IARIW 2024

IARIW 2024

Thursday, August 22 – Friday, August 30

Estimating Equivalence Scales from Satisfaction Data with Endogenous Household Size and Income

Susanne Elisabeth Elsas

(University of Bamberg)

Melanie Borah

(University of Magdeburg)

Paper prepared for the 38th IARIW General Conference
August 26-30, 2024

THEME 12: Measuring Well-being for Policy: Theory and Application

Time: Friday, August 30, 2024 [13:00-14:00 GMT]

Estimating equivalence scales from satisfaction data with endogenous income and household size data

Susanne Elsas
State Institute for Family Research at the University of Bamberg, Germany

Melanie Borah
Otto von Guericke University Magdeburg, Germany

July 31, 2024

Abstract

Analyses of income inequality across households crucially depend on equivalence scales. They define income increments necessary to keep a household's living standard constant as it is joined by additional adults or children. Such scales have frequently been estimated using income satisfaction data, yet under the assumption that household income, size and structure are exogenous. The present paper is the first to relax this assumption and consider the possible endogeneity of income and family size in income satisfaction. This involves an empirical analysis of SOEP data using fixed-effects regressions with heteroscedasticity-based instruments. Our results confirm that endogeneity is relevant in the regression; equivalence weights, however, appear to be not generally biased from endogeneity.

1 Introduction

Equivalence scales are an important tool used in the analysis of economic inequality and poverty. They summarize percentage increments in expenditure needed to keep a household's welfare constant as it is joined by additional members. Among the various approaches to estimating equivalence scales, one strand in the literature uses subjective evaluations of satisfaction with own household income to proxy households' living standards. In the last 20 years, various studies using panel data on income satisfaction have shown that equivalence weights for adults and children are typically smaller than those suggested by the commonly applied OECD and square-root scale (see e.g., Schwarze (2003), van Praag and Ferrer-i-Carbonell (2004), Rojas (2007), Biewen and Juhasz (2017), Buetikofer and Gerfin (2017), Borah et al. (2019), Rapp (2021)). Regressing satisfaction on household income, family size and composition along with other explanatory variables, all of these studies assume that both, income and the number of household members, are exogenous.

In this paper, we question this assumption arguing that household income as well as family size may not only be a determinant but also an outcome of income satisfaction. Several empirical studies have provided evidence for a positive effect running from satisfaction to income (see Graham et al. (2004), De Neve and Oswald (2012), Mishra and Smyth (2014) and Elsas (2021) Prati (2017)). This link may exist because of actual causality but also because of systematic misreporting. Similarly, marital status, fertility and hence family size and composition seem to be explained by satisfaction (see Stutzer and Frey (2006), Parr (2010), Le Moglie et al. (2015), Cetre et al. (2016) and Mencarini et al. (2018)). These studies suggest that happier individuals are more likely to get married and have children. Thus, one may expect a positive effect of satisfaction on household size.

These findings call for an assessment of the consequences of endogeneity of income and household size in satisfaction. To check if regression coefficients and derived equivalence scale parameters are biased when exogeneity is assumed but actually is not met, we compare results from the conventional model to results from estimations in which instruments for household income and family size and structure are being used. Because plausible external instruments are absent, we apply internal instruments as proposed by Lewbel (2012). In an empirical application using data from the German Socio-Economic Panel (SOEP) v38.1, we find that there is considerable heteroskedasticity in income, family size and family structure across the respondent's and interviewer's age distribution. This can be used to construct strong internal instruments that we apply in the identification of household equivalence scales.

Our estimation results suggest that endogeneity of household income and of household size and structure in income satisfaction attenuates the estimated effects. Therefore the true costs of additional household members and the benefits of higher household income are understated. Effects on the estimated equivalence scale, though, appear only in samples where household composition does not change only from children passing the scale relevant age threshold. Equivalence weights that are estimated from the entire population sample are almost not biased from endogeneity.

The paper is structured as follows. Section 2 introduces the conventional model used to estimate equivalence scales from income satisfaction data. It discusses the likely consequences of endogeneity of income and household size in the satisfaction regression and presents our approach of using internal instruments to estimate unbiased coefficients. Section 3 introduces the data and reports descriptive statistics, before Section 4 shows our estimation results. In Section 5, we illustrate that the sort of endogeneity we focus on has only little implications with respect to the income distribution, but notable implications for satisfaction regressions. Section 6 concludes.

2 Model

There is a large strand in the economic literature that is concerned with the assessment of household economies of scale and the cost of children. Over the last five decades, contributions to this strand have made increasing use of subjective survey data to estimate income equivalence scales. Especially in the analysis of income satisfaction, important advances have been made with respect to estimation methods and model specification (see e.g., Biewen and Juhasz (2017), Borah et al. (2019), Rapp (2021)). Despite of differences in each of the underlying research designs, all of these recent studies support earlier findings of additional household members' consumption needs being relatively small (Schwarze (2003), van Praag and Ferrer-i-Carbonell (2004), Rojas (2007)). Borah et al. (2019) and Rapp (2021), however, disagree with the previous result that children's needs are significantly lower than those of additional adults. A very influential paper of the earlier mentioned ones is that by Schwarze (2003). The author follows Coulter et al. (1992) in proposing a linear model of latent income satisfaction that distinguishes the equivalence scale by the number of adults and children. Given that latent satisfaction cannot be observed, Schwarze (2003) estimates an ordered probit model on pooled cross-sections and a binary probit model with individual fixed effects. Its panel application has made the model an appealing starting point for many subsequent studies. In this paper, we will use a linear model specification closely corresponding to Schwarze (2003) to derive our baseline results. We will then illustrate the consequences of endogeneity in household size, structure and income in this very framework. Throughout, we will assume cardinality of the income satisfaction data.[1]

To clarify the basis of our analysis, we will now briefly introduce the fixed-effects model that Schwarze (2003) employed, where we make the additional assumption that latent equals stated satisfaction. Suppose that income satisfaction s stated by respondent i at time t represents an evaluation of this household member's consumption possibilities, i.e. equivalent income, rather than unadjusted household income y. If we assume the marginal utility of equivalent income to be decreasing, income satisfaction can be expressed as a function of log equivalent income and other potentially important control variables X_{it}. α_i represents the individual-fixed effect.

$$s_{it} = \beta_1 \ln\left(\frac{y_{it}}{EqSc_{it}}\right) + X_{it}'\gamma + \alpha_i + \varepsilon_{it} \tag{1}$$

Let the equivalence scale $EqSc$ be of the functional form $EqSc_{it} = h_{it}^{(a-bk_{it})}$, where h is the number of household members, k is the number of children and a and b are the equivalence scale parameters of interest. Parameter a represents the equivalence scale elasticity for households whose members are all adults. It determines the percentage increase in income necessary to keep the household members financial welfare constant as the household experiences a relative increase in the number of adult household members. Parameter b describes a linear decline of the equivalence scale elasticity in the number of children at a given household size. It thus captures possibly lower needs of children compared to adults. With this formulation of the equivalence scale, the above equation can be translated into the following linear equation.

$$s_{it} = \beta_1 \ln(y_{it}) - \beta_1 a \ln(h_{it}) + \beta_1 b k_{it} \ln(h_{it}) + X_{it}'\gamma + \alpha_i + \varepsilon_{it} \tag{2}$$

Equivalence scale parameters a and b can thus be identified by dividing the negative of the coefficient on log household size and the coefficient on the interaction between log household size and the number of children by the coefficient on log household income, respectively.

In the first step of our empirical analysis, we will conduct linear fixed effects regressions of Eq. (2). We argue that the results from this replication may be biased, however. This is because endogeneity of income, household size and the number of children in income satisfaction im-

[1] Ferrer-i-Carbonell and Frijters (2004) found that results from happiness regressions are not particularly sensitive to the choice of estimators for discrete data versus those for continuous data. Their work has set the standard for estimations in empirical happiness research and our subsequent empirical analysis.

plies that the crucial explanatory variables will be correlated with the error term in our regression model. Individual fixed effects and relevant control variables could help to mitigate this problem, but fixed effects can only cure time-constant endogeneity and observable controls could themselves be endogenous. Crucial coefficients may thus be expected to suffer from simultaneity bias, which could potentially affect the estimated equivalence scales as well. In what follows, we will discuss possible reasons and consequences of endogeneity and present our hypotheses regarding the direction of bias in the respective coefficients and the resulting equivalence scale. After that, we will introduce our approach to empirically estimate the extent of this bias by the help of internal instruments.

2.1 Endogeneity Bias

2.1.1 Endogeneity of household size changes

Suppose that income satisfaction is an appropriate measure of individual household members' consumption possibilities and that there are no perfect economies of scale. Under these conditions, an increase in the number of household members at given income causes a decrease in satisfaction. Hence, the coefficient on log household size ($-\beta_1 a$) in Eq. (2) can be expected to be negative. Now suppose that income satisfaction not only depends on household size but that it also predicts it. This may be the case (1) when there are true causal effects of income satisfaction on household size changes, (2) when income satisfaction adjustments precede household size changes due to anticipation effects and/or (3) when there is a spurious relationship via individual characteristics affecting both, verbal responses and household size changes. The latter possibility is unproblematic as long as time-invariant characteristics are concerned and the estimation controls for individual fixed effects. If, for instance, more optimistic individuals generally report higher levels of satisfaction, ceteris paribus, and are also more likely to form a couple or decide to have children, this will not pose a problem in fixed effects estimations of the effect of household size on income satisfaction. But if certain traits become apparent in changes rather than in levels of response and household formation behavior, Eq. (2) will not be able to account for the spurious relation from household size to income satisfaction. Impulsive individuals or enthusiasts, for example, may react to changes in their living conditions more strongly in terms of both, expressed satisfaction levels and family formation behavior. Household size will thus partly be endogenous. If unobserved personality traits drive income satisfaction and family size changes in the same direction, higher income satisfaction will predict household size increases. We cannot know with certainty, however, that there are no other unobserved stressors or individual characteristics that neutralize or reverse this relation.

More interesting in terms of theoretical predictions is the case where income satisfaction has a direct, causal effect on household formation or where anticipation effects lead income satisfaction to change well before household size adjustments are realized. We will now consider the possible changes in household size that are relevant to our estimation sample and see how these may be predicted/preceeded by income satisfaction changes.

Consider first two partners moving in together. While it is conceivable that especially financially dissatisfied individuals would choose to cohabit to benefit from economies of scale, it is much more likely that the opposite is true (Stutzer and Frey 2006). It might be easier for financially satisfied individuals to date and engage in a serious relationship. They may perceive the costs of moving or getting married to be less pressing. Furthermore, an anticipation effect is very likely. The prospect of enjoying economies of scale when living together may affect income satisfaction before the new household is actually formed. All these effects point towards higher income satisfaction predicting growth in household size by one adult. The negative effect of an additional adult on income satisfaction is thus very likely to be underestimated.

Similarly, the separation of two partners appears more likely when individuals are less satisfied with their income. The anticipation of having scale economies vanish in the event of separation or union dissolution may further add to a decline in income satisfaction preceding the actual change household size. Thus again, the link from satisfaction to household size seems to be positive, with lower income satisfaction preceding a decline in household size by one adult member.

This also applies to cases of one partner's death. Had this event been preceded by severe illness, it is likely that increased medical and care expenses had affected the income satisfaction of all adults in the household negatively before. Again, we would find that shrinking households are comparatively less satisfied, such that the positive effect of having fewer household members at given incomes would be underestimated.

The second type of household size changes crucial to our estimations involves the number of children. An increase in this number may be caused by the birth of a child or by older children moving in. A decrease will be observed as a child moves out from the parental home (primarily when parents separate) or as it turns 14 years old and thus is considered an adult.

We suspect a causal relationship between lagged income satisfaction and the birth of a child. Given that parents wish to provide a financially stable and secure environment to their offspring, being satisfied with own economic conditions may be one prerequisite for planned pregnancies (research on life satisfaction an fertility suggests a positive effect Le Moglie et al. (2015), Mencarini et al. (2018), Parr (2010), Cetre et al. (2016)). With the considerable delay of births after fertility decisions, it is a priori unclear if current period satisfaction upon birth remains to be higher. Autoregressive processes may induce satisfaction to be persistently higher, whereas regression to the individual-specific mean or the anticipation of the child's cost may imply significant declines in income satisfaction from the previous to the current period. The predictive content of current-period income satisfaction for the birth of a child thus remains unclear.

While there is a biologically dictated considerable time lag between changes in income satisfaction and the birth of a child, changes in satisfaction may be much more contemporaneously linked to children moving in and out. Similar to a partner joining the household, having children move in may be especially likely in times of great income satisfaction, especially so when accompanying an adult. Similarly, children will be more likely to leave the household when income satisfaction is low, because most children leave the household (instead of passing the age threshold to count as an adult) when parents separate. We can therefore reasonably expect greater current period satisfaction to predict increases in number of children through this channel.

When the number of children declines mechanically only because of them aging, the corresponding change in the number of children at a given household size clearly is not caused by changes in income satisfaction.

Overall, we therefore expect a small, but positive effect of income satisfaction on the observed number of children.

Taking our expectations regarding changes in the number of adults and children together, we thus derive the following hypothesis:

Hypothesis 1 *Higher income satisfaction predicts increases in the number of household members. If this link is not accounted for, the negative effect of household size on income satisfaction (coefficient* $\beta_1 a$ *in Eq.* (2)*) will be underestimated.*

The number of children enters our regression also via its interaction with log household size. The associated coefficient will also be biased if the extent of endogeneity in household size and the number of children differs.

To see this, suppose for a moment that the number of adults is exogenous but that higher income satisfaction leads to an increase in the number of children. If a child enters the household through birth, the increase in household size and the number of children due to higher income satisfaction will thus be identical. In this case, all bias will be captured by the coefficient on household size. Increases in satisfaction due to a higher number of children at given household size, as indicated by the coefficient of the interaction term, will be correctly estimated. If, however, income satisfaction had a positive effect on the number of individuals in the household but no distinct effect on the number of children in the household, the extent of endogeneity in household size and the number of children would differ. The negative impact of additional household members would be underestimated and part of that bias would be captured by the interaction effect, thus indicating smaller than actual satisfaction increases to be obtained from additional children compared to additional adults (or at given household size).

In reality, we expect both, the number of adults and the number of children, to be endogenous. Number of children is by definition less endogenous, because children "become" adults when they pass the age threshold; this alters the number of children exogenously. There is no reason to believe that given income satisfaction changes will increase the number of children more than it will increase the total number of household members as this would imply a replacement of adults by children. Hence, we expect that endogeneity in the number of children is smaller than endogeneity in household size. This leads us to the following hypothesis to be tested empirically.

Hypothesis 2 *Higher income satisfaction predicts increases in the number of children. If this link is not accounted for, the positive interaction effect of the number of children with household size on income satisfaction (coefficient* $\beta_1 b$ *in Eq.* (2)*) will be underestimated.*

2.1.2 Endogeneity of income

Just as we can expect family size and structure to be predicted by income satisfaction, we must account for the possibility that reported household income may be endogenous.

Most recent studies that analyze endogeneity of income in income satisfaction resume that reporting behavior or recall ability (Prati 2017, Elsas 2021) are causing this endogeneity. Comparing survey to register data, Prati (2017) shows that people who are satisfied with their wage over-report, whereas those relatively less satisfied tend to under-report their wage. Elsas (2021) finds a similar effect for annual household income and life satisfaction. Her analysis is based on an income measure that refers to the year before the interview and is generated from detailed data on all household members' income over the whole previous year. For our estimation of equivalence scales, another income measure is preferable, because it refers to the same point in time when the number of individuals in the household is surveyed. This so-called income screener measures monthly net household income in one single item that is reported only by the household head. We therefore expect measurement error or recall bias to be even more pronounced in this income measure. Assuming the existence of classical measurement error in monthly net household incomes, Borah and Knabe (2018) propose a correction using an alternative, constructed income measure in their estimation of equivalence scales using income satisfaction data. Their results suggest that the bias in estimated equivalence scales introduced by measurement error is small but their finding depends crucially on the measurement error being non-systematic. This assumption is challenged by the results by Prati (2017) and hence a re-examination of the effect of endogeneity in income, in particular due to measurement error, in the estimation of equivalence scales from income satisfaction data is required.

If incomes were indeed over-/under-reported in times of greater/smaller income satisfaction, we would expect the following hypothesis to hold.

Hypothesis 3 *Higher income satisfaction predicts increases in measured household incomes. If this link is not accounted for, the coefficient* β_1 *in Eq.* (2) *will be overestimated.*

Endogeneity of household income may not only be the outcome of systematic measurement error, it could also stem from an actual causal effect, whereby individuals that are more satisfied with their income strive less for further income. Less satisfied individuals, on the other hand, may be more motivated to strive for growing income, thus making them more productive. If this was the source of endogeneity, the positive effect of income on satisfaction would be biased downwards. Another source of endogeneity of income in the satisfaction regression could be confounding factors, such as the experienced disutility of labor. People who perceive strong disutility of labor will c.p. work less and earn less and will on the other hand enjoy their income less, because it is earned at a higher price. Disutility of labor would hence impact negatively on income and income satisfaction and thus also bias estimates of the income effect downwards.

In case of such a negative causal link or spurious relation between income satisfaction and productivity or effort (over-weighting the hypothesized measurement error), we will expect the following alternative hypothesis to be confirmed.

Hypothesis 4 *Lower income satisfaction predicts increases in measured household incomes. If this link is not accounted for, the coefficient* β_1 *in Eq.* (2) *will be underestimated.*

2.1.3 Bias in equivalence scale parameters

As mentioned above, the equivalence scale parameters will be identified by dividing the coefficients on family structure by the coefficient household income. Given that Hypothesis 1 is true, the negative effect of household size on income satisfaction (coefficient $\beta_1 a$ in Eq. (2)) will be underestimated. If the coefficient on income was without bias, this would mean that we would also underestimate the equivalence scale parameter a. As outlined in the previous subsection, this may be naive, however (especially when taking into account that household income and family size and structure are naturally correlated). Depending on the direction and size of this bias, the equivalence scale elasticity in adult equivalents will be over- or underestimated. If the positive effect of income on satisfaction was overestimated (Hypothesis 3 was true), parameter a would clearly be biased downwards. The consequences of an underestimation of the positive effect of household income (as formulated by Hypothesis 4) would depend on the relative size of this bias. Only if it was small compared to the bias in the coefficient on household size, parameter a would be underestimated. Else, the bias could also run in the opposite direction.

Hypothesis 2 suggests that the positive interaction effect of the number of children with household size (coefficient $\beta_1 b$ in Eq. (2)) will be underestimated. Again, suppose there was no bias in the income coefficient. In this case, the true linear decline of the equivalence scale elasticity in the number of children was larger, suggesting lower material needs of children than conventionally estimated. This result would also hold when the income effect was overestimated (Hypothesis 3 was true). With a negative bias in the income coefficient, the bias in parameter b would again depend on its relative size.

Given the absence of a clear prior regarding the bias in the income coefficient, we hence cannot predict the direction of bias in the equivalence scale parameters. This calls for an empirical assessment, in which we employ internal instruments to solve the problem of endogeneity. The construction of these instruments is outlined in the following subsection.

2.2 Internal instruments approach

A first cure for endogeneity problems in satisfaction regressions is the application of panel fixed effects, which we also apply. Yet, if endogeneity is not time-constant at the individual level but time-varying, further steps will be needed. A standard approach would be to find suitable external instruments for household size, the number of children and income. A obvious instrument for income in satisfaction equations is windfall income, such as lottery wins or inheritances (Meer et al. 2003), yet the lottery wins are rare events and inheritances typically occur in the later family life cycle, when children already moved out of the household. Another typical instrument is industry- and occupation-wide variation in earnings (Luttmer 2005, Vendrik 2013, Kaiser 2018). This instrument, however, is not very suitable for our purpose, as it is valid especially for permanent income rather than for shorter period incomes (Vendrik 2013). The number of children has been instrumented using twinning (Rosenzweig and Wolpin 1980, Rosenzweig and Schulz 1987) and the sex of the first child (Lee 2008). Finding external instruments for both income and household size for the same observations is very demanding, if possible at all.

If no external instruments are available, an alternative strategy can be employed, where instruments are constructed from a readily available subset of the observed exogenous variables in X. In the context of life satisfaction measures, for instance, this approach has been applied by Le Moglie et al. (2015) and Otrachshenko et al. (2023). The approach was developed by Lewbel (2012) and identifies structural parameters by exploiting heteroscedasticity in the exogenous variables in X. More precisely, given that exogenous regressors are correlated with the variance of residuals in the first- and second-stage equation (errors are heteroscedastic), identification is achieved by having (a subset of) the exogenous regressors uncorrelated with the product of these heteroscedastic errors (i.e., $\mathrm{Cov}(Z, \varepsilon_1 \varepsilon_2) = 0$). Because the parameters of interest are identified via a second-moment condition, efficiency will be lower as compared to an external instrument approach.

The intuition of this identification strategy follows from linear regression mechanics: If the model in Eq.(2) is correctly specified, residuals are by construction exogenous to the the right-hand-side variables. If satisfaction is independent of the residuals of an auxiliary regression of the endogenous regressors on only exogenous controls (see Eq. (3) in the Appendix), and if the variables that were used for instrument construction are exogenous in the structural equation, then the instruments are exogenous. In that case, the instruments affect the outcome only via the endogenous regressor. To be relevant, the instrument must be related to the endogenous regressor. If residuals are heteroscedastic, they capture the variation of the outcome variable, which makes the instruments relevant.

The strength of these instruments depends on the degree of heteroscedasticity of the errors in the pre stage. When heteroscedasticity exists in the linear projection of each, household size in log values, the interaction effect of number of children and log. household size and household income on the exogenous variables Z, and when this heteroscedasticity is related to the exogenous variables, the instruments' are correlated with the three endogenous variables. The instruments' relevance can therefore be tested by checking the sample covariance between variables contained in Z and the squared residuals from linear regressions of the endogenous on the exogenous variable. In Section A3 in the Appendix, we will apply a Breusch-Pagan-Tests to verify the presence of heteroscedasticity.

3 Data

3.1 Sample

Data for this analysis come from the German SOEP release v38.1. The sample covers data from the years 1985 to 2021 (doi:10.5684/soep.core.v38.1eu). Subsamples that oversample high income households and refugees have been excluded.

We do not use each respondent's first year in the panel because Frick et al. (2006) demonstrate that data quality increases with the second year in panel, and that this is especially true for income data. Our sample is further restricted to respondents in private households, and further to households with no other adults than household head, partner and adult children[2]. For the estimations, however, we use only data from the household head and his/her partner. Only for analyzing the income distribution computed from the newly estimated equivalence income, we expand the sample sample according to household size (in Section 5). Households with more than seven children are also dropped, likewise families with members who need special care, because this usually causes extraordinary financial need. Finally, we excluded household in the outer percentiles of each year's per capita income distribution. This leads to a sample of 364,038 observations of 47,395 individuals in 32,449 households.

3.2 Variables

As income measure we use the current net monthly household income as stated by the household head, deflated by CPI to warrant the inter-temporal comparability of income.

Household composition at the time of the interview is captured with two variables: Number of individuals in the household and number of children, younger than 14. We chose this age threshold because the OECD scale applies the same. This will allow comparison of our results with the most widely used equivalence scale.

For construction of the instruments we use interviewer's working experience and respondent's age and age squared.

Region fixed-effects are defined according to the broader categories East and West Germany.

Other individual characteristics contained in vector X in Eq. 2 are survey wave squared and cubed, the number of hospital overnight stays in the previous year, the marital status (single, married, separated or widowed) and home ownership.

[2] Households with children older than 30 years are excluded from the analysis.

The set of control variables was chosen according to the Hansen J Test and the Kleibergen-Paap F statistic in the instrumtental variable estimation.

4 Results

Estimations are run in a fixed effects set up, i.e. with demeaned data. Estimations are run with the Stata-ado ivreg2h, written by Baum and Schaffer (2012) calling the GMM estimator and using clustered standard errors, since we use several observations per person.

The Kleibergen-Paap rank F statistic is used to assess if instruments are sufficiently strong. We followed the rule-of-thumb that it should exceed 50 (Keane and Neal 2022). Since j instruments are used for each endogenous regressor it is possible to apply the Hansen J test to indicate if instruments (including also all other variables of the estimation) are endogenous. It tests the joint null that that all instruments (the constructed instruments and the exogenous covariates) are uncorrelated with the errorterm of the second stage regression. Hansens J-Test should be insignificant; if it's significant some of the regressors (the instrumented or other) are still endogenous. For each instrumental variable estimation these statistics are listed in the bottom of the table.

4.1 Baseline results

Table 1 presents our baseline results for the standard model presented in Eq. (2).

Column 1 reports results for a benchmark model that takes into account individual fixed effects and some typical controls in income satisfaction regressions but does not use instruments account for endogeneity.

In Section 2.1 we argued why changes in household size and composition could be endogenous in income satisfaction. To account for this endogeneity we estimated the model using internal instruments for number of household members and number of children (transformed according to the model equation so that the scale parameters are easily obtained from the regression coefficients). We further elaborated on endogeneity of the income measure and therefore also instrumented income in our estimation. Results from this estimation are presented in column 2 in Table 1.

These results support our hypotheses 1 and 2 that ignoring endogeneity of household size and composition yields underestimated coefficients of the respective effects in income satisfaction regressions. These findings hence suggest that satisfaction has some positive effect on number of household members, and that ignoring endogeneity in satisfaction regressions will lead to biased estimates. With respect to potential effects of endogeniety in the income measure, we developed

Table 1
FE-OLS and FE-IV regressions of income satisfaction

	Whole sample				Reduced sample (see section 4.2)			
	(1)		(2)		(3)		(4)	
Log(y)	1.912	***	3.168	***	1.922	***	2.847	***
	(0.012)		(0.147)		(0.013)		(0.132)	
Log(h)	–0.868	***	–1.433	***	–0.948	***	–1.574	***
	(0.016)		(0.130)		(0.018)		(0.148)	
k*log(h)	0.085	***	0.193	***	0.111	***	0.254	***
	(0.004)		(0.023)		(0.005)		(0.033)	
Nights in hospital	–0.002	***	–0.001	***	–0.002	***	–0.001	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Divorced/separated	–0.080	***	0.003		–0.124	***	–0.107	**
	(0.027)		(0.047)		(0.028)		(0.049)	
Married	0.093	***	–0.062		0.091	***	0.003	
	(0.021)		(0.054)		(0.022)		(0.054)	
Widowed	0.192	***	0.070		0.131	***	–0.066	
	(0.032)		(0.066)		(0.033)		(0.074)	
Home owner	–0.035	***	–0.127	***	–0.051	***	–0.113	***
	(0.012)		(0.022)		(0.013)		(0.022)	
Interviewer birthyear	0.003	***	0.003	***	0.003	***	0.003	***
	(0.000)		(0.001)		(0.000)		(0.001)	
Age	–0.082	***	–0.160	***	–0.078	***	–0.136	***
	(0.004)		(0.012)		(0.005)		(0.011)	
Age, squared	0.000	***	0.001	***	0.000	***	0.001	***
	(0.000)		(0.000)		(0.000)		(0.000)	
East Germany	–0.220	***	–0.038		–0.205	***	–0.065	
	(0.044)		(0.086)		(0.047)		(0.087)	
Survey wave, squared	–0.000	**	0.002	***	–0.000	**	0.001	***
	(0.000)		(0.000)		(0.000)		(0.000)	
Survey wave, cubed	0.000	***	–0.000		0.000	***	0.000	
	(0.000)		(0.000)		(0.000)		(0.000)	
Constant	–11.001	***			–11.411	***		
	(0.814)				(0.885)			
K-P F-Statistic			61.2				68.0	
Hansen J			4.126				4.829	
Prob>Chi squared			0.660				0.566	
Observations	364,038		364,038		313,127		313,127	
Individuals	47,395		47,395		44,970		44,970	

Source: SOEP v38.1, own calculations.

Notes: Significance levels * 0.10 ** 0.05 *** 0.01. Cluster-robust standard errors in parentheses.

Table 2
Scale parameters from estimates in Table 1 with bootstrapped confidence intervals

	OLS FE			IV FE		
	Observed	Confidence interval		Observed	Confidence interval	
	Whole sample estimation (364,038 observations of 47,395 individuals.)					
Scale parameter a	0.454	0.431	0.477	0.452	0.375	0.529
Scale parameter b	0.045	0.039	0.050	0.061	0.046	0.075
	Reduced sample estimation (313,127 observations of 44,970 individuals.)					
Scale parameter a	0.493	0.468	0.519	0.553	0.457	0.649
Scale parameter b	0.058	0.050	0.066	0.089	0.066	0.113

Source: SOEP v38.1, own calculations.

Notes: 500 replications clustered by person. Confidence intervals are normal-based.

two opposing hypotheses and our estimation results clearly support hypothesis 4. This is in contradiction to findings from Prati (2017) and Elsas (2021) who found the reverse effect. In contrast to Prati's (2017) study, our study design focuses on another question and thus uses another setup: Prati (2017) analyzes only data from employed individuals and their earnings, while we refer to household income and also include individuals who do not earn own income but answer the income satisfaction question. Compared to Elsas (2021) the satisfaction measure that we employ here is focused on income while Elsas' (2021) analysis refers to life satisfaction that encompasses more aspects of life and is typically only weakly associated with income. The finding that endogneity leads to an underestimation of the true income effect in income satisfaction is, though, inline with (Vendrik 2013, Kaiser 2018, Luttmer 2005).

Further, and this might be the most important point, when individuals evaluate their household income they refer to both income and needs, and the household's financial needs depend on the number of household members and the number of children. Meaning that when instrumentation changes the estimated effects of household composition on income satisfaction, the estimate for the income effect adapts to this change.

We therefore suggest to focus on the estimated scale parameters a and b in Table 2. Compared to the baseline FE-OLS based scale parameters the FE-IV based parameters in the upper half of Table 2 show a larger decline in financial needs for additional children compared to additional adults, i.e. scale parameter b from the IV estimation is slightly larger than from the baseline estimation. Confidence intervals from bootstrapped standard errors confirm that the IV-based scale parameter b is higher than the OLS-based one.

Financial needs of adults appear to be less affected by endogeneity, the scale parameter a is nearly identical in the IV and the OLS estimation.

4.2 More endogeneity - more bias

One important change in household composition, though, occurs due to the aging of children, which is obviously not endogenous, but just as obviously changes the household composition in the sense of our analysis. When a child passes the age threshold of 13 years, the child-discount is no longer applied, i.e. estimated. This implies that the 14 year old child counts as an adult in the household, and more important for our analysis, that the number of children changes. The probabilities of changes in household composition, given in the Appendix in Table A4.3, show that from one year to the next 6% of all parent couples with children younger than 14 become parent couple without children younger than 14. Similar for single parents, from one to the next year 12% of those with children younger than 14 become single parents without children younger than 14. These transitions through aging are the most frequent transitions in the sample.

Since these transitions are clearly not endogenous and therefore mitigate endogeneity in the number of children, we use these considerations as a starting point to challenging our Hypothesis 2: If endogeneity in the number of children was a problem in the income satisfaction regression, it should be more pronounced in an estimation where households are excluded when the number of children decreases only due to their aging. Therefore we rerun the estimations on a sample without households with children between 14 and 17 years.

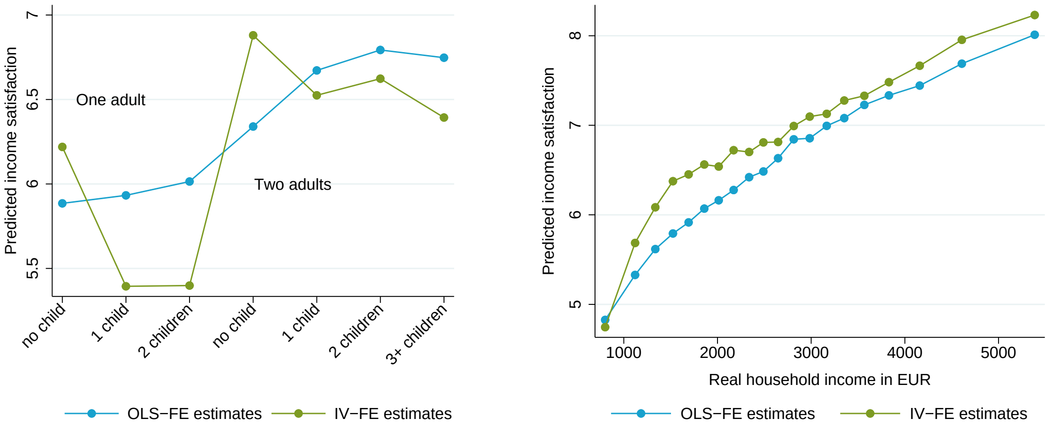
Estimation results in column 3 and 4 in Table 1 and in the lower half of Table 2 confirm this expectations. The financial need for additional adults is higher when households with children between 14 and 17 are excluded from the estimation. This is evident already in the OLS-based scale parameter $a = .493$ and more pronounced in the IV-based parameter $a = .553$. It was expected that the scale parameter b would be higher in the reduced sample than in the whole sample. This expectation is met, $b_{IV} = .089$ while $b_{OLS} = .058$.

5 Implications

The results we presented here have implications for research on income satisfaction and for equivalence scale estimations. Implications for research on satisfaction refer to the fact that endogeneity of income and household composition affected the fixed effects estimation results.

As shown in the left panel in Figure 1, living with children comes with lower income satisfaction than one would expect if neglecting endogeneity in income and household size and composition. Likewise, the income effect in income satisfaction is underestimated if endogeneity in income and household size and composition is ignored, see right panel in Figure 1. Future

Figure 1
Binned scatterplots of predicted income satisfaction against household composition and income

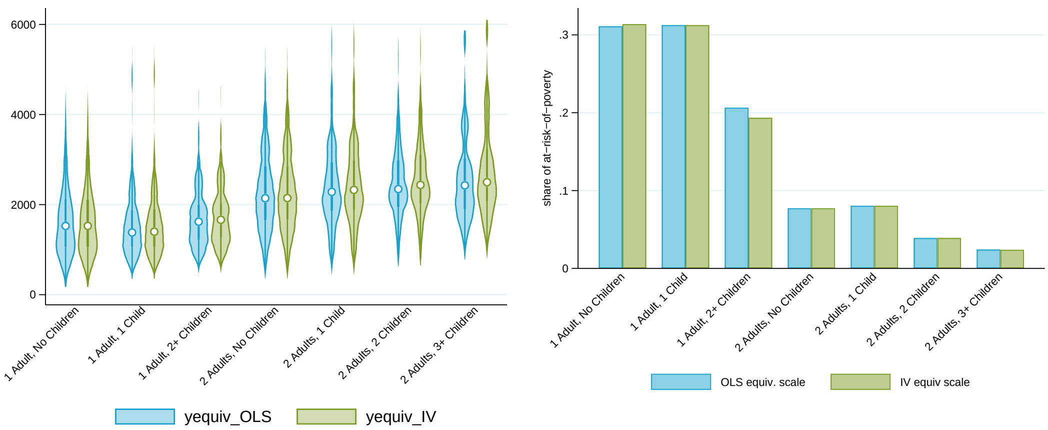


Source: SOEP v38.1, own calculations. 364,038 person year observations of 47,395 individuals.
Notes: Means of predicted values of income satisfaction from baseline and IV estimations in Table 1.

research on income satisfaction should take that into account and carefully check whether other determinants of satisfaction are themselves determined by satisfaction.

Implications on the income distribution are very moderate: Households with children are slightly better off which also affects the share of individuals at risk of poverty, see Figure 2.

Figure 2
Mean equivalized income by household type and share of individuals at risk of poverty in 2021



Source: SOEP v38.1, own calculations. 567,283 observations.

6 Conclusion

In this study, we have estimated equivalence scale parameters under consideration of potentially endogenous income, household size and structure. We have applied the income satisfaction approach, which estimates the effect of income, the number of household members and the number of children on income satisfaction and computes the equivalence weights for additional household members in a household and the linear decline of the equivalence scale elasticity in the number of children at a given household size.

Using a regression approach that has been frequently applied in the previous literature, we have found that endogeneity causes downward biases in the estimated effect of income, the number of household members and the number of children in the household. Our results thus support earlier studies that have found endogeneity of household composition in satisfaction data. With respect to endogeneity of income, our results do not confirm the hypothesis that systematic measurement or reporting causes an overestimation of the true income effect. We found, inline with Luttmer (2005), Vendrik (2013), Kaiser (2018), the opposite, i.e. that the income effect is higher when endogeneity of income is taken into account.

We have also provided new insights on the impact of income satisfaction on household composition and income data.

Moreover, we have contributed to research on the empirical estimation of equivalence scales from satisfaction data in general, and the implications of endogeneity, more specifically.

Scale parameters rely on the income coefficient as well as the estimated effect of the number of household members (and children). It appears that equivalence weights for additional adults do not change very much when using the novel estimation method, while the downward adjustment in equivalence weights for children does. Endogeneity in household composition is predominantly relevant for increases in the number of household members, especially children.

This result could serve as a starting point to analyze how satisfaction affects adult children's decision to move out of their parental homes, and of course partners or other adults moving in and out.

References

Baum, C. and Lewbel, A. (2019). Advice on using heteroskedasticity-based identification. *Stata Journal*, 19(4):757–767.

Baum, C. F. and Schaffer, M. E. (2012). IVREG2H: Stata module to perform instrumental variables estimation using heteroskedasticity-based instruments. Statistical Software Components, Boston College Department of Economics.

Biewen, M. and Juhasz, A. (2017). Direct estimation of equivalence scales and more evidence on independence of base. *Oxford Bulletin of Economics and Statistics*, 79(5):875–905.

Borah, M., Knabe, A., and Keldenich, C. (2019). Reference income effects in the determination of equivalence scales using income satisfaction data. *Review of Income and Wealth*, 65(4):736–770.

Buetikofer, A. and 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.

Cetre, S., Clark, A. E., and Senik, C. (2016). Happy people have children: Choice and self-selection into parenthood. *European Journal of Population*, 32(3):445–473.

Coulter, F. A. E., Cowell, F. A., and Jenkins, S. P. (1992). Differences in needs and assessment of income distributions. *Bulletin of Economic Research*, 44:77–124.

De Neve, J.-E. and Oswald, A. J. (2012). Estimating the influence of life satisfaction and positive affect on later income using sibling fixed effects. *Proceedings of the National Academy of Sciences of the United States of America*, 109(49):19953–19958.

Elsas, S. (2021). Causality in the link between income and satisfaction: IV estimation with internal instruments. *SOEP Papers*, 1143:1–39.

Ferrer-i-Carbonell, A. and Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *The Economic Journal*, 114:641–659.

Frick, J. R., Goebel, J., Schechtman, E., Wagner, G. G., and Yitzhaki, S. (2006). Using analysis of Gini (AnoGi) for detecting whether two sub-samples represent the same universe: The German Socio-Economic Panel Study (SOEP) experience. *Sociological Methods & Research*, 34(4):427–468.

Graham, C., Eggers, A., and Sukhtankar, S. (2004). Does happiness pay? In *Challenges for Quality of Life in the Contemporary World*, pages 179–204. Springer.

Kaiser, C. (2018). People do not adapt to income changes: A re-evaluation of the dynamic effects of (reference) income on life satisfaction with gsoep and ukhls data. *MPRA Papers*, 89867.

Keane, M. P. and Neal, T. (2022). A practical guide to weak instruments. *UNSW Economics Working Paper*, 2021(05d).

Le Moglie, M., Mencarini, L., and Rapallini, C. (2015). Is it just a matter of personality? on the role of subjective well-being in childbearing behavior. *Journal of Economic Behavior & Organization*, 117:453–475.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1):67–80.

Luttmer, E. (2005). Neighbours as negatives: relative earnings and well-being. *Quarterly Journal of Economics*, 120(3):963–1002.

Meer, J., Miller, D. L., and Rosen, H. S. (2003). Exploring the health–wealth nexus. *Journal of Health Economics*, 22(5):713–730.

Mencarini, L., Vignoli, D., Zeydanli, T., and Kim, J. (2018). Life satisfaction favors reproduction. the universal positive effect of life satisfaction on childbearing in contemporary low fertility countries. *PLOS one*, 13(12):e0206202.

Mishra, V. and Smyth, R. (2014). It pays to be happy (if you are a man): Subjective wellbeing and the gender wage gap in urban China. *International Journal of Manpower*, 35(3):392–414.

Otrachshenko, V., Nikolova, M., and Popova, O. (2023). Double-edged sword: persistent effects of communist regime affiliations on well-being and preferences. *Journal of Population Economics*, 36:1139–1185.

Parr, N. (2010). Satisfaction with life as an antecedent of fertility: Partner + happiness = children? *Demographic Research*, 22:635–662.

Prati, A. (2017). Hedonic recall bias. why you should not ask people how much they earn. *Journal of Economic Behavior & Organization*, 143:78–97.

Rapp, S. (2021). Asset bias in household needs measurement. *INEQ Working Paper Series*, (22).

Rojas, M. (2007). A subjective well-being equivalence scale for mexico: Estimation and poverty and income-distribution implications. *Oxford Development Studies*, 35:273–293.

Schwarze, J. (2003). Using panel data on income satisfaction to estimate equivalence scale elasticity. *Review of Income and Wealth*, 49(3):359–372.

Stutzer, A. and Frey, B. S. (2006). Does marriage make people happy, or do happy people get married? *Journal of Socio-Economics*, 35(2):326–347.

van Praag, B. M. and Ferrer-i-Carbonell, A. (2004). *Happiness Quantified: A Satisfaction Calculus Approach*. Oxford University Press, Oxford.

Vendrik, M. (2013). Adaptation, anticipation and social interaction in happiness: An integrated error-correction approach. *Journal of Public Economics*, 105:131–149.

Appendix

A1 Sample descriptives

Table A1.1
Descriptive Statistics

	Mean	Standard deviation between	within	Min	Max
Satisfaction with household income	6.525	1.893	1.399	0	10
Real monthl. household income	2736.65	1382.07	674.70	253.31	24509.8
No. of persons	2.697	1.272	0.553	1	9
No. of children	0.518	0.857	0.464	0	7
Nights in hospital	1.600	5.639	6.570	0	365
Interviewer birthyear	1945.257	10.999	5.935	1907	2004
Respondent's age	49.969	15.818	5.163	18	100
Survey wave	22.985	9.815	5.163	2	38
	Percent overall	between	within	Min	Max
Single	13.74	22.77	83.32	0	1
Separated	10.66	14.67	74.78	0	1
Married	69.00	70.30	91.05	0	1
Widowed	6.61	7.69	78.65	0	1
Home owner	47.19	49.92	86.91	0	1
Living in East Germany	22.56	22.12	97.32	0	1

Source: SOEP v38.1, own calculations. 364,038 person-year observations of 47,395 individuals in 32,449 households.

Note: Children are here only children younger than 14 living in the respondent's household.

A2 Heteroscedasticity-based instrumentation

For our application, it can be summarized as follows.

To construct instruments for h_{it} in Eq. (2), the following auxiliary regression is run in a first step:[3]

$$h_{it} = Z'_{it}\delta + \nu_{it}, \tag{3}$$

where Z is a subset of J variables in X that satisfy the exogeneity assumption, i.e. $E(Z_{it}\nu_{it}) = 0$. Here, we regress household size on *age*, age^2, and interviewer experience, which we deem to be sufficiently exogenous to be considered in Z. We then predict the residuals $\hat{\nu}_{it}$ from this

[3] Actually we do not conduct the estimation step by step, but use the Stata-ado 'ivreg2h' by Baum and Schaffer (2012).

regression. The residuals $\hat{\nu}_{1it}$ and sample-centered values of each variable z_j in Z can then be used to calculate J instruments for h_{it} in the second step:

$$hinst_{jit} = (z_{jit} - \bar{z}_j)\hat{\nu}_{it} \tag{4}$$

We thus construct one instrument per exogenous variable contained in Z representing the product of the observation-specific residual and the demeaned respective variable. Instruments for our other two potentially endogenous regressors k_{it} and y_{it} are generated analogously.

These instruments are then - as in conventional instrumental variable estimations - used on the first stage to predict the endogenous regressors in the structural equation Eq. (2). In short, the estimator thus relies on a linear two-stage least squares regression of income satisfaction on exogenous controls X and on h_{it}, k_{it} and y_{it}, using X and estimates of $(Z - E(Z))\varepsilon_2$ as instruments.

The intuition of this identification strategy follows from linear regression mechanics: If the model in Eq.(2) is correctly specified, residuals are by construction exogenous to the the right-hand-side variables. If satisfaction is independent of the residuals of the auxiliary regression, Eq. (3), and if the variables in Z are exogenous in the structural equation, then the instruments are exogenous. In that case, the instruments affect the outcome only via the endogenous regressor. To be relevant, the instrument must be related to the endogenous regressor. If residuals are heteroscedastic, they capture the variation of the outcome variable, which makes the instruments relevant.

Instruments constructed by Eq. (4) are uncorrelated with the error in the income satisfaction equation (second stage) because of $\text{Cov}(Z, \varepsilon_1\varepsilon_2) = 0$), making them valid instruments. The strength of the instruments depends on the degree of heteroscedasticity of the error in the pre stage. The greater heteroscedasticity of the errors in a linear projection of household size, the number of children or household income on the exogenous variables contained in Z, the greater the instruments'correlation with these three endogenous variables. The instruments' relevance can therefore be tested by checking the sample covariance between variables contained in Z and the squared residuals from linear regressions of the endogenous on the exogenous variable. In Section A3, we will apply a Breusch-Pagan-Tests to verify the presence of heteroscedasticity.

As noted in Lewbel (2012) and Baum and Lewbel (2019), for each instrument $hinst_j$ to be valid, it is required that the errors of each auxiliary regression and the structural equation are uncorrelated, i.e., $\text{Cov}(z_j, \varepsilon_1\varepsilon_2) = 0$ for the instruments for h and $\text{Cov}(z_j, \varepsilon_1\varepsilon_3) = 0$ for instruments for y. This will be the case if both household composition and income are each endogenous in satisfaction because of a common factor, τ_{it}. This could be something like mindset or the ability to enjoy. In contrast to the individual fixed effect α_i this common factor τ_{it} can vary over time. Thus we reformulate the satisfaction equation Eq. (2)

$$s_{it} = \beta_1 \ln(y_{it}) - \beta_1 a \ln(h_{it}) + \beta_1 b k_{it} \ln(h_{it}) + X_{it}'\gamma_1 + \alpha_{1i} + \tau_{it} + \varepsilon_{1it} \tag{5}$$

and analogously a household size equation can be stated:

$$h_{it} = +\beta_2 \ln(y_{it}) + \delta_2 s_{it} + X_{it}'\gamma_2 + \alpha_{2i} + \tau it + \varepsilon_{2it} \tag{6}$$

and analogously for income and number of children in the household.

If endogeneity occurs due to such a common factor the models fulfill the assumption that the ideosyncratic errors of the satisfaction equation is independent of those of the endogenous regressor. [4]

Secondly, the estimator relies on heteroscedasticity in the endogenous regressors. Heteroscedasticity must be related to the exogenous controls, which are used for instrument construction.[5]

A3 Heteroscedasticity in the model's data

The relevance of an instrument constructed from a particular variable z_j depends on the strength of heteroscedasticity with respect to z_j, this section therefore inspects heteroscedasticity in the data with regressions estimates, statistical tests and residual plots that should be indicative of the patterns of heteroscedasticity.

Table A3.2 presents estimates of the association between exogenous controls and the squared residuals of the three endogenous regressors, the household income in log space, the log of the number of individuals in the household, and the number of children in the household multiplied with the log of the number of individuals in the household. The bottom lines present Breusch-Pagan test statistics, which assumes a normal distribution of the error terms.

The Breusch-Pagan test indicates that the data is heteroscedastic, all estimated coefficients are significantly different from zero, even though the absolute values are visually zero. This could be circumvented by rescaling the data, but this would not substantially change anything.

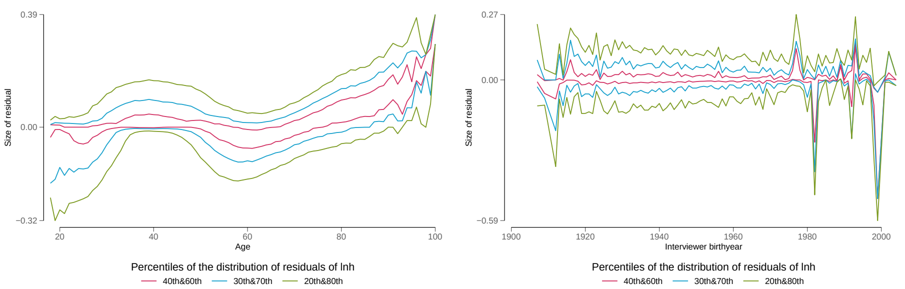
[4] Formally: $\text{Cov}(z_j, \varepsilon_1 \varepsilon_2) = 0$, and analogously for income and number of children in the household.
[5] Formally: $\text{Cov}(z_j, \varepsilon_{2it}^2) \neq 0$, and analogously for income and number of children in the household.

Table A3.2
OLS regression of squared residuals of endogenous regressors on exogenous controls

	Income		Ln(Persons)		Kids*Ln(Pers)	
Interviewer birthyear	0.000	***	0.000	***	0.004	***
	(0.000)		(0.000)		(0.000)	
Age	–0.004	***	–0.007	***	–0.099	***
	(0.000)		(0.000)		(0.002)	
Age, squared	0.000	***	0.000	***	0.001	***
	(0.000)		(0.000)		(0.000)	
Constant	0.053	***	0.041	***	0.489	***
	(0.000)		(0.000)		(0.003)	
Chi squared	1385		2953		3953	
Prob>Chi squared	0.000		0.000		0.000	

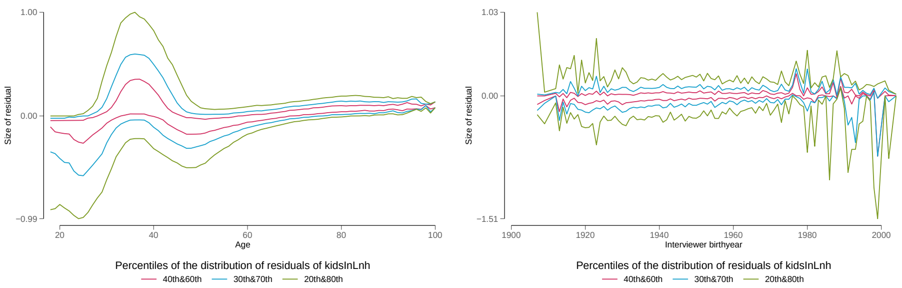
Source: SOEP v38.1, own calculations. 366,616 observations of 47,764 individuals
Notes: Significance levels * 0.10 ** 0.05 *** 0.01. Standard errors in parantheses. Dependent variables are squared residuals from fixed effects regressions, exogenous variables are within transformed.

Figure A3.1
Age- and interviewer birthyear related heteroscedasticity in number of household members (log.)



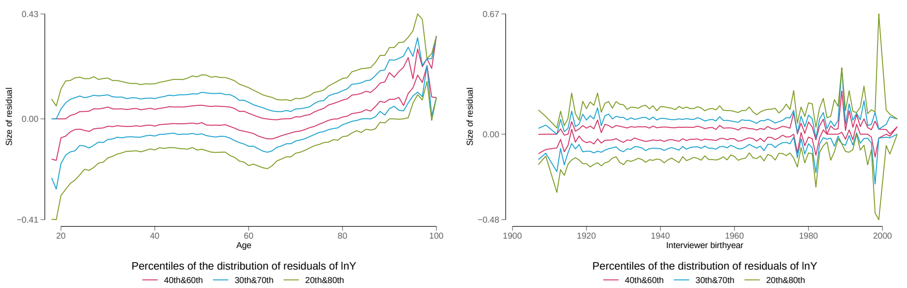
Source: SOEP v38.1, own calculations. 366,616 person year observations of 47,764 individuals.
Notes: Residuals from fixed effects regressions of log.household size on age, age^2, interviewer experience

Figure A3.2
Age- and interviewer birthyear related heteroscedasticity in number of children



Source: SOEP v38.1, own calculations. 366,616 person year observations of 47,764 individuals.
Notes: Residuals from fixed effects regressions of $k_{it} \ln(h_{it})$ on age, age^2, interviewer experience. Children are here children of age<14.

Figure A3.3
Age- and interviewer birthyear related heteroscedasticity in household income (log.)



Source: SOEP v38.1, own calculations. 366,616 person year observations of 47,764 individuals.
Notes: Residuals from fixed effects regressions of log.household income on age, age^2, interviewer experience

Figure A3.1, A3.2 and A3.3 are plots of the distribution of residuals that are used to construct the instruments for household size, number of young children and income.[6] The plots underline the message of the Breusch-Pagan test in Table A3.2. The distribution of the residuals of all three endogenous regressors narrows over the interviewers working experience. The distribution of the residuals of number of children also narrows over respondent's age, for household size and income the tendency is less clear. It thus appears that the exogenous controls and squared residuals are sufficiently associated to allow for identification based on Lewbel's instruments.

A4 Probabilities of changes in household composition

Table A4.3
Year-to-year transition probabilities for household types

Rows show initial and columns show final values.

	Household Typology						
Household Typology	1	2	3	4	5	6	Total
1 One-person HH	94.65	4.09	0.23	0.25	0.68	0.11	100.00
2 Couple w/o children	2.43	94.35	0.04	0.02	2.88	0.29	100.00
3 Single parent w/ children<14	0.70	0.29	78.83	10.97	8.43	0.79	100.00
4 Single parent w/o children<14	9.59	1.24	0.30	85.69	0.19	2.99	100.00
5 Parent couple w/ children<14	0.78	0.19	1.31	0.17	91.44	6.11	100.00
6 Parent couple w/o children<14	0.54	8.40	0.01	1.02	0.26	89.77	100.00
Total	25.24	30.58	3.42	4.00	23.76	13.00	100.00

Source: SOEP v38.1, own calculations. 230,874 household-year observations of 32,449 households.
Note: Children are not restricted to minors. Probabilities are not normalized for missing periods.

[6] Demeanded endogenous regressors were regressed on demeaned exogenous controls, i.e. age, age^2 and interviewer's experience.