

Beznoska, Martin

Article — Published Version

Testing the Income Pooling Hypothesis and its Link to the Taxation of Couple Households: Evidence from Demand System Estimation for Germany

Journal of Family and Economic Issues

Provided in Cooperation with:

Springer Nature

Suggested Citation: Beznoska, Martin (2023) : Testing the Income Pooling Hypothesis and its Link to the Taxation of Couple Households: Evidence from Demand System Estimation for Germany, Journal of Family and Economic Issues, ISSN 1573-3475, Springer US, New York, NY, Vol. 45, Iss. 3, pp. 687-719,
<https://doi.org/10.1007/s10834-023-09914-y>

This Version is available at:

<https://hdl.handle.net/10419/309455>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Testing the Income Pooling Hypothesis and its Link to the Taxation of Couple Households: Evidence from Demand System Estimation for Germany

Martin Beznoska¹ 

Accepted: 31 May 2023 / Published online: 15 July 2023
© The Author(s) 2023, corrected publication 2023

Abstract

Whether couples pool their resources and behave like a unit or spend their income individually is crucial for social and tax policy. In this paper, I provide a test of the income pooling hypothesis using administrative cross-sectional survey data on expenditures and individual incomes of couple households in Germany. The test is performed within the quadratic almost ideal demand system framework, which allows for an endogenous expenditure budget and endogenous individual income contribution shares in an instrumental variables approach. Although perfect income pooling is broadly rejected, there are significant differences regarding the marital status, the presence of at least one child in the household and whether the household is located in a former West or East German federal state. Married and unmarried couples with children are closer to the acceptance of the hypothesis than couples without children. The approach allows to calculate justifiable differentials of the marginal tax rates within the household if income pooling is rejected.

Keywords Income Pooling · Intra-Household Allocation · Income Taxation · QUAIDS

JEL Classification D12 · D13 · H24 · J12

Introduction

Income pooling within the household means that the individual consumption of each member does not depend on the individual income contributions to the household budget. Therefore, a shift in the income contribution share should not alter individual consumption—given that the household budget stays constant. Income pooling within couple households does not only have important implications for the consumer behavior and the intra-household labor supply allocation, but it also determines the needs and the equivalence scale of the household. These issues may have consequences for the design of social benefits and for poverty politics. But the presence or absence of income pooling is also relevant for income taxation. In Germany, married couples are treated as a single tax unit with joint assessment and

the offsetting of income differences between spouses. This treatment implies perfect income pooling of the spouses and grants, in a progressive tax system, a lower average tax rate for the couple than individual tax assessment (in the presence of income differences between the spouses). In consequence, marginal tax rates are also shared within the couple, which would be justified in terms of the ability-to-pay principle of taxation (“Leistungsfähigkeitsprinzip”) in the presence of income pooling. If income pooling is violated, the marginal tax rate for the second earner would be too high, while it would be too low for the first earner, which would also excessively distort the individual leisure-consumption decision of both partners.

The research question is particularly relevant for Germany, as the marginal net income surplus for the second earner (which is predominantly the woman) from additional labor supply is very low compared to other OECD countries (Becker, 2022). But many countries have similar joint taxation elements in their income tax assessment for married couples, for example France, Poland and up to a certain income threshold also the United States. Other countries like Austria or Sweden switched more or less completely

✉ Martin Beznoska
beznoska@iwkoeln.de

¹ German Economic Institute (IW), Konrad-Adenauer-Ufer
21, Cologne, Germany

to an individual tax assessment with the motivation of both labor market participation of women and individual income considerations. This topic has not yet been examined under an income pooling perspective but there is a large literature with respect to optimal labor supply conditions for married couples (e.g., Bick & Fuchs-Schündeln, 2017).

Income pooling in the theoretical economic framework means in its simplest form that the household faces a single utility function which is maximized by its members. This so-called “unitary” model (Becker, 1991) implies that total household budget is the relevant determinant for the individual consumption in the household while the individual contribution shares to the household income do not matter. However, Browning et al. (2010) showed theoretically that local income pooling can even exist in non-cooperative household models, when both partners contribute to a public good. Therefore, it is an empirical question which type of household dominates in reality and to which extent income pooling exists.

Since expenditures are typically observed only on the household level in common household surveys, data on personally allocable consumption expenditures is rarely available. So, many studies testing the income pooling hypothesis, also the one at hand, rely on expenditures on clothing and footwear (see Section “Literature Background”). These categories are often separately available for women, men and children in the data and can therefore be attributed to persons. Although the demand for clothing and footwear can only be a proxy for individual (private) consumption, a non-zero marginal effect with respect to changes in individual’s income share of total household income would reject income pooling within the household in a model setting which also provides the demand for a composite (public) good. The limitation of the empirical results is that these expenditure categories account only for a small share of the consumption budget, on average 4.6 percent in the used data set. Endogeneity issues in this context and possible confounders that limit the explanatory power of clothing and footwear consumption are discussed in this study.

On this basis, the paper contributes new evidence on testing the income pooling hypothesis for Germany with survey data on household expenditures within the framework of a Quadratic Almost Ideal Demand System (QUAIDS) (see Banks et al., 1997). The paper adds evidence from structural demand system estimation applied to pooled data of the income and consumption survey for Germany (*Einkommens- und Verbrauchsstichprobe, EVS*) for the years 2008 and 2013. Endogeneity issues of the household budget and women’s income contribution share are handled in an instrumental variables setting. The analysis also tries to identify price effects despite limited price variation in the data. This structural consumer demand model setup also allows to calculate justified tax

differentials for the individual marginal income tax rates in terms of deviations from income pooling within the household (in which case the tax differential would be zero). Under the assumption that the relative marginal deviation between the individual’s income effect from household’s income effect on individual consumption is the same for all private goods, the estimated effect on clothing and footwear consumption could be used to calculate the tax differential. However, as these commodities represent only a small share of the consumption budget, derived tax differentials for policy recommendations should be based on broader measures of individual consumption that also involve durable goods and intertemporal consumption decisions. Despite this highly relevant question for tax policy, this specific connection between the estimable degree of income pooling and taxation has not been addressed yet in the literature to the best of my knowledge. Additionally, heterogeneity in the effects regarding the marital status and the presence of children in the household, as well as between former East and West German federal states is examined.

Similar to former results in the literature, the income pooling hypothesis is broadly rejected, which implies a relationship between the individual income contribution share and individual consumption (Bourguignon et al., 1993; Lundberg et al., 1997; Phipps & Burton, 1998; Ward-Batts, 2008). The magnitude of differences for married couples in the spouses’ income effects on individual consumption is in the range of the one found in Phipps and Burton (1998). However, I find significant differences regarding the marital status, the presence of at least one child in the household and whether the household is located in a former West or East German federal state. Married couples and couples with children are more closely related to the acceptance of the hypothesis than unmarried couples without children, a result that was previously only reported descriptively in a survey question study (Bonke & Uldall-Poulsen, 2007). Unmarried couples in former East German federal states are closer to income pooling than in former West German states. A negative effect of women’s income contribution on men’s clothing and footwear consumption is confirmed in all specifications, which in turn means a positive effects on women’s consumption and the composite good. While these results provide new insights to consumption behavior in couple households, the external validity transferred to total consumption must be interpreted with caution given the mentioned assumptions.

Transferred to total consumption the findings would suggest that the marginal tax rate differentials for couples with children should not be significantly different from zero, which means that joint tax assessment can be justified in terms of the ability-to-pay principle of taxation for these types of households. However, as the result is based only

on clothing and footwear consumption, its application to actual tax policy is limited. Further research using broader measures of individual consumption is therefore necessary.

Literature Background

There is a class of papers that uses structural household consumption models to test the income pooling hypothesis with micro data. Bourguignon et al. (1993) tested the hypothesis within a structural model of consumption functions on French survey data and rejected it. Browning et al. (1994) reject the unitary household model with Canadian survey data in a structural framework and identify a household sharing rule of resources. They find that personal expenditures is significantly affected by the share of income a spouse contributes to total household income. Phipps and Burton (1998) test the pooling hypothesis in a demand system also with Canadian data and find mixed results for different expenditure data. Income pooling cannot be rejected e.g. for housing but on the other side, wives' clothing consumption increases more strongly with their personal income and wives are more likely to spend their income on childcare than husbands. Bütikofer and Gerfin (2017) find that the sharing rule is significantly influenced by the ratio of wife's hourly wage relative to husband's hourly wage.

The study of Lundberg et al. (1997) belongs to a class of papers that uses a policy change as natural experiment. A child allowance was transferred to wives in the UK starting in 1977. The authors find strong evidence of a shift toward greater expenditures on women's and children's clothing due to the reform which is not in line with the pooling hypothesis. A more recent reform of child and working tax credits in 2003 in the UK was used by Fisher (2016) to analyze the effects on spending patterns. He finds significant positive effects on expenditures related to children. Ward-Batts (2008) combines a structural model with the exogenous variation of the UK reform in 1977 and confirms the findings of Lundberg et al.

Another strand in the literature uses survey questions that are directly related to pooling in the household. Bonke and Uldall-Poulsen (2007) exploit Danish survey data and find that most couples fully or partly pool their income. They also show that the probability of income pooling depends on several household characteristics as e.g. the duration of marriage and the existence of children in the household. Bonke and Browning (2009a) use the same data and report that two-thirds of couple households answers that they pool their resources. However, a small part of them indicates inconsistency if other answers are taken into account. Bonke and Browning (2009b) show that the most important correlate with the relative financial satisfaction of women in couples is their income share of household income.

Intra-household allocation of resources is also examined in experimental settings. Attanasio and Lechene (2002) use a welfare program designed as a field experiment in Mexico, which transferred money to mothers, to look at the outcomes of the correspondents. They find that women gained more influence in the decision-making process of the household due to the shift of resources. Beblo and Beninger (2017) use experimental data on 95 German couples and conclude that the hypothesis is rejected for more than a half of the couples, also noting that couples with higher household income and higher education are more likely to pool their resources.

The link between income pooling and taxation has not yet been directly examined empirically in the literature to the best of my knowledge. Related studies are e.g. Büttner et al. (2019) who focused on how the intra-household income distribution affects tax planning of a married couple. They find that couples tend to resign from a tax-minimizing treatment if it negatively affects the individual net income of the second earner (which is often the woman). There is also a vast indirectly related literature, which examines optimal income taxation for married couples but under the aspect of labor supply and time-allocation. An important result in this literature is the optimality of negative tax jointness, which means that the marginal tax rate of an individual should fall with its spouse's income (e.g., Alesina et al., 2011; Gayle & Shephard, 2019).

Model and Empirical Strategy

An example for consistent income pooling within the household applies if one partner reduces working time for childcare, which reduces his or her income, while the other partner increases working time to compensate the income reduction. Individual consumption measured as expenditures for goods and services solely consumed by one partner should not be affected by this shift in income contribution (given that preferences for the goods do not change, which also means that childcare expenditure should not change because of the shift). This general test is embedded in a structural household demand system, controlling for the total consumption budget, prices and taste shifters. Additionally, the model is extended to allow for endogeneity of the individual income contribution share and the budget.

The Model

The structural framework for the test of the income pooling hypothesis is the Quadratic Almost Ideal Demand System (QUAIDS) (see Banks et al., 1997). The QUAIDS is often used in the literature to model consumer demand with household data and is based on price-independent generalized logarithmic (PIGLOG) preferences with Engel curves

that are modeled as budget shares being a quadratic function of the log-budget.¹ It has the advantage of a flexible underlying utility function and allows imposing the restrictions of a consistent demand system like homogeneity and symmetry.

By applying PIGLOG preferences and separability, household's decision over budget and demand can be theoretically modeled as a two-stage budgeting process (Deaton & Muellbauer, 1980). Therefore, household's consumption decisions are made by setting the budget at the first stage and then consumption demand is chosen at the second stage. Household's total labor supply, which is necessary to generate the consumption budget, is not modeled explicitly in this setting. Instead, education and wages of both partners determine the total budget and the individual contributions to it at the first stage. At the second stage, consumption demand respectively households' preferences for each good are determined. This setting is econometrically modeled as an instrumental variables setting, where the second stage is the QUAIDS demand system. Endogeneity issues and the first stage is described in Subsection "Endogeneity".

Introducing the demand system: For each $i = 1, \dots, N$ goods and the corresponding budget shares w_i , the QUAIDS forms the following non-linear system of equations:

$$w_i = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln \left[\frac{m}{a(p)} \right] + \frac{\lambda_i}{b(p)} \left\{ \ln \left[\frac{m}{a(p)} \right] \right\}^2 + u_i \quad (1)$$

for $i = 1, \dots, N$ goods and $j = 1, \dots, N$ with consumption budget m , prices p_i and price indices

$$a(p) = \alpha_0 + \sum \alpha_i \ln p_i + \frac{1}{2} \sum \sum \gamma_{ij} \ln p_i \ln p_j$$

$$b(p) = \prod p_i^{\beta_i}$$

The model allows computing budget and (un)compensated price elasticities, as well as cross-price elasticities between different good prices. Although the system is non-linear due to its price indices, it can be estimated easily using the method of the Iterated Linear Least Squares Estimator (ILLE) which imposes conditional linearity on the parameters (Blundell & Robin, 1999).² Further explanatory variables that account as taste shifters for household consumption like demographic characteristics can be added to the equations.

¹ Ward-Batts (2008) also uses the QUAIDS to test the income pooling hypothesis. The indirect utility function of PIGLOG preferences can be found for instance in Poi (2012).

² The Stata routine "aidsills" by Lecocq & Robin (2015), which implements the Blundell & Robin (1999) estimator, is used in this paper to estimate the QUAIDS model.

Application

The test of the income pooling hypothesis is performed with administrative household microdata containing consumption expenditures (see Section "Data and Descriptives"). The budget shares w_i in the QUAIDS model are calculated by dividing the expenditures of consumption good i by the total consumption budget within each household. Expenditures are in general observed on household level in the data and therefore not assignable to the individuals in couple households. The only expenditure categories which allow the assignment to individuals in the household are those for clothing and footwear. These categories are observed for adults by gender (explicitly for persons aged 14 or older) and additionally for children. Therefore, clothing and footwear consumption is the private good separately available for each partner, while all non-durable consumption will enter the model as the public good. Since clothing and footwear categories have only a small share in the consumption budget (see Section "Data and Descriptives"), this is a limitation of the analysis and the following policy implications. However, this limitation has been widely accepted in the literature as other personally assignable expenditures are typically not observed in classic consumption surveys (see e.g. Bourguignon et al., 1993; Browning et al., 1994; Lundberg et al., 1997; Phipps & Burton, 1998; Ward-Batts, 2008; Bose-Duker et al., 2021). Clothing and footwear consumption is therefore only a proxy for private consumption, but under the assumption that the deviation from household's marginal income effect due to the individual marginal income effect is the same for all private goods, this allows to conclude about the degree of income pooling.

Note additionally that the preferences for the public good can vary between the partners. The marginal income effects in the model are always a results from mixed preferences in the household. The only additional assumption for the public good is, that both partners receive utility from all included goods, so that household's utility is a weighted sum from individual utility functions. In general, aggregation of the included goods and services to a commodity group requires the assumption that preferences for them are weakly separable or the generalized composite commodity theorem is fulfilled (Lewbel, 1996). In each case, under income pooling, there should also be no effect of the individual income contribution on public good consumption.

Although the demand system can be modeled as detailed as the expenditure categories in the data allow, the underlying utility function with weakly separable preferences makes it possible to aggregate the single goods to commodity groups. This attribute is useful to keep the estimation feasible by reducing the number of price effects in the model because of the general problem with small variation in prices, which occurs in demand system

estimation on pooled cross-sectional data. As the focus is on testing the income pooling hypothesis, there is also no major objection to restrict the demand system to three commodity groups.

I use quarterly data from two survey years, 2008 and 2013, which leaves the price variation to eight points in time. The aggregation of single expenditure categories to commodity groups features the attribute of computing Stone-Lewbel prices for the groups to increase the variation in prices (Lewbel, 1989). With the assumption of constant expenditure shares within a commodity group (implying Cobb–Douglas preferences in the group), the prices of the single goods are weighted with their expenditure shares in the commodity group. Since these shares vary for every household, price variation increases with the use of Stone-Lewbel prices which enables to identify price effects despite the inclusion of quarterly and yearly time dummies in the estimation. However, the small variation over time remains a challenge for the estimation, especially if the commodity groups consist of only a few goods. As a results, the standard errors of the estimated price effects are expected to be rather high.

For the unambiguous assignment of the spending, the approach focuses on mixed-gender couple households with no further adults and uses the personal expenditures on clothing and footwear to test the hypothesis. To avoid a wrong assignment of clothing and footwear expenditures for older children in the household to the categories for adults, the sample is restricted to households with children aged below 14 (or without children).

The budget m therefore contains the spending on non-durable consumption including the expenditures on clothing and footwear of both partners. The share of gross income contributed by the woman to the household gross income is introduced as s . Accordingly, the share of income contributed by the man to the household income is $1 - s$. Income can thereby stem from different income sources, not only labor but also transfers, pension income, business income and so on. If income is pooled and individual consumption only depends on the household budget, commodity prices and taste shifters but not on the individual income contribution, then the parameter on s should not be significantly different from zero. The variable can be added to the system of equations (1) in the same way as taste shifters and other control variables x_k by entering equation (1) in:

$$\alpha_i = \alpha_{i,0} + \sum_k \alpha_{i,k} x_k + \varphi_i s \quad (2)$$

The hypothesis is obviously rejected if $\varphi_{\text{women}} \neq 0$ in the equation for women's clothing and footwear. But as the adding-up restriction of demand systems is imposed in

the estimation ($\alpha_{\text{women}} + \alpha_{\text{men}} + \alpha_{\text{composite}} = 0$), this would in principle allow $\varphi_{\text{women}} = 0$, while φ_{men} in the equation for men's clothing and footwear and $\varphi_{\text{composite}}$ in the composite good equation are different from zero. But this result would imply that household consumption patterns depend on the income contribution share of women, which would also reject the hypothesis and therefore has to be tested. The parameters φ_i measure therefore the deviation from perfect income pooling.

The demographic control variables consist of dummies for the number of children in the household, quartic polynomials of age of both partners, a dummy for marriage, dummies for agglomeration level of the place of residence, time dummies (quarter, year), dummies for the federal state and a dummy for owner-occupied housing.

Endogeneity

There are some potential sources of endogeneity in the model, which are addressed in an instrumental variables (IV) approach. A classic endogeneity issue in demand systems is related to the budget m , which stands in the denominator of the expenditure share on the left-hand side of the equations and depends on the consumption preferences. The common solution for this issue can be implemented by using the disposable household income and its quadratic term as instruments in a Two-Stage Least Square (2SLS) type of estimator for a system of equations (see e.g., Blundell & Robin, 1999). The basic idea follows the augmented regression framework. In the first stage, budget m is regressed on the exogenous control variables x_k and the instruments. Then, the residuals of this regression are added to every equation in the system via (2) as additional control variables. Blundell and Robin (1999) show that under the assumption that the error term u_i of (1) can be orthogonally decomposed into the residuals from stage one and a white noise term, the augmented regression estimator is identical to the 2SLS estimator. Since the assumption of exogenous labor supply within the household must be somehow relaxed in the approach at hand, the disposable income is not an appropriate instrument. Instead, the gross wages of both partners are assumed to be exogenous (given completely inelastic labor supply of the household) and taken as instruments. The specific modeling will be discussed later.

The exogeneity of the commodity prices may be challenged in an analysis like the one at hand. The small but existent time variation of the consumer prices used in the model stem from the years 2008 and 2013. There was a reform of the standard rate of value-added tax in Germany in 2009, which affected many commodities including the expenditures on clothing and footwear. The rate was increased by three percentage points from 16 to 19 percent,

which can be seen as an exogenous variation in prices given an elastic supply curve. Despite this, the prices can at least be seen as measured with error with regard to the individual facing them and the construction of Stone-Lewbel prices with constant within-group shares to increase price variation may introduce endogeneity and measurement error. However, Hoderlein and Mihaleva (2008) showed that the use of Stone-Lewbel prices provides precise and economically plausible results compared to the use of standard consumer prices in demand models.

Another potential endogenous regressor can be seen in the women's share of income contribution s , as the preferences for clothing and footwear and therefore the household consumption pattern as a whole could affect the labor supply decision of the couple. For example, if women with strong preferences for clothing and footwear work more compared to their partners than women with lower preferences for these goods, the coefficient on the share of income contribution would be upwards biased and reject the hypothesis although the household members pool their income. Thus, women's income contribution is endogenous to the labor supply allocation of both partners, which in turn can be endogenous to consumer preferences.³ Another potential endogeneity issue stems from the matching of the couples. The preferences for clothing and footwear of partner A may influence the match with partner B, for example, because of partner B's income. This could also distort the test on income pooling as the considered couples may systematically vary in their unobserved characteristics (see Lundberg et al., 1997).

The standard approach in household demand analysis assumes separability between consumer demand and labor supply (e.g., Banks et al., 1997). This is also a useful assumption in the type of literature, which examines the identification of the sharing rule—the shares of resources that are jointly or privately consumed in the household (e.g., Browning et al., 2013). Separability can be theoretically modeled as a two-stage budgeting process (Deaton & Muellbauer, 1980). I assume a constant labor supply of the household, but an endogenous distribution within the household. So, at the first stage, the within household labor supply decision is made, which determines leisure, non-durable and durable consumption of the household members (and savings, which are future consumption). At the second stage, non-durable consumption is allocated on goods and services. The separability assumption allows focusing on non-durable

consumption and to treat the labor supply decision of the first stage independently. This assumption does not have to be relaxed in the analysis at hand to test the hypothesis. Total labor supply of couple households remains separable from consumption but the separability from the distribution of labor supply within couple households is relaxed. Therefore, it is needful to tackle the endogeneity issues linked to the share of gross income contributed by the woman s .

At the first stage, I estimate the two equations:

$$s = \Phi^{11}X_1 + \Theta^{11}Z_1 + v_1 \quad (3)$$

$$\ln(m) = \Phi^{21}X_1 + \Phi^{22}X_2 + \Theta^{21}Z_1 + \Theta^{22}Z_2 + v_2$$

where X_1 and X_2 are vector-subsets of the exogenous variables, which also enters the demand system at the second stage in (2). The vectors Z_1 and Z_2 are subsets of the instrumental variables (excluded in the demand system). Φ and Θ are parameter vectors.

In the first equation, women's contribution share s depends on the subset of control variables X_1 and the instruments Z_1 . The instruments in Z_1 are dummies for the type of school graduation of the man and the woman, interaction terms between them, as well as dummies for the type of highest educational/vocational graduation of both partners and again their interactions.⁴ The idea here is that education is separable from the preferences for non-durable consumption and can be left out in the demand system. However, it influences the share of income contribution *ex-ante* by bargaining position of the partners in the household labor supply decision and is also assumed to be correlated with the match of couples apart from preferences for consumption. The vector-subset X_1 contains all exogenous variables of the demand system except for the marriage dummy, the dummies for the number of children and the dummy for owner-occupied housing. These variables are denoted as vector X_2 and only appear in the second equation. The reason is that they are assumed to be potentially endogenous to the share of income contribution e.g. via the tax benefits of joint assessment of married couples in Germany if the share is far away from 0.5, which is also part of the research question and will be further examined in the heterogeneity analysis. Therefore, vector X_2 only appears in the budget equation at the first stage as the variables are in principle important for attributes that influence the household income and ultimately the consumption budget.

³ A collective household labor supply model that also incorporates consumer demand is examined theoretically for example in Blundell et al. (2005) and in Cherchye et al. (2012), who extend the former and apply it empirically. Applications within the QUAIDS framework that focus on environmentally relevant consumer goods can be found in West & Williams (2004) and Beznoska (2014).

⁴ Maitra & Ray (2006) also use the highest level of education to instrument the incomes of the household members to evaluate household expenditure allocation in South Africa.

The instruments in Z_2 are man's and woman's gross wages in logs, which are empirically derived from the data on individual gross income and working time. A classic Heckman model (Heckman, 1979) is estimated separately for men and women to impute the wages.⁵ The wages are left out of the first equation because women's income contribution share and the wages are all derived from the information on individual gross income, which creates a dependency by construction.

The two equations are overidentified as there are much more instruments than endogenous variables and can be estimated by seemingly unrelated regressions (SUR) to have efficiently estimated standard errors for the F -tests of the instruments. The first stage can then be linked to the second stage, which is the demand system, in an augmented regression framework like the one used in Blundell and Robin (1999). Thus, the predicted residuals from (3), \hat{v}_1 and \hat{v}_2 , are included in (2) to account for the endogeneity of s and m . Tests on the exogeneity of s and m can be derived from the estimated coefficients on \hat{v}_1 and \hat{v}_2 in the demand system. This test on exogeneity is combined with a test for overidentifying restrictions and Shea's partial R^2 to further check for the validity of the instruments.

Data and Descriptives

The model is estimated with two pooled cross-sections of data from the income and consumption survey for Germany (*Einkommens- und Verbrauchsstichprobe, EVS*) for the years 2008 and 2013. This administrative data set is a representative sample of households in Germany containing detailed information on income and expenditures. Each survey year features about over 40,000 households. The households are observed for one quarter equally distributed over all four quarters and with quarterly income and expenditure information. While consumption expenditures are only reported at the household level, income information is available individually for every household member. Very rich households are not included in the data as it prevents households with a quarterly household net income of more than 18,000 euro per quarter to enter the sample. However, this should not have a great impact on the average marginal effects regarding a consumption analysis.

As already described in the previous section, I focus on the demand analysis of non-durable consumption and explicitly the expenditures for clothing and footwear expenditures, which are observed for women and men separately. Expenditures on durables are manually compiled from all

goods and excluded from the budget m . The non-durable expenditures contain the categories food, drinks, tobacco, heating and electricity, mobility, articles of daily use, health expenditures, childcare, spending for leisure activities and other smaller items. Housing expenditures are also included and can either be actually paid rents without heating and electricity costs or imputed rents for owner-occupied houses and flats. The imputed rents are calculated by the German Federal Statistical Office (Statistisches Bundesamt) and already implemented in the EVS data sets.

Price data is supplemented to the survey data with official consumer price indices provided by the German Federal Statistical Office. While most expenditure categories in the EVS data refer to the two-digit and the four-digit price indices, especially the categories for men's and women's clothing and footwear are on a more disaggregated level but without an exact match in the available price data. I therefore use available ten-digit prices as proxies for these categories.⁶ The monthly price indices are averaged over the quarter to fit the quarterly expenditure data. Afterwards, household-specific Stone-Lewbel prices are constructed with the price data by weighting the prices with the respective expenditure shares for every commodity group to increase price variation (see previous section for details). There are eight points in time, which create price variation (quarterly data for two years). Additional regional price variation only comes from the prices for housing, which are differentiated available by federal states.

The basic sample of the analysis is restricted to mixed-gender couple households (who declare in the data to be a couple) with exactly two adult persons and optionally children below the age of 14. Additional criterion is the presence of income from occupation in the household from at least one partner. This restriction excludes the households from the analysis, which completely rely on transfer income. The reason is that these households could be systematically different in their preferences and consumption decision-making from households with at least one occupied partner. I end up with 29,461 households, 15,367 from the year 2008 and 14,094 from the year 2013.

The existence of zero expenditures in households could potentially be seen as a data issue. In the crucial equations for clothing and footwear expenditures, the share of zeros is 6.9 percent for women, while it is 15.4 percent in the

⁵ See Appendix A for detailed results of the estimated Heckman model and the distributions of imputed wages.

⁶ I use the prices for men's trousers and women's trousers as proxies for men's clothing and women's clothing, respectively. Revenue statistics for 2013 from data provider *Statista* suggest that expenditures for trousers are the most relevant expenditure group within clothing expenditures for men and women. For the price of men's footwear, I use the price category "classic or casual shoes for men". The respective price for women's footwear is called "pumps or casual shoes for women".

Table 1 Sample descriptives

Variables	Mean	Standard Deviation
Budget	2,607.91	1,042.45
Budget (log)	7.80	0.37
Expenditures for the commodity groups		
Women's clothing and footwear	77.31	84.96
Men's clothing and footwear	45.69	59.23
Composite good	2,484.91	992.21
Budget shares		
Women's clothing and footwear	0.029	0.027
Men's clothing and footwear	0.017	0.019
Composite good	0.954	0.038
Prices		
Women's clothing and footwear	100.03	3.74
Men's clothing and footwear	100.46	4.55
Composite good	102.65	4.07
Income		
Couple's gross income	4,626.62	2,331.98
Women's income contribution share	0.342	0.235
Selected control variables		
Women's age	45.02	12.78
Men's age	47.89	13.15
Married (dummy)	0.831	0.375
Children in household (dummy)	0.360	0.480
Number of children (below 14 years)	0.612	0.922
Owner occupied housing (dummy)	0.647	0.478
Women's wage per hour (log)	2.63	0.56
Men's wage per hour (log)	2.89	0.56
Number of observations: 29,461		

Expenditures and income in euro per month

Sources: EVS data 2008, 2013; own calculations

equation for men and the composite good has no zero-expenditures. A large share of zeros of the dependent variables because of a corner solution (zero consumption) can result in a biased estimation that can be fixed with a censored regression model. But in this case, the existence of zero expenditures in clothing and footwear is rather a problem of errors in variables due to infrequency of purchase. Everybody purchase clothing sometime, but some individuals are not observed purchasing. So, if the error term in the IV approach of the demand system is uncorrelated with the explanatory

⁷ The problem should therefore be rather minor, which is also confirmed by comparing the results without correction with the estimator for censored demand systems proposed by Shonkwiler & Yen (1999) in Appendix D. This also adds confidence to the assumption that the demand for clothing and footwear behaves more or less like other non-durable consumption, even though it can be seen as a semi-durable good.

variables, this measurement error in the dependent variable would yield an imprecise but still consistent estimate of the parameters.⁷

Table 1 shows the descriptives of the sample. The total budget of non-durable consumption expenditures is about 2608 euro per month at the mean, which is 73.6 percent of total spending on consumption in the sample. The average gross income of the couple households is 4627 euro per month, of which the women's contribution share is 34.2 percent.

A women's contribution share of zero is found in about ten percent of the households. In principle, this should not be a problem in the first stage of the 2SLS approach if there is sufficient correlation between s and the instruments. However, since the households with a women's income contribution share of zero may systematically vary in their preferences for consumption, the model is also run solely with the sample, in which both partners work, to compare the estimated coefficients as a robustness check, which can be found in Appendix E.⁸

Results

The QUAIDS model is firstly estimated with the ILLE ignoring endogeneity issues regarding the women's income contribution share s and the budget m . The results are discussed and compared to those of the 2SLS implementation in the augmented regression setting, which is presented secondly. Parameters of interest are, besides the one for s , the price and budget elasticities for the demand system, which can be derived from the estimated parameters.

Results for the QUAIDS Model (Neglecting Endogeneity)

Table 2 shows the result for the demand system by using the ILLE ignoring endogeneity issues and imposing homogeneity in prices. While budget effects are highly significant in all three equations, the price effects are only significant in the equations for women's clothing and footwear and for the composite commodity group. Importantly, the coefficients φ_i for women's income contribution share are significant in all equations with a positive sign for women's clothing and footwear and a negative sign for both other commodity groups. The system-wide joint test of the coefficients is also highly significant with a chi-squared statistic of 211

⁸ The expenditures for clothing and footwear may depend on the occupational status in case of specific jobs. Therefore, the sample is also restricted on households with a more similar income share of both partners limiting the variation from couple households with larger income differences and heterogeneous jobs.

Table 2 Estimation results for the demand system neglecting endogeneity

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
Price women's clothing and footwear (log)	– 0.0307* (0.0148)	0.0051 (0.0109)	0.0255 (0.0208)
Price men's clothing and footwear (log)	0.0051 (0.0148)	– 0.0076 (0.0109)	0.0024 (0.0209)
Price composite good (log)	0.0255** (0.0095)	0.0024 (0.0070)	– 0.0280* (0.0134)
Budget (β)	0.0758*** (0.0046)	0.0385*** (0.0034)	– 0.1142*** (0.0064)
Budget squared (λ)	– 0.0097*** (0.0007)	– 0.0048*** (0.0005)	0.0145*** (0.0010)
Women's income contribution share (φ)	0.0073*** (0.0007)	– 0.0032*** (0.0005)	– 0.0042*** (0.0010)
One child (baseline category: no children)	– 0.0099*** (0.0005)	– 0.0071*** (0.0004)	0.0170*** (0.0007)
Two children	– 0.0130*** (0.0005)	– 0.0092*** (0.0004)	0.0222*** (0.0008)
Three children	– 0.0152*** (0.0009)	– 0.0115*** (0.0007)	0.0268*** (0.0013)
Four or more children	– 0.0198*** (0.0019)	– 0.0126*** (0.0014)	0.0323*** (0.0027)
Married	– 0.0013** (0.0005)	– 0.0007* (0.0003)	0.0020** (0.0007)
Owner-occupied housing	– 0.00306*** (0.00036)	– 0.00222*** (0.00027)	0.00527*** (0.00051)
Constant	– 0.0066 (0.0188)	– 0.0011 (0.0139)	1.0077*** (0.0265)
Polynomials for the age of both partners, dummies of German citizenship, dummies for the federal state, level of agglomeration, quarter and survey year	Yes	Yes	Yes
N	29.461	29.461	29.461
R^2	0.090	0.063	0.108
Test on joint significance of coefficients φ , Chi^2 -statistic (p -value)	210.5 (0.000)		

Standard errors in parentheses, asymptotic variance–covariance matrix according to Blundell and Robin (1999). Homogeneity and symmetry restrictions imposed

Sources: EVS data 2008, 2013; own calculations

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(with two degrees of freedom because one equation must be dropped). This result implies a rejection of the income pooling hypothesis because a higher income contribution share of the woman means a higher consumption of women's clothing and footwear and a lower consumption of men's for a given household budget.

To classify the quantity of the effect, it can be evaluated at the mean of the expenditure shares. A switchover from zero income contribution to being the sole income earner would increase the expenditures on women's clothing and footwear

by 25 percent or nearly 20 euro per month. Simultaneously, consumption of men's clothing and footwear drops by about 19 percent and the composite good by 0.4 percent. So, the substitution happens mostly between the two private goods in this model.

Other control variables are left out in Table 2 but are included in all equations.⁹ Importantly, the controls for the

⁹ See Appendix B for the complete estimation results.

presence of children show consistent signs in a way that more children in the household reduce the private consumption of both partners. A hint on the differences in preferences between married couples and unmarried couples is the significant coefficient of the dummy for marriage. In this model, private consumption is lower for married couples, but further heterogeneity is evaluated in Subsection “[Heterogeneous Effects for Married Couples and the presence of Children](#)”.

Price and budget elasticities of the demand system can also be derived from the estimated parameters of Table 2.¹⁰ They are presented in Table 10 in the Appendix.

Results for the QUAIDS Model with Endogeneity

The model can be augmented by allowing for endogeneity of the women’s income contribution share and the expenditure budget. Following the approach presented in Section “[Model and Empirical Strategy](#)”, the two endogenous variables are regressed in a first-stage SUR model on the instruments. The predicted residuals from the first stage are subsequently inserted in the QUAIDS model.

The estimation results from the first stage are presented in Table 3. Since there are in total 48 dummies and interaction terms of the instruments *school graduation* and *high-est educational/vocational graduation* of both partners, the table is shortened by leaving out the results for the interaction terms.¹¹ Most instrumental dummies are clearly significant, although an interpretation is not meaningful without the interaction effects. The wages, which only appear in the budget equation, are also strongly significant. The *F*-tests in both equations on joint significance of the instruments do not indicate a weak instruments problem. Additionally, *Shea’s Partial R²* (Shea, 1997) for the women’s income contribution share *s* is about 0.086 suggesting a properly high correlation with the instruments.

The test of the income pooling hypothesis in the demand system with endogenous regressors remains a significant rejection, although the effect of women’s income contribution share on women’s clothing and footwear consumption is smaller compared to the first model (Table 4). The direct effect from a shift in the contribution share from zero to one is now 17.4 percent more consumption compared to 25

percent in the first model but still significant at the 5 percent level. The coefficients are also still jointly significant at the 1 percent level in the system of equations. While the significant effect on the consumption of the composite good vanishes, the negative one on men’s clothing and footwear becomes even more negative inducing a strong rival relationship between the two private goods.

However, the coefficients φ_i do not differ substantially from those of the first model. Accordingly, the test on exogeneity of women’s income contribution share, which is the test on the joint significance of the included residuals, is only significantly rejected at the 10 percent level with a *p*-value of 0.06. A somewhat different picture shows the test on exogeneity of the budget, which is strongly rejected.

The corresponding elasticities are presented in Table 5, where the dependent variables (quantities of demand) can be found in the lines and the columns refer to the exogenous variables (the budget and the prices). The budget elasticities are highly significant in all equations indicating that the demand for clothing and footwear is budget elastic. They also vary significantly from those of the model without endogeneity (see Table 10 in the Appendix). In contrast to the parameters, the own-price elasticities have to be tested against the null of -1 , which would indicate an estimated own-price effect of zero and Cobb–Douglas preferences. The compensated own-price elasticities are consistently negative for all commodity groups with a high own-price elasticity for women’s footwear and clothing of -2.0 . Substitutional relationships are found between all goods, whereas significant symmetric effects are only confirmed between women’s clothing and footwear and the composite good.¹²

Heterogeneous Effects for Married Couples and the Presence of Children

While the income pooling hypothesis is rejected by estimating one marginal effect for all couple households, there are still important questions left open regarding the heterogeneity in the household context. To deal with these questions, the model will be extended with interaction effects for married status and the presence of children in the household. Additionally, differences in the effects between the former East and West German federal states are explored. The underlying model is the augmented regression, which accounts for potential endogeneity and allows in principle for adding interaction terms without estimating a different first-stage equation.¹³ An important limitation of the sub-group results

¹⁰ The formulas for the elasticities in the QUAIDS can be derived according to Banks et al. (1997) from the estimated parameters. The budget elasticities are obtained by $\eta_i = 1 + \frac{\mu_i}{w_i}$ where $\mu_i = \beta_i + \frac{2\lambda_i}{b(p)} \ln\left(\frac{m}{a(p)}\right)$ is the marginal effect with respect to a change in the budget. The uncompensated price elasticities can be calculated $\epsilon_{ij}^u = \frac{\mu_{ij}}{w_i} - \delta_{ij}$ where $\mu_{ij} = \gamma_{ij} - \mu_i(\alpha_j + \sum_k \gamma_{jk} \ln p_k) - \frac{\lambda_i \beta_j}{b(p)} \left[\ln\left(\frac{m}{a(p)}\right) \right]^2$ is the marginal effect with respect to a change of price *j* and δ_{ij} is the Kronecker delta, which is 1 if *i* = *j* and 0 otherwise. The compensated price elasticities can be derived via Slutsky equation $\epsilon_{ij}^c = \epsilon_{ij}^u + \eta_i w_j$.

¹¹ See Appendix C for the complete estimation results.

¹² Symmetry cannot be rejected at the 10 percent level in this model and is therefore imposed.

¹³ A different first-stage SUR model is only estimated if the control variables are changed. For example, the dummies for the federal states had to be excluded in the interaction model with a dummy for East Germany. Therefore, this was also done in the first stage to have an equivalent specification.

Table 3 Estimation results for the simultaneous equation model (first stage)

	Women's income contribution share		Budget (log)	
Price women's clothing and footwear (log)	– 0.094	(0.149)	– 0.471*	(0.199)
Price men's clothing and footwear (log)	– 0.119	(0.182)	0.442†	(0.244)
Price composite good (log)	– 0.103	(0.086)	– 1.119***	(0.115)
Woman's wage per hour (log)			0.098***	(0.003)
Man's wage per hour (log)			0.167***	(0.004)
Baseline category: no school graduation (woman)				
Secondary school (woman)	– 0.129*	(0.065)	– 0.099	(0.087)
Intermediate school-leaving certificate (woman)	– 0.055	(0.060)	– 0.063	(0.081)
Specialized A-levels (woman)	– 0.025	(0.089)	0.157	(0.120)
A-levels (woman)	0.068	(0.072)	0.060	(0.097)
Baseline category: no school graduation (man)				
Secondary school (man)	– 0.030	(0.068)	– 0.230*	(0.092)
Intermediate school-leaving certificate (man)	0.006	(0.072)	0.097	(0.097)
Specialized A-levels (man)	– 0.110	(0.092)	– 0.088	(0.124)
A-levels (man)	0.070	(0.086)	0.071	(0.116)
Baseline category: no vocational training (woman)				
Vocational training (woman)	0.081***	(0.015)	0.028	(0.020)
Technician (woman)	0.059*	(0.026)	0.033	(0.036)
College (woman)	0.194***	(0.026)	0.168***	(0.035)
University (woman)	0.156***	(0.022)	0.090**	(0.029)
Baseline category: no vocational training (man)				
Vocational training (man)	– 0.090***	(0.013)	0.026	(0.018)
Technician (man)	– 0.090***	(0.018)	0.033	(0.024)
College (man)	– 0.162***	(0.017)	0.123***	(0.023)
University (man)	– 0.173***	(0.019)	0.067**	(0.026)
German citizenship (man)	– 0.079***	(0.009)	0.059***	(0.012)
German citizenship (woman)	0.103***	(0.008)	0.113***	(0.011)
One child (baseline category: no children)			0.068***	(0.006)
Two children			0.144***	(0.006)
Three children			0.198***	(0.010)
Four or more children			0.246***	(0.021)
Married			– 0.004	(0.005)
Owner-occupied housing			0.152***	(0.004)
Constant	3.917***	(1.123)	10.249***	(1.511)
Interaction terms between the type of school graduation of both partners and between the type of highest educational graduation of both partners	Yes		Yes	
<i>N</i>	29.461		29.461	
<i>R</i> ²	0.168		0.393	
<i>F</i> -statistic (<i>p</i> -value)	52.3 (0.000)		128.2 (0.000)	

Notes: Bootstrapped standard errors in parentheses

Sources: EVS data 2008, 2013; own calculations

†*p* < 0.10, **p* < 0.05, ***p* < 0.01, ****p* < 0.001

Table 4 Estimation results for the demand system with endogenous budget and endogenous women's income contribution share

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
Price women's clothing and footwear (log)	− 0.0427** (0.0149)	− 0.0032 (0.0109)	0.0459* (0.0210)
Price men's clothing and footwear (log)	− 0.0032 (0.0150)	− 0.0139 (0.0111)	0.0171 (0.0212)
Price composite good (log)	0.0459*** (0.0099)	0.0171* (0.0073)	− 0.0630*** (0.0141)
Budget (β)	0.0834*** (0.0047)	0.0436*** (0.0034)	− 0.1270*** (0.0066)
Budget squared (λ)	− 0.0097*** (0.0007)	− 0.0048*** (0.0005)	0.0146*** (0.0010)
Women's income contribution share (φ)	0.0051* (0.0024)	− 0.0070*** (0.0017)	0.0019 (0.0033)
\hat{v}_1 -Womens income contribution share	0.0026 (0.0025)	0.0042* (0.0018)	− 0.0068† (0.0035)
\hat{v}_2 -Budget	− 0.0091*** (0.0011)	− 0.0062*** (0.0008)	0.0154*** (0.0016)
One child (baseline category: no children)	− 0.0102*** (0.0005)	− 0.0073*** (0.0004)	0.0175*** (0.0007)
Two children	− 0.0139*** (0.0005)	− 0.0098*** (0.0004)	0.0237*** (0.0008)
Three children	− 0.0166*** (0.0009)	− 0.0125*** (0.0007)	0.0290*** (0.0013)
Four or more children	− 0.0214*** (0.0019)	− 0.0137*** (0.0014)	0.0351*** (0.0027)
Married	− 0.0013** (0.0005)	− 0.0007† (0.0003)	0.0020** (0.0007)
Owner-occupied housing	− 0.0045*** (0.0004)	− 0.0032*** (0.0003)	0.0077*** (0.0006)
Constant	0.0068 (0.0197)	0.0138 (0.0145)	0.9794*** (0.0279)
Polynomials for the age of both partners, dummies of German citizenship, dummies for the federal state, level of agglomeration, quarter and survey year	Yes	Yes	Yes
<i>N</i>	29.461	29.461	29.461
Shea's partial R^2 between s and its instruments	0.086		
Test on joint significance of coefficients of \hat{v}_1 , Chi^2 -statistic (p -value)	5.6 (0.061)		
Test on joint significance of coefficients of \hat{v}_2 , Chi^2 -statistic (p -value)	91.4 (0.000)		
Test on joint significance of coefficients of φ , Chi^2 -statistic (p -value)	28.9 (0.000)		

Adjusted standard errors in parentheses, asymptotic variance–covariance matrix according to Blundell and Robin (1999). Homogeneity and symmetry restrictions imposed

Sources: EVS data 2008, 2013; own calculations

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

appears regarding the interpretation. The differences must be interpreted rather descriptive than causal as the endogeneity of fertility choices, or the marital status is not addressed. So, if a couple gets a child, it does not necessarily change pooling behavior with the estimated coefficients. In general, different

coefficients between the groups only signal different behavior but the groups could also differ before the event of a birth.¹⁴

¹⁴ The interaction models could also be estimated with sub-sample analyses. However, the coefficients do not vary much.

Table 5 Estimated elasticities for the demand system with endogeneity

	Budget elasticity	Uncompensated price elasticity Change in the price of...			Compensated price elasticity Change in the price of...		
		(1)	(2)	(3)	(1)	(2)	(3)
Women's clothing and footwear (1)	1.72*** (0.035)	– 2.06* (0.494)	0.09 (0.500)	0.24 (0.317)	– 2.00* (0.494)	0.12 (0.500)	1.89*** (0.324)
Men's clothing and footwear (2)	1.73*** (0.049)	0.15 (0.618)	– 1.61 (0.626)	– 0.27 (0.397)	0.20 (0.618)	– 1.58 (0.626)	1.38*** (0.405)
Composite good (3)	0.96*** (0.002)	0.03 (0.022)	0.01 (0.022)	– 1.000 (0.014)	0.06** (0.022)	0.03 (0.022)	– 0.09*** (0.014)

Notes: Standard errors in parentheses. Homogeneity and symmetry restrictions imposed. Elasticities evaluated at sample means. Own-price elasticities tested against the null of – 1

Sources: EVS data 2008, 2013; own calculations

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Four different interaction models are estimated with different specifications and presented in Table 6. The first model incorporates an interaction term for married couples allowing women's income contribution to have varying effects for married and unmarried couples. The effects on both “private” goods are significantly higher for unmarried couples. This result is confirmed with the system-wide chi-squared test which has a much higher statistic for unmarried couples (main coefficient), although the hypothesis remains rejected for married ones (main coefficient and interaction effect combined). Married couples are thus nearer to the theoretical construct of pooling.

Interestingly, there are substantial differences between the former East and West German federal states. The effect on women's clothing and footwear is smaller for unmarried couples in East Germany compared to West Germany, while the effect on men's clothing and footwear is much larger. Consequently, there is more substitution between the composite good and men's private good. The test signalizes a less strong rejection in the East. However, the status of marriage reduces the effects in both regions bringing about a low chi-squared statistic.

A similar effect compared to the one found for married couples appears in the model with an interaction term for the presence of at least one child in the household (below 14 years old). There is no effect left on women's clothing and footwear consumption but a high effect on men's. Without children the substitution happens almost exclusively between the private goods. But contrary to the model with a term for marriage, the expenditure shares of the private goods are both significantly lower with the presence of at least one child which indicates a large preference shift toward the composite commodity group. This is plausible because the expenditures on goods for children are contained in this group. However, it could be the case that all pure privately consumed goods for the adults are equally devaluated with the presence of children, which means there is still explanatory power to the test indicating substitution between men's consumption and the

composite good dependent on women's income contribution. The chi-squared statistic has a similarly low value as in the models with an interaction term for marriage.

The combination of marriage and the presence of at least one child confirms the found results. Interestingly, the constellation which is nearest to perfect income pooling according to the chi-squared test is an unmarried couple with at least one child. Though, this result is mainly driven by higher standard errors, as the differences between this case and the combination with marriage are not significant.

The Link to the Taxation of Couple Households

The estimated parameters from the structural demand system on income pooling can be used to evaluate a justified differential in the marginal income tax rates between the partners. If income pooling is hurt, marginal tax rates should not be equal in terms of the ability-to-pay principle of taxation (“Leistungsfähigkeitsprinzip”). The marginal tax rate for the second earner would be too high, which would also distort the individual leisure-consumption decision. This follows from the fact that the second earner does not participate accordingly from an income increase of the first earner as what household's utility function would suggest.

Under the assumption that parameters β_i and λ_i measure household's “true” income effects on consumption of the private goods and the public good under the assumption of perfect income pooling, the parameters φ_i measure its deviations from this assumption. Under perfect income pooling, equitable income taxation with joint assessment yields a shared marginal tax rate. In this case, individual as well as

Table 6 Heterogeneous effects for married couples and the presence of children

	Dependent variables: Expenditure shares			χ^2 -statistic on system significance
	Women's clothing and footwear	Men's clothing and footwear	Composite good	
Interaction with <i>Married</i> dummy				
Women's income contribution share <i>s</i>	0.0121*** (0.0029)	− 0.0096*** (0.0021)	− 0.0025 (0.0041)	55.2, 0.000
Interaction effect: <i>Married</i> × <i>s</i>	− 0.0083*** (0.0019)	0.0031* (0.0014)	0.0052† (0.0027)	21.9, 0.000 (Combined effect)
<i>Married</i>	0.0022* (0.0010)	− 0.0020** (0.0007)	− 0.0002 (0.0013)	19.1, 0.000
Interactions with <i>Married</i> and <i>East Germany</i> dummies				
Women's income contribution share <i>s</i>	0.0140*** (0.0031)	− 0.0076*** (0.0023)	− 0.0064 (0.0043)	45.8, 0.000
Interaction effect: <i>Married</i> × <i>s</i>	− 0.0115*** (0.0023)	0.0006 (0.0017)	0.0109*** (0.0032)	21.1, 0.000 (Combined effect)
Interaction effect: <i>East Germany</i> × <i>s</i>	− 0.0084* (0.0040)	− 0.0074* (0.0029)	0.0158** (0.0057)	34.9, 0.000 (Combined effect)
Int. effect: <i>East Germany</i> × <i>Married</i> × <i>s</i>	0.0112** (0.0043)	0.0087** (0.0032)	− 0.0199** (0.0061)	17.9, 0.000 (Combined effect)
<i>Married</i>	0.0028* (0.0011)	− 0.0015† (0.0008)	− 0.0012 (0.0016)	14.1, 0.001
<i>East Germany</i>	0.0039* (0.0020)	0.0031* (0.0014)	− 0.0070* (0.0028)	6.3, 0.038
Int. effect: <i>East Germany</i> × <i>Married</i>	− 0.0023 (0.0021)	− 0.0019 (0.0015)	0.0042 (0.0030)	2.1, 0.352
Interaction with <i>At least one child in the household</i> dummy				
Women's income contribution share <i>s</i>	0.0073** (0.0024)	− 0.0065*** (0.0018)	− 0.0008 (0.0034)	33.1, 0.000
Interaction effect: <i>At least one child</i> × <i>s</i>	− 0.0071*** (0.0014)	− 0.0014 (0.0010)	0.0086*** (0.0020)	20.1, 0.000 (Combined effect)
<i>At least one child</i>	− 0.0077*** (0.0007)	− 0.0068*** (0.0005)	0.0144*** (0.0010)	218.3, 0.000
Interactions with <i>Married</i> and <i>At least one child in the household</i> dummies				
Women's income contribution share <i>s</i>	0.0139*** (0.0031)	− 0.0109*** (0.0022)	− 0.0030 (0.0043)	64.0, 0.000
Interaction effect: <i>At least one child</i> × <i>s</i>	− 0.0104* (0.0043)	0.0021 (0.0032)	0.0083 (0.0061)	10.3, 0.006 (Combined effect)
Interaction effect: <i>Married</i> × <i>s</i>	− 0.0081*** (0.0023)	0.0051** (0.0017)	0.0030 (0.0032)	23.2, 0.000 (Combined effect)
Int. effect: <i>At least one child</i> × <i>Married</i> × <i>s</i>	0.0045 (0.0046)	− 0.0038 (0.0034)	− 0.0007 (0.0064)	17.1, 0.000 (Combined effect)
<i>At least one child</i>	− 0.0073*** (0.0019)	− 0.0102*** (0.0014)	0.0174*** (0.0027)	53.5, 0.000
<i>Married</i>	0.0017 (0.0012)	− 0.0037*** (0.0009)	0.0020 (0.0017)	26.9, 0.000
Int. effect: <i>At least one child</i> × <i>Married</i>	− 0.0005 (0.0020)	0.0041** (0.0015)	− 0.0036 (0.0029)	9.0, 0.011

Adjusted standard errors in parentheses, asymptotic variance–covariance matrix according to Blundell and Robin (1999). Homogeneity and symmetry restrictions imposed. The right column contains *p*-values of the chi-squared test

Sources: EVS data 2008, 2013; own calculations

†*p* < 0.10, **p* < 0.05, ***p* < 0.01, ****p* < 0.001

composite good consumption does not depend on who contributed to the household income but only on the aggregate. If budget $m = (y_W + y_M)(1 - t)$, where y_W is woman's gross income, y_M is the man's gross income and t is the average tax rate, then the marginal effect of one more euro of woman's income on expenditure for good i in the QUAIDS is¹⁵:

$$\frac{\partial p_i q_i}{\partial y_W} = \mu_i + w_i \text{ with } \mu_i = \beta_i + \frac{2\lambda_i}{b(p)} \ln \left(\frac{m}{a(p)} \right) \quad (4)$$

If income pooling does not hold, then expression (4) applied for woman's private good becomes:

$$\frac{\partial p_W q_W}{\partial y_W} = \mu_W + w_W + \varphi_W \frac{y_M}{y_W + y_M} \text{ where } \frac{y_M}{y_W + y_M} = s_M \quad (5)$$

Note that s_M is men's share of gross income and coefficient φ_W is here the marginal effect of the woman's income share in the demand equation for women's private consumption. The marginal effect of one more euro of man's income on the expenditure of woman's private good would just be:

$$\frac{\partial p_W q_W}{\partial y_M} = \mu_W + w_W - \varphi_W \frac{y_W}{y_W + y_M} \text{ where } \frac{y_W}{y_W + y_M} = s_W = 1 - s_M \quad (6)$$

So, parameter φ represents the wedge in the marginal propensity to consume between the partners compared to the unitary model with perfect income pooling. The justifiable relative differential in marginal income tax rate can then be calculated in the two private-good-equations by dividing the estimated values $\hat{\varphi}_i$ with $i \in W, M$ by the respective marginal income effects given the income pooling assumption, which can be derived from Eq. (4). Given $\hat{\varphi}_W \neq -\hat{\varphi}_M$, there is also a distortion of the contribution to the composite (public) good compared to the situation with income pooling, which could also be targeted instead of the distortion of private good. This distortion is expected to be reduced with the respective tax differentials concerning the private goods but does not necessarily have to disappear. However, firstly, since the composite good is no pure public good, this measure is less precise in detecting a violation of income pooling. Secondly, the assumed policy goal in this context is only an equitable taxation in terms of the ability-to-pay principle, thus further violations of income pooling are not policy relevant.¹⁶

A positive and significant $\hat{\varphi}_W$ would request a lower marginal tax rate for the woman compared to the shared tax rate

because of the on average lower individual gross income of women. This would yield a reallocation of the leisure-consumption decisions of both partners within the household (total labor supply remains unchanged per assumption).¹⁷ If this adjustment of the individual marginal tax rates also results in a lower average tax for the second earner and a higher one for the first earner, it would imply an external redistributive effect that has also consequences for the intra-household allocation of resources.

The correspondent tax differentials for the estimated coefficients of the analysis can be found in Appendix F. Accordingly to the heterogeneous effects found in the previous section, the findings would suggest that the marginal tax rate differentials for couples with children should not be significantly different from zero, which means that joint tax assessment can be justified in terms of the ability-to-pay principle of taxation for these types of households. However, as clothing and footwear consumption represents only a small share of total private consumption, it serves only as proxy for the latter. Therefore, policy recommendations based on the estimated parameters in this analysis have limitations. Further research using broader measures of private consumption that also involve intertemporal consumption decisions is necessary.

Conclusion

The validity of the income pooling hypothesis has important implications for social and tax policy as well as for inequality research. In this paper, I provide a test of the income pooling hypothesis using administrative cross-sectional survey data on German couple households. I use information on expenditures and individual incomes to test the hypothesis in a structural consumer demand system. While most expenditures are only observed at the household level in the survey, expenditures on clothing and footwear are separately available for women and men and can be taken as proxies for individual consumption within the couple household. The limitation is that these expenditure categories account only for a small share of the consumption budget, on average 4.6 percent in the used data set.

According to the hypothesis, household consumption decisions should only depend on the household budget, prices and taste shifters. The individual income contribution share should therefore have no effect on consumption patterns, which can be tested within the framework of a Quadratic Almost Ideal Demand System (QUAIDS). I expand the model by controlling for endogeneity of the expenditure

¹⁵ This holds under the assumptions that savings and durable consumption are separable and the derivative of the average tax rate with respect to income is zero. The latter is obviously not true with progressive marginal tax rates, but the effect is expected to be neglectable and therefore left out for simplification.

¹⁶ This perspective could alter if family policy is concerned, as expenditures for children are contained in the composite good.

¹⁷ Since total labor supply of the household is not explicitly modelled in this approach, this is a potential limitation of the results.

budget and the individual income contribution shares in an instrumental variables approach. Additionally, heterogeneous effects are evaluated according to household attributes.

Although the hypothesis is broadly rejected, which implies a relationship between the individual income contribution share and individual consumption, there are significant differences regarding the marital status, the presence of at least one child in the household and whether the household is located in a former West or East German federal state. Married couples and couples with children are more closely related to the acceptance of the hypothesis than unmarried couples without children. Unmarried couples in former East German federal states are closer to income pooling than in former West German states. A negative effect of women's income contribution on men's clothing and footwear consumption is confirmed in all specifications, which in turn means a positive effects on women's consumption and the composite good.

The approach in principle allows to calculate tax differentials for the individual marginal income tax rates accounting for the distortion of the private good consumption. The findings would suggest that the marginal tax rate differentials for couples with children should not be significantly different from zero, which means that joint tax assessment can be justified in terms of the ability-to-pay principle of taxation for these types of households. However, as the result is based only on clothing and footwear

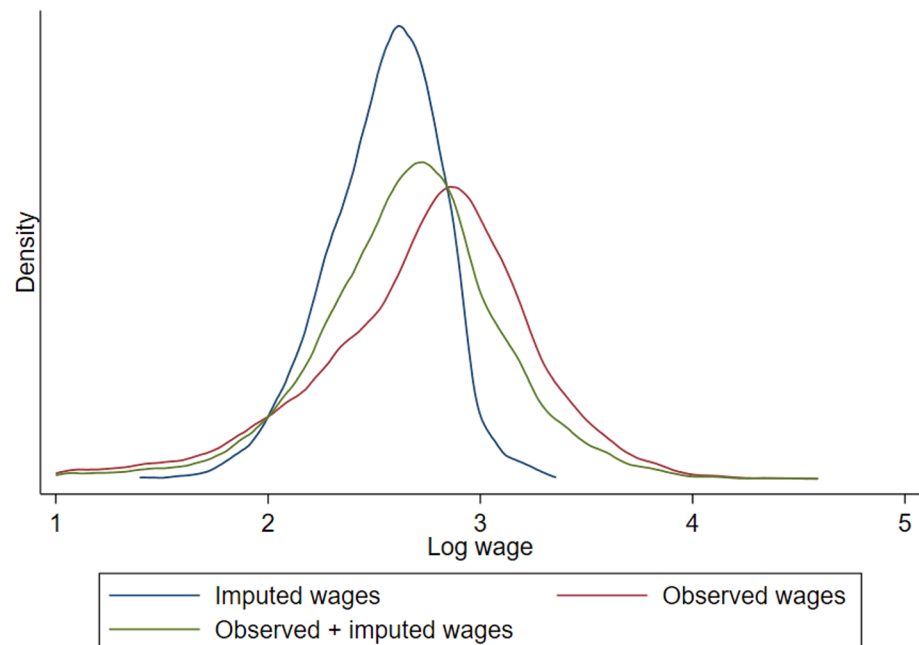
consumption, its application to actual tax policy is limited. Further research using broader measures of individual consumption is therefore necessary.

Appendix A

The wages for women and men are used in the first stage of the instrumental variables approach. Since the wages are not observed for those that are not working, they have to be imputed to estimate the model on all observations in the sample. For this purpose, a classic Heckman model is estimated for women and men separately to impute wages for the missing cases. The underlying sample of couple households is the same as the one for the demand system estimation. The sample size varies between men and women because the samples are restricted to persons below the age of 65 years.

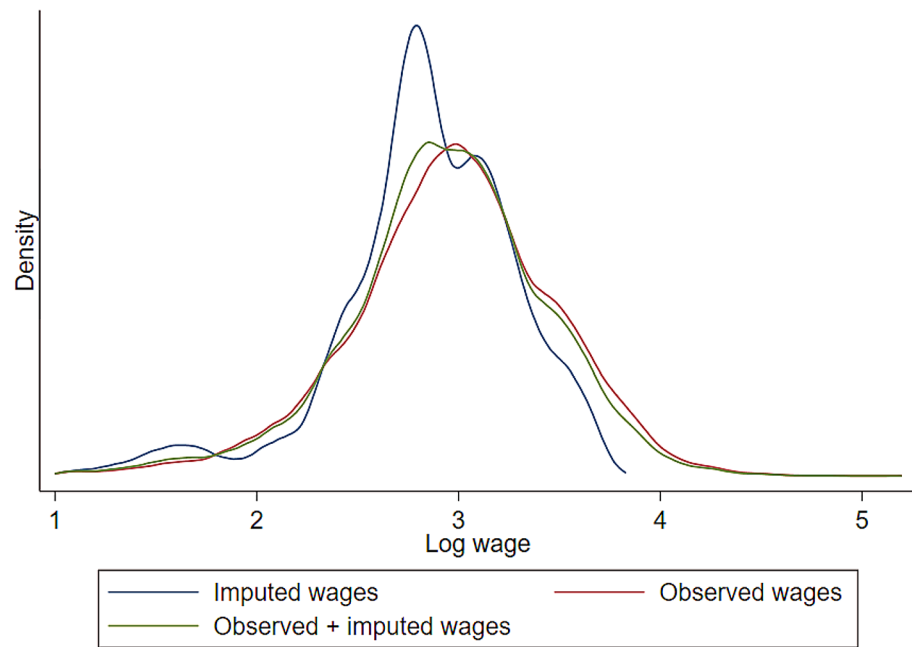
Figures 1 and 2 show the distribution of observed and imputed hourly wages for women and men. The distributions of imputed wages are left skewed compared to the observed ones for both men and women. For men, the combined distribution from observed and imputed wages varies only slightly from the observed one because only 5938 observations from 27,950 get an imputed value. The imputation rate is higher for women, for whom the wage is missing in 11,831 observations out of 31,260. Therefore, the combined wage distribution moves somewhat more to the left (Tables 7 and 8).

Fig. 1 Observed versus imputed distribution of women's hourly wages. Kernel density estimation. Distribution of imputed wages for those with missing observed wages. Sources: EVS data 2008, 2013; own calculations



Notes: Kernel density estimation. Distribution of imputed wages for those with missing observed wages.
Sources: EVS data 2008, 2013; own calculations.

Fig. 2 Observed versus imputed distribution of men's hourly wages. Kernel density estimation. Distribution of imputed wages for those with missing observed wages. Sources: EVS data 2008, 2013; own calculations



Notes: Kernel density estimation. Distribution of imputed wages for those with missing observed wages.
Sources: EVS data 2008, 2013; own calculations.

Table 7 Estimation results of the Heckman wage equation: Women

	Log Wage per hour		Work = 1	
Age	0.06189***	(0.01211)	0.09793	(0.15118)
Age squared	− 0.00050***	(0.00014)	− 0.00448	(0.00550)
Age cubic			0.00011	(0.00009)
Age quartic			− 0.00000**	(0.00000)
Partner's age			0.00562	(0.00759)
Partner's age squared			− 0.00013†	(0.00007)
Baseline category: no vocational training				
Vocational training (2)	0.58934*	(0.24270)	0.15086†	(0.08512)
Technician (3)	0.65796*	(0.31513)	0.01815	(0.15085)
College (4)	0.43722	(0.30037)	0.81753***	(0.18008)
University (5)	0.08044	(0.29231)	0.37141**	(0.13057)
Interaction terms				
Interaction: Vocational training × age	− 0.01201	(0.01260)		
Interaction: Technician × age	− 0.01471	(0.01555)		
Interaction: College × age	− 0.00087	(0.01519)		
Interaction: University × age	0.01970	(0.01483)		
Interaction: Vocational training × age squared	0.00005	(0.00015)		
Interaction: Technician × age squared	0.00011	(0.00018)		
Interaction: College × age squared	− 0.00002	(0.00018)		
Interaction: University × age squared	− 0.00022	(0.00017)		
Baseline category: partner has no vocational training				
Partner's graduation: Vocational training (2)			− 0.05971	(0.07541)
Partner's graduation: Technician (3)			− 0.04921	(0.10751)
Partner's graduation: College (4)			0.04234	(0.10368)
Partner's graduation: University (5)			0.12450	(0.11685)

Table 7 (continued)

	Log Wage per hour		Work = 1	
Interaction terms				
Interaction educational graduation 2 (own) × 2 (partner)			0.24223*	(0.09498)
Interaction educational graduation 2 × 3			0.24185†	(0.12386)
Interaction educational graduation 2 × 4			0.13018	(0.12091)
Interaction educational graduation 2 × 5			− 0.00339	(0.13454)
Interaction educational graduation 3 × 2			0.62338***	(0.16047)
Interaction educational graduation 3 × 3			0.57513***	(0.17778)
Interaction educational graduation 3 × 4			0.42714*	(0.17941)
Interaction educational graduation 3 × 5			0.30978	(0.18852)
Interaction educational graduation 4 × 2			0.03282	(0.19034)
Interaction educational graduation 4 × 3			0.14430	(0.21149)
Interaction educational graduation 4 × 4			− 0.16620	(0.20127)
Interaction educational graduation 4 × 5			− 0.31622	(0.21066)
Interaction educational graduation 5 × 2			0.21249	(0.14565)
Interaction educational graduation 5 × 3			0.28633*	(0.17237)
Interaction educational graduation 5 × 4			0.19310	(0.16231)
Interaction educational graduation 5 × 5			0.08174	(0.16619)
German citizenship	0.05153	(0.03356)	0.59728***	(0.04746)
Partner has German citizenship			− 0.02569	(0.05573)
Married (baseline category: single)	− 0.03672*	(0.01587)	− 0.29201***	(0.03151)
Widowed	− 0.06036	(0.07458)	− 0.41168***	(0.11547)
Divorced	− 0.06045*	(0.03027)	− 0.13179*	(0.05932)
Permanently separated	− 0.17778*	(0.07407)	0.29775*	(0.17948)
East Germany	− 0.18964***	(0.03462)	0.23317***	(0.06077)
Baseline category: population < 5000				
Population between 5,000 and 20,000	0.04916***	(0.01432)	0.03000	(0.02455)
Population between 20,000 and 100,000	0.05697***	(0.01488)	0.02177	(0.02566)
Population over 100,000	0.07945***	(0.01618)	0.03745	(0.02818)
Student	0.04225	(0.05127)	− 1.66974***	(0.05905)
Civil servant	0.10634***	(0.01485)		
One child (baseline category: no children)			− 0.63297***	(0.02797)
Two children			− 0.79272***	(0.02925)
Three children			− 1.26564***	(0.04592)
Four or more children			− 1.47939***	(0.09707)
Constant	0.96547***	(0.23956)	− 1.05446	(1.50307)
Dummies for the federal state, quarter and year	Yes		Yes	
Observations	31.260			
Selected	19.429			
Nonselected	11.831			
Lambda	− 0.38597***	(0.02491)		

Standard errors in parentheses

Sources: EVS data 2008, 2013; own calculations

†p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001

Table 8 Estimation results of the Heckman wage equation: Men

	Log Wage per hour		Work = 1	
Age	0.08111***	(0.01127)	1.45959***	(0.20562)
Age squared	– 0.00087***	(0.00014)	– 0.05801***	(0.00734)
Age cubic			0.00098***	(0.00011)
Age quartic			– 0.00001***	(0.00000)
Partner's age			0.00879	(0.01003)
Partner's age squared			– 0.00016	(0.00011)
Baseline category: no vocational training				
Vocational training (2)	0.60998**	(0.22967)	0.25738***	(0.08319)
Technician (3)	0.58928*	(0.28047)	0.11924	(0.12145)
College (4)	0.73314**	(0.25665)	0.65054***	(0.12519)
University (5)	0.16217	(0.26030)	0.48295***	(0.13400)
Interaction terms				
Interaction: Vocational training × age	– 0.01227	(0.01166)		
Interaction: Technician × age	– 0.00104	(0.01365)		
Interaction: College × age	– 0.00430	(0.01276)		
Interaction: University × age	0.02412†	(0.01290)		
Interaction: Vocational training × age squared	0.00010	(0.00014)		
Interaction: Technician × age squared	– 0.00006	(0.00016)		
Interaction: College × age squared	0.00005	(0.00015)		
Interaction: University × age squared	– 0.00024	(0.00015)		
Baseline category: partner has no vocational training				
Partner's graduation: Vocational training (2)			0.07127	(0.09212)
Partner's graduation: Technician (3)			0.16231	(0.16752)
Partner's graduation: College (4)			– 0.04199	(0.16293)
Partner's graduation: University (5)			0.23993†	(0.13764)
Interaction terms				
Interaction educational graduation 2 (own) × 2 (partner)			0.16175	(0.10458)
Interaction educational graduation 2 × 3			0.13729	(0.17883)
Interaction educational graduation 2 × 4			0.33672†	(0.17727)
Interaction educational graduation 2 × 5			– 0.08784	(0.15557)
Interaction educational graduation 3 × 2			0.30183*	(0.13940)
Interaction educational graduation 3 × 3			0.22471	(0.19962)
Interaction educational graduation 3 × 4			0.46965*	(0.20546)
Interaction educational graduation 3 × 5			0.05474	(0.18783)
Interaction educational graduation 4 × 2			0.00523	(0.14366)
Interaction educational graduation 4 × 3			– 0.13876	(0.20839)
Interaction educational graduation 4 × 4			0.24727	(0.20028)
Interaction educational graduation 4 × 5			– 0.35416†	(0.18346)
Interaction educational graduation 5 × 2			0.12427	(0.15439)
Interaction educational graduation 5 × 3			0.12369	(0.21558)
Interaction educational graduation 5 × 4			0.31263	(0.20862)
Interaction educational graduation 5 × 5			– 0.07990	(0.18355)
German citizenship	0.19657***	(0.02489)	0.44816***	(0.05983)
Partner has German citizenship			0.15327**	(0.05671)
Married (baseline category: single)	0.07515***	(0.01036)	0.13149***	(0.03523)
Widowed	– 0.38595***	(0.11491)	– 0.16545	(0.23720)
Divorced	– 0.06948***	(0.02201)	– 0.07127	(0.06066)
Permanently separated	0.03579	(0.05815)	0.05902	(0.15817)
East Germany	– 0.24787***	(0.02556)	– 0.02326	(0.07392)
Baseline category: population < 5000				

Table 8 (continued)

	Log Wage per hour		Work = 1	
Population between 5,000 and 20,000	0.03407***	(0.01038)	0.06368*	(0.02955)
Population between 20,000 and 100,000	0.03976***	(0.01084)	0.06067*	(0.03066)
Population over 100,000	0.03159***	(0.01178)	– 0.01317	(0.03346)
Student	– 0.64078***	(0.05072)	– 1.98636***	(0.06914)
Civil servant	– 0.05653***	(0.00960)		
One child (baseline category: no children)			– 0.07814*	(0.03275)
Two children			– 0.03939	(0.03484)
Three children			– 0.21018***	(0.05655)
Four or more children			– 0.35377***	(0.11213)
Constant	0.39361†	(0.22616)	– 13.09894***	(2.09765)
Dummies for the federal state, quarter and year	Yes		Yes	
Observations	27.950			
Selected	22.012			
Nonselected	5.938			
<i>Lambda</i>	0.111***	(0.03559)		

Standard errors in parentheses

Sources: EVS data 2008, 2013; own calculations

†p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001

Appendix B

See Tables 9 and 10.

Table 9 Estimation results for the demand system without endogeneity

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
Price women's clothing and footwear (log)	– 0.0307* (0.0148)	0.0051 (0.0109)	0.0255 (0.0208)
Price men's clothing and footwear (log)	0.0051 (0.0148)	– 0.0076 (0.0109)	0.0024 (0.0209)
Price composite good (log)	0.0255** (0.0095)	0.0024 (0.0070)	– 0.0280* (0.0134)
Budget (β)	0.0758*** (0.0046)	0.0385*** (0.0034)	– 0.1142*** (0.0064)
Budget squared (λ)	– 0.0097*** (0.0007)	– 0.0048*** (0.0005)	0.0145*** (0.0010)
Woman's income contribution share (α)	0.0073*** (0.0007)	– 0.0032*** (0.0005)	– 0.0042*** (0.0010)
One child (baseline category: no children)	– 0.0099*** (0.0005)	– 0.0071*** (0.0004)	0.0170*** (0.0007)
Two children	– 0.0130*** (0.0005)	– 0.0092*** (0.0004)	0.0222*** (0.0008)
Three children	– 0.0152***	– 0.0115***	0.0268***

Table 9 (continued)

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
Four or more children	(0.0009) – 0.0198***	(0.0007) – 0.0126***	(0.0013) 0.0323***
Married	(0.0019) – 0.0013**	(0.0014) – 0.0007*	(0.0027) 0.0020**
Woman's age	(0.0005) – 0.0052**	(0.0003) – 0.0006	(0.0007) 0.0057*
Woman's age squared	(0.0018) 0.0002**	(0.0013) 0.0000	(0.0026) – 0.0002*
Woman's age cubic	(0.0001) – 0.0000**	(0.0000) – 0.0000	(0.0001) 0.0000*
Woman's age quartic	(0.0000) 0.0000*	(0.0000) 0.0000	(0.0000) – 0.0000*
Man's age	(0.0000) – 0.0031†	(0.0000) – 0.0030*	(0.0000) 0.0061*
Man's age squared	(0.0019) 0.0001	(0.0014) 0.0001†	(0.0026) – 0.0001†
Man's age cubic	(0.0001) – 0.0000	(0.0000) – 0.0000	(0.0001) 0.0000
Man's age quartic	(0.0000) 0.0000	(0.0000) 0.0000	(0.0000) – 0.0000
Man German citizenship	(0.0000) 0.0000	(0.0000) 0.0004	(0.0000) – 0.0005
Woman German citizenship	(0.0009) 0.0003	(0.0007) – 0.0005	(0.0013) 0.0002
Owner-occupied housing	(0.0011) – 0.0031***	(0.0008) – 0.0022***	(0.0015) 0.0053***
Year 2013 (baseline category: year 2008)	(0.0004) 0.0016***	(0.0003) 0.0010***	(0.0005) – 0.0026***
2nd quarter (baseline category: 1st quarter)	(0.0004) 0.0027***	(0.0003) 0.0005	(0.0006) – 0.0032***
3rd quarter	(0.0005) 0.0008†	(0.0003) 0.0001	(0.0006) – 0.0009
4th quarter	(0.0004) 0.0045***	(0.0003) 0.0032***	(0.0006) – 0.0077***
Constant	(0.0005) – 0.0066	(0.0004) – 0.0011	(0.0007) 1.0077***
Dummies for the federal state and level of agglomeration	(0.0188) Yes	(0.0139) Yes	(0.0265) Yes
<i>N</i>	29.461	29.461	29.461
<i>R</i> ²	0.090	0.063	0.108
Test on joint significance of coefficients α , <i>Chi</i> ² -statistic (<i>p</i> -value)	210.5 (0.000)		

Standard errors in parentheses, asymptotic variance–covariance matrix according to Blundell and Robin (1999). Homogeneity and symmetry restrictions imposed

Sources: EVS data 2008, 2013; own calculations

†*p* < 0.10, **p* < 0.05, ***p* < 0.01, ****p* < 0.001

Table 10 Estimated elasticities for the demand system without endogeneity

	Budget elasticity	Uncompensated price elasticity Change in the price of...			Compensated price elasticity Change in the price of...		
		(1)	(2)	(3)	(1)	(2)	(3)
Women's clothing and footwear (1)	1.48*** (0.017)	– 1.72 (0.493)	0.32 (0.495)	– 0.08 (0.312)	– 1.68 (0.492)	0.35 (0.495)	1.33*** (0.314)
Men's clothing and footwear (2)	1.45*** (0.025)	– 0.55 (0.615)	– 1.30 (0.618)	– 0.70† (0.392)	– 0.59 (0.615)	– 1.27 (0.618)	0.68† (0.392)
Composite good (3)	0.98*** (0.001)	0.01 (0.022)	– 0.01 (0.022)	– 0.98 (0.014)	0.04† (0.022)	0.01 (0.022)	– 0.06*** (0.014)

Standard errors in parentheses. Homogeneity and symmetry restrictions imposed. Elasticities evaluated at sample means. Own-price elasticities tested against the null of – 1

Sources: EVS data 2008, 2013; own calculations

†p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001

Appendix C

See Tables 11 and 12.

Table 11 Estimation results for the first-stage-SUR model

	Women's income contribution share		Budget (log)	
Price women's clothing and footwear (log)	– 0.094	(0.149)	– 0.471*	(0.199)
Price men's clothing and footwear (log)	– 0.119	(0.182)	0.442†	(0.244)
Price composite good (log)	– 0.103	(0.086)	– 1.119***	(0.115)
Woman's wage per hour (log)			0.098***	(0.003)
Man's wage per hour (log)			0.167***	(0.004)
Baseline category: no school graduation (woman)				
Secondary school—2 (woman)	– 0.129*	(0.065)	– 0.099	(0.087)
Intermediate school-leaving certificate—3 (woman)	– 0.055	(0.060)	– 0.063	(0.081)
Specialized A-levels—4 (woman)	– 0.025	(0.089)	0.157	(0.120)
A-levels—5 (woman)	0.068	(0.072)	0.060	(0.097)
Baseline category: no school graduation (man)				
Secondary school—2 (man)	– 0.030	(0.068)	– 0.230*	(0.092)
Intermediate school-leaving certificate—3 (man)	0.006	(0.072)	0.097	(0.097)
Specialized A-levels—4 (man)	– 0.110	(0.092)	– 0.088	(0.124)
A-levels—5 (man)	0.070	(0.086)	0.071	(0.116)
Interaction terms				
Interaction school graduation 2 (woman) × 2 (man)	0.114	(0.078)	0.270*	(0.105)
Interaction school graduation 2 × 3	0.058	(0.082)	0.012	(0.110)
Interaction school graduation 2 × 4	0.193†	(0.100)	0.226†	(0.135)
Interaction school graduation 2 × 5	0.020	(0.094)	0.087	(0.127)
Interaction school graduation 3 × 2	0.091	(0.075)	0.286**	(0.100)
Interaction school graduation 3 × 3	0.038	(0.078)	– 0.000	(0.105)
Interaction school graduation 3 × 4	0.164†	(0.097)	0.195	(0.131)
Interaction school graduation 3 × 5	– 0.014	(0.091)	0.069	(0.123)
Interaction school graduation 4 × 2	0.074	(0.100)	0.113	(0.134)
Interaction school graduation 4 × 3	0.033	(0.102)	– 0.175	(0.137)
Interaction school graduation 4 × 4	0.146	(0.118)	– 0.018	(0.158)
Interaction school graduation 4 × 5	– 0.035	(0.113)	– 0.148	(0.151)
Interaction school graduation 5 × 2	– 0.003	(0.085)	0.217†	(0.114)
Interaction school graduation 5 × 3	– 0.062	(0.088)	– 0.070	(0.118)
Interaction school graduation 5 × 4	0.050	(0.105)	0.112	(0.141)

Table 11 (continued)

	Women's income contri- bution share		Budget (log)	
Interaction school graduation 5 × 5	– 0.124	(0.099)	– 0.033	(0.134)
Baseline category: no vocational training (woman)				
Vocational training—2 (woman)	0.081***	(0.015)	0.028	(0.020)
Technician—3 (woman)	0.059*	(0.026)	0.033	(0.036)
College—4 (woman)	0.194***	(0.026)	0.168***	(0.035)
University—5 (woman)	0.156***	(0.022)	0.090**	(0.029)
Baseline category: no vocational training (man)				
Vocational training—2 (man)	– 0.090***	(0.013)	0.026	(0.018)
Technician—3 (man)	– 0.090***	(0.018)	0.033	(0.024)
College—4 (man)	– 0.162***	(0.017)	0.123***	(0.023)
University—5 (man)	– 0.173***	(0.019)	0.067**	(0.026)
Interaction terms				
Interaction educational graduation 2 (woman) × 2 (man)	– 0.004	(0.016)	0.014	(0.022)
Interaction educational graduation 2 × 3	– 0.040†	(0.021)	0.017	(0.028)
Interaction educational graduation 2 × 4	– 0.018	(0.020)	– 0.030	(0.027)
Interaction educational graduation 2 × 5	– 0.020	(0.022)	0.038	(0.030)
Interaction educational graduation 3 × 2	0.076**	(0.028)	0.028	(0.038)
Interaction educational graduation 3 × 3	0.054†	(0.031)	0.014	(0.041)
Interaction educational graduation 3 × 4	0.040	(0.031)	– 0.057	(0.041)
Interaction educational graduation 3 × 5	0.040	(0.032)	0.033	(0.043)
Interaction educational graduation 4 × 2	– 0.009	(0.027)	– 0.080*	(0.037)
Interaction educational graduation 4 × 3	– 0.027	(0.031)	– 0.060	(0.042)
Interaction educational graduation 4 × 4	– 0.038	(0.030)	– 0.130**	(0.040)
Interaction educational graduation 4 × 5	– 0.068*	(0.031)	– 0.067	(0.042)
Interaction educational graduation 5 × 2	0.010	(0.024)	– 0.020	(0.032)
Interaction educational graduation 5 × 3	– 0.023	(0.028)	– 0.020	(0.038)
Interaction educational graduation 5 × 4	0.024	(0.026)	– 0.084*	(0.036)
Interaction educational graduation 5 × 5	– 0.001	(0.027)	0.009	(0.036)
German citizenship (man)	– 0.079***	(0.009)	0.059***	(0.012)
German citizenship (woman)	0.103***	(0.008)	0.113***	(0.011)
One child (baseline category: no children)			0.068***	(0.006)
Two children			0.144***	(0.006)
Three children			0.198***	(0.010)
Four or more children			0.246***	(0.021)
Married			– 0.004	(0.005)
Woman's age	– 0.112***	(0.015)	0.051*	(0.021)
Woman's age squared	0.004***	(0.001)	– 0.001†	(0.001)
Woman's age cubic	– 0.000***	(0.000)	0.000	(0.000)
Woman's age quartic	0.000***	(0.000)	– 0.000	(0.000)
Man's age	– 0.065***	(0.016)	0.060**	(0.022)
Man's age squared	0.001*	(0.000)	– 0.002**	(0.001)
Man's age cubic	– 0.000	(0.000)	0.000**	(0.000)
Man's age quartic	– 0.000	(0.000)	– 0.000*	(0.000)
Owner-occupied housing			0.152***	(0.004)
Constant	3.917***	(1.123)	10.249***	(1.511)
Dummies for the federal state and level of agglomeration				
Yes			Yes	
<i>N</i>	29.461		29.461	
<i>R</i> ²	0.168		0.393	
<i>F</i> -statistic (<i>p</i> -value)	52.3 (0.000)		128.2 (0.000)	

Bootstrapped standard errors in parentheses

Sources: EVS data 2008, 2013; own calculations

†*p* < 0.10, **p* < 0.05, ***p* < 0.01, ****p* < 0.001

Table 12 Estimation results for the demand system with endogenous budget and endogenous women's income contribution share

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
Price women's clothing and footwear (log)	– 0.0427** (0.0149)	– 0.0032 (0.0109)	0.0459* (0.0210)
Price men's clothing and footwear (log)	– 0.0032 (0.0150)	– 0.0139 (0.0111)	0.0171 (0.0212)
Price composite good (log)	0.0459*** (0.0099)	0.0171* (0.0073)	– 0.0630*** (0.0141)
Budget (β)	0.0834*** (0.0047)	0.0436*** (0.0034)	– 0.1270*** (0.0066)
Budget squared (λ)	– 0.0097*** (0.0007)	– 0.0048*** (0.0005)	0.0146*** (0.0010)
Women's income contribution share (α)	0.0051* (0.0024)	– 0.0070*** (0.0017)	0.0019 (0.0033)
$\hat{\nu}_1$ -Women's income contribution share	0.0026 (0.0025)	0.0042* (0.0018)	– 0.0068† (0.0035)
$\hat{\nu}_2$ -Budget	– 0.0091*** (0.0011)	– 0.0062*** (0.0008)	0.0154*** (0.0016)
One child (baseline category: no children)	– 0.0102*** (0.0005)	– 0.0073*** (0.0004)	0.0175*** (0.0007)
Two children	– 0.0139*** (0.0005)	– 0.0098*** (0.0004)	0.0237*** (0.0008)
Three children	– 0.0166*** (0.0009)	– 0.0125*** (0.0007)	0.0290*** (0.0013)
Four or more children	– 0.0214*** (0.0019)	– 0.0137*** (0.0014)	0.0351*** (0.0027)
Married	– 0.0013** (0.0005)	– 0.0007† (0.0003)	0.0020** (0.0007)
Woman's age	– 0.0065*** (0.0019)	– 0.0016 (0.0014)	0.0081** (0.0026)
Woman's age squared	0.0002*** (0.0001)	0.0001 (0.0000)	– 0.0003** (0.0001)
Woman's age cubic	– 0.0000** (0.0000)	– 0.0000 (0.0000)	0.0000** (0.0000)
Woman's age quartic	0.0000** (0.0000)	0.0000 (0.0000)	– 0.0000** (0.0000)
Man's age	– 0.0043* (0.0019)	– 0.0041** (0.0014)	0.0084** (0.0027)
Man's age squared	0.0001† (0.0001)	0.0001* (0.0000)	– 0.0002* (0.0001)
Man's age cubic	– 0.0000 (0.0000)	– 0.0000† (0.0000)	0.0000* (0.0000)
Man's age quartic	0.0000 (0.0000)	0.0000 (0.0000)	– 0.0000 (0.0000)
Man German citizenship	– 0.0004 (0.0010)	0.0004 (0.0007)	0.0000 (0.0014)
Woman German citizenship	– 0.0007 (0.0011)	– 0.0014† (0.0008)	0.0020 (0.0015)
Owner-occupied housing	– 0.0045*** (0.0004)	– 0.0032*** (0.0003)	0.0077*** (0.0006)
Year 2013 (baseline category: year 2008)	0.0017*** (0.0004)	0.0010*** (0.0003)	– 0.0027*** (0.0006)
2nd quarter (baseline category: 1st quarter)	0.0028***	0.0006†	– 0.0034***

Table 12 (continued)

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
3rd quarter	(0.0005) 0.0003 (0.0004)	(0.0003) – 0.0002 (0.0003)	(0.0006) – 0.0001 (0.0006)
4th quarter	0.0045*** (0.0005)	0.0033*** (0.0004)	– 0.0077*** (0.0007)
Constant	0.0068 (0.0197)	0.0138 (0.0145)	0.9794*** (0.0279)
Dummies for the federal state and level of agglomeration	Yes	Yes	Yes
<i>N</i>	29,461	29,461	29,461
Shea's partial R^2 between s and its instruments	0.086		
Test on joint significance of coefficients of \hat{v}_1 , χ^2 -statistic (p -value)	5.6 (0.061)		
Test on joint significance of coefficients of \hat{v}_2 , χ^2 -statistic (p -value)	91.4 (0.000)		
Test on joint significance of coefficients of α , χ^2 -statistic (p -value)	28.9 (0.000)		

Adjusted standard errors in parentheses, asymptotic variance–covariance matrix according to Blundell and Robin (1999). Homogeneity and symmetry restrictions imposed

Sources: EVS data 2008, 2013; own calculations

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix D—Correction of Censored Dependent Variables

In the equation for men's clothing and footwear, there are 15.4 percent of households, which have zero expenditures. In principle, this could be a problem for estimation in terms of a censored depended variable and therefore a Shonkwiler and Yen (1999) correction of the linearized version of the demand system is done in this robustness check. The correction in the non-linear demand system would be hard to compute but the differences between linear and non-linear demand system mainly relate to the estimated price and budget parameters (as shown below) but not the parameters of interest ϕ_i . So, the impact of selection correction on the parameter of interest can be evaluated in the linearized model. In the first step, a probit regression is run for the indicator if men's clothing and footwear expenditures are > 0 . All explanatory variables of the demand system are included in this regression except for the parameters of interest. Additionally, some excluded variables from the IV-specification are also included, as endogeneity is neglected in this robustness check for selection bias. In the second step, all explanatory variables in the correspondent equation of the demand system are multiplied by the predicted probabilities from the probit. Furthermore, the linear predictions from the probit inserted in the normal density function are added to the equation as a selection term.

Table 13 shows that correcting for censored demand of men's clothing and footwear has neglectable effects on the parameter of interest nor is the correction term in the equation significant. There are only small changes for the coefficients of the price effects which are however sensitive to estimate due to the low-price variation and the imposed restrictions.

Appendix E—Robustness Checks

A person's preference for clothing and footwear may depend on the occupational status in a way that expenditures could potentially be higher for occupied persons than for unemployed, or that they could be higher in higher paid jobs than in low-income jobs. As the main specification of the model features couples with the presence of income from occupation in the household from at least one partner, this could be a factor in the hypothesis test. The instrumentation of the income contribution share should tackle the problem sufficiently as it only depends on educational backgrounds and not on actual occupational status in consequence. However, there could potentially be a correlation left between the preferences for clothing/footwear and unobserved factors regarding the occupational status because educational information is too broadly defined (e.g. not specifying the potential job positions).

Table 13 Estimation results for the demand system controlling for censored demand for men's clothing and footwear (linear demand system)

	Dependent variables: Expenditure shares			
	Women's clothing and footwear	Men's clothing and footwear	Women's clothing and footwear	Men's clothing and footwear
	Original specification		Correcting for censored men's clothing and footwear demand	
Price women's clothing and footwear (log)	– 0.0322** (0.0125)	– 0.0184† (0.0097)	– 0.0353** (0.0130)	– 0.0218* (0.0108)
Price men's clothing and footwear (log)	0.0184† (0.0097)	0.0175† (0.0105)	0.0218* (0.0108)	0.0155 (0.0123)
Price composite good (log)	0.0138 (0.0088)	0.0009 (0.0066)	0.0135 (0.0088)	0.0063 (0.0079)
Budget (β)	0.0759*** (0.0046)	0.0387*** (0.0034)	0.0762*** (0.0045)	0.0397*** (0.0065)
Budget squared (λ)	– 0.0097*** (0.0007)	– 0.0048*** (0.0005)	– 0.0097*** (0.0007)	– 0.0051*** (0.0009)
Women's income contribution share (α)	0.0073*** (0.0007)	– 0.0032*** (0.0005)	0.0073*** (0.0007)	– 0.0039*** (0.0006)
Selection correction term				– 0.0011 (0.0084)
Constant	– 0.0062 (0.0188)	– 0.0011 (0.0139)	– 0.0046 (0.0182)	0.0074 (0.0055)
Polynomials for the age of both partners, dummies of German citizenship, dummies for the federal state, level of agglomeration, quarter and survey year	Yes	Yes	Yes	Yes
N	29.461	29.461	29.461	29.461
R^2	0.091	0.064	0.091	0.065

Adjusted standard errors in parentheses. Homogeneity and symmetry restrictions imposed

Sources: EVS data 2008, 2013; own calculations

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Therefore, two robustness checks are done for the IV specification: Firstly, the sample is restricted to both partners having market income (and thus being occupied) and secondly to women's income contribution shares lying between 0.2 and 0.8, which gives both partners a significant contribution to the household budget. The first check should clarify the influence of a preference shift from (non-)occupation of a partner, while the second one should evaluate the situation in which both incomes are not too far different from each other. The first stage remains in both checks the same as in the main specification to avoid selection effects in the coefficients of the instruments.

The results for the parameters α_i are presented in Table 14. For a better comparison with the initial parameters from the IV model of Table 4, these are repeated in the first rows. In the first robustness check with both partners having positive market income, the effect on women's clothing and footwear is smaller and insignificant while the one on

men's expenditures is larger. In consequence, there is more substitution between men's consumption and the composite good, but a rejection of the hypothesis. This result is even stronger if the sample is restricted to s lying between 0.2 and 0.8. Since the average share of women's income contribution is increasing going from the original model to the first and second robustness check, there is a possibly nonlinear effect on consumption pattern. For low values of s , there is a positive effect on women's clothing and footwear consumption. For higher values, the effect becomes stronger for other goods contained in the composite good and lower for clothing and footwear. However, the rejection of the income pooling hypothesis persists in any case, since the negative relationship between women's share and men's private consumption holds (Table 14).

Detailed estimation results for these models can be found on the following pages (Tables 15 and 16).

Table 14 Robustness checks: Different samples

	Dependent variables: Expenditure shares			Chi^2 -statistic on system significance
	Women's clothing and footwear	Men's clothing and footwear	Composite good	
Original effects (IV), $N=29,461$				
Women's income contribution share (α)	0.0051* (0.0024)	− 0.0070*** (0.0017)	0.0019 (0.0033)	28.9, 0.000
\hat{v}_1	0.0026 (0.0025)	0.0042* (0.0018)	− 0.0068† (0.0035)	5.6, 0.061
Sample restriction on both partners having positive market incomes, $N=19,322$				
Women's income contribution share (α)	0.0041 (0.0030)	− 0.0098*** (0.0023)	0.0057 (0.0043)	26.6, 0.000
\hat{v}_1	0.0034 (0.0032)	0.0060* (0.0024)	− 0.0094* (0.0045)	6.4, 0.041
Sample restriction on s lying between 0.2 and 0.8, $N=19,257$				
Women's income contribution share (α)	0.0030 (0.0032)	− 0.0124*** (0.0023)	0.0094* (0.0045)	36.0, 0.000
\hat{v}_1	0.0037 (0.0033)	0.0067** (0.0024)	− 0.0103* (0.0046)	8.0, 0.018

Adjusted standard errors in parentheses, asymptotic variance–covariance matrix according to Blundell and Robin (1999). Homogeneity and symmetry restrictions imposed. The right column contains p -values of the chi-squared test

Sources: EVS data 2008, 2013; own calculations

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15 Robustness check: Sample restriction on both partners having positive market incomes

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
Price women's clothing and footwear (log)	– 0.0350† (0.0190)	– 0.0088 (0.0141)	0.0438 (0.0270)
Price men's clothing and footwear (log)	– 0.0088 (0.0192)	– 0.0123 (0.0143)	0.0211 (0.0273)
Price composite good (log)	0.0438*** (0.0128)	0.0211* (0.0095)	– 0.0650*** (0.0182)
Budget (β)	0.0812*** (0.0064)	0.0433*** (0.0048)	– 0.1245*** (0.0091)
Budget squared (λ)	– 0.0094*** (0.0010)	– 0.0047*** (0.0007)	0.0141*** (0.0014)
Women's income contribution share (α)	0.0041 (0.0030)	– 0.0098*** (0.0023)	0.0057 (0.0043)
\hat{v}_1 -Women's income contribution share	0.0034 (0.0032)	0.0060* (0.0024)	– 0.0094* (0.0045)
\hat{v}_2 -Budget	– 0.0092*** (0.0014)	– 0.0066*** (0.0011)	0.0158*** (0.0020)
One child (baseline category: no children)	– 0.0102*** (0.0006)	– 0.0077*** (0.0004)	0.0178*** (0.0009)
Two children	– 0.0138*** (0.0007)	– 0.0103*** (0.0005)	0.0242*** (0.0009)
Three children	– 0.0174*** (0.0012)	– 0.0135*** (0.0009)	0.0309*** (0.0017)

Table 15 (continued)

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
Four or more children	– 0.0201*** (0.0028)	– 0.0141*** (0.0021)	0.0342*** (0.0040)
Married	– 0.0021*** (0.0006)	– 0.0008* (0.0004)	0.0029*** (0.0008)
Woman's age	– 0.0009 (0.0029)	– 0.0020 (0.0022)	0.0029 (0.0041)
Woman's age squared	0.0000 (0.0001)	0.0001 (0.0001)	– 0.0001 (0.0001)
Woman's age cubic	0.0000 (0.0000)	– 0.0000 (0.0000)	0.0000 (0.0000)
Woman's age quartic	– 0.0000 (0.0000)	0.0000 (0.0000)	– 0.0000 (0.0000)
Man's age	– 0.0099*** (0.0029)	– 0.0049* (0.0022)	0.0148*** (0.0042)
Man's age squared	0.0003** (0.0001)	0.0001† (0.0001)	– 0.0004** (0.0001)
Man's age cubic	– 0.0000** (0.0000)	– 0.0000 (0.0000)	0.0000** (0.0000)
Man's age quartic	0.0000* (0.0000)	0.0000 (0.0000)	– 0.0000* (0.0000)
Man German citizenship	– 0.0001 (0.0014)	0.0010 (0.0010)	– 0.0009 (0.0019)
Woman German citizenship	– 0.0003 (0.0015)	– 0.0007 (0.0011)	0.0010 (0.0021)
Owner-occupied housing	– 0.0043*** (0.0005)	– 0.0032*** (0.0004)	0.0074*** (0.0007)
Year 2013 (baseline category: year 2008)	0.0022*** (0.0005)	0.0010** (0.0004)	– 0.0032*** (0.0007)
2nd quarter (baseline category: 1st quarter)	0.0025*** (0.0006)	0.0004 (0.0004)	– 0.0029*** (0.0008)
3rd quarter	– 0.0000 (0.0005)	– 0.0007 (0.0004)	0.0007 (0.0008)
4th quarter	0.0042*** (0.0006)	0.0031*** (0.0005)	– 0.0073*** (0.0009)
Constant	0.0153 (0.0290)	0.0276 (0.0216)	0.9572*** (0.0412)
Dummies for the federal state and level of agglomeration	Yes	Yes	Yes
<i>N</i>	19.322	19.322	19.322

Adjusted standard errors in parentheses, asymptotic variance–covariance matrix according to Blundell and Robin (1999). Homogeneity and symmetry restrictions imposed

Sources: EVS data 2008, 2013; own calculations

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 16 Robustness check:
Sample restriction on s lying
between 0.2 and 0.8

	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
Price women's clothing and footwear (log)	– 0.0467* (0.0193)	– 0.0035 (0.0139)	0.0502† (0.0271)
Price men's clothing and footwear (log)	– 0.0035 (0.0196)	– 0.0181 (0.0141)	0.0216 (0.0275)
Price composite good (log)	0.0502*** (0.0130)	0.0216* (0.0093)	– 0.0718*** (0.0183)
Budget (β)	0.0895*** (0.0061)	0.0502*** (0.0044)	– 0.1397*** (0.0085)
Budget squared (λ)	– 0.0104*** (0.0009)	– 0.0058*** (0.0007)	0.0161*** (0.0013)
Women's income contribution share (α)	0.0030 (0.0032)	– 0.0124*** (0.0023)	0.0094* (0.0045)
\hat{d}_1 -Women's income contribution share	0.0037 (0.0033)	0.0067** (0.0024)	– 0.0103* (0.0046)
\hat{d}_2 -Budget	– 0.0117*** (0.0015)	– 0.0077*** (0.0011)	0.0194*** (0.0021)
One child (baseline category: no children)	– 0.0108*** (0.0006)	– 0.0080*** (0.0005)	0.0188*** (0.0009)
Two children	– 0.0147*** (0.0007)	– 0.0109*** (0.0005)	0.0256*** (0.0010)
Three children	– 0.0203*** (0.0016)	– 0.0148*** (0.0011)	0.0351*** (0.0022)
Four or more children	– 0.0241*** (0.0037)	– 0.0150*** (0.0027)	0.0391*** (0.0052)
Married	– 0.0017** (0.0006)	– 0.0008* (0.0004)	0.0025** (0.0008)
Woman's age	– 0.0071** (0.0024)	– 0.0008 (0.0018)	0.0079* (0.0034)
Woman's age squared	0.0002** (0.0001)	0.0000 (0.0001)	– 0.0002* (0.0001)
Woman's age cubic	– 0.0000* (0.0000)	– 0.0000 (0.0000)	0.0000† (0.0000)
Woman's age quartic	0.0000* (0.0000)	0.0000 (0.0000)	– 0.0000† (0.0000)
Man's age	– 0.0062* (0.0025)	– 0.0050** (0.0018)	0.0111** (0.0035)
Man's age squared	0.0002* (0.0001)	0.0001* (0.0001)	– 0.0003** (0.0001)
Man's age cubic	– 0.0000† (0.0000)	– 0.0000* (0.0000)	0.0000* (0.0000)
Man's age quartic	0.0000† (0.0000)	0.0000† (0.0000)	– 0.0000* (0.0000)
Man German citizenship	0.0006 (0.0015)	0.0009 (0.0011)	– 0.0015 (0.0022)
Woman German citizenship	– 0.0026† (0.0015)	– 0.0017 (0.0011)	0.0043* (0.0022)
Owner-occupied housing	– 0.0052*** (0.0005)	– 0.0035*** (0.0004)	0.0087*** (0.0007)
Year 2013 (baseline category: year 2008)	0.0019***	0.0010**	– 0.0029***

Table 16 (continued)

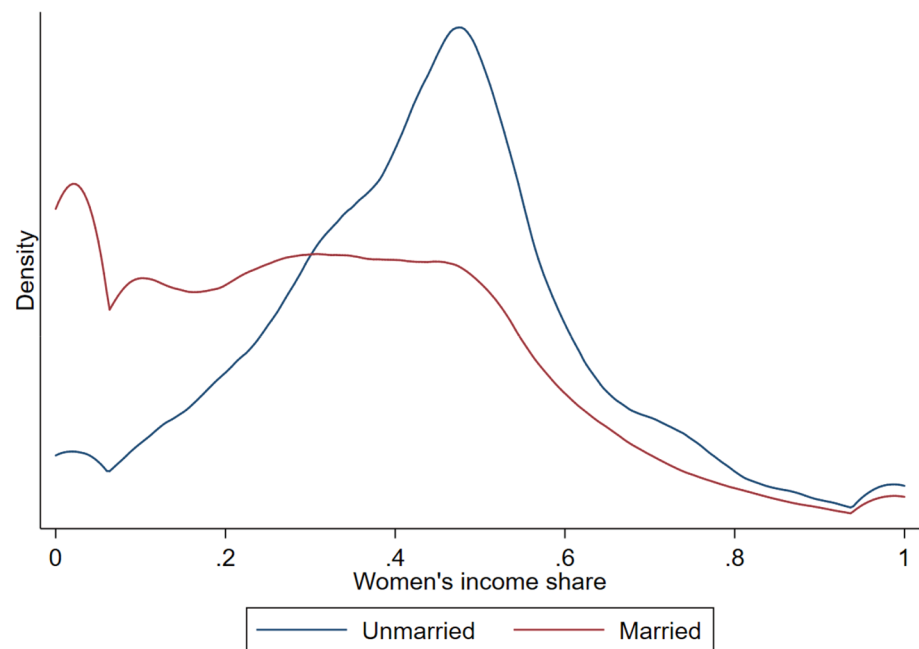
	Dependent variables: Expenditure shares		
	Women's clothing and footwear	Men's clothing and footwear	Composite good
2nd quarter (baseline category: 1st quarter)	(0.0005) 0.0026***	(0.0004) 0.0002	(0.0007) – 0.0028***
3rd quarter	(0.0006) – 0.0005	(0.0004) – 0.0008*	(0.0008) 0.0013†
4th quarter	(0.0006) 0.0045***	(0.0004) 0.0031***	(0.0008) – 0.0076***
Constant	(0.0007) 0.0247	(0.0005) 0.0068	(0.0009) 0.9685***
Dummies for the federal state and level of agglomeration	(0.0256) Yes	(0.0184) Yes	(0.0359) Yes
<i>N</i>	19.257	19.257	19.257

Adjusted standard errors in parentheses, asymptotic variance–covariance matrix according to Blundell and Robin (1999). Homogeneity and symmetry restrictions imposed

Sources: EVS data 2008, 2013; own calculations

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Fig. 3 Estimated density of women's income share. Kernel density estimation. Sources: EVS data 2008, 2013; own calculations



Notes: Kernel density estimation.

Sources: EVS data 2008, 2013; own calculations.

Appendix F

The distribution of women's income share varies systematically between married and unmarried couples (see Fig. 3). The conditional mean for married couples is 0.322, while it is 0.439 for unmarried couples. However, women remain the

second earners at the mean even for the conditional cases and therefore for the results of the interaction models. Based on the estimated coefficients and the formulas given in Section “[The Link to the Taxation of Couple Households](#)”, Table 17 shows the justifiable relative differentials of the marginal income tax rates between first and second earner. Because

Table 17 Relative differentials of the marginal income tax rate between first and second earner in couple households

Model	Relative differential of the marginal tax rate	\bar{s}_w (average of women's income share in the subpopulation)
Pooled model	0.098 (0.046)	0.342
Interaction model 1		
Married	0.075 (0.047)	0.322
Unmarried	0.202 (0.048)	0.439
Interaction model 2		
Married—West Germany	0.052 (0.050)	0.286
Unmarried—West Germany	0.232 (0.062)	0.435
Married—East Germany	0.100 (0.050)	0.431
Unmarried—East Germany	0.098 (0.071)	0.448
Interaction model 3		
No child	0.131 (0.043)	0.387
At least one child	0.005 (0.055)	0.262
Interaction model 4		
Unmarried—no child	0.220 (0.048)	0.458
Unmarried—at least one child	0.072 (0.091)	0.380
Married—no child	0.106 (0.045)	0.369
Married—at least one child	– 0.003 (0.061)	0.247

Evaluated at the conditional mean. Standard errors in parentheses

Sources: EVS data 2008, 2013; own calculations

these results are based only on clothing and footwear consumption taken as a proxy for total private consumption, the relevance for actual tax policy is limited. Further analysis should incorporate broader measures of consumption that also allow for durables and intertemporal consumption choices.

The calculated tax differentials have in general the lowest values for married couples. In interaction model 1, women in married couples should have a 7.5 percent lower marginal tax rate than the men (but not significant). So, if the shared marginal tax rate for the couple in case of joint assessment is for instance 0.3, women's marginal tax rate should be 2.3 percentage points lower than the one of the men. The tax differential for unmarried couples is higher with 20.2 percent. This value can overshoot the spread in the marginal tax rate even for individual tax assessment, for example at the mean of women's income share where gross incomes of the partners are similar, but it is a plausible result because of the distorted allocation of the private goods. It would thus be justified in terms of the ability-to-pay principle of taxation to lower the tax rate for women further at the expense of the one for men compared to individual taxation. However, since the debate in practice moves within the borders of individual versus joint tax assessment, the tax differential would be capped by the one resulting from individual taxation.

The differences in the results for West and East German couples are also reflected in the tax differentials. There is a large gap between the tax differential in West Germany depending on the marital status, while the marital status in

East Germany makes no significant difference. For married couples in West Germany, the resulting tax differential is very low coinciding with a low-income share of women at the mean. In interaction model 4, the same is true for unmarried couples with at least one child although women's income share is much higher. This demonstrates that, given the underlying assumptions, joint tax assessment can be justified independently of the income share and consequently also independently of the labor market participation. A significant tax differential for married couples without children remains in this model.

Acknowledgements I would like to thank two anonymous referees, José Alberto Molina, Viktor Steiner, Elena Del Rey, Jochen Pimpertz, Tobias Hentze, Susanna Kochskämper, Judith Niehues, Maximilian Stockhausen and participants of the 75th Annual Congress of the IIPF and VfS-Jahrestagung 2019 for helpful comments.

Funding Open Access funding enabled and organized by Projekt DEAL. Financial support by the Carl-Deilmann foundation (C. D.-Stiftung) in Bad Bentheim, Germany, through the research project 'Family Taxation in Germany' is gratefully acknowledged.

Data Availability The used microdata was provided by the Research Data Center of the Federal Statistical Office and Statistical Offices of the Länder. The anonymized data can be requested for research use (<http://forschungsdatenzentrum.de>).

Declarations

Conflicts of interest The author has no competing interests to declare that are relevant to the content of this article.

Ethical Approval This research article is a secondary data analysis and involves no own human or animal participants.

Consent for Participant Not applicable.

Consent for Publication The used secondary data only involves anonymized data. Therefore, participant consent was not required.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alesina, A., Ichino, A., & Karabarbounis, L. (2011). Gender-based taxation and the division of family chores. *American Economic Journal: Economic Policy*, 3(2), 1–40. <https://doi.org/10.1257/pol.3.2.1>
- Attanasio, O., & Lechene, V. (2002). Tests of income pooling in household decisions. *Review of Economic Dynamics*, 5(4), 720–748. <https://doi.org/10.1006/redy.2002.0191>
- Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *The Review of Economics and Statistics*, 79(4), 527–539.
- Beblo, M., & Beninger, D. (2017). Do husbands and wives pool their incomes? A couple experiment. *Review of Economics of the Household*, 15(3), 779–805. <https://doi.org/10.1007/s11150-016-9342-0>
- Becker, G. (1991). *A treatise on the family* (Enlarged). Harvard University Press.
- Becker, J. (2022). *Ehegattenbesteuerung in Deutschland*. Bertelsmann Stiftung.
- Beznoska, M. (2014). Estimating a consumer demand system of energy, mobility and leisure: A microdata approach for Germany. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2433025>
- Bick, A., & Fuchs-Schündeln, N. (2017). Quantifying the disincentive effects of joint taxation on married women's labor supply. *American Economic Review*, 107(5), 100–104.
- Blundell, R., Chiappori, P.-A., & Meghir, C. (2005). Collective labor supply with children. *Journal of Political Economy*, 113(6), 1277–1306. <https://doi.org/10.1086/491589>
- Blundell, R., & Robin, J.-M. (1999). Estimation in large and disaggregated demand systems: An estimator for conditionally linear systems. *Journal of Applied Econometrics*, 14(3), 209–232.
- Bonke, J., & Browning, M. (2009a). Allocation of expenditures within the household: A new Danish survey. *Fiscal Studies*, 30(3/4), 461–481.
- Bonke, J., & Browning, M. (2009b). The distribution of financial well-being and income within the household. *Review of Economics of the Household*, 7(1), 31–42. <https://doi.org/10.1007/s11150-008-9044-3>
- Bonke, J., & Uldall-Poulsen, H. (2007). Why do families actually pool their income? Evidence from Denmark. *Review of Economics of the Household*, 5(2), 113–128. <https://doi.org/10.1007/s11150-007-9010-5>
- Bose-Duker, T., Henry, M., & Strobl, E. (2021). Children's resource shares: male versus female-headed households. *Journal of Family and Economic Issues*, 42(4), 573–585. <https://doi.org/10.1007/s10834-020-09743-3>
- Bourguignon, F., Browning, M., Chiappori, P.-A., & Lechene, V. (1993). Intra household allocation of consumption: A model and some evidence from French data. *Annales D'économie Et De Statistique*, 29, 137–156. <https://doi.org/10.2307/20075899>
- Browning, M., Bourguignon, F., Chiappori, P.-A., & Lechene, V. (1994). Incomes and outcomes: A structural model of intrahousehold allocation. *Journal of Political Economy*, 102(6), 1067–1096.
- Browning, M., Chiappori, P.-A., & Lechene, V. (2010). Distributional effects in household models: Separate spheres and income pooling. *The Economic Journal*, 120(545), 786–799. <https://doi.org/10.1111/j.1468-0297.2009.02311.x>
- Browning, M., Chiappori, P.-A., & Lewbel, A. (2013). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. *Review of Economic Studies*, 80(4), 1267–1303. <https://doi.org/10.1093/restud/rdt019>
- Bütikofer, A., & Gerfin, M. (2017). The economies of scale of living together and how they are shared: Estimates based on a collective household model. *Review of Economics of the Household*, 15(2), 433–453. <https://doi.org/10.1007/s11150-014-9255-8>
- Büttner, T., Erbe, K., & Grimm, V. (2019). Tax planning of married couples and intra-household inequality. *Journal of Public Economics*. <https://doi.org/10.1016/j.jpubeco.2019.104048>
- Cherchye, L., De Rock, B., & Vermeulen, F. (2012). Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. *American Economic Review*, 102(7), 3377–3405.
- Deaton, A., & Muellbauer, J. (1980). *Economics and consumer behavior*. Cambridge University Press.
- Fisher, P. (2016). British tax credit simplification, the intra-household distribution of income and family consumption. *Oxford Economic Papers*, 68(2), 444–464. <https://doi.org/10.1093/oepp/gpv068>
- Gayle, G.-L., & Shephard, A. (2019). Optimal taxation, marriage, home production, and family labor supply. *Econometrica*, 87(1), 291–326. <https://doi.org/10.3982/ECTA14528>
- Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161. <https://doi.org/10.2307/1912352>
- Hoderlein, S., & Mihaleva, S. (2008). Increasing the price variation in a repeated cross section. *Journal of Econometrics*, 147(2), 316–325. <https://doi.org/10.1016/j.jeconom.2008.09.022>
- Lecocq, S., & Robin, J.-M. (2015). Estimating almost-ideal demand systems with endogenous regressors. *Stata Journal*, 15(2), 554–573. <https://doi.org/10.1177/1536867X1501500214>
- Lewbel, A. (1989). Identification and estimation of equivalence scales under weak separability. *The Review of Economic Studies*, 56(2), 311–316. <https://doi.org/10.2307/2297464>
- Lewbel, A. (1996). Aggregation without separability: A generalized Composite commodity theorem. *American Economic Review*, 86(3), 524–543.
- Lundberg, S. J., Pollak, R. A., & Wales, T. J. (1997). Do husbands and wives pool their resources? Evidence from the U.K. child benefit. *Journal of Human Resources*, 32(3), 463–480. <https://doi.org/10.2307/146179>
- Maitra, P., & Ray, R. (2006). Household expenditure patterns and resource pooling: Evidence of changes in post-apartheid South Africa. *Review of Economics of the Household*, 4(4), 325–347. <https://doi.org/10.1007/s11150-006-0011-6>
- Phipps, S., & Burton, P. (1998). What's mine is yours? The influence of male and female incomes on patterns of household expenditure. *Economica*, 65(260), 599–613.
- Poi, B. (2012). Easy demand-system estimation with quads. *The Stata Journal*, 12(3), 433–446. <https://doi.org/10.1177/1536867X1201200306>

- Shea, J. (1997). Instrument relevance in multivariate linear models: A simple measure. *Review of Economics and Statistics*, 79(2), 348–352. <https://doi.org/10.1162/rest.1997.79.2.348>
- Shonkwiler, J. S., & Yen, S. T. (1999). Two-step estimation of a censored system of equations. *American Journal of Agricultural Economics*, 81(4), 972–982. <https://doi.org/10.2307/1244339>
- Ward-Batts, J. (2008). Out of the wallet and into the purse: Using micro data to test income pooling. *Journal of Human Resources*, 43(2), 325–351.
- West, S., & Williams, R. (2004). Estimates from a consumer demand system: Implications for the incidence of environmental taxes. *Journal of Environmental Economics and Management*, 47(3), 535–558. <https://doi.org/10.1016/j.jeem.2003.11.004>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.