Discussion Paper No. 03-31

Welfare Analysis of Fiscal Reforms: Does the Representation of the Family Decision Process Matter? Evidence for Germany

Denis Beninger, François Laisney and Miriam Beblo

ZEW
Zentrum für Europäische Wirtschaftsforschung GmbH
Centre for European Economic Research

Discussion Paper No. 03-31

Welfare Analysis of Fiscal Reforms: Does the Representation of the Family Decision Process Matter? Evidence for Germany

Denis Beninger, François Laisney and Miriam Beblo

Non-technical summary

We compare the results of a tax reform analysis obtained with the collective model of household labour supply and consumption behaviour with results obtained with the unitary model, in which the household is assumed to maximise a single utility function, even if it consists of several individuals. The theoretical implications of this simplifying assumption turn out to be rather restrictive and are most often rejected when tested. The collective approach to household behaviour can be seen as a substantial improvement over the unitary model as it explicitly accounts for the existence of several decision makers. The main assumption is that the intrahousehold bargaining process results in Pareto efficient outcomes. Depending on the choice of the representation of household behaviour one may expect welfare analysis of policy reforms to yield different results. Here we focus on the quantitative discrepancies entailed by the use of a unitary model when assessing the labour supply response to a tax reform and the subsequent impacts on individual welfare.

The first objective is the calibration of a collective model of household labour supply. The approach focuses on the direct modelling of the household's position on the Pareto frontier, which determines the labour supply and consumption of household members. This allows to account for the distinction between participation and hours of work, for non-linear and non-convex budget sets, and for non-egoistic individual preferences. The question addressed by our second objective is: 'how large are the distortions arising from the use of a unitary model in the assessment of a tax reform, when household behaviour follows the collective model?' To do so, we simulate data by a compound procedure of estimation and calibration based on the 1998 wave of the German socio-economic panel. A unitary model is then estimated on this 'collective data'. Investigating the effects of a move from joint to individual taxation, we obtain important discrepancies between predicted adjustments to labour supply, and distortions in the welfare analysis of the reform on the basis of unitary estimates. Apart from possible misspecification of the unitary model, discrepancies between the predictions from both representations can only be due to the wrong assumption that (collective) households behave as if they were single decision makers. Our results hence suggest that increased efforts should be devoted to the estimation of collective models with taxation.

Welfare analysis of fiscal reforms: does the representation of the family decision process matter? Evidence for Germany

Denis Beninger*, François Laisney† and Miriam Beblo‡

May 2003

Abstract

This paper compares predictions obtained for the analysis of tax reforms with collective and unitary models of household labour supply and consumption behaviour. We simulate real world microdata by means of a collective approach, using a compound procedure of estimation and calibration based on the 1998 wave of the German socio-economic panel. We estimate a unitary model on this 'collective' data set. Investigating a move from joint to individual taxation on the basis of both models, we obtain important discrepancies between predicted adjustments to labour supply and distortions in the welfare analysis of the reform on the basis of unitary estimates. These results suggest that increased efforts should be devoted to the estimation of collective models with taxation.

Key Words: Collective model, household labour supply, intra household allocations
JEL Classification: D11, D12, J22

Acknowledgements:
We are grateful for comments and advice from Richard Blundell, Martin Browning, Pierre-André Chiappori, Valérie Lechene, Costas Meghir, Nathalie Picard, Frederic Vermeulen, and all participants in the one-year project "Welfare analysis of fiscal and social security reforms in Europe: does the representation of family decision processes matter?", partly financed by the EU, General Directorate Employment and Social Affairs, under grant VS/2000/0778. Thanks also to Olivier Bargain, Stefan Boeters, Bernhard Boockmann, Hermann Buslei, Peter Jacobebbinghaus, Nicolas Moreau and Elke Wolf for careful reading of preliminary versions. All errors remain our own.

*ZEW, Mannheim. Beninger@zew.de
†Université Louis Pasteur, Strasbourg, and ZEW, Mannheim. Laisney@zew.de
‡ZEW, Mannheim. Beblo@zew.de

1 Introduction

Reforms of the tax and social benefit system generally have an impact on individuals' and families' living conditions and on their behaviour with regard to labour supply and consumption. Moreover, such reforms may affect the intra-household distribution of resources in ways that differ from the intentions behind them. In this paper we analyse the impact a particular tax reform would (possibly) have in Germany, on households' consumption behaviour, on the labour supply of women and men, and on the within-family distribution of wealth.

In investigating these questions, an appealing representation of the decision process of the household is the collective framework, introduced by Chiappori (1988, 1992), and Apps and Rees (1988). This type of model accounts for the presence of multiple decision makers within the household, in contrast to the traditional, or unitary, representation. Unitary models consider household behaviour as resulting from the decisions of a single unit, obscuring the fact that households are most of the time composed of several members. They treat the family as a 'black box', and thus the within-family reallocation of resources resulting from a policy reform cannot be reconstructed. It is assumed that the household maximises a unique utility function, independent of prices and incomes. With the unitary models, only inter-household income inequality can be studied. Yet the question of intra-family redistribution can be crucial in determining household choices. These issues are important if policy makers want to conduct efficient and fair economic and social policies.

Drawing on the collective framework, on the contrary, opens up the possibility to infer aspects of the within-household welfare implications of policy changes. Collective models basically assume only Pareto-optimality of intra-household allocations, and this assumption defines the collective rationality of households. Chiappori (1988) distinguishes two cases, according to whether agent's preferences are egoistic or altruistic. If the agents have egoistic preferences and face a linear budget constraint or a convex budget set (Donni, 2003 and Beninger, 2000), their behaviour can be represented sequentially, using an explicit income sharing rule. Otherwise, a sequential representation of household behaviour is no longer possible, but this behaviour still results from the sharing of consumption between agents (Chiappori, 1988).

A main contribution of this study will be a concrete comparison of the implications of the choice between the two representations mentioned above. The question is whether the predictions of labour supply responses to a tax reform and the changes of the distribution of welfare vary substantially with the representation chosen. Our goal is to *quantify* the distortions that may affect policy recommendations obtained with the unitary representation.

In order to compare the two settings, we cannot rely on an estimation strategy, in contrast to Moreau and Donni (2002), as the full estimation of a collective model with *nonconvex* budget sets and a participation decision for both spouses is a difficult task which has not yet been achieved. To circumvent this problem, we simulate real world microdata by means of a 'deterministic' collective labour supply model. A unitary model is then estimated on this 'collective' data set where households behave according to the collective rationality. The evaluation of a tax reform based on these two models allows to compare the performance of the collective and the unitary model of family behaviour.[1]

Previous work by Beninger and Laisney (2002) addressed these issues on the basis of purely synthetic data. Their study reveals important discrepancies (1) in the incentive and distribution effects of revenue-neutral reforms based on unitary estimates rather than on the true collective parameters, (2) in predicted adjustments of labour supplies following a switch between two tax regimes (individual and joint taxation for couples), and (3), it also provides evidence of conflicting results when welfare analyses of policy reforms are based on unitary estimates rather than collective estimates. The aim of the present paper is to check the robustness of these results when the 'collective baseline situation' does not result from synthetically simulated households but rather is generated by a model based on reasonable assumptions, giving results close to reality, thanks to the use of calibration.

2 Brief outline

Our real world collective data are calibrated so as to reflect the characteristics (labour supply, household income/consumption level) of couples of the 1998 wave of the German Socio-Economic Panel (GSOEP) in a realistic way. Our strategy relies on the assumption that some aspects of individual preferences – but not all – concerning own consumption and leisure are the same for single and married women or men. To take up possible utility-producing complementarity effects of leisure, a term accounting for interaction between spouses' leisures, namely a 'cross-leisure effect' is also considered. Under these assumptions we estimate individual preference parameters for singles and use these estimates to simulate collective couple data simply by exploit-

[1]Here, in order to save space, we focus on a single reform, namely a revenue neutral move from joint to individual taxation. Further evidence for Germany concerning a revenue neutral move to a linear tax system, and the 'reform 2000', as well as results for other European countries (Belgium, France, Italy, Spain and the UK) can be found in Laisney (2002, ed.).

ing the Pareto-optimality assumption for the households' decision process; this allows us to retrieve both the cross-leisure effect and a power index for each couple, by calibration on GSOEP couple data. We then regress this calibrated power index on a set of variables, including some original variables describing the way in which the tax and benefit system affects the relative net earning potential of the spouse. We subsequently predict the power index for different tax situations and thereby simulate the effects of a tax reform on spouses' consumption levels and labour supplies. The reform studied here consists in a revenue neutral switch from joint to individual taxation.

In a further step, the baseline simulated data are used to estimate a unitary model for consumption and household labour supply. This allows us to compare the results of policy evaluations obtained with both types of models on a number of dimensions, including both positive (impact on behaviour) and normative aspects (impact on individual welfare and on household welfare, and impact on inequality measures). Thus, we could analyse the impacts of a given tax reform on both efficiency (incentive effects on joint decisions of individual labour) and equity (distributional and welfare impacts on both an individual and household basis). The question we address here is: how large are the errors committed when basing policy recommendations on estimations from the unitary model when the collective model is the true one? If these distortions turn out to be large, more effort should be put into the estimation of collective models in complex settings, including non-convex budgets and possible non-participation for both spouses.

The outline of the paper is the following. In Section 3 we describe the simplified version of the 1998 German tax-benefit system that will be used in the sequel. Section 4 introduces the characteristics of the samples from the GSOEP data used to construct our collective data set; this includes the description of the samples of single men and women that will be used not only for calibration of the preferences of couples but also for designing revenue neutral reforms, and more generally for the analysis of the reforms (the reader may first skip Sections 3 and 4, referring back to them when required by the context). Section 5 presents the steps used in the calibration procedure. These include the estimation of preference parameters for singles, the calibration of leisure interaction coefficients, one for each spouse, and of a power index. The latter turns out to be related to variables reflecting the impact of the tax system on the relative net earnings capacity of husband and wife; through that channel, reforms of the tax-benefit system will affect the bargaining power of the spouses. In Section 6 we describe the reform considered, and analyse it on the basis of the collective model. The estimation and simulation results for the unitary model are presented in Section 7, where we also compare the simulated effects of the tax reform in the collective world with those predicted by the unitary model. Section 8 concludes.

Figure A.1, placed on the last page for easy reference, gives a road map through the paper, visualising the links between the subsections.

3 The 1998 German tax-benefit system

Germany has a personal income tax system administered at the federal level and regulated by the Personal Income Law (Einkommensteuergesetz). The German tax system is characterized by a comprehensive tax that covers labour earnings as well as income from other sources such as capital investment, housing rents etc. and by joint taxation for married couples. For our exercise we use a simplified form of the 1998 German tax and benefit system described in detail in Table 1.[2] Gross income is the sum of income from different sources: income from employment, capital investment, rental and leasing and maintenance payments received from an ex-partner. Gross taxable income is equal to gross income minus income-related expenses. The standard deductions are listed in Table 1.

The function applied to the tax base is smoothly progressive. In 1998 the top rate applied was 53% for yearly earnings in excess of DM 120,041.[3] Earnings below the basic personal allowance of DM 12,365 are tax free. The tax schedule used is the same for singles and for couples. However, the 'Ehegattensplitting' method (marital splitting) is used for couples: the tax rate is applied to half of the joint taxable income, and the outcome is doubled in order to obtain the total income tax liability of the spouses. Tax rate progressivity and marital splitting lead to a relative advantage for married couples if spouses have unequal incomes.

Parents can opt for either a child benefit (DM 220 for the first and the second child, DM 300 for the third and DM 350 from the fourth child on) or a child allowance, that is a lump-sum deduction of DM 6,912 for each child up to age 27, if still in education or doing military or civil service. Due to the progressive tax scheme the child benefit is less, and the tax deduction is more, favourable for high-income households.

Social benefits are means-tested and depend on the number of people in the household.[4] As a simplification, we assume that the maximum social

[2] This simplification is an adaptation, to our particular sample and emphasis, of the microsimulation program developed at ZEW (for a description see Jacobebbinghaus and Steiner, 2003).

[3] For convenience, the tax rules, and also our tax program, are written in DM rather than in euro, as the non-linear 1998 German tax scheme is only available in DM (1 euro = 1.95583 DM). All other nominal magnitudes in the paper will be given in euro.

[4] In our static setting we ignore unemployment insurance and unemployment benefits

benefit (including housing benefit and special payments) a person can receive is DM 1,000 a month and DM 700 for the partner.[5] In addition there are age-dependent supplementary payments for children. The amount of the transfer depends on the level of earned income ('anrechnungsfreies Erwerbseinkommen') and is degressive, depending on the relevant income measure. In addition, social benefits are related to the geographical location, since they are paid by the local governments and housing benefits depend on the average rent of the locality. We distinguish only between East and West Germany, and approximate that social benefits are 10% lower in the East. The difference between East and West Germany stems from the lower costs of living in the East: a substantial fraction of the social benefit is the housing benefit. Finally, social benefit payments depend on the wealth situation of the household, and child benefits are deducted from social benefit payments.

As a graphical illustration of the tax-benefit system described above, Figure 1 depicts a typical situation for a couple with two children. The husband has an hourly wage rate of 25 euro, the wife earns 18 euro per hour. The household does not have any capital inflows or income from rental or leasing. It is therefore eligible for means-tested social benefits at low labour income. The parents receive child benefit for both children. From a yearly gross income of just above 80,000 euro they will opt for child allowance instead, as the tax relief exceeds the lump-sum benefit payment. Figure 1 also reveals the non-convexity of the resulting budget constraint when labour earnings are high enough for social benefit payments to cease.

4 Data

We use the 1998 wave from the German Socio-Economic Panel (GSOEP), a representative panel data sample of households and individuals living in Germany. The GSOEP started in 1984 with annual interviews in the Federal Republic of Germany and was extended to the former Germany Democratic Republic from 1990 on. The panel gives a wealth of information on the labour market status of individuals and on the various income sources of families.

We selected German nationals aged between 25 and 55 years. All are employees with a contractual labour supply of at least 10 hours per week or individuals who are voluntarily out of employment. The restriction on hours

which are both related to former earnings. Both transfers actually require the search of, and the willingness to take up, a job.

[5]The maximum social allowances we apply for both parts of Germany are based on the average effective maximum social benefits paid in 1998 (see Statistisches Bundesamt, 2001).

Table 1: Simplified tax-benefit system for household taxation, Germany 1998 singles (married couples)

Taxable income

earnings

capital income

income from rental and leasing

maintenance payments from ex-partner

Tax reliefs

DM 2,000 (4,000) standard deduction for earnings[(a)]

DM 6,000 (12,000) standard deduction for capital income

DM 108 (216) standard special expense deduction

child allowance (or child benefit alternatively)

exemption for social security contribution ('Vorsorge')

maintenance payments to ex-partner

Tax base: Taxable income - Tax reliefs

Tax schedule: Tax rate applied to (half) the tax base:

Income (X) bracket	Income tax liability
0-12,365	0
12,366-58,643	(91.19·Y+2,590)·Y
58,644-120,041	(151.96·Z+3,343)·Z+13,938
>120,041	0.53·X-22,843

X=rounded taxable income, Y=(X-12,312)/10,000, Z=(X-58,590/10,000)

solidarity supplement: tax scaled up by a factor of 1.055 [(b)]

Net Income: gross income - (twice) the tax liability

Benefits

child benefit: DM 220 for 1st and 2nd child, DM 300 for 3rd,
DM 350 from 4th child on (or child allowance alternatively)

means tested social benefits (incl. housing benefit and special payments):[(c)]
DM 1,000 in the West, DM 900 in the East.[(d)]

means tested social benefits for partner and children, depending on age

Maintenance

maintenance payments to children, ex-partners
or parents outside the household

Disposable income: net income + benefits - maintenance[(e)]

Notes: *(a)* The tax scheme is given in DM because of the non-linearity of the tax function. Since 2000, the tax scheme is given in DM and euro by the Federal Government. 1 euro = 1.95583 DM. Time unit is the year. *(b)* The solidarity supplement for the reconstruction of East Germany ('Solidaritätszuschlag') is based on a measure of taxable income that includes the child allowance whether or not parents opt for it. *(c)* For lack of information on the stock of savings etc., we assume that couples reporting more than DM 600 capital income or more than DM 4,800 rental income per year are not eligible for social benefit payments. *(d)* These numbers are based on the average effective maximum social benefit paid in 1998 Statistisches Bundesamt (2001). *(e)* Social security contributions, although largely compulsory, are taken as consumption expenditures and are not deducted from disposable income. Admittedly, the different types of social security contributions paid in Germany have different consumptive aspects, and our assumption is probably more appropriate for payments to the pension system than for health insurance contributions.

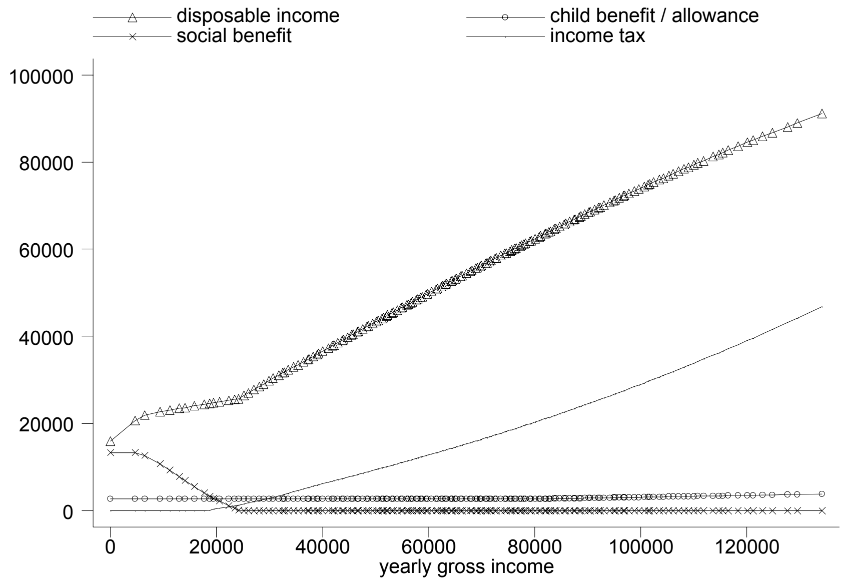

Figure 1: The 1998 German tax-benefit system.
The figure illustrates the situation of a couple with two children. The wife and the husband earn 18 and 25 euro per hour, respectively. At low levels of labour income they are eligible to means-tested social benefit.

is introduced to avoid the occurrence of extraordinarily high wage rates as the ratio of earnings over hours for people with less than 10 hours. We excluded the self-employed from the sample, as well as individuals in parental leave or in education or registered as unemployed.

The sample of singles consists of 488 individuals: 208 women and 280 men. A 'single' is defined as an individual with the above characteristics and living alone. He or she may have dependent children living outside the household. We also selected 1332 families composed of a married couple and, possibly, dependent children.[6] As for singles, dependent children may live outside the household.

Table 2 shows descriptive statistics for single and married women and men, respectively. The regional indicator East Germany should capture social, educational, economic and cultural differences between both parts of the country, at least partially. Given our selection, single women from Eastern Germany are under-represented in the sample (approximately 20% of the population live in the East, which is correctly reflected for men). In general singles are characterised by a younger age, higher schooling level, more university graduates and a remarkably higher participation rate, particularly among women. Single men have on average a higher level of vocational education than women: while 13% of the women have no vocational training, the corresponding figure for men is only 6%. This concerns essentially West Germans, but the proportion of people with polytechnic or university degree is lower in the East. Wives have on average a lower schooling and vocational education level than husbands. For example 15% of the husbands in the sample have a polytechnic or university degree, but only 8% of the wives. The main financial resource of single men and women is earnings. The unearned income can be negative because of maintenance payments to children, parents or ex-partner.[7] Men pay on average more maintenance than women. A high proportion of singles has positive capital income. In contrast, only few of the singles in our sample receive social benefits.

More than half of the married couples live with at most 1 child. Similarly to singles, only few couples receive housing or social benefits. Most of them have capital income. As is the case for singles, married men pay on average a little more maintenance than married women.

There are no large discrepancies in the distributions of hours of work between men and single women. Both present a sharp mode around 40 hours. But the distribution of hours of married women is very different.

---

[6]This means that we excluded households conmprising other adults than those of the married couple.

[7]Admittedly, the fact that maintenance payments may depend on income creates an endogeneity problem, which we shall neglect here.

Married women have a significantly lower participation rate than married men or singles, and the distribution of their labour supply is spread more evenly. Married women work more often in part-time jobs, and their weekly working time distribution has a mode at 20 hours.

5 A 'collective' baseline situation for Germany

To obtain a data set representing the collective world for Germany, we proceed as follows. In a first step we estimate preference parameters for single men and women and then predict their labour supply and the corresponding consumption level. This involves also the estimation of wage equations. In the second step we use a calibration method to determine the partners' relative weights in the bargaining processes taking place in the household, and to compute a leisure interaction coefficient describing the effect of the leisure of the spouse on each individual's utility. The calibration is done by optimising the fit of predicted to observed hours of work. By introducing a cross leisure term we relax the strong assumption of separability of individual preferences in the pairs (c_f, l_f) , (c_m, l_m) which is usually made in the empirical literature on collective models. Apart from this cross leisure term we assume that married individuals have the same preference parameters as singles.

We resort to calibration in order to avoid the difficult task of estimating a collective model with non-convex budget sets and non-participation, but the path followed here could eventually be extended from calibration to estimation, the crucial identifying assumption being the similarity of preferences of individuals before and after marriage. In the simpler context of two earner households and linearised budget restrictions, this approach was also followed by Barmby and Smith (2001).

5.1 Step 1: Estimation of preference parameters for singles

5.1.1 Wage equations

In order to obtain wage rates for the whole population, including those not working, we estimate wage equations separately for women and men. We posit a linear normal selection model and use the maximum likelihood method for women and the two-step Heckman procedure for men.[8] We tried

[8]A linear normal selection model combines two latent linear regression models with jointly normal errors independent of the regressors. The observation rule states that only

Table 2: Descriptive statistics on the selected samples

	single w.		single men		mar. w.		mar. men	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
participation	.94		.96		.71		.97	
age	37.2	9.44	36.7	7.98	39.5	7.04	41.9	7.34
schooling$^{(a)}$								
no degree	.03	$^{(e)}$	.04		.02		.01	
short secondary	.34		.35		.41		.31	
long secondary	.23		.21		.12		.19	
other	.03		.02		.06		.06	
voc. training$^{(b)}$								
no training	.13		.06		.10		.05	
technical training	.21		.18		.24		.25	
univ. or polytech.	.14		.18		.08		.15	
East Germany	.14		.20		.28		.28	
separ. or divorced	.27		.32					
never married	.66		.68					
widowed	.07		.00					
# children	0		0		1.40	1.06	1.40	1.06
no children	1		1		.22		.22	
1 child	0		0		.31		.31	
2 children	0		0		.36		.36	
$\geq$ 3 children (#)$^{(c)}$					3.39	.28	3.39	.28
hours work (week)	36.8	5.65	38.9	6.40	30.1	10.4	39.6	5.53
gross wage (hour)	13.1	4.75	14.6	5.75	11.6	5.17	15.6	6.56
capital income	40.1	72.1	42.4	74.9	52.1	143.5	52.1	143.5
child benefit	131.2	45.9	147.6	64.2	206.0	124.8	206.0	124.8
total unearned inc. (month)$^{(f)}$	49.5	195.5	30.4	234.2	321.9	459.1	321.9	459.1
observations	208		280			1332		1332

Notes: *(a)* reference category: primary school; short and long secondary school correspond to 'Realschulabschluß' and 'Abitur'; *(b)* reference category: apprenticeship; technical training and polytechnic correspond to 'Fachschule' and 'Fachhochschule'; *(c)* for this and the four subsequent variables, statistics concern only positive values; *(d)* nominal variables in euro; *(e)* variables for which only the mean is shown are indicator variables; *(f)* total unearned income also includes rental and leasing, social benefit, housing benefit, incomes of children and net maintenance payments; highly negative net maintenance may lead to negative total unearned income.

Table 3: Wage equations for singles

	women			men		
	coef.	std.err.	t-value	coef.	std.err.	t-value
East	-.35	.06	-5.6	-.41	.05	-8.6
$a40$	.37	.12	3.0	.01	.002	4.2
$a40\_child$	-.14	.06	-2.3			
sch_reabi	.11	.04	2.5			
job_noap	-.16	.04	-3.7			
job_fach				.17	.05	3.0
job_uni	.21	.06	3.4	.28	.06	4.5
constant	2.22	.11	19.8	2.22	.10	22.8
(uncensored) obs.		178			222	
	-76.03 (log likelihood)			.31 (adj. R-squared)		

Notes: dependent variable: log hourly wage rate; *East*: person lives in Eastern Germany. $a40$: age divided by 40. $a40\_child$: interaction $a40 \times child$ where child indicates that a woman has born at least one child. sch_reabi: at least middle secondary school. job_noap: no vocational training. job_fach: technical training. job_uni: polytechnic or university degree.

a number of different estimation methods, including also two step methods with other regressors than the predicted normal hazard used in the Heckman approach, but the choices mentioned above gave the most accurate predictions for working singles.

The results reported in Table 3 were obtained with a descending specification search. The variable ensuring identification of the selection process is gross unearned income, which is significant in the participation equation for both men and women. The results indicate that wages are on average lower in East Germany. Mothers have lower wages too, even though the children do not live in the household in our sample. This effect may be related to the negative impact of career interruptions (see e.g. Beblo and Wolf, 2002). Not surprisingly, a higher level of education has a positive impact on wages. In Table 4 we see that the estimated gross hourly wage of German single men is on average almost one euro higher than for single women.

the sign of the first variable is observed, and that the second variable is observed only when the first one is positive. Regressors are always observed.

Table 4: Wages for singles

	no.	mean	sd.d.	min.	10%	50%	90%	max.
w_f	208	13.29	5.46	6.02	8.88	12.10	17.84	56.29
w_m	280	14.32	5.52	3.42	8.59	13.11	21.69	35.67

Notes: w_i represents i's wage rate, with $i = f, m$. These are the actual wage rate for those people the information in available in the GSOEP, and predictions for all other observations, obtained taking account of the estimated variance of the log wage.

5.1.2 Preferences of single women and men

We estimate preference parameters separately for women and men. We assume linear expenditure system (LES)-type preferences:

$$U_i\left(c_i, l_i\right) = \beta_c^i \ln\left(c_i - \bar{c}_i\right) + \beta_l^i \ln\left(l_i - \bar{l}_i\right) \qquad i = f, m, \tag{1}$$

where c_i represents consumption (i.e. disposable income in this static model) and l_i demand for leisure. $\bar{c}_i$ and $\bar{l}_i$ are respectively the 'minimum' requirements in consumption and leisure. Instead of estimating these, which proved difficult, we chose to calibrate them, as done also by Barmby and Smith (2001). Details are given in the next subsection. We do not impose the constraint $\beta_c^i + \beta_l^i = 1$ in the estimation, but check that the estimates are positive, which allows a posteriori to rescale the utility function by $\beta_c^i + \beta_l^i$. An alternative specification with $\beta_c^i = F(z_i\gamma)$, where F denotes the logistic cumulative distribution function, and $z_i\gamma$ a linear index depending on characteristics for individual i, used for instance by Hoynes (1996), led to much lower likelihood values.[9]

The budget constraint is defined as:

$$c_i = g\left(l_i, w_i, y_i, \phi_i\right) \qquad i = f, m, \tag{2}$$

where w_i and y_i are i's wage rate and i's unearned income, respectively, ϕ_i represents a vector of characteristics relevant to the tax system, and the function g expresses the German tax and redistribution system.

For the estimation, we use a multinomial logit model with mass points on the consumption coefficients in order to account for unobserved heterogeneity (see Heckman and Singer, 1984, and Hoynes, 1996).[10] We suppose that each

[9] It is worth noting that (a) the two specifications are non-nested, and (b) the cardinalisation associated with division by $\beta_c^i + \beta_l^i$ depends on individual characteristics.

[10] We also estimated a random parameter logit model (RPL, see e.g. Mc Fadden and

person has K alternative values h^k for his/her weekly labour supply, leading to leisure choices $l^k = T - h^k$, where T is the total time available: 168 hours a week. We choose $K = 7$, and the following set of possible values for h^k: 0, 10, 20, 30, 40, 50, 60.

The contribution to the likelihood for person i choosing combination $\left(c_i^j, l^j\right)$ is:

$$L = \sum_{r=1}^{R} p_r \frac{\exp\left[\beta_{cr}^i \ln\left(c_i^j - \overline{c}_i\right) + \beta_l^i \ln\left(l^j - \overline{l}_i\right)\right]}{\sum_{k=1}^{K} \exp\left[\beta_{cr}^i \ln\left(c_i^k - \overline{c}_i\right) + \beta_l^i \log\left(l^k - \overline{l}_i\right)\right]} \tag{3}$$

where R denotes the number of mass points, p_r the probability associated with mass point (or regime) r in the mixture. Estimation results will be produced for three mass points: this appears sufficient for our data, given the heavy dominance of one of the regimes for both preference estimations, single women as well as single men.

In order to ensure that the probabilities p_r do lie between 0 and 1, we adopt the following 'logit' parameterisation:

$$\begin{aligned} p_r &= \exp(e_r) / \left[1 + \textstyle\sum_{s=1}^{R-1} \exp(e_s)\right] \qquad r = 1, \ldots, R-1, \\ p_R &= 1 - \textstyle\sum_{s=1}^{R-1} p_s. \end{aligned} \tag{4}$$

5.1.3 Calibration of minimum consumption and leisure for singles

The minimum consumption, $\overline{c}_i$, is calibrated as the lowest disposable income $(\widetilde{c}_i)$ over the hours choices listed above, over the whole sample, minus $c_i^0 = 1$. The latter number was obtained by grid search for the value of c_i^0 that maximises the likelihood. We have $\overline{c}_f = \widetilde{c}_f - c_f^0 = -118.61$ and $\overline{c}_m = \widetilde{c}_m - c_m^0 = -144.57$ for women and men, respectively. The minimum consumption is negative because of maintenance payments, even if the actual disposable income corresponding to the optimal or observed choice is positive. Minimum leisure, $\overline{l}_i$, is set uniformly to 94 hours a week for women ($\overline{l}_f = 94$), that is 10.2 hours of daily physiological regeneration and almost 3.3 hours for household work. It is set to 87 hours for single men ($\overline{l}_m = 87$) weekly, that is 10 hours a day of physiological regeneration and 2.4 hours for household work. These numbers refer to results drawn from the German Time Budget Survey 1991/1992 (see Beblo, 2001, and Statistisches Bundesamt, 1995).

Train, 2000) with a normal distribution for the constant terms in β_c^i and β_l^i. We obtained significant dispersion for the consumption term, but not for the leisure term, both for men and women. The specification with mass points on β_c^i alone strongly dominated the RPL specification, both in terms of likelihood and in terms of accuracy of predictions.

Note that an ordered model would be out of place here, because of the complex budget restriction.

Table 5: Mixed multinomial logit estimates for single women: three mass points

		coef.	robust std.err.	t-value
β^f_{l0}	$\ln\left(l_f - \bar{l}_f\right)$	38.29	4.11	9.3
β^f_{lsra}	$\ln\left(l_f - \bar{l}_f\right) \times sch\_reabi$	4.69	1.24	3.8
β^f_{lE}	$\ln\left(l_f - \bar{l}_f\right) \times East$	-20.31	5.16	-3.9
β^f_{cK}	$\ln\left(c_f - \bar{c}_f\right) \times kid$	-10.58	3.41	-3.1
β^f_{cjn}	$\ln\left(c_f - \bar{c}_f\right) \times job\_noap$	6.48	3.75	1.7
β^f_{cE}	$\ln\left(c_f - \bar{c}_f\right) \times East$	-22.88	11.1	-2.1
β^f_{c1}	$\ln\left(c_f - \bar{c}_f\right)$, regime 1	31.25	5.64	5.5
β^f_{c2}	$\ln\left(c_f - \bar{c}_f\right)$, regime 2	75.99	8.16	9.3
β^f_{c3}	$\ln\left(c_f - \bar{c}_f\right)$, regime 3	13.96	3.34	4.2
e_1	'logit' regime 1	-1.44	0.73	-2.0
e_2	'logit' regime 2	2.55	0.24	10.5
log lik. $R = 3$			-170.91	
log lik. $R = 2$			-172.44	
log lik. mult. logit			-287.97	

Notes: see Table 3. kid has value 1 if the woman has born a child, pays maintenance to a child, or receives child benefit. e_1 and e_2 parameterise the probabilities of the three mass points (regimes). See equation (4).

5.1.4 Estimated preference parameters for singles

Tables 5 and 6 report maximum likelihood estimation results based on equation (3) for three mass points for single women and men. These results were obtained using analytical gradients and hessian.[11] The last row of each table shows the log likelihood value obtained with the multinomial logit model, which does not allow for unobserved heterogeneity, and for two and three mass points. The improvement when allowing for unobserved heterogeneity is very large, especially for women. Moving from two to three mass points leads to a significant improvement in both cases, but a larger one for men.

The interpretation of the coefficients is not straightforward, since the propensities to consume and to demand leisure are related to them in a non-linear way. These propensities are obtained as $B^i_j = \beta^i_j / \left(\beta^i_c + \beta^i_l\right)$ for

[11] Obtaining these estimates proved difficult. We used sequentially numerical gradients, then analytical gradients, and finally the analytical hessian also, and iterated until estimates were identical with all three methods.

Table 6: Mixed multinomial logit estimates for single men: three mass points

		coef.	robust std.err.	t-value
β^m_{l0}	$\ln\left(l_m - \bar{l}_m\right)$	10.70	5.62	1.9
β^m_{la}	$\ln\left(l_m - \bar{l}_m\right) \times \ln(age)$	4.52	1.55	2.9
β^m_{ljn}	$\ln\left(l_m - \bar{l}_m\right) \times job\_noap$	-13.70	4.94	-2.8
β^m_{lE}	$\ln\left(l_m - \bar{l}_m\right) \times East$	-17.45	4.01	-4.4
β^m_{cjn}	$\ln\left(c_m - \bar{c}_m\right) \times job\_noap$	-18.24	8.32	-2.2
β^m_{cE}	$\ln\left(c_m - \bar{c}_m\right) \times East$	-24.96	7.25	-3.4
β^m_{c1}	$\ln\left(c_m - \bar{c}_m\right)$, regime 1	93.09		
β^m_{c2}	$\ln\left(c_m - \bar{c}_m\right)$, regime 2	44.03	2.58	17.1
β^m_{c3}	$\ln\left(c_m - \bar{c}_m\right)$, regime 3	11.71	1.99	5.9
e_1	'logit', regime 1	-.45	.38	-1.2
e_2	'logit', regime 2	2.71	.21	12.5
log lik. $R = 3$			-322.03	
log lik. $R = 2$			-340.62	
log lik. mult. logit			-405.13	

Notes: see Table 5. *age* is age in years. No variance was obtained for parameter β^m_{c1} but the improvement over $mp = 2$ in both the log likelihood value and the accuracy of predictions was sufficient to let us prefer estimates obtained with three mass points; results for other coefficients are close enough to those obtained for two mass points to be trusted.

Table 7: Marginal propensities for singles

	no.	mean	std. dev.	min.	10%	50%	90%	max.
B_c^f	208	.62	.10	.19	.42	.64	.66	.75
B_l^f	208	.38	.11	.25	.32	.36	.58	.81
B_c^m	279	.64	.09	.29	.61	.63	.71	.90
B_l^m	279	.36	.10	.10	.22	.37	.39	.71

Note: One of the men had a negative predicted propensity to consume and was ignored in the sequel.

Table 8: Elasticities for single women

	no.	mean	s.d.	min.	10%	50%	90%	max.
price (c)	208	-1.15	.15	-1.90	-1.20	-1.11	-1.08	-1.03
income (c)	208	2.40	.33	1.94	2.17	2.30	2.61	4.06
income (h)	208	-2.01	1.52	-13.17	-2.36	-1.65	-1.51	-1.06
wage (h)	208	.30	.23	-.21	.18	.26	.42	2.33

Note: c denotes the consumption and h the working hours.

$i = f, m$, and $j = c, l$, and their distributions are described in Table 7.

We use the propensities reported in Table 7 to compute elasticities at observed hours for the predicted regime (see next subsection), by linearising the budget constraint at that point for each individual. The elasticities are given in Tables 8 and 9.

While the price and wage elasticities are not unusual compared to the literature, the income elasticities are large in absolute value, possibly because of our inclusion of maintenance payments in unearned income.

Table 9: Elasticities for single men

	no.	mean	s.d.	min.	10%	50%	90%	max.
price (c)	279	-1.15	.13	-2.00	-1.19	-1.13	-1.08	-1.04
income (c)	279	2.16	.29	1.43	1.96	2.14	2.33	3.81
income (h)	279	-1.73	.74	-5.97	-2.10	-1.74	-.92	-.37
wage (h)	279	.37	.27	.01	.20	.32	.48	2.63

Note: see Table 8.

Table 10: Frequencies and estimated probabilies for the regimes

	women		men	
regime	est. prob.	frequency	est. prob.	frequency
1	.02	.08	.04	.14
2	.91	.88	.90	.82
3	.07	.04	.06	.04

Notes: the estimated probabilities result from the estimation procedure. The frequencies correspond to the regime which gives the best labour supply prediction.

5.1.5 Predicted hours for singles

In the computation of the optimal predicted labour supply of each single in the sample, for each individual we choose the regime (or mass point) r which gives the best prediction. When several regimes give predictions with the same absolute deviation from the observed labour supply, we keep the regime which yields normalised marginal propensity B_c^i (see definition above) closest to 0.5, in order to avoid the ocurrence of values near 0 or 1. Table 10 reports the resulting frequencies for the 'chosen' regimes.

From the estimates for e_1 and e_2 we see (but this can be read more clearly from Table 10) that regime 2, which corresponds to the highest value of β_c^f and to the second highest value of β_c^m, has the highest probability in both cases (men and women).

The regime frequencies obtained are not in stark contradiction with the estimated probabilities, although regime 1 appears to be chosen too often. Tables 11 and 12 show cross tabulations of predicted discretised hours (columns) against actual discretised hours (rows).

For both men and women, the observed marginal distribution of hours worked is fairly accurately reproduced, with the exception of some non-working men predicted to work full-time.

5.2 Step 2: Calibration of the cross-leisure term and the power index

Once we have obtained estimates of the preference parameters for women and men, we calibrate the couple-specific parameters, that is the power index and the cross-leisure term, since these parameters cannot yet be estimated.

Table 11: Actual versus predicted labour supply for single women

	0	10	20	30	40	50	60	total
0	**5.29**				.48			5.77
10		**.96**	.96					1.92
20	.48	.48	**.96**		.48			2.40
30			2.40	**.96**	**4.33**			7.69
40		1.92		1.92	**76.44**	.96		**81.25**
50					.96			.96
60								.00
total	5.77	3.37	4.33	2.88	**82.69**	.96	.00	208

Notes: rows: actual (discretised) labour supply, columns: predicted labour supply. Entries in the body of the table and in the margins give frequencies (in %), except the last cell which gives the number of observations. The proportion of exact predictions is 84.61%, and the proportion of 'adjacent' predictions is 12.03%. The bad predictions (the rest) represent 3.36% of all observations.

Table 12: Actual versus predicted labour supply for single men

	0	10	20	30	40	50	60	total
0	**2.15**	.36			1.43	.36		4.30
10								.00
20		1.43	**.36**	.36	1.79			3.94
30		.36			2.51	.36		3.23
40			.36		**67.74**	8.96	1.79	**78.85**
50					.36	**1.43**	4.30	6.09
60						1.08	**2.51**	3.59
total	2.15	2.15	.72	.36	**73.83**	12.19	8.60	279

Notes: see Table 11. The proportion of exact predictions is 74.19%, and the proportion of 'adjacent' predictions is 19.36%. The bad predictions (the rest) represent 6.45% of all observations.

5.2.1 Wage equations

As for singles, we estimate wage equations separately for wives and husbands. The following conceptual difficulty arises here due to selectivity: a participation model would need to be based on the collective framework, which is difficult. Lewbel (2000) proposes an estimation method for the selection model which does not require the specification of the selection mechanism. The method relies on the existence of a variable which is monotonically related with the selection variable: in the case of participation, household unearned income is a plausible candidate. For the wives, we use the simplest of the estimators proposed by Lewbel. For men we apply OLS, as the selectivity problem is much less severe for them, so that the OLS predictions are more accurate than those based on the Lewbel estimator. Results are gathered in Appendix A.

5.2.2 Preferences of individuals in couples

We assume that a person, once married, retains basically the same preferences as in the single status. An additional term in the utility function is the interaction in log leisures. We thus assume the following functional form for the utility function of spouse i in household h, parameterised by a vector z_h of variables related to children in the household, and by the spouse's leisure l_j:

$$\begin{aligned} U_i\left(c_i, l_i; z_h, l_j\right) &= \beta_c^i \ln\left[c_i - \bar{c}_i\left(z_h\right)\right] + \beta_l^i \ln\left[l_i - \bar{l}_i\left(z_h\right)\right] \\ &+ \delta_i\left(z_h\right) \ln\left[l_f - \bar{l}_f\left(z_h\right)\right] \ln\left[l_m - \bar{l}_m\left(z_h\right)\right] \quad \forall\, i, j = f, m, \end{aligned} \tag{5}$$

where $\delta_i\left(z_h\right)$ represents the cross leisure effect on the spouses' utilities. The presence of δ means that we do not restrict our attention to 'egoistic' or 'caring' agents (see Chiappori 1988). $U_i\left(c_i, l_i; z_h, l_j\right)$ is increasing in c_i and l_i if

$$\begin{aligned} \beta_c^i &> 0, \\ \beta_l^i + \delta_i\left(z_h\right) \ln\left[l_j - \bar{l}_j\left(z_h\right)\right] &> 0. \end{aligned} \tag{6}$$

This utility function is concave in $\left(c_i, l_i\right)$ if it is increasing in these arguments. As mentioned above, we assume that the β coefficients are unaffected by marriage and obtain them from the estimates for singles. In a first step we will calibrate a δ^h for each household (temporarily adopting the restriction $\delta_f\left(z_h\right) \equiv \delta_m\left(z_h\right)$) and a 'power index' ω_m^h which will be defined precisely in Subsection 5.2.4, so as to optimize the fit of predicted hours to observed hours. In the same way as for the singles, we calibrate the minimum requirements in consumption and leisure $\bar{c}_i\left(z_h\right)$ and $\bar{l}_i\left(z_h\right)$.

5.2.3 Calibration of minimum consumption and leisure for couples

In order to account for the impact of children on time use, as well as on consumption, we calibrate minimum expenditures and leisure times for couples, using equivalence scales as used for social benefit payments in Germany. We thereby draw on information from the German Ministry of Labour. As to the minimum time needs of children of different ages we refer to the results of the German time budget survey already used for singles.

The minimum consumption, $\overline{c}$, is calibrated as for the singles: for all possible choices of working hours of the couple, the household's weekly disposable income is computed, and the sample minimum is taken ($\widetilde{c} = 292.52$ euro per week). To account for the minimum needs of children in the household, the minimum consumption for each partner, $\overline{c}_i$, is then set equal to half of $\widetilde{c}$ minus $c_i^0 = 153$ divided by the household equivalence number heq, i.e.:[12]

$$\overline{c}_i = \widetilde{c}/2 - c_i^0/heq. \tag{7}$$

The household equivalence number varies with the number and age of children, and thereby approximates the social benefit equivalence scale in Germany. Each child under age 7 is assigned a weight of 0.4, children between 7 and 15 are weighted by 0.5 and older children still living in the household are given a weight of 0.6. We set $c_i^0 = 153$ so that the father and mother of one ten year old child will both have subtracted $c_i^0/heq = 153/1.5 = 102$, instead of $c_i^0 = 153$ in the childless household. As a result, their minimum consumption is 51 euro higher per week, which corresponds to the increase in minimum needs caused by an additional child as officially assigned to low-income households by the governmental authorities (Statistisches Bundesamt, 2003). Note that minimum consumption in a household without children is negative, just as for singles.

The minimum amount of leisure in the absence of children in the household is set to 87 and 94 hours for husband and wife, respectively, as for singles. To account for some minimum requirement of household production associated with child care, these hours are altered depending on the age of children living in the household. According to the results of a study on the leisure time of German couples (see Beblo, 2001, and Statistisches Bundesamt, 1995), for a full-time working man household time increases by 0.8 hours per day if he is living together with at least one child up to age 6. For a full-time working woman this rise in time requirement amounts to about 2 hours. Therefore, in our calculation, the father's weekly minimum leisure is increased by 6 hours, the mother's by 14 hours, if at least one child up to

[12] The justification for the choice of the number 153 is given below.

age 6 is present. For older children, that is, if all children are aged between 7 and 12, 3 hours are added to the father's (7 times 0.4 hours per day) and 7 hours are added to the mother's minimum leisure (7 times 1 hour per day) – again as suggested by the respective changes of household time for this group in the German time use data.

5.2.4 Determining the Pareto frontier

In earlier stages of this work we obtained the optimal choice of hours associated with a (δ, ω) for each household by maximising a convex combination of individual utilities, ω being the parameter of the convex combination. This approach is adapted to the case where the household's Pareto frontier is concave to the origin (or the utility set is convex). But as a consequence of the nonconvexity of many budget sets, the utility sets of many of the concerned households turn out not to be convex (see e.g. Mas-Colell et al., 1995). Thus, in order to capture all behavioural possibilities, we must adopt another approach than the simple maximisation of a household social welfare function defined as a convex combination of individual utility functions. The procedure adopted consists in determining the Pareto frontier and, for a given δ, searching for that point of the frontier which corresponds best to observed behaviour.[13]

The details of our version of the procedure are as follows:

1. Define by $U_m^{\max}$ the maximum utility the husband can reach, considering all labour supply combinations of husband and wife, and all consumption shares with one decimal point between 0.1 and 0.9.[14] $U_m^{\min}$ is the husband's utility level when his wife has $U_f^{\max}$. $U_m^{\max}$ and $U_m^{\min}$ can be considered as the utility levels the husband reaches when he or his wife is in a dictatorial position, respectively.

2. Select $K+1$ points on the Pareto frontier with the following coordinates on the horizontal axis:

$$U_m^k = U_m^{\min} + \frac{k}{K}\left(U_m^{\max} - U_m^{\min}\right) \quad \text{for } k = 0, ..., K.$$

   We choose $K = 50$. A simple index of the male's bargaining power is then: $\omega_m = k/K$. Recall that this concept is central in our approach to the collective model.

[13] The procedure was elaborated jointly by the authors and by Frederic Vermeulen. See also the theoretical description in Varian (1992, p.562). However, to the best of our knowledge, graphs like those reproduced in Figures 2 and 3 are the first *empirical* examples in the literature.

[14] Moving to two decimal points barely changed the results.

3. For each point k, the wife maximises her utility U_f^k given her husband's U_m^k and the budget constraint:

$$\max_{c,p,l_f,l_m} U_f^k \tag{8}$$
$$s.t. \begin{cases} U_m \geq U_m^k \\ c \leq g\left(l_f, l_m, w_f, w_m, y, \phi\right) \end{cases},$$

where c is the couple's consumption.

4. The couple's choice, given a value of the power index ω_m, a value of cross leisure effect δ and values for the regimes r_f and r_m, is:

$$\left(l_f\left(\omega_m, \delta, r_f, r_m\right),\ l_m\left(\omega_m, \delta, r_f, r_m\right),\ c\left(\omega_m, \delta, r_f, r_m\right),\ p\left(\omega_m, \delta, r_f, r_m\right)\right),$$

where p represents the wife's consumption share, i.e. the part of c the wife consumes.

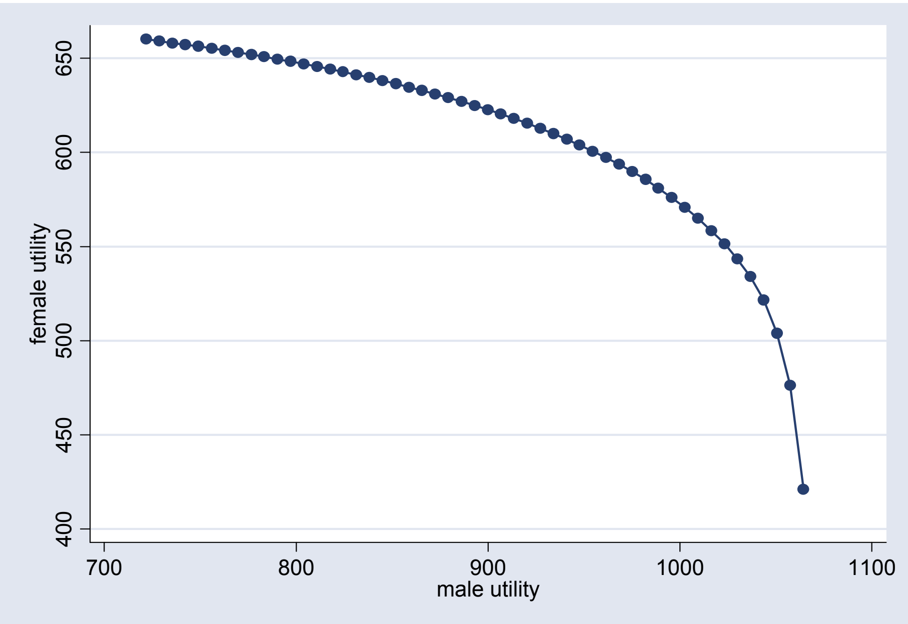

Figure 2: A convex utility set (U_m on the horizontal axis, U_f on the vertical axis)

Approximately only 42% of the households in our sample turn out to have a convex utility set, or, in other words, a concave Pareto frontier. Figures 2

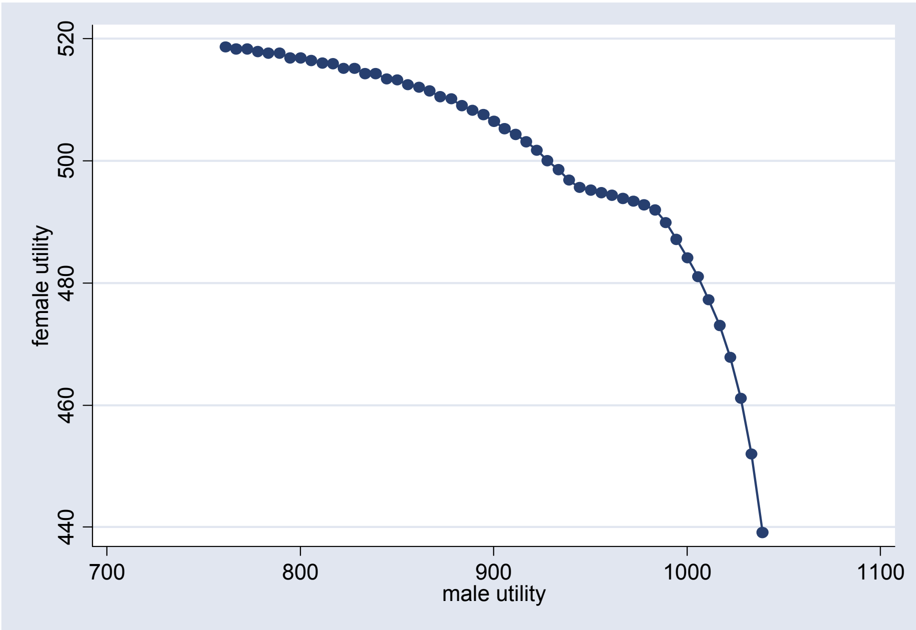

Figure 3: A non convex utility set (U_m on the horizontal axis, U_f on the vertical axis)

and 3 document two Pareto frontiers obtained in our sample, a concave one and non-concave one.[15]

5.2.5 Calibration of the cross-leisure term and the power index

Denoting W the set of possible values for the combination $(\omega_m, \delta, r_f, r_m)$, taking account of the restriction on δ ensuring regularity of preferences, i.e. condition (6), an optimal value of the combination satisfies:

$$\left(\overline{\omega}_m, \overline{\delta}, \overline{r}_f, \overline{r}_m\right) \in \underset{W}{\operatorname{argmin}} \left\{ \left[h_f^* - h_f\left(\omega_m, \delta, r_f, r_m\right)\right]^2 + \left[h_m^* - h_m\left(\omega_m, \delta, r_f, r_m\right)\right]^2 \right\}, \tag{9}$$

where h_i^* denotes i's actual labour supply. We only consider relative integer values of δ. In case of multiple solutions, we select the combination for which, lexicographically, $\overline{\omega}_m$ is closest to $\overline{\omega}_f$ (where $\overline{\omega}_f$ is the female's power index, defined as $\overline{\omega}_f = \left(\bar{U}_f - U_f^{\min}\right) / \left(U_f^{\max} - U_f^{\min}\right)$), $\overline{\delta}$ is closest to 0 and the sum of the regime probabilities estimated for the singles is the largest.

[15] The empirical Pareto frontiers (Figures 2 and 3) has been obtained in solving system (8) for each $k = 0, ..., K$ $(K = 50)$.

In order to normalise the power indices of the spouses so that they sum to one, while taking account of the degree of concavity of the Pareto-frontier at the optimum, we compute the variable α solving $\overline{\omega}_f^\alpha + \overline{\omega}_m^\alpha = 1$ and use $\overline{\omega}_m^\alpha$ as normalised power index for the male. The distribution of α is documented in Table 15.

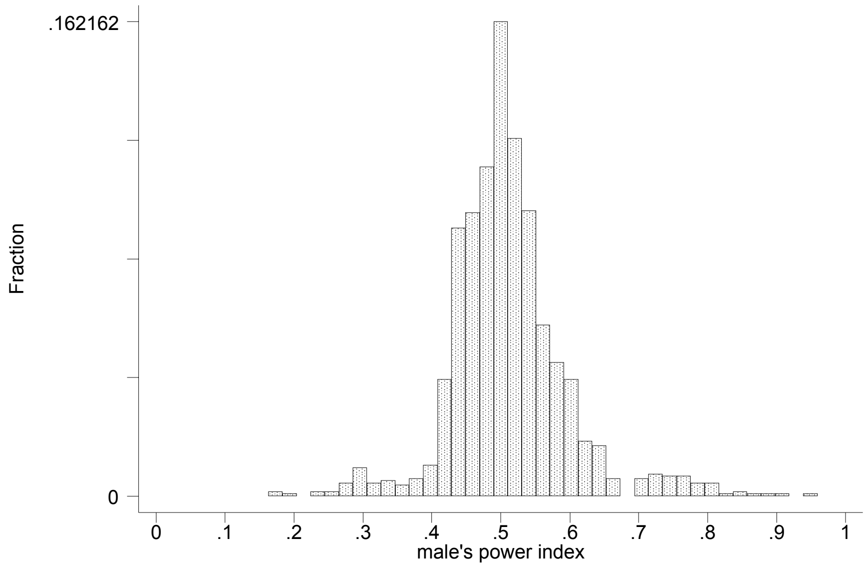

Figure 4: Distribution of the calibrated normalised power index of the man

In order to obtain hours predictions from the collective model, we solve the problem (8) for each household, for given preferences (including $\overline{\delta}$, $\overline{r}_f$ and r_m) and power index ($\overline{\omega}_m$).

5.2.6 Predicting the power index

We now turn to the estimation of a logistic equation relating the calibrated normalised power index $\bar{\omega}_m^\alpha$ to a set of explanatory variables, that is, an equation of the type $\ln \left[\bar{\omega}_m^\alpha / \left(1 - \bar{\omega}_m^\alpha\right)\right] = x\gamma + \epsilon$, which will allow us to obtain predicted values $\hat{\omega}_m^\alpha$ between 0 and 1 given x. Important variables to include in x are variables capturing the way in which the tax benefit system influences the relative earning potential of the spouses. If these turn out to contribute significantly to the prediction of $\bar{\omega}_m^\alpha$, they will allow us to describe changes in the power index induced by tax reforms. Here we consider two such variables, y_d^{40} and y_d^{20} defined as follows.[16] Let p_f^k and p_m^k denote the observed sample

[16] Thanks to Richard Blundell for this idea.

frequencies of (discretised) weekly labour supplies h^k of wives and husbands, respectively. Denote $R_{mk}^{fk'}$ the household disposable income when the husband works h^k hours and the wife works $h^{k'}$ hours. Variable y_f^{40} defined as:

$$y_f^{40} = \sum_{k=1}^{K} p_m^k \left(R_{mk}^{f40} - R_{mk}^{f0} \right) \tag{10}$$

then measures the expected increase in the household disposable income if the wife switches from 0 to 40 hours, the expectation being taken over the male hours distribution. Defining y_f^{20} and y_m^{40} similarly, we then consider the two ratios $y_d^{40}{=}y_f^{40}/y_m^{40}$ and $y_d^{20}{=}y_f^{20}/y_m^{40}$ which we term 'relative earning potential of the wife at 40 hours' (resp. 20).

The initial list of explanatory variables x we started with in the logistic regression included, beside y_d^{20}, y_d^{40} and the age difference variable reported in the table, income variables like capital income and child benefits, and a whole set of socio-demographic variables. But the only variables surviving the descending specification searches conducted separately for East and West Germany are given in Table 13. Using variables y_d^{20}, y_f^{20}, y_f^{40} and y_m^{40} instead or on top of y_d^{40} proved inferior. An optimistic interpretation of the result that basically only the relative earnings potential variable plays a role, would be that observable heterogeneity has been satisfactorily dealt with in the wage and preference specifications.

5.3 Step 3: Re-calibration of the individual cross-leisure terms

Hours predictions obtained with the predicted $\hat{\omega}_m^{\alpha}$ and calibrated $\bar{\delta}$ turn out not to be very accurate. We thus choose to re-calibrate separate leisure interaction terms δ_m and δ_f for husband and wife given the predicted power index $\hat{\omega}_m^{\alpha}$. The couple solves the following program:

$$\max_{c,p,l_f,l_m} U_f \qquad s.t. \begin{cases} U_m \geq \tilde{U}_m \\ c \leq g\left(l_f, l_m, w_f, w_m, y, \phi\right) \end{cases}, \tag{11}$$

where $\tilde{U}_m$ is such that:

$$\tilde{U}_m = U_m^{\min} + \hat{\omega}_m \left(U_m^{\max} - U_m^{\min} \right),$$

and the utility function U_i is given as follows:

$$\begin{aligned} U_i &= \beta_c^i \log\left(c_i - \bar{c}_i\left(z_h\right)\right) + \beta_l^i \log\left(l_i - \bar{l}_i\left(z_h\right)\right) \\ &+ \delta_i\left(z_h\right) \log\left(l_f - \bar{l}_f\left(z_h\right)\right) \log\left(l_m - \bar{l}_m\left(z_h\right)\right) \quad \forall\ i = f, m. \end{aligned} \tag{12}$$

Table 13: Logistic regression results for the power index

	East Germany			West Germany		
parameter	coef.	s.e.	t-value	coef.	s.e.	t-value
y_d^{40}				-.16	.02	-6.20
age	-.01	.00	-2.81	-.01	.00	-2.89
dage				.01	.01	3.69
sch_real	.11	.04	2.98			
sch_abi	.13	.05	2.73			
job_noap	-.56	.15	-3.66			
job_fach				.06	.02	3.37
nchild	.06	.02	3.19			
*child*0	-.37	.08	-4.40	-.24	.02	-9.95
*child*3	-.16	.05	-3.05	-.11	.02	-5.49
*child*7	-.04	.03	-1.56			
*child*16				-.05	.02	-2.73
constant	-.01	.11	-.02	.22	.07	3.31
adj. R-squared		.126			.196	

Note: y_d^{40}: her relative earning potential at 40 hours, taking account of the tax system, as explained in the text; *age*: her age, *dage*: her age minus his age; *sch_real, sch_reabi, job_noap* and *job_fach*: indicator variables for short secondary schooling, long secondary schooling, no vocational degree, technical training, and polytechnic or university degree, respectively; *nchild*: number of children; *child0, child3, child7* and *child16*: number of children in the household less than 3 years old, between 3 and 6 years, between 7 and 12 years and over 16 years, respectively.

Table 14: Regime frequencies for couples after re-calibration

	regime 1	regime 2	regime 3
$\bar{r}_f$	.025	.880	.095
$\bar{r}_m$	.057	.938	.005

We allow for different cross leisure terms for both partners. Thus the couple's choice (l_f, l_m, c, p) now depends on the combination $(\delta_f, \delta_m, r_f, r_m)$. As for the first computation, the optimal combination $\left(\bar{\delta}_f, \bar{\delta}_m, \bar{r}_f, \bar{r}_m\right)$ satisfies

$$\left(\bar{\delta}_f, \bar{\delta}_m, \bar{r}_f, \bar{r}_m\right) \in \underset{W}{\operatorname{argmin}} \left\{ \left[h_f^* - h_f\left(\delta_f, \delta_m, r_f, r_m\right)\right]^2 + \left[h_m^* - h_m\left(\delta_f, \delta_m, r_f, r_m\right)\right]^2 \right\}. \tag{13}$$

Table 14 shows the new regime (or mass point) frequencies. The most frequently chosen regime is the second one for each spouse (like for singles). The results are very similar to those for the regimes picked in the first calibration, although the exotic regimes are chosen a little more often.

Table 15 documents the results of the re-calibration of $\bar{\delta}$ and $\hat{\omega}_m^{\alpha}$. The cross leisure effect $\bar{\delta}_f$ turns out to be negative for most women. This was not what we expected when introducing the leisure interaction term, we rather expected to find a majority of households with complementary leisures. The result that leisures are substitutes (in direct utility) for so many women may be partly due to the inappropriateness of the way in which we import the β coefficients estimated for singles into the preferences of individuals in couples and to the insufficiency of our efforts to distinguish between domestic production and leisure time. Further research is needed here. $\bar{\delta}_f$ increases with the number of children, the exception being that mothers of a single child present the largest average. The cross leisure effect for men, $\bar{\delta}_m$, is positive on average. As for women, it increases with the number of children, except that childless men have the same average coefficient as fathers of two children. In our model, the asymmetry in the labour supplies of husbands and wives thus translates into an asymmetry in the impact of the spouse's leisure on utility, rather than in asymmetry in bargaining power: indeed, the power index values obtained suggest a balanced power sharing in German households.

Table 15: Statistics on re-calibrated cross leisure effects and $\hat{\omega}_m$

	no child	1 child	2 child.	$\geq$ 3 child.	total			
	mean	mean	mean	mean	mean	s.d.	min.	max.
$\bar{\delta}_f$	-4.17	-3.34	-3.63	-3.43	-3.63	3.22	-12	6
$\bar{\delta}_m$	.21	.02	0.21	.62	.19	2.23	-8	6
$\hat{\omega}_m^{\alpha}$	.52	.49	.51	.53	.51	.08	.33	.93
no.	296	410	478	148	1332			

Table 16: Female actual vs. predicted labour supply using $\hat{\omega}_m$, $\overline{\delta}_f$ and $\overline{\delta}_m$

	0	10	20	30	40	50	60	total
0	**28.38**	.15						28.53
10	.30	**5.63**						5.93
20		1.43	**15.09**					16.52
30			2.10	**10.21**	.08			12.39
40				2.93	**33.18**	.08		**36.19**
50						**.30**		.30
60							**.15**	.15
total	28.68	7.21	17.19	13.14	**33.26**	.38	.15	1332

Note: rows: actual, columns: predicted.

5.4 Hours predictions and baseline situation for couples

Tables 16 and 17 show cross tabulations of actual versus predicted hours obtained with the estimated $\hat{\omega}_m$ and the calibrated $\overline{\delta}_f$ and $\overline{\delta}_m$.

The predictions are very accurate: both for wives and for husbands almost 90% of the observations are located on the diagonal. However, the fit deteriorates considerably if we use predicted rather than calibrated values of the cross leisure effects δ_i.[17] In view of these results, we decided to use the calibrated values of the leisure interaction terms, rather than their predictions, in the definition of the baseline situation to be used in simulating the effects of fiscal reforms.

[17] Less than 29% of the labour supplies of wives are correctly predicted. The percentage for the husbands is much higher, but still, we obain very bad predictions for some men: for example, 2.4% of the husbands are predicted to work full time although they actually do not work. These deviations may be due to factors we are not able to control for, such as health condition, or wealth.

Table 17: Male actual vs. predicted labour supply using $\hat{\omega}_m$, $\overline{\delta}_f$ and $\overline{\delta}_m$

	0	10	20	30	40	50	60	total
0	**3.15**							3.15
10		**.38**						.38
20		.08	**.23**					.30
30			.15	**1.50**				1.65
40			.08	8.56	**76.50**			**85.14**
50				.08	2.25	**3.30**		5.63
60					.75	2.85	**.15**	3.75
total	3.15	.45	.45	10.14	**79.50**	6.16	.15	1332

Note: see Table 16.

6 Analysis of a tax reform with the collective model

We analyse the welfare effects of the switch from joint to individual taxation. This reform entails a replacement of joint with individual taxation on the basis of the 1998 tax schedule. Tax liabilities are scaled down by a factor $f = .942$ in order to obtain revenue neutrality, at least approximately. The computation takes account of singles and of sampling weights. Specifically, denoting R_o the baseline government tax revenue, and $R(f)$ the post-reform tax revenue for factor f given behavioural adjustments, we solve equation $R(f) = R_o$ in f. This is a conceptually simple problem, but given the complexity of the function $R(f)$ it is numerically burdensome, and we chose to stop the iterative algorithm after a small number of iterations, yielding the solution up to the third decimal position.

It is important to note that the definition of a revenue neutral reform will differ if it is based on the unitary rather than on the collective model, because predicted behavioural adjustments will differ between the two models.

6.1 Simulation results

6.1.1 Changes in the power index

Recall that the predicted power index depends on a variable which reflects the wife's relative earning potential at 40 hours, taking account of the complete tax system. Thus the power index is potentially affected by a tax reform. Table 18 summarises its distribution for the two tax-benefit systems

Table 18: Estimated normalised power index for the two tax-benefit systems

	no.	mean	s.d.	min.	10%	50%	90%	max.
$\hat{\omega}_m^{\alpha joint}$	1332	.510	.076	.333	.428	.496	.609	.927
$\hat{\omega}_m^{\alpha ind}$	1332	.506	.076	.330	.426	.492	.604	.927

Note: $\hat{\omega}_m^{\alpha joint}$and $\hat{\omega}_m^{\alpha ind}$: normalised power index of male for 1998- and individual taxation, respectively.

we consider, based on the estimates given in Table 13.[18]

There is a slight reduction in the mean and in all quantiles of $\hat{\omega}_m^{\alpha}$, and Figure 5 shows that changes are small for all individuals, and that there are few increases. The reason for this change lies in their improved relative earning potential y_d^{40} connected to the fact that they typically have lower wages than their husbands. Individual taxation then lowers their marginal tax rate. Because of this shift of the bargaining position in favour of women, we expect the reform to be relatively more beneficial for married women than for married men.

6.1.2 Changes in tax revenues

Table 19 summarises tax revenues from different population subgroups under the two tax regimes. Marital status is an important determinant of the direction of change in tax liabilities. The individual tax reform is beneficial to the singles. Indeed the singles' tax liability is scaled down by the factor $f = .942$. On the other hand, the status quo (splitting system) is favourable to couples, and even when the spouses adapt their labour supply to the new situation of individual taxation, couples have higher tax liabilities under individual taxation.

6.1.3 Changes in participation and labour supply

Tables 20 and 21 compare labour supplies pre an post-reform for wives and husbands, respectively. The participation rate of married females increases from 71% to 81% under individual taxation.

Under individual taxation, the labour supply of married women becomes more concentrated on part-time jobs: 56% are predicted to work between 10

[18]These predictions are obtained using the baseline distributions of male and female hours.

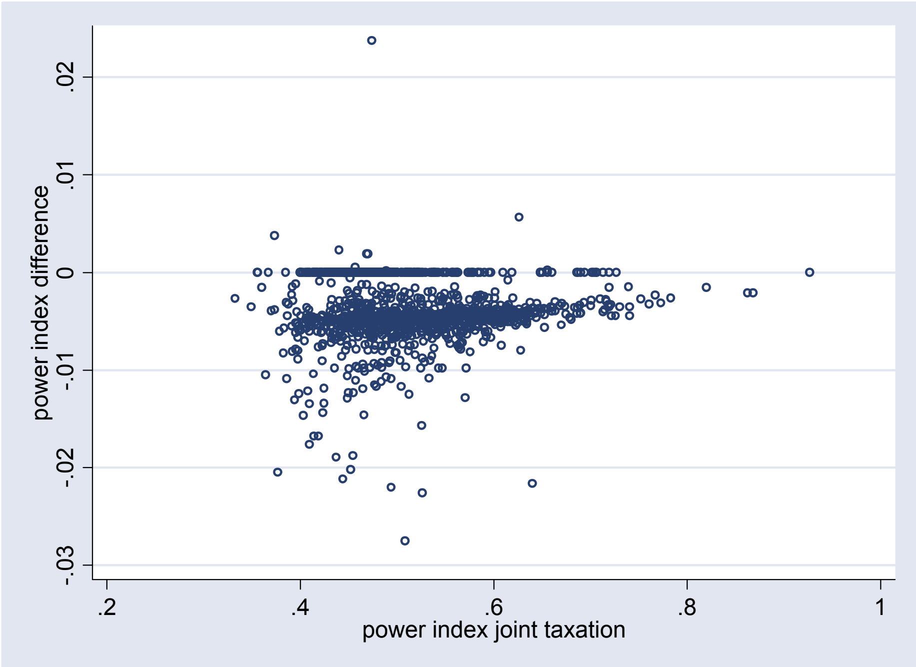

Figure 5: Difference in power index, after - before the reform

Table 19: Tax revenues in billion euro

	collective			unitary		
	joint tax	ind. tax	% change	joint tax	ind. tax	% change
single women	8.92	8.39	-5.9%	8.92	7.99	-10.4%
single men	15.58	14.68	-5.8%	15.58	14.01	-10.1%
couples	44.10	45.54	+3.3%	44.71	47.22	+5.6%
total	68.59	68.61		69.21	69.22	

Notes: approximate revenue neutrality is obtained by multiplying the tax liability in the case of the individual taxationby a factor $f = .942$ for the collective model and $f = .894$ for the unitary model. Tax revenues in billion euro

Table 20: Wives' labour supply pre- and post-move to individual taxation

	0	10	20	30	40	50	60	total
0	**17.64**	7.88	2.63	.38	.08	.08		28.68
10	1.13	**2.70**	2.55	.60	.15		.08	7.21
20	.15	4.65	**9.31**	2.40	.60		.08	17.19
30		.23	4.88	**6.38**	1.35	.30		13.14
40		.08	1.13	9.83	**18.99**	2.70	.53	**33.26**
50					.15	**.23**		.38
60						.08	**.08**	.15
total	18.92	15.54	20.50	19.59	**21.32**	3.38	.75	1332

Note: rows: pre-reform, columns: post-reform.

Table 21: Husbands' labour supply pre- and post-move to individual taxation

	0	10	20	30	40	50	60	total
0	**2.63**	.53						3.15
10		**.45**						.45
20		.08	**.38**					.45
30		.15	1.35	**8.48**	.15			10.14
40		.15	.08	7.28	**71.85**	.15		**79.50**
50					.45	**5.63**	.08	6.16
60							**.15**	.15
total	2.63	1.35	1.80	15.77	**72.45**	5.78	.23	1332

Note: see Table 20.

and 30 hours, against only 38% pre-reform. However the mode of the hours distribution remains at 40 hours.

Husbands react less to the reform than their wives do. The participation rates of married males remain almost the same in all situations (around 97%), but their labour supply decreases on average: e.g. about 7% are predicted to move from full time to 30 hours (see Table 21).

Table 22 documents joint variations in labour supply within the household. A salient feature of this table is that the most frequent cell is (0,0), for about half of the cases. For 3.3% of the households, there is substitution of his for her hours of work (his hours decrease and hers increase) while there is complementarity for 1.4%.

Table 23 summarises the labour supply and consumption effects of the reform, and Tables 24 and 25 show the change in the situation of households without children and with three children or more, respectively. The overall

Table 22: Δ labour supply of couples, joint - ind. tax.: wives vs. husbands (collective model)

	-30	-20	-10	0	10	total
-30			.08			.08
-20				1.50		1.50
-10			**.83**	19.89		20.72
0	.08	.23	5.18	**49.47**	.38	**55.33**
10	.08		2.18	14.26	**.38**	16.89
20			.60	3.90	.15	4.65
≥30			.30	.53		.83
total	.15	.23	9.16	**89.56**	.90	1332

Note: rows: wives, columns: husbands.

Table 23: Labour supply and consumption - couples

	joint taxation	individual taxation
c_f	596.2	613.0
h_f	21.7	22.2
p_f	.71	.81
c_m	349.2	310.7
h_m	38.1	37.2
p_m	.97	.97

Notes: *c, p* and h represent average consumption, participation, and hours given participation, for wives and husbands. 1332 observations.

observation is that women are better off with individual taxation in terms of consumption and men are worse off, no matter whether children are present or not.

Singles are less affected by the reform than couples. The corresponding tables are not shown to save on space.

7 Unitary model: estimation and predictions

For the specification of the unitary model, we adopt the analogue to the individual utility functions used in the collective model (equation (5)), that

Table 24: Labour supply and consumption - couples without children

	joint taxation	individual taxation
c_f	679.8	689.2
h_f	30.7	28.9
p_f	.86	.94
c_m	405.0	368.2
h_m	37.2	36.3
p_m	.96	.97

Notes: see Table 23. 296 observations.

Table 25: Labour supply and consumption - couples with more than 3 children

	joint taxation	individual taxation
c_f	518.5	539.9
h_f	10.7	13.5
p_f	.41	.58
c_m	344.3	291.6
h_m	37.0	35.4
p_m	.95	.96

Notes: see Table 23. 148 observations.

is:

$$\begin{aligned}\hat{U}\left(c,l_f,l_m;z\right) &= \beta_c\left(z\right)\ln\left[c-\bar{c}\left(z\right)\right] \\ &+ \beta_f\left(z\right)\ln\left[l_f-\bar{l}_f\left(z\right)\right]+\beta_m\left(z\right)\ln\left[l_m-\bar{l}_m\left(z\right)\right] \\ &+ \delta\left(z\right)\ln\left[l_f-\bar{l}_f\left(z\right)\right]\ln\left[l_m-\bar{l}_m\left(z\right)\right].\end{aligned} \tag{14}$$

Note that, by using aggregate consumption as argument in the utility function, rather than introducing the individual consumptions of husband and wife as separate argument, we place ourselves in the common situation of consumption surveys, where only aggregate consumption of the household is typically available.

Appendix B gives the conditions under which this utility function is increasing in its arguments and concave. Actually, we also experimented with direct translog utility functions along the lines of Van Soest (1995), but with a quadratic form in logs of departures from minimal requirements, as in equation (14). Although several specifications were superior to this one in terms of likelihood, all led to utility functions that were non-increasing in at least one argument for a majority of observations.

The β and δ functions of characteristics z are assumed to be linear in z, and the minimum requirements in consumption and leisure are set to the values calibrated for the collective model. Of course, the budget constraint remains the same as for the collective model.

Since each spouse has K labour supply choices, the couple has K^2 possible combinations. If $\hat{U}_j=\hat{U}\left(c^j,l_f^j,l_m^j;z\right)$ denotes the utility generated by combination j of the set of combinations $\left\{\left(c^j,l_f^j,l_m^j\right)_{j=1}^{K^2}\right\}$, adding an error term ε_j to the utility derived from combination j, we have:

$$\hat{U}_j=\hat{U}\left(c^j,l_f^j,l_m^j;z\right)+\varepsilon_j \qquad \forall\ j=1,...,K^2. \tag{15}$$

The distribution of ε_j is assumed to be the extreme value distribution defined by:

$$\Pr\left[\varepsilon_j<\varepsilon\right]=\exp\left(-\exp\left(-\varepsilon\right)\right), \qquad \varepsilon\in\mathbb{R}. \tag{16}$$

If combination j turns out to be the best possible choice for the family, we have:

$$\Pr\left[\hat{U}_j>\hat{U}_k,\ \forall\ k\neq j\right]=\frac{\exp\left[\hat{U}\left(c^j,l_f^j,l_m^j;z\right)\right]}{\sum_{k=1}^{K^2}\exp\left[\hat{U}\left(c^k,l_f^k,l_m^k;z\right)\right]}. \tag{17}$$

The above expression corresponds to the density function of the multinomial logit model. As for the singles, we also estimate a discrete mixture of such models, with two to three mass points on the coefficient of $\ln\left[c-\bar{c}\left(z\right)\right]$.

7.1 Estimation results

Table 26 presents estimation results for a model with three mass points. Although the coefficient corresponding to regime 3 is poorly determined, an LR test rejects the specification with two mass points, which itself rejects the MNL specification. However, the gains are lower here than for the estimation of the preferences of singles. The specification search started with a full set of houshold and personal characteristics for z in β_c and δ and with only own characteristics plus characteristics of children in β_f and β_m. Among the variables z we include information concerning the regimes 'chosen' in the calibration of the collective model (see Table 14). This can be seen here as a sort of 'observed unobservable heterogeneity', and the corresponding variables turn out to be highly significant (variables $reg3_f$ and $reg1_m$). This may explain why the gain of accounting for unobservable heterogeneity is lower here than for singles.

Given these parameter estimates, all couples present a positive marginal utility of the female's leisure. The concavity condition is fulfilled too. Table 27 gives statistics on the distribution of marginal utilities and of the cross leisure effect δ. Contrary to the collective model, the latter is positive on average, although it is negative for 3 couples. Living in Eastern Germany has a negative effect on this interaction term. Children have a negative impact on the consumption coefficient. This effect increases if the household has children less than 6 years old.

7.2 Base case predictions with the unitary model

We assume again that the regime corresponding to a couple is the one which gives the best labour supply predictions under the condition that it leads to an increasing utility function. Table 28 shows that the frequencies of the regimes are more or less in accordance with the estimated probabilities, although regime 2 is chosen more often than it should. The third regime is never chosen, as it never satisfies the positivity restrictions on the marginal utility of consumption.

Tables 29 and 30 show the predictions obtained for wives and husbands. The unitary model performs less well as the collective model in predicting labour supplies. Predictions are correct only for a third of the wives and for 45% of the husbands. The margins of the tables are not very well predicted, except as regards the participation rate. But results for cells within the tables are very bad. Some large discrepancies occur: for instance, over 3% of the wives are predicted to work fulltime although they actually do not work. The unitary model tends to smooth the distribution of the labour supply.

Table 26: Mixed multinomial logit estimates of preferences for couples: three mass points

		coef.	s.e.	t-value
β_K^f	$\ln\left(l_f - \bar{l}_f\right) \times kid$	5.27	1.15	4.6
β_{K6}^f	$\ln\left(l_f - \bar{l}_f\right) \times kid6$	1.01	.30	3.4
β_{sr}^f	$\ln\left(l_f - \bar{l}_f\right) \times sch\_real_f$	-.55	.25	-2.2
β_{ju}^f	$\ln\left(l_f - \bar{l}_f\right) \times job\_uni_f$	.92	.40	2.4
$\beta_{r_3^f}^f$	$\ln\left(l_f - \bar{l}_f\right) \times reg3_f$	5.02	.52	9.6
β_0^f	$\ln\left(l_f - \bar{l}_f\right)$	-.80	.70	-1.1
β_K^m	$\ln\left(l_m - \bar{l}_m\right) \times kid$	3.15	1.21	2.6
β_{ju}^m	$\ln\left(l_m - \bar{l}_m\right) \times kid6$	1.12	.31	3.7
β_0^m	$\ln\left(l_m - \bar{l}_m\right)$	1.40	.76	1.9
$\delta_{a_f^2}$	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times (\ln age_f)^2$	.12	.01	9.5
δ_K	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times kid$	-1.38	.31	-4.5
δ_{K6}	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times kid6$	-.34	.11	-3.0
δ_{sr_f}	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times sch\_real_f$	.56	.11	5.0
δ_{sa_f}	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times sch\_abi_f$	.49	.14	3.5
δ_{sa_m}	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times sch\_abi_m$	.15	.06	2.6
δ_{jn_m}	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times job\_noap_m$	-.18	.07	-2.6
δ_E	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times East$	-.56	.04	-12.4
$\delta_{r_3^f}$	$\ln\left(l_f - \bar{l}_f\right) \times \ln\left(l_m - \bar{l}_m\right) \times reg3_f$	-1.49	.13	-11.7
β_K^c	$\ln(c - \bar{c}) \times kid$	-10.82	1.92	-5.6
β_{K6}^c	$\ln(c - \bar{c}) \times kid6$	-3.77	1.19	-3.2
$\beta_{sr_f}^c$	$\ln(c - \bar{c}) \times sch\_real_f$	5.56	1.20	4.6
$\beta_{sa_f}^c$	$\ln(c - \bar{c}) \times sch\_abi_f$	5.28	1.46	3.6
$\beta_{r_3^f}^c$	$\ln(c - \bar{c}) \times reg3_f$	-16.94	1.25	-13.5
$\beta_{r_1^m}^c$	$\ln(c - \bar{c}) \times reg1_m$	4.21	.97	4.3
β_1^c	$\ln(c - \bar{c})$, regime 1	26.94	1.92	14.0
β_2^c	$\ln(c - \bar{c})$, regime 2	10.81	3.01	3.6
β_3^c	$\ln(c - \bar{c})$, regime 3	-22.39	44.67	-0.5
e_1	'logit', regime 1	6.36	.85	7.5
e_2	'logit', regime 2	1.58	1.05	1.5
log likelihood R=3			-4218.75	
log likelihood R=2			-4224.75	
log lik. multinomial logit			-4245.56	

Notes: kid and $kid6$ are indicator variables with value 1 if the couple has at least one child and one child less than 6 years old. sch_real_f, sch_real_m, sch_abi_f and sch_abi_m are indicator variables with value 1 if the wife (f) or the husband (m) have short and long secondary schooling. job_uni_f and job_noap_m are indicator variables with value 1 respectively if the wife has university or college degree and if the husband has no training. $East$ is an indicator variable with value 1 if the couple lives in East Germany. $reg3_f$ and $reg1_m$ are indicator variables with value 1 if respectively the third consumption regime was the best one for the wife and the first consumption regime was the best one for the husband in the calibration procedure. $(\ln age_f)^2$ is the square of the logarithmed wife's age in years.

Table 27: Marginal utilities and $\delta(z)$ coefficients

	mean	s. d.	min.	10%	50%	90%	max.
U_f	7.07	1.46	4.12	5.32	6.86	9.04	12.96
U_m	7.33	3.13	.82	3.83	6.53	11.82	19.96
U_c	20.06	10.99	.88	9.78	15.77	38.48	59.11
δ	1.92	1.43	-.23	.56	1.27	4.00	6.42

Note: 1332 observations.

Table 28: Regimes: estimated probabilities and predicted frequencies

	regime 1	regime 2	regime 3
estimated probability	.990	.008	.002
predicted frequency	.947	.053	

The mode on full time working is significantly lower for the unitary model, both for women and men: 80% of the husbands work full time, but only 49% are predicted to work 40 hours. The labour supplies are underpredicted on average. This points to the mis-specification of the model, at least concerning the particular unitary model estimated here, but possibly of any unitary model.

7.3 Tax reform

We adopt as baseline situation for the unitary model its own predicted labour supplies. Alternatively, we could have calibrated some preference parameters in order to obtain predictions coinciding with the collective baseline. Both approaches have advantages and drawbacks.

We start with the impact of the choice of the unitary representation on the definition of the reform itself. For the reform to be revenue neutral, we now have to multiply the tax burden by a factor $f = .894$. This already indicates that the unitary model overestimates the reaction of couples to the tax reform, with regard to labour supplies. The use of the unitary model to evaluate the tax revenues introduces distortions: while by construction the overall level is unchanged (because the reform was designed in both cases to be revenue neutral), the tax burden is predicted to shift from singles to couples in a much more pronounced way when using the unitary model (See Table 19). We now examine in turn the *positive* aspects of the reform (its impact on observable behaviour) and its *normative* aspects (its impact on

Table 29: Collective versus unitary female labour supply, joint taxation

	0	10	20	30	40	50	60	total
0	**15.99**	6.98	3.30	2.10	.30			28.68
10	4.13	**.90**	1.28	.90				7.21
20	6.76	3.75	**3.45**	2.48	.75			17.19
30	2.63	1.88	2.78	**2.93**	2.70	.23		13.14
40	3.30	2.63	3.90	9.23	**11.49**	2.70		**33.26**
50	.23			.08	.08			.38
60	.08			.08				.15
total	**33.11**	16.14	14.71	17.79	15.32	2.93	.00	1332

Note: rows: wives' collective labour supplies, columns: unitary.

Table 30: Collective versus unitary male labour supply, joint taxation

	0	10	20	30	40	50	60	total
0	**.45**	.60	.30	.23	1.35	.23		3.15
10		**.08**		.08	.23	.08		.45
20				.15	.30			.45
30	.30	.83	.45	**1.20**	4.58	2.70	.08	10.14
40	1.35	8.93	3.30	8.71	**40.47**	16.14	.60	**79.50**
50		.23	.08	.15	1.73	**3.68**	.30	6.16
60						.15		.15
total	2.10	10.66	4.13	10.51	**48.65**	22.97	.98	1332

Note: rows: husbands' collective labour supplies, columns: unitary.

Table 31: Participation rates

	collective model		unitary model	
	joint	individual	joint	individual
wives	71.33	81.09	76.89	84.68
husbands	96.85	97.38	97.90	98.87
single women	94.23	94.23	94.23	94.23
single men	97.85	97.85	97.85	97.85

welfare).

7.4 Positive aspects of the reform

Recall that the baseline situation used for the unitary model consists of the predictions obtained from that model. Table 31 compares participation rates before and after the reform, as predicted by the collective and the unitary models. Note that for singles the predictions for individual taxation only differ because of rescaling of the tax scheme (to ensure revenue neutrality). The largest discrepancies between collective and unitary predictions are obtained for wives.

Table 32 compares collective and unitary labour supplies after the reform, for wives. The table for men is omitted to save on space. The predictive power of the unitary model is limited. For instance 5.48% of the wives are predicted to work 10 hours under individual taxation, and 6.23% to work 20 hours, whereas they do not work in the collective baseline situation. Overall, only a third of the wives' post-reform hours of work are well predicted by the unitary approach.

Table 33 documents joint variations in the labour supply of husbands and wives as predicted by the unitary model. As with the actual variations in the collective setting (see corresponding Table 22) more than half of the couples do not adjust their hours of work under the new tax rule. Both women and men are predicted to increase hours more than we would expect from the results obtained with the collective model. The summary Table 34 compares the female labour supply adjustments of both models. Whereas the unitary model seems to predict that women, if any, will rather increase their labour supply under individual taxation, the collective setting predicts labour supply changes in both directions. Overall, the reaction in hours worked is slightly underestimated for the wives - almost 60% of the wives have an unchanged labour supply according to the unitary model but only 55% in the collective setting (and overpredicted for the husbands, table not presented here).

Table 32: Collective vs. unitary female labour supply, individual taxation

	0	10	20	30	40	50	60	total
0	**5.63**	5.48	6.23	1.43	.15			18.92
10	3.08	**2.25**	6.83	2.85	.53			15.54
20	2.63	3.00	**9.01**	4.05	1.80			20.50
30	2.40	1.73	3.30	**5.71**	5.48	.98		19.59
40	1.28	2.40	1.80	4.50	**8.41**	2.93		**21.32**
50	.23	.08	.53	1.13	1.05	**.38**		3.38
60	.08	.08		.30	.23	.08	**.00**	.75
total	15.32	15.02	**27.70**	19.97	17.64	4.35	.00	1332

Note: rows: wives' collective labour supplies, columns: unitary.

Table 33: Δ labour supply of couples, joint - ind. tax.: wives vs. husbands (unitary model)

	≤-10	0	10	20	≥30	total
-10		.15	.23			.38
0	3.15	**51.73**	4.43	.08		**59.38**
10	6.31	26.28	**.90**			33.48
20	2.10	2.25		**.15**	1.13	5.63
≥30				.30	**.83**	1.13
total	11.56	80.41	5.56	.53	1.96	1332

Note: rows: wives, column: husbands.

Table 34: Δ female labour supply, joint - ind. tax., collective vs. unitary

	-10	0	10	20	≥30	total
-30	.08					.08
-20		1.20	.60	.08		1.88
-10		13.21	5.48	1.58	.08	20.35
0	.30	**33.71**	18.32	2.03	.90	**55.26**
≥10		11.26	**9.08**	1.96	.15	22.45
total	.38	**59.38**	33.48	5.63	1.13	1332

Note: rows: collective, column: unitary.

7.5 Normative aspects of the reform

We describe the welfare effects of the reform at the household level, as measured by the unitary model, by showing the distribution of percentage changes in household utility for every decile of the pre-reform distribution of the household equivalent disposable income.[19] This graph requires cautious interpretation: we do not mean to suggest that considering percentual changes permits interhousehold welfare comparisons. But given that the composition of deciles refers to the baseline situation, we expect such a graph to convey a feeling for the importance of welfare effects. What is well defined in the graph, in terms of welfare comparisons, is the information on proportions of winners and losers by decile. The graph shows the quartiles of the distribution (box). The lines emerging from the box extend upwards to the largest utility change smaller than $Q_{75} + 1.5(Q_{75}$ - $Q_{25})$ and downwards to the smallest utility change larger than $Q_{25} - 1.5(Q_{75}$ - $Q_{25})$. Points outside this range, if any, are plotted individually.

Individual welfare effects of the reform, measured for husbands and wives separately within the collective model framework, are described by showing the distribution of percentage changes in individual utility for every decile of the pre-reform distribution of the wives' or husbands' equivalent disposable income.[20]

A direct comparison of the welfare analysis based on the two models is made on the basis of cross-tabulation of the positions of households (winner, indifferent, loser) with the pairs of positions of the spouses, whereby a cut-off $\pm 0.1\%$ change has been adopted to define indifference.

The following figures show the unitary (Figure 6), and the collective predictions (Figure 7 and 8) for individual taxation. Table 35 provides a comparison between the two models. Figures 7 and 8 show that, in the collective case, individual taxation is only advantageous for women in the higher equivalent income deciles. Men show some large gains and losses at all income levels. The unitary model shows a balanced distribution of losers and winners, and some very large losses and gains, especially at high income levels. Surprisingly the welfare gain from individual taxation increases in average with equivalent income.

Table 35 indicates that 36% of the couples lose, as well as 29% of the

[19] The equivalence scale for the household disposable income is: 1 for the first parent, 0.7 for the second, 0.6 for each child more than 16 years old, 0.5 for each child between 7 and 15 years and 0.4 for each child younger than 6 years, see equation (7).

[20] The equivalence scale for the individual disposable income is 1 for the parent (the wife or the husband) and the same as above for the children. It corresponds to variable *heq* in equation (7).

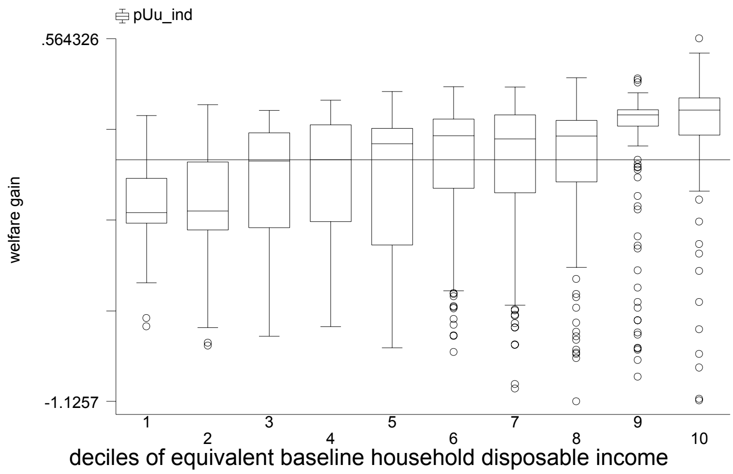

Figure 6: Relative welfare gains for households

women, but that for men, the proportion of losers is higher, and attains 57%. The percentage of *Pareto winning* households (i.e. combinations ++, +0, 0+) is only 18%; there are 27% *Pareto losers* (- -, -0, 0-), and 22% *contradictory* entries (h+,-0), (h+,0-), (h+,- -), (h+,00), (h0,++) etc.. The percentage of households for which the move to individual taxation generates *conflicting* effects for husband and wife is as high as 39%. Other results, not illustrated here, show that the collective model predicts richer women without children to fare better under individual taxation, whereas the unitary model suggests that couples without children are disadvantaged by the reform.

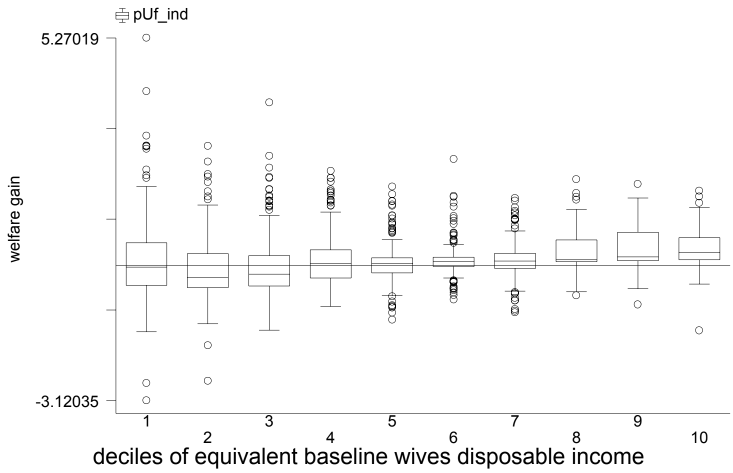

Figure 7: Relative welfare gains for married women

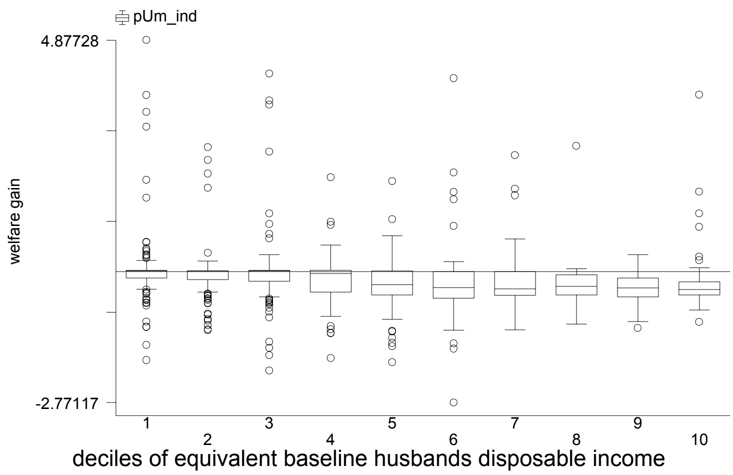

Figure 8: Relative welfare gains for married men

Table 35: Winners and losers: collective vs.unitary model, individual taxation

	f_-	f_-	f_-	f_0	f_0	f_0	f_+	f_+	f_+	
	m_-	m_0	m_+	m_-	m_0	m_+	m_-	m_0	m_+	total
hous_-	11.71	3.08	3.38	2.10	3.38	.45	10.89	1.05		36.04
hous_0	3.75	.68	1.28	.68	5.86	.15	6.01	1.28		19.67
hous_+	3.60	1.05	.45	.90	5.41	.38	17.42	14.79	.30	44.29
total	19.07	4.80	5.11	3.68	14.64	.98	34.31	17.12	.30	1332
total f	(f_-)	28.98		(f_0)	19.29		(f_+)	51.73		1332
total m	(m_-)	57.06		(m_0)	36.56		(m_+)	6.38		1332

Notes: move from the 1998 joint system to individual taxation. Rows correspond to winning (hous_+), indifferent (hous_0) and losing (hous_-) couples, on the basis of the estimated coefficients of the unitary model. Households are considered indifferent if their post-reform utility level is the same $\pm$ 0.1% as before the reform. Columns correspond to the winning, indifferent or losing wives and husbands on the basis of the simulated "collective" data. Spouses are considered indifferent if their post-reform utility level reform is the same $\pm$ 0.1% as before the reform.

8 Conclusion

The aim of this study was to illustrate the distortions arising in policy evaluation on the basis of a unitary model when the real world, in fact, follows collective rationality. We have tackled this question by simulating real world microdata within a collective set up and by estimating a unitary model on these data. A comparison of the collective data and the unitary estimation results showed that in the baseline situation, on average, labour supply is under-predicted by the unitary model. In total, only a third of female labour supply decisions are correctly predicted as well as 40% of the male.

In terms of policy evaluation the distortions entailed by the use of the wrong model may be even more interesting. A first distortion shows up in the *design* of the tax reform. While a revenue neutral move from joint to individual taxation, realised by proportional scaling of tax liabilities, leads to a factor of .942 with the collective model, the unitary model leads to a much lower factor of .894. The predictions concerning changes in the distribution of the tax burden on couples and singles, while going in the same direction of a shift from singles to couples, are of starkly different magnitudes.

As regards changes in labour supplies, the unitary predictions do well for less than half of the wives. Overall, in the collective setting, wives' labour supply is more responsive to the reform, whereas using the unitary model, more men are predicted to alter their labour supply. That is, when basing policy evaluations on estimations from the unitary model while living in a collective world, we underestimate the change in hours for wives and overestimate the change for husbands.

As for the normative aspects of the reform, we have looked at changes in household utility (unitary model) and changes in individual utility (collective model). In a comparison of the predictions of both set-ups, 22% of the conclusions at the household level turn out to be *contradictory*. That is, both spouses are predicted to be affected in one way by the collective model, and household is predicted to be affected in the opposite way by the unitary model.

Finally, an aspect which is totally ignored by the unitary model turns out to be quantitatively important. The collective model reveals that the reform has *conflicting* effects for 39% of the households: that is, a welfare gain is predicted for the wife and a loss for the husband or vice versa. Note that these within-household implications of a policy measure can only be uncovered by using a collective model of household behaviour where husband and wife are considered as distinct decision makers with individual preferences.

References

[1] Apps, P. and R. Rees (1988), Taxation and the household, *Journal of Public Economics*, 35, 355-369.

[2] Barmby, T. and N. Smith (2001), Household labour supply in Britain and Denmark: some interpretations using a model of Pareto optimal behaviour, *Applied Economics*, 33, 1109-1116.

[3] Beblo, M. (2001), Die Freizeitlücke zwischen erwerbstätigen Müttern und Vätern - Ein ökonomischer Erklärungsversuch mit Daten der deutschen Zeitbudgeterhebung 1991/92, in: Ehling, M., Merz, J. et al. (ed.): Zeitbudget in Deutschland - Erfahrungsberichte der Wissenschaft*,* Spektrum Bundesstatistik, Vol. 17, Wiesbaden.

[4] Beblo, M. and E. Wolf (2002), How much does a year off cost? Estimating the wage effects of employment breaks and part-time periods, *Cahiers Économiques de Bruxelles,* 45(2), 191-217.

[5] Beninger, D. (2000), Collective models: introduction of the taxation, *mimeo,* ZEW Mannheim.

[6] Beninger, D. and F. Laisney (2002), Comparison between unitary and collective models of household labor supply with taxation, *ZEW Discussion Paper 02-65*, ZEW, Mannheim.

[7] Chiappori, P.-A. (1988), Rational household labour supply, *Econometrica*, 56, 63-89.

[8] Chiappori, P.-A. (1992), Collective labour supply and welfare, *Journal of Political Economy*, 100, 437-467.

[9] Donni, O. (2003), Collective household labor supply: Nonparticipation and income taxation, *Journal of Public Economics*, 87(5-6), 847-1334.

[10] Heckman, J. and B. Singer (1984), A method for minimizing the impact of distributional assumptions in econometric models of duration data, *Econometrica*, 52, 271-320.

[11] Hoynes, H. (1996), Welfare transfers in two-parent families: labor supply and welfare participation under AFDC-UP, *Econometrica*, 64, 295-332.

[12] Jacobebbinghaus, P. and V. Steiner (2003), Dokumentation des Steuer-Transfer-Mikrosimulationsmodells STSM, *ZEW Dokumentation* 03-06, Mannheim.

[13] Laisney, F. (2002, ed.), *Final Report on the EU Project "Welfare analysis of fiscal and social security reforms in Europe: Does the representation of family decision processes matter?"* Contract No: VS/2000/0776, Written for the European Commission, DG Employment and Social Affairs.

[14] Lewbel, A. (2000), Two stage least squares estimation of endogeneous sample selection models, *mimeo,* Boston College.

[15] Mas-Colell, A., M.D. Whinston and J.R. Green (1995), *Microeconomic Theory*, Oxford University Press.

[16] McFadden, D. and K. Train (2000), Mixed MNL model for discrete response, *Journal of Applied Econometrics,* 15, 447-470.

[17] Moreau, N. and O. Donni (2002), Une estimation d'un modèle collectif d'offre de travail avec taxation, *Annales d'Économie et de Statistique*, 65, 55-83.

[18] Statistisches Bundesamt (1995, ed.), *Die Zeitverwendung der Bevölkerung: Ergebnisse der Zeitbudgeterhebung 1991/92*, Tabellenbände I - IV.

[19] Statistisches Bundesamt (2003): *Sozialhilfe in Deutschland: Entwicklung und Strukturen*, Internet:

http://www.bma.de/downloads/sozialhilfe_bericht_entwicklung_struktur.html, Status 15/02/03.

[20] Van Soest, A. (1995), Structural models of family labor supply: a discrete choice approach, *Journal of Human Resources* 30, 63-88.

[21] Varian, H. (1992), *Microeconomic Analysis*, third edition, Norton.

9 Appendix A: Estimation results: wage equations and predictions for couples

Table A1: Wage equations for couples

	wives			husbands		
	coef.	std.err.	t-value	coef.	std.err.	t-value
East	-.21	.08	-2.61	-.27	.12	-2.3
*a*40	.79	3.52	.23	.38	.06	6.5
$a40^2$	-.32	1.58	-.20			
*a*40_*East*				-.15	.11	-1.4
sch_real	.07	.10	.75	.10	.02	4.7
sch_abi	.22	.11	1.88	.22	.03	6.8
job_noap	-.32	.08	-3.97	-.06	.04	-1.6
job_fach	.07	.07	.93	.03	.02	1.6
job_uni	.50	.11	4.69	.28	.04	7.8
*child*0	.13	.24	.53			
*child*3	-.16	.21	-.74			
*child*7	-.07	.07	-1.02			
*child*16	-.02	.05	-.47			
constant	1.96	1.94	1.01			
(uncensored) obs.		809			1113	
adj. R-squared					0.416	

Notes: *East* : person lives in Eastern Germany. *a*40: age divided by 40. $a40^2$ square of *a*40. *a*40_*East* : interaction $a40 \times East$,*sch_real, sch_abi*, *job_noap*, *job_fach* and *job_uni*: indicator variables for short secondary schooling, long secondary schooling, no vocational degree, technical training, and polytechnic or university degree, respectively. *child*0, *child*3, *child*7 and *child*16: number of children in the household less than 3 years old, between 3 and 6 years, between 7 and 12 years and over 16 years, respectively.

Table A2: Wages for couples

	no.	mean	s.d.	min.	10%	50%	90%	max.
w_f	1332	11.66	4.48	1.81	7.17	11.09	16.07	74.32
w_m	1332	15.57	6.16	4.46	9.22	14.61	23.65	59.45

Note: see note to Table 4.

10 Appendix B: Concavity condition of the unitary utility function

The utility function is:

$$\begin{aligned} U(c, l_f, l_m) &= \beta_c \ln(c - \bar{c}) + \beta_f \ln(l_f - \bar{l}_f) + \beta_m \ln(l_m - \bar{l}_m) \\ &+ \delta \ln(l_f - \bar{l}_f) \ln(l_m - \bar{l}_m) \end{aligned}$$

We assume that all the differences $c - \bar{c}$, $l_f - \bar{l}_f$, $l_m - \bar{l}_m$ are positive, as well as the coefficients β_c, β_f and β_m. Since the gradient of U is:

$$\partial U = \begin{bmatrix} \dfrac{\beta_c}{c - \bar{c}} \\ \dfrac{1}{l_f - \bar{l}_f} \left[\beta_f + \delta \ln\left(l_m - \bar{l}_m\right)\right] \\ \dfrac{1}{l_m - \bar{l}_m} \left[\beta_m + \delta \ln\left(l_f - \bar{l}_f\right)\right] \end{bmatrix},$$

a sufficient condition for U to be increasing in its arguments is

$$\delta > \min\left\{-\frac{\beta_f}{\ln\left(l_m - \bar{l}_m\right)}, -\frac{\beta_m}{\ln\left(l_f - \bar{l}_f\right)}\right\}. \tag{18}$$

We assume that this condition is satisfied. The hessian of U is:

$$\partial^2 U = \begin{bmatrix} -\dfrac{\beta_c}{(c - \bar{c})^2} & 0 & 0 \\ 0 & -\dfrac{\beta_f + \delta \ln\left(l_m - \bar{l}_m\right)}{\left(l_f - \bar{l}_f\right)^2} & \dfrac{\delta}{\left(l_f - \bar{l}_f\right)\left(l_m - \bar{l}_m\right)} \\ 0 & \dfrac{\delta}{\left(l_f - \bar{l}_f\right)\left(l_m - \bar{l}_m\right)} & -\dfrac{\beta_m + \delta \ln\left(l_f - \bar{l}_f\right)}{\left(l_m - \bar{l}_m\right)^2} \end{bmatrix}. \tag{19}$$

Calling A the second diagonal block, concavity of U is then equivalent with A negative, thus, given condition 18, with $\det A > 0$. Thus the concavity condition is

$$\left[\beta_f + \delta \ln\left(l_m - \bar{l}_m\right)\right]\left[\beta_m + \delta \ln\left(l_f - \bar{l}_f\right)\right] > \delta^2. \tag{20}$$

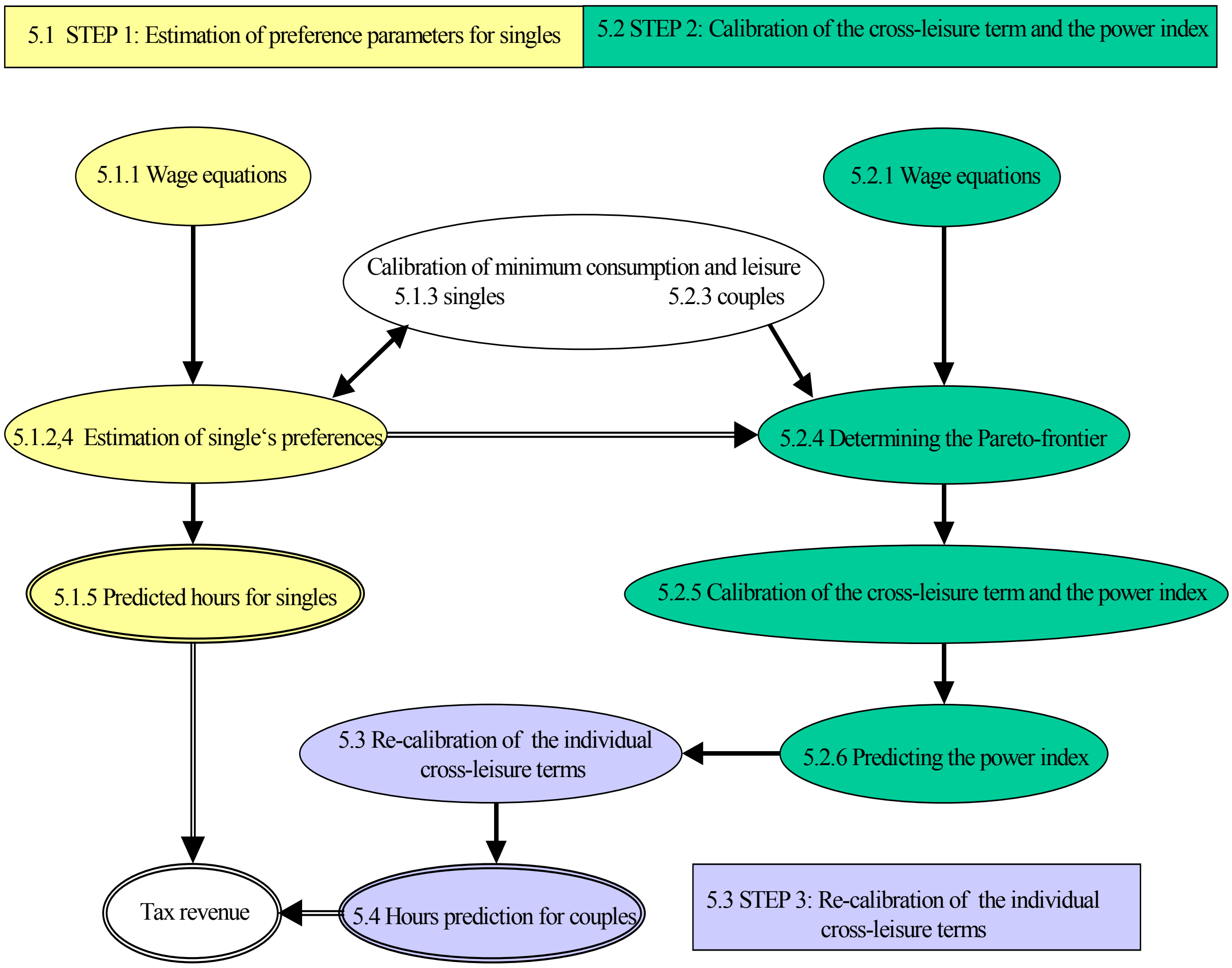

Figure A.1: Road map through the paper