at within-household changes, and their method does not allow for any comparison between households.

12 Note that this recently developed notion of "equivalent income" differs from "equivalent income" as defined in Killingsworth (1983), where money metrics (in the forms of expenditure functions) are used in the classical framework of a representative agent with reference preferences and with reference prices to compare individual welfare levels.

underpinning the use of this specific individual welfare metric explicit. As shown in Fleurbaey (2006), e.g., when using the rent criterion one offers maximal protection for people who have a larger distaste for working. Otherwise stated, choosing the rent criterion holds people with a strong aversion to work minimally responsible for that preference. This of course suggests that other choices are possible, each embedding different normative priors. In this paper, we used one other individual welfare metric, which we christen the FT criterion (for "full-time"). It also consists of an intersection of the indifference curve with a vertical axis, but now at the level of full-time (or maximal) labor supply. This criterion definitely treats individuals characterized by a high preference for work more favorably, in the sense that, *ceteris paribus*, individuals with less "industrious" preferences are now considered as less badly off than under the rent criterion.

Heterogeneity in preferences shows up in crossing indifference curves. Therefore, the probability that the rent criterion and the FT criterion will rank two individuals differently in terms of individual welfare is high (depending of course also on the specific bundles chosen). However, this explicit dependence on the normative priors is not to be considered as a drawback of the approach, but quite the contrary, as a trump card if one analyses the reform using the different metrics. Moreover, the approach gets rid of the need to use a reference preference ordering of some kind of representative agent and retains the full preference heterogeneity of the population in the normative step of the analysis.

Since in this paper we use a sample of couple households, we had to extend the welfare metrics to be applicable to households with two agents who decide on labor supply. Indeed, so far the literature has mainly discussed the choice of individual welfare metrics in the two-dimensional single-agent model; see, e.g., Preston and Walker (1999), Fleurbaey (2006), or Decoster and Haan (2010). But exactly because we opt for two intercept incomes to span the normative space, the extension to two spouses who can choose their labor supply was straightforward. For the rent criterion we assume that the spouses have both maximal leisure time (hence behaving as if neither worked). For the FT criterion we calculate welfare by determining the consumption level that would yield the same utility level as the currently chosen bundle in the counterfactual situation where both spouses would work full time. To make the equivalent incomes comparable across household sizes, we applied the standard OECD equivalence scale (1 for the first adult, 0.5 for additional adults, and 0.3 for children).

The results are in tables 5 and 6. We compare the effect of the tax reform by means of three different individual welfare measures: disposable income, rent criterion, and FT criterion. All measures include the labor supply responses induced by the tax reforms. For the welfare analysis, the measures

Table 5

Change in Individual Welfare Measures (in Euro per Month): Strict Individual Taxation

	Deciles of Income			*Deciles of Rent*			*Deciles of FT*		
	Δ Income	Δ Rent	Δ FT	Δ Income	Δ Rent	Δ Work	Δ Income	Δ Rent	Δ FT
1	−16	−15	−34	−26	−23	−51	−13	−16	−22
2	−37	−43	−101	−37	−39	−85	−21	−26	−40
3	−50	−63	−145	−46	−54	−118	−31	−38	−66
4	−54	−66	−139	−54	−67	−146	−41	−55	−101
5	−67	−82	−180	−61	−79	−176	−53	−67	−124
6	−62	−85	−179	−63	−85	−180	−61	−79	−150
7	−67	−89	−189	−68	−94	−193	−71	−92	−174
8	−76	−100	−200	−74	−99	−202	−84	−112	−216
9	−88	−122	−246	−78	−117	−245	−77	−114	−232
10	−69	−109	−205	−79	−118	−222	−136	−178	−495
Average	−59	−78	−162	−59	−77	−162	−59	−78	−162

Note: Own calculations based on SOEP and STSM. The changes (Δ) are always defined as the difference between the pre-reform and the post-reform metric. All income information is per month. Inequality of earnings between spouses is calculated as the absolute distance of the wife's labour earning to the average labour earnings of the household. The Rent criterion is calculated as monthly net income assuming that both spouses are not working and the FT criterion as monthly net income assuming that both spouses are working full time, i.e 40 hours per week.

Table 6

Change in Individual Welfare Measures (in Euro per Month): Quasi Individual Taxation

	Deciles of Income			*Deciles of Rent*			*Deciles of FT*		
	Δ Income	Δ Rent	Δ FT	Δ Income	Δ Rent	Δ FT	Δ Income	Δ Rent	Δ FT
1	−3	−3	−7	−5	−5	−11	−2	−3	−4
2	−9	−10	−25	−10	−10	−23	−5	−6	−9
3	−14	−17	−40	−14	−16	−35	−9	−10	−17
4	−18	−20	−44	−18	−20	−44	−13	−16	−29
5	−25	−29	−63	−24	−27	−62	−19	−23	−42
6	−26	−33	−70	−26	−33	−70	−25	−30	−56
7	−31	−38	−81	−31	−38	−79	−32	−39	−73
8	−39	−48	−96	−38	−47	−97	−42	−53	−100
9	−53	−69	−139	−46	−64	−134	−43	−60	−115
10	−46	−67	−125	−51	−73	−136	−74	−94	−245
Average	−26	−33	−69	−26	−33	−69	−26	−33	−69

Note: Own calculations based on SOEP and STSM. The changes (Δ) are always defined as the difference between the pre-reform and the post-reform metric. All income information is per month. Inequality of earnings between spouses is calculated as the absolute distance of the wife's labour earning to the average labour earnings of the household. The Rent criterion is calculated as monthly net income assuming that both spouses are not working and the FT criterion as monthly net income assuming that both spouses are working full time, i.e 40 hours per week.

play a double role. First, we can use the levels of the metrics to construct the individual welfare ordering. This leads to the three parts of the table, with deciles constructed by means of disposable income, of the rent criterion,

and of the FT criterion respectively. For each of these welfare orderings, we present the change in individual welfare due to the introduction of strict individual taxation. This allows us to analyze the sensitivity of the distributional effects of the tax reform to the different individual welfare measures. All measures can be interpreted in monetary terms (euros per month). But of course, their different definitions and normative content make a direct comparison difficult. Moreover, in tables 5 and 6 we have not imposed revenue neutrality. The reason is that the design chosen to implement revenue neutrality (e.g., by a lump sum or a proportional transfer) will of course strongly affect the distributional pattern. In the next section, where we look at the effects at the aggregate level in terms of social welfare, we will also integrate revenue neutrality into the analysis.

Starting with the traditional analysis of changes in disposable income across deciles of (equivalised) disposable income, the first column of table 5 shows that the income loss is bigger for the income groups in the middle and in the upper half of the distribution. This is due to the fact that the splitting gain, which is lost under individual taxation, depends on the inequality between the spouses' labor incomes within households. This inequality is largest in the middle of the income distribution, and hence the losses are the largest for the deciles in the upper middle. The loss for the top decile is significantly lower. In terms of percentage changes (not shown in the table), the picture is U-shaped and the Gini coefficient increases after the reform (see upper panel in table 7). In line with the results in Decoster and Haan (2010), we find that using another welfare concept to calculate welfare changes, but staying within the same welfare ordering, does not affect this conclusion much. The same U-shaped pattern is displayed in columns 2 and 3 of table 5. What does matter, however, is how we construct the welfare ordering, i.e., whom we consider to be the worse off and the better off in the initial distribution. In that case, whereas the use of the rent criterion closely resembles the use of the household income, the introduction of the FT criterion definitely changes the picture. Independently of whether we use income, rent, or FT as the criterion by which we calculate the welfare change, the three rightmost columns of table 5 reveal that the transition to strict individual taxation now imposes a larger loss on better-off households and hence decreases inequality. In case we use the FT criterion, e.g., inequality as measured by the Gini coefficient goes down after the reform (see upper panel in table 7).

For all three welfare concepts used in the rightmost part of table 5, where we have ordered individuals on the basis of the FT criterion, we even find Lorenz dominance of the postreform distribution w.r.t. the prereform one. That means that the conclusion that the reform reduces inequality when the welfare distribution is constructed on the basis of the FT criterion holds not only for the Gini coefficient, but for any relative inequality measure that is

Table 7
Effect on Social Welfare Induced by Individual Taxation

	Strict individual Taxation	Strict individual Taxation with lump sum transfer	Quasi individual Taxation	Quasi individual Taxation with lump sum transfer
		Changes in the Gini coefficient (G)		
Net income	+	–	–	–
Rent criterion	+	–	–	–
FT criterion	–	–	–	–
		Changes in average welfare metric (μ)		
Net income	–	0	–	0
Rent criterion	–	–	–	–
FT criterion	–	–	–	–
		Changes in social welfare ($W = \mu(1 - G)$)		
Net income	–	+	–	+
Rent criterion	–	–	–	+
FT criterion	–	–	–	+

Note: Own calculations based on SOEP and STSM. Social welfare: $W = \mu(1 - G)$, in which μ stands for the average level of the individual welfare measure and G for the inequality in it, as measured by the Gini coefficient. The changes are always defined as the difference between the pre-reform and the post-reform metric. A '+' implies a positive, a '–' a negative change.

Lorenz-consistent. Comparing table 5, where we show the results of strict individual taxation, with table 6, where the results of quasi-individual taxation are shown, reveals that the distributional patterns and the sensitivities for the chosen welfare concepts are fairly similar. Households in the highest decile face lower losses than households in lower deciles when we use the income or rent criterion, but when using the FT criterion we find a monotone distributional pattern where the households in the highest deciles suffer the largest negative effect. Interestingly, however, according to the Gini coefficients (see upper panel in table 7), inequality decreases irrespective of the chosen metric. This is related to the fact that households in the middle of the income distribution benefit particularly from the joint exemption level in the case of quasi-individual taxation.

5.3. Effects on Social Welfare

In the previous section we focused on the distributional pattern of gains and losses. However, when we attempt to aggregate the individual welfare changes into one overall welfare effect, we have to impose revenue neutrality, i.e., we require that disposable income be the same before and after the reform. If not, it is not surprising that a reform that leads to a sizable increase in tax revenue creates overall welfare losses. As mentioned above, and not unexpectedly, the design of revenue neutrality might strongly affect

the overall result. In order to make our analysis as transparent as possible, we decided to redistribute the additional revenue through a lump-sum transfer. This lump-sum transfer of course induced further labor supply reactions of the spouses, which are taken into account in our calculation. In fact, a lump-sum transfer induces negative labor supply reactions on the basis of the income effect. This effect can be quite sizable, particularly for low-income households.

We present the combined effect of distributional changes and changes in the average level of welfare, by means of what is called by Lambert (2001) an abbreviated social welfare function. We use the following form to represent social welfare: $W = \mu(1 - G)$, in which μ stands for the average level of the individual welfare measure (see second panel in table 7), and G for the inequality in it, as measured by the Gini coefficient (see upper panel in table 7). The last panel of table 7 displays whether this measure W is increased or decreased by the reform of the tax system, both for the move to strict and the move to quasi-individual taxation, and both without and with revenue neutrality.[13]

Starting with the first column, we find that, not surprisingly, the reform towards strictly individual taxation lowers social welfare. For the income and rent criteria this is explained by the fact that inequality increases and also average individual welfare goes down. For the FT criterion we find that inequality is reduced by the reform; however, for the social welfare this is overcompensated by the decrease in average individual welfare. Redistributing the additional tax revenue as a lump sum is inequality-decreasing (for all three measures). Since, by definition, average disposable income is now unchanged, this leads to an increase of social welfare when one uses disposable income as the individual welfare criterion. For the other two criteria, the decrease in inequality is not large enough to compensate for the decrease in average individual welfare (column 2). Without revenue neutrality, the results for quasi-individual taxation are similar. But the move towards quasi-individual taxation with a full recycling of the additional revenue through a lump-sum redistribution is remarkable. For all three individual welfare measures, we now obtain a social welfare improvement. In fact, with this scenario we find, for all three measures, Lorenz dominance of the situation after the reform over the baseline situation. This is equivalent with general-

13 We also checked for generalized Lorenz dominance, to investigate whether the results in table 7 hold for a broader class of social welfare functions than the one based on the Gini coefficient (viz., for all Paretian, scale-invariant, inequality-averse SWFs). In case we do not implement revenue neutrality, though, the result is trivial. The income or welfare loss at all income (or welfare) levels evidently leads to generalized Lorenz dominance of the actual tax system over the two proposed reforms. In the case of the two revenue-neutral tax reforms, we did not find generalized Lorenz dominance, either for strict individual taxation or for quasi-individual taxation.

ized Lorenz dominance (and hence welfare dominance) in the case where we opt for income as a welfare measure, since disposable income is unchanged. In the cases of the rent and the FT criterion, there is a decrease in average individual welfare, which prevents us from finding generalized Lorenz dominance in this case. However, the decrease in average welfare is not large enough to annihilate the social welfare gain of the lower inequality as measured by the Gini coefficient. The result is that for the specific social welfare function $W = \mu(1 - G)$ and a lump-sum redistribution, we assess the move to quasi-individual taxation as social-welfare-improving. It is necessary to stress again that this strong result depends on the way we impose revenue neutrality. For strict individual taxation it would certainly be possible to redistribute the additional revenue (e.g., by providing a larger share for households in the lower deciles) so that the reduction in inequality would overcompensate the decrease in average individual welfare.

6. Conclusion

In this paper we provide an empirical welfare evaluation of replacing joint taxation of couple households with individual taxation.

Our findings in terms of labor supply support previous results for Germany – e.g., Steiner and Wrohlich (2004) or Bach et al. (2011) – that have studied a replacement of joint taxation using the same empirical method that takes into account labor supply effects by means of a discrete-choice model. However, we go beyond this literature in that, in addition to labor supply and income effects, we calculate the welfare effects induced by this tax reform. The welfare effects are particularly informative for the political debate, because they take into account the change in leisure time induced by a tax reform.

For empirical welfare evaluations individual welfare metrics, such as equivalent or compensating variations, have often been used, and it is well known that in a context of individuals with heterogeneous preferences, both the interpretation of these welfare metrics and especially their aggregation quickly face serious difficulties. In this paper we overcome these difficulties and pursue a different empirical strategy that is based on normative welfare concepts proposed in Fleurbaey (2006). The advantage of these welfare measures is that they try to unveil more explicitly the normative assumptions on which the interpersonal comparability rests.

Our results of the welfare analysis suggest that the choice of the welfare metric can be crucial, particularly for the simple question who are the gainers and who are the losers. In our specific application we find that for the case of quasi-individual taxation, *any* social planner who expresses social welfare by

means of a social preference ordering defined over these individual welfare metrics would come to the same conclusion: a policy change that replaces joint taxation with full splitting by individual taxation with a joint exemption level would be welfare-improving for both individual welfare metrics. The results for strict individual taxation are less clear cut.

7. Appendix – Gross Hourly Wages

In order to calculate the employment-state-specific gross labor earnings it is necessary to derive the gross wage distribution for the working and non-

Table 8
Wage Regression for West Germany

	(Men) log wage		(Women) log wage	
Medium education	0.00871	(0.0104)	0.0226*	(0.0114)
Higher education	0.0797***	(0.0123)	0.0798***	(0.0136)
High education	0.219***	(0.0126)	0.227***	(0.0145)
Unemployment history	−0.0809***	(0.00508)	−0.0419***	(0.00582)
Experience	0.0272***	(0.00108)	0.0184***	(0.00128)
$Experience^2$	−0.0482***	(0.00248)	−0.0331***	(0.00312)
Tenure	0.0138***	(0.000998)	0.0136***	(0.00117)
$Tenure^2$	−0.0204***	(0.00261)	−0.0204***	(0.00331)
Migration Background	−0.00755	(0.00789)	−0.0170	(0.00875)
Mills Ratio	−0.0234	(0.0175)	0.0217	(0.0123)
Constant	2.351***	(0.0217)	2.193***	(0.0245)
	Selection Equation			
Child younger 3	−0.237**	(0.0739)	−1.835***	(0.0482)
Child younger 6	−0.265***	(0.0773)	−1.081***	(0.0467)
Child younger 16	−0.186***	(0.0415)	−0.508***	(0.0299)
Child aged 17 or 18	−0.0924	(0.0532)	−0.161***	(0.0384)
Other household income	−0.000213***	(0.00000931)	−0.0000436***	(0.00000470)
Constant	−0.819***	(0.243)	0.922***	(0.173)
Observations	19193		22724	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Note: Own calculations based on SOEP. Estimation includes data for years 2005–2009. In addition to the presented coefficients we include time, regional and industry dummies. The Selection Equation includes all variables from the main specification and in addition the exclusion restrictions which are presented in the table.

Table 9
Wage Regression for East Germany

	(Men) log wage		(Women) log wage	
Medium education	−0.00155	(0.0252)	0.0307	(0.0289)
Higher education	0.0352	(0.0292)	0.0803*	(0.0326)
High education	0.103***	(0.0304)	0.193***	(0.0312)
Unemployment history	−0.125***	(0.00668)	−0.0803***	(0.00619)
Experience	0.0276***	(0.00206)	0.0196***	(0.00231)
$Experience^2$	−0.0591***	(0.00482)	−0.0432***	(0.00527)
Tenure	0.00851***	(0.00195)	0.0178***	(0.00213)
$Tenure^2$	−0.0153**	(0.00525)	−0.0278***	(0.00569)
Mills Ratio	0.00239	(0.0241)	−0.0173	(0.0224)
Constant	2.260***	(0.0362)	2.111***	(0.0412)
	Selection Equation			
Child younger 3	−0.177	(0.0985)	−1.313***	(0.0788)
Child younger 6	0.0363	(0.102)	−0.519***	(0.0805)
Child younger 16	−0.116	(0.0665)	−0.264***	(0.0581)
Child aged 17 or 18	−0.164*	(0.0832)	−0.213**	(0.0751)
Other household income	−0.000225***	(0.0000188)	−0.0000885***	(0.0000147)
Constant	−1.652***	(0.385)	−1.016**	(0.338)
Observations	6563		7407	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Own calculations based on SOEP. Estimation includes data for years 2005–2009. In addition to the presented coefficients we include time, regional and industry dummies. The Selection Equation includes all variables from the main specification and in addition the exclusion restrictions which are presented in the table.

working populations. This is the distribution of the offered marked wages, which measures the signaled productivity of an individual. For the working individuals we define their observed wage as their individual draw from the offered wage distribution.

For individuals belonging to the non-working population the draw from the offered wage distribution cannot be observed. Therefore, it is necessary to estimate their expected gross hourly wages, that is, the person-specific distribution of offered wages W_i. As W_i is not observable, we estimate $\hat{W}_i$, a measure of the mean of the distribution of offered wages. More precisely, we estimate a Mincer wage equation with selection effects using the information on the working population and interpret the predicted hourly wages of the

non-working individuals as the mean of the distribution of offered wages. The specification includes educational and experience information, and in addition time-, region-, and industry-specific effects. As exclusion restriction for the selection equation we use, as is common, non-labor income and the presence of children in different age categories. The wage estimation is performed separately for east and west Germany and by gender. In the following tables we present the results for the central coefficients.

As expected, we find a negative effect of the exclusion restrictions on the participation decisions, particularly for women in west Germany. Yet, we only find a modest effect of the Mills ratio on the wages. Moreover, in line with the literature, our results indicate that wages *ceteris paribus* increase with eduction, experience, and tenure, while periods in unemployment have a negative effect on wages.

References

Aaberge, R., and Colombino, U. (2013), Designing Optimal Taxes with a Microeconometric Model of Household Labour Supply, Scandinavian Journal of Economics 115, 449–475.

Aaberge, R., Colombino, U., and Stroem, S. (2004), Do More Equal Slices Shrink the Cake? An Empirical Investigation of Tax-Transfer Reform Proposals in Italy, Journal of Population Economics 17, 767–785.

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

Alesina, A., Ichino, A., and Karabarbounis, L. (2011), Gender Based Taxation and the Division of Family Chores, American Economic Journal: Economic Policy 3, 1–40.

Apps, P., and Rees, R. (1999), Individual Versus Joint Taxation in Models with Household Production, Journal of Political Economy 107, 393–403.

Auerbach, A. (1985), The Theory of Excess Burden and Optimal Taxation, in: Auerbach, A., and Feldstein, M. (eds.), Handbook of Public Economics 1, 61–127.

Bach, S., Geyer, J., Haan, P., and Wrohlich, K. (2011), Reform des Ehegattensplittings: Nur eine reine Individualbesteuerung erhöht die Erwerbsanreize deutlich, DIW Wochenbericht 78, 13–19.

Beblo, M., Beninger, D., and Laisney, F. (2007), Welfare Analysis of Fiscal Reforms: Does the Representation of the Family Process Matter? Evidence from Germany, Journal of Population Economics 20, 869–893.

Boadway, R., and Bruce, N. (1984), Welfare Economics, Basil Blackwell, Oxford, UK.

Boskin, M. J., and Sheshinski, E. (1983), Optimal Tax Treatment of the Family: Married Couples, Journal of Public Economics 20, 281 – 297.

Creedy, J., and Kalb, G. (2005), Discrete Hours Labour Supply Modelling: Specification, Estimation and Simulation, Journal of Economic Surveys 19, 697–734.

Decoster, A., and Haan, P. (2010), Empirical Welfare Analysis in Random Utility Models of Labour Supply, IZA Discussion Paper No. 5301.

Eissa, N., Kleven, H., and Kreiner, C. (2008), Evaluation of Four Tax Reforms in the United States: Labor Supply and Welfare Effects for Single Mothers, Journal of Public Economics 92, 795–816.

Fleurbaey, M. (2006), Social Welfare, Priority to the Worst-off and the Dimensions of Individual Well-Being, in: Farina, F., and Savaglio, E. (eds.), Inequality and Economic Integration.

Fleurbaey, M., and Blanchet, D. (2013), Beyond GDP Measuring Welfare and Assessing Sustainability, Oxford University Press, Oxford, UK.

Haan, P. (2006), Much Ado About Nothing: Conditional Logit vs. Random Coefficient Models for Estimating Labour Supply Elasticities, Applied Economics Letters 13, 251 –256.

Heckman, J. J., and Singer, B. (1984), A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, Econometrica 52, 271–320.

King, M. (1983), Welfare analysis of tax reforms using household data, Journal of Public Economics 21, 183–214.

LaLumia, S. (2008), The Effects of Joint Taxation of Married Couples on Labor Supply and Non-Wage Income, Journal of Public Economics 92, 1698–1719.

Lambert, P. (2001), The Distribution and Redistribution of Income, Manchester University Press, Manchester, UK.

Mroz, T. (1987), The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions, Econometrica 55, 765–799.

Preston, I., and Walker, I. (1999), Welfare Measurement in Labour Supply Models with Nonlinear Budget Constraints, Journal of Population Economics 12, 343–361.

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

Soest, A. van, and Stancanelli, E. G. F. (2010), Does Income Taxation Affect Partners' Household Chores?, IZA Discussion Paper No. 5036.

Steiner, V., and Wrohlich, K. (2004), Household Taxation, Income Splitting and Labor Supply Incentives. A Microsimulation Study for Germany, CESifo Economic Studies No. 50, 541–568.

Steiner, V., Wrohlich, K., Haan, P., and Geyer, J. (2008), Documentation of the Tax-Benefit Microsimulation Model STSM: Version 2008, DIW Data Documentation 31.

Wagner, G., Frick, J., and Schupp, J. (2007), The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements, Schmollers Jahrbuch 127, 139–169.

Wrohlich, K. (2006), Labor Supply and Child Care Choices in a Rationed Child Care Market, IZA Discussion Paper No. 2053.