

**DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS**



**ON THE INCOME DEPENDENCE OF EQUIVALENCE
SCALES**

**Christos Koulovatianos, Carsten Schröder and
Ulrich Schmidt**

Discussion Paper 2004-01

On the Income Dependence of Equivalence Scales

Christos Koulovatianos^{a,*}, Carsten Schröder^b and Ulrich Schmidt^c

First version: November 12, 2001

This version: February 6, 2004

Abstract

We suggest a simple survey method for obtaining direct subjective estimates of equivalence scales, also appropriate for testing whether equivalence scales depend on reference-household income. We implement our approach in two countries, Germany and France. In both countries independence of base is rejected. In particular, we find that equivalence scales depend negatively on reference income, an indication of increasing economies of scale in household consumption as living standards go up. Our estimation method is non-parametric, and it allows us to test generalized equivalence-scale exactness, which is not rejected in any of our samples.

Key Words: equivalence scales, survey method, independence of base

JEL Classification: C42, C90, D31, D63, I31

^a Dept. of Economics, Univ. of Cyprus

^b Institute of Public Economics, Univ. of Kiel

^c Department of Financial Economics, Univ. of Hannover

* Corresponding author, Dept. of Economics, Univ. of Cyprus, P.O. Box 20537, Nicosia CY 1678, Cyprus. E-mail: chrk@ucy.ac.cy, Tel: +357-22-892451, Fax:+357-22-892432. We are indebted to Francois Bourguignon and Olivier Bargain for kindly providing us with data on the French distribution from the 1999 household census in France. We also thank Krishna Pendakur and two anonymous referees for their valuable comments and suggestions. This paper is a revision of the two shorter mimeos by Schröder and Schmidt (1999) and (2000) and also of the working paper by Koulovatianos, Schröder and Schmidt (2001). Financial support from the TMR network "Living Standards, Inequality and Taxation," contract No ERBFMRXCT980248, is gratefully acknowledged. Koulovatianos thanks the Leventis foundation and the RTN project on "The Economics of Ageing in Europe" for financial support. Schmidt acknowledges financial support from the Deutsche Forschungsgemeinschaft, contract No Schm1396/1-1.

1. Introduction

How do household costs and welfare change as household income and demographic composition vary? Our study explores this question through a survey method: we ask people to tell us about the relationship among household income, family demographic composition and the well-being of a household.

We ask our subjects questions as: “which family-income level can make a household with one adult and two children achieve the same well-being as a household with a single adult only and a monthly family income of \$2,000, according to your opinion?” In this way we collect a sample of subjective “equivalent incomes”: incomes that make the well-being of households with different demographic composition equal.¹ Dividing equivalent incomes by the income of a household with a specific demographic composition (reference household) gives an equivalence scale.

In our questionnaire we give to our subjects a specific income level (reference income) for a single-childless-adult household (our reference household). We ask them to think of the well-being of the reference household at this reference income and to give us equivalent incomes for seven other family types, according to their own perception of utility and existing markets. We ask our respondents to repeat the same procedure for five different reference incomes for the imaginary single-childless-adult (reference) household. In this way we collect five sample equivalence scales corresponding to five different reference-income levels. The database we construct provides a range of subjective household welfare evaluations that enables us to test for a possible dependence of equivalence scales on reference incomes, the central issue of this paper.

A reason we adopt a survey method for estimating equivalence scales is the poorly un-

¹ By identifying subjective equivalent incomes for many household types, we obtain subjective “equivalent-income functions”: functions that give equivalent income for all household types, all household incomes and any commodity price vector. A significant body of literature attempts to estimate equivalent-income functions using consumer-expenditure data. See, for example, Donaldson and Pendakur (2003).

derstood determinants of a key unobservable factor: within-household economies of scale in consumption. Plausibly, as the number of family members increases in a household, the sharing of goods such as housing, furniture, household appliances or means of transportation, also increases. Testing for the dependence of equivalence scales on reference incomes is another way of asking: do within-household consumption economies of scale vary as we move from households with low family income and a poor living standard to households with higher income and living standard? Income dependence of consumption sharing in the household may be a key determinant of the distribution of household well-being across family types and household incomes.

If equivalence scales are negatively correlated with reference incomes and utility levels, then within-household economies of scale in consumption increase with rising household income. For example, according to the model of Barten (1964), the expenditure share on food and clothing may be higher for households with low income. As the number of family members increases, the expenditure share on food and clothing is likely to increase even more for households with lower family income. This may happen because economies of scale in food and clothing consumption are not likely to be important.²

As another example, considering housing, rich single adults may have larger houses that will not be congested much by adding one or two extra people. On the contrary, poor single adults, demanding smaller houses in size, may have to bear high disutility of congestion as household members are added.

An alternative reasoning for increasing within-household economies of scale with income is that families with low income may be credit constrained. Credit constraints may shift the

² Deaton and Paxson (1998) provide evidence that the food expenditure share decreases as the household size increases, keeping per capita household income constant. Thus, food may contain significant sharing possibilities. In their comment to Deaton and Paxson (1998), Gan and Vernon (2003) argue that, at least compared to housing, in a two-good framework (food and housing) food exhibits increasing expenditure shares with increasing family size, so food has comparatively lower sharing possibilities to housing. Independently from these empirical findings, the main point of our argument about why equivalence scales may decrease with income, is that there might be goods with comparatively low potential for sharing that take the biggest part of total household expenditures in low-income families.

chosen family consumption bundle towards lower expenditure shares on durables with high within-household sharing potential than otherwise preferred.

On the other hand, if richer single adults have a very high expenditure share on goods with little or no scale economies, like expensive vacation travels, keeping the same high living standard may require that additional household members also spend a lot on traveling. Thus, equivalence scales may remain constant, or even increase, as income and utility increase.

Do our examples about food, housing or shared durables dominate in the aggregated equilibrium consumption demands of a household, so that reference incomes and equivalence scales are negatively related? Or is it that equivalence scales are positively related with reference incomes, indicating that within-household economies of scale decrease with household income? Is a higher family income directed to larger expenditure shares for shared or non-shared goods in the household?

We provide evidence from two countries, Germany and France. We find the same qualitative results in both countries, namely, (i) independence of base is rejected, (ii) equivalence scales are significantly decreasing with reference income, and (iii) a linear relationship between the log of equivalence scales and the log of reference incomes, a property implied by generalized equivalence scale exactness (see Donaldson and Pendakur (2003)), is not rejected.

Our key message is that we find strong evidence for a negative dependence of equivalence scales on reference income. This finding has direct implications for the building of consumer-demand systems that aim at estimating equivalence scales from consumer-expenditure data. So far, a large body of literature estimating equivalence scales through parametric demand systems puts the following assumption in the structure of these models: that household expenditure functions across families with different demographic composition are proportional with respect to reference income, hence equivalence scales are a-priori independent from reference income, or “Independent of Base” (IB).³

³ Independence of Base is also named “Equivalence Scale Exactness” (ESE), see, for example, Blackorby and

In contrast to the usual IB hypothesis that provides convenience to econometric approaches, our study strongly encourages using parametric or semi-parametric demand systems producing equivalence scales that are decreasing in reference income. An example of an econometric approach relaxing IB and testing a non-IB framework is that of Donaldson and Pendakur (2003). They use a generalized parametric demand system and reach conclusions similar to ours.⁴

Our methodology is borrowed from other experimental literatures that target revealing behaviors blurred by the presence of several statistically unobserved factors. A classic example is literature studying “willingness to pay versus willingness to accept.”⁵ With respect to previous methodologies for measuring subjective scales as the one pioneered by Kapteyn and van Praag (1976), we differ in two aspects: (i) both the stimulus and response variable in our questionnaire is the same (income), as opposed to being two different variables (income and verbal characterizations of well-being); and (ii) we do not use any functional or parametric utility system in order to elicit equivalence-scales from our database.

Completing the questionnaire requires some effort from each respondent. For this reason, at this stage, our samples consist of about two hundred respondents in each country. Each respondent gives us 35 equivalent incomes, referring to hypothetical demographic and income situations of households in their country. So, a plausible question arises: “do people understand sharing costs of households with a different living standard from their own?” We test this question, by comparing answers on hypothetical household arrangements with answers that pertain the living standard of each respondent. We find that, in both Germany and France, respondents understand satisfactorily well the needs of households with different

Donaldson (1993).

⁴ We provide a discussion of how our findings compare to Donaldson and Pendakur (2003) in section 5.

⁵ As an example, Knetsch and Sinden (1984) use questionnaires similar to ours in order to assess discrepancies between maximum desired payments for avoiding a loss and a minimum desired compensation for accepting the loss. Waldfogel (1993) uses the same methodology for uncovering discrepancies between the personal utility that Christmas gift givers assess for gift receivers and the personal utility experienced by gift receivers, in order to finally estimate the deadweight loss of gift mismatches in the Christmas gift market.

living standards.⁶ This finding is encouraging for using small samples and it gives more flexibility to our method with respect to data-collection costs.

The plan of the paper is as follows. In section 2 we present and explain the structure of the questionnaire and the samples from the two countries. In section 3 we present the average equivalence scales and tests of IB. In section 4 we test the robustness of our method, by examining the possibility of framing effects, comparing our results with these from questionnaires with slightly different structure. In section 5 we compare our findings with these of previous studies and we suggest new directions. In section 6 we conclude.

2. Methodology, sampling, and data

2.1 Structure of the questionnaire

Our questionnaire is divided into two parts. The first part of the questionnaire is the experiment: we give questions to our subjects about hypothetical situations referring to relationships between income, family demographic composition and well-being. In the second part we ask for our respondents' personal characteristics. Our questionnaire appears in the Appendix, in section A1.

In the first part of the questionnaire we ask the respondents to evaluate five different incomes which describe five different welfare levels of the reference household. Each situation corresponds to a separate small table. Within each small table we provide eight hypothetical families of different size and composition (we tell our respondents to assume that adults are of age between 35 and 55, and children between 7 and 11). Only for one of these family types, a single adult without children (our reference household), we provide a monetary value that gives this household's after-tax income (the reference income). We leave gaps next to the remaining seven family types. We ask our respondents to fill in the gaps, putting the after-tax *family* income that brings the other household types to the same living standard as

⁶ In Koulovatianos et. al. (2001) we conclude the same for Cyprus as well, with a sample of 130 respondents.

the one of the single childless adult (with the given reference income), according to their own perception. There are five tables with identical structure, each of them providing a different reference income for the single-adult (reference) household.

We have selected five monetary values that match approximately the income levels of income quintiles for single childless adults at the time of sampling for each country. In particular, the reference-income level of the poorest single-childless adult that we provide is the poverty line and we proceed by adding 150% of the poverty line as we move upwards to a higher reference-income level.⁷

We emphasize that varying (gradually) the reference-income level is one of the crucial characteristics of our questionnaire. We want to test whether equivalence scales, as perceived by our subjects, depend on reference income in a systematic way.

In the second part of the questionnaire we ask our subjects to state several of their personal characteristics: gender, whether they have a partner, the number of children in the household, their after-tax personal income, their educational level (taking into account the differences in educational systems across the two countries), whether they had siblings during their childhood and, finally, their occupation.

As it is outlined by Bradbury (1989), in previous literature survey studies measuring equivalence scales, stimulus variables in the questionnaires have been qualitatively different from response variables: the past-literature questionnaires gave either an income stimulus

⁷ The definition of our family-income classes for Germany is based on the German Microcensus 1999: it uses the definition of the German poverty line in order to define certain income brackets. In particular, while the poverty line of the single-childless-adult household is about DM1000 per month (which is our starting value for our reference incomes in the questionnaire), the poverty line for a two-adult household in Germany is about DM1750. (See “Übersicht über das Sozialrecht” (Overview of Social Law) 1998). For Germany we provided monthly incomes of DM1000, DM2500, DM4000, DM5500, and DM7000 as reference incomes for single adults belonging to the five quintiles of the income distribution in 1999.

For France, our definition of reference income classes was motivated by our German definitions, in order to make the two databases directly comparable. Thus, we provided the following single-adult reference incomes: FF3000, FF7500, FF12000, FF16500, FF21000 for the year 2002, the amounts that are analogous to these defined for Germany.

If we use the PPP prices from the World Bank, these values correspond to year-2000 US after-tax annual incomes of \$7500, \$18750, \$30000, \$41250, and \$52500 for single childless adults.

and asked for a verbal proxy of welfare, or the other way around. This qualitative difference of stimulus/response variables in the previous surveys obliged scholars to use a utility theory in order to elicit equivalence scales from their survey databases. On the contrary, our questionnaire gives an income stimulus and receives an income response. In this way we require from our respondents to use tacitly “their own utility theory” and to reach conclusions by themselves about welfare-equivalent incomes across families with different demographic composition. In addition, we avoid noise stemming from different perceptions of verbal proxies for welfare by different respondents.

A plausible question about our method is whether the information contained in the equivalence scales that we report may be a superset of the information contained in scales elicited from consumer data. Our method’s equivalence scales may also contain information about the fertility preferences of our respondents. Pollak and Wales (1978) argue that equivalence scales obtained by consumer expenditure data are logically distinct from equivalence scales that also contain fertility information. They call the first category “conditional” equivalence scales and the second “unconditional” scales, arguing that the latter are appropriate for welfare comparisons. A cost function $C(u, p, z)$, depending on *exogenous* household characteristics, z , yields the minimum expenditure for reaching utility level u with a price vector p . The conditional equivalence scale is the ratio $\frac{C(u, p, z)}{C(u, p, \bar{z})}$, where \bar{z} is the exogenous characteristics of the reference household. Now, if household characteristics are *endogenous*, a cost function $C(u, p, p_z)$ yields the minimum expenditure for reaching utility level u with a commodity price vector p and a price vector p_z for household characteristics. For identifying the ratio that determines the unconditional equivalence scales, a good hypothesis might be that $C(u, p, p_{\bar{z}}) = C(u, p, \bar{z})$, with \bar{z} being a single adult household, assuming that being a single adult is an unavoidable choice in the beginning of each person’s life cycle. Then, by convention, unconditional scales will be given by the ratio $\frac{C(u, p, p_z)}{C(u, p, p_{\bar{z}})}$, where \bar{z} is the single

adult household.⁸ The Leyden school approach asks people their utility level on scale (e.g. 1 to 7), and the researcher asks how much money is needed to reach a given level of utility (1 to 7) for different household types. The equivalence scale is the ratio of these costs.⁹ Instead of the strong assumption made by the Leyden approach that all respondents understand the same while stating “my utility level is 4,” we assume that all our respondents understand the given household characteristics \bar{z} and all hypothetical z 's in the same way.

Which kind of equivalence scales do we obtain through our survey? We believe that we obtain “conditional” equivalence scales, i.e. scales that do not contain fertility information, or, at least, very little fertility information. While eliciting information about equivalence scales from consumer expenditure data, children expenditures are “de facto” conditional upon the actual family choice of having children. When single adults respond to our questionnaires, they examine a hypothetical situation of others having children as a given fact. We ask: “given that someone has an extra child, how much would they need to reach the same level of well-being?” Our questions are logically distinct from questions of the form “what monetary value would you place on having a child?”

It can be argued that someone who may not like having children may report lower costs for an extra child, or, alternatively, higher scales as a compensation for living with a child. Especially after having completed this study, and all formal tests, we believe that bias from people who are single or people who have chosen not to have children, is not significant. As we show in section 3, the scales we have collected have low standard errors and they are similar across two different countries. In all regressions appearing in section 3, the personal characteristics of our respondents, and especially whether our respondents have children, are insignificant. So, even if there are extra factors affecting our results, these are not too noisy.

Our consistent results across two countries make us believe that significant biases generated

⁸ We thank Krishna Pendakur for guiding us through the literature of conditional versus unconditional scales, and an anonymous referee for making several definitional distinctions about these concepts clear to us.

⁹ We report scales using the Leyden method in the section where we make a comparison of scales obtained by other approaches to ours, see Table 7.b.

by fertility preferences are unlikely. Our findings that follow make us suggest that our method leads to the measurement of *conditional* equivalence scales. For this reason we believe that our survey can be fruitfully combined with econometric approaches using demand data.

2.2 Sampling and Data

Our German sample consists of 167 respondents. We collected this sample in August 1999 mainly from the area of Schleswig-Holstein and especially from the city of Kiel. We approached people directly at their work places (companies, stores etc.) or at their leisure places (e.g. at parks or cafés). All our German subjects responded in written and received a compensation of about \$5. We did not hand out the questionnaire in a university classroom. By approaching people in person, we could identify more easily potential respondents with families and children.

Our French sample consists of 223 respondents. The sample is practically from all regions of France and we collected it in August 2002 through October 2002. A hundred of the respondents from this sample responded in written at a camping place in Bordeaux during their summer vacation time, and received a compensation of about \$5. By sampling at a camping place we were able to locate more easily people in households with more than one adults and with children, originating from many regions of France and from different family-income classes. The other 123 respondents responded electronically to randomly selected e-mails. All subjects who responded electronically participated into a lottery with expected payoff of about \$5.¹⁰

In Table 1 we present an outline of the personal characteristics of our respondents for both countries. We have collected personal features that could be important in affecting people's perceptions about equivalence scales. In particular, we asked for the respondent's

¹⁰The response rate of people who received e-mails was about 1.2%. We have found no statistically significant differences in the responses of the two groups of respondents, (i.e. the 100 respondents from the Bordeaux camping and the 123 people who responded through e-mail). A formal test can be provided by the authors upon request.

gender and their current family demographic composition.

We present two categories of income classes. The first is the family “after-tax income class” for all families of our data, independently of the family demographic composition. The after-tax income brackets are the same as these used in the German Microcensus for 1999. The income level “P” is the German poverty line for single-childless adults (see “Übersicht über das Sozialrecht” (Overview of Social Law) 1998), and the first after-tax income bracket is below $1.75 \times P$. The 1999 German Microcensus starts from this threshold in order to define the lowest-income class and then adds increments such that the mean of the third income class is about the mean German household income. Each increment is $1.5 \times P$.

In order to have a taste of our sampling efficiency, we compare the income and household-type distributions with other surveys with larger samples, or, whenever possible, with the whole household population. Sample frequencies from studies with larger samples are presented in the two columns named “Pop.” of Table 1, next to our own sample’s distribution frequencies.

Data on the non-adjusted German income distribution come from the 1999 German Microcensus (about 64 thousand households). Data on the French income and household type distribution refer to the whole French population (23.3 million households) in 1999.¹¹ Data on the German household-type distribution come from the 1998 Income and Expenditure Survey for Germany (EVS98) that is conducted only every fifth year. As we can see, our sample’s non-adjusted after-tax (and after-benefit) family income distributions do not differ significantly from these of the German Microcensus and these of the whole French household population.

The second category of income classes is the “adjusted after-tax income class” which is the

¹¹We are indebted to Francois Bourguignon for providing us with this database and to Olivier Bargain for retrieving the variables of interest for us. In all our calculations we adjusted the 1999 nominal values for inflation and PPP changes in order to make the 1999 database comparable to our 2002 sample. For our calculations we fitted a cubic spline in the cumulative French distribution.

family income per equivalent adult. In this second case we convert each respondent’s stated family income to their equivalent childless-single-adult household income. This conversion relies on the assumption that each respondent recognizes his/her own income situation when confronted with family types that correspond to his/her own demographic status. We find each respondent’s stated equivalent income for his/her own family type that is closest to his/her own family income.¹² Then, we divide this income with the respondent’s stated equivalence scale. In other words, instead of using the mean equivalence scales that appear later (in Table 2) for calculating our adjusted-income frequencies, we present distributions in which each person of our sample places (without noticing it explicitly) himself/herself in the adjusted (single-adult household) income classes they believe they belong to. These person-by-person adjusted-income distributions appear in Table 1 in the category “adjusted after-tax income class,” which reflects our sample’s distribution of living standards.

In order to calculate the frequencies of “adjusted after-tax income class,” that come from the larger samples on the income distributions of Germany and France (the numbers appearing in parentheses in the column “Pop.”), we used the OECD equivalence scales.¹³

In Table 1 we report data on the distribution of occupational characteristics from a larger sample only for Germany. In France we have a relatively high share (45.7%) of students who mainly responded through e-mail. With respect to educational categories, we take into account the differences in the two educational systems. The French education distribution appearing in column “Pop.” comes from a sample of 7602 heads of French households.

Finally, we ask our subjects how many siblings they had during their childhood. We

¹²We do not ask our subjects to make any explicit statement about their own family when asked about equivalence scales. We find out about it after looking at their personal characteristics.

¹³These are a 0.5 weight for each additional adult and a 0.3 weight for each additional child for all income categories. The fact that in Germany the fifth income category has very few or close to zero observations for most multi-person household types, may come partly from the fact that the German Microcensus provides only the 5 income categories that we define as well. These are very few data points for fitting a cubic spline, so we anticipate some error in the German distribution appearing in parentheses. On the contrary, the French data, provided to us by Francois Bourguignon, are split into 20 different income categories, enabling us to make more reliable calculations.

examine whether this is an important factor in forming people’s perceptions around possible household-production economies of scale. The corresponding distributions are shown at the bottom of Table 1.

3. Means of Equivalence Scales and Tests of the IB hypothesis

A direct way of evaluating our results is to look at the sample means of the equivalence scales we collected and their standard deviations. In Table 2 we give an outline of our sample means per hypothetical household composition and reference income.¹⁴ The symbol “A” stands for one adult and “C” for one child in the household.¹⁵ We remind that these sample means correspond to 167 observations for Germany and 223 observations for France, since in both countries our respondents gave a complete set of answers to the hypothetical situations that we asked them to evaluate. Underneath each of the sample means is the corresponding sample standard deviations appearing in parentheses.

We also give a visual outline of Table 2 in Figure 1, where we plot the average sample equivalence scales against the reference-income classes. The preliminary message of Figure 1 is clear: for all hypothetical household types, in both countries, equivalence scales fall with rising reference income. Moreover, our estimates for both countries are even quantitatively close.

It is easily seen by Figure 1 that the most intense decline in average equivalence scales occurs as we move from the lowest income class to the next. Is it the only statistically significant one?

In order to test the statistical significance of the overall picture in Figure 1 we perform tests of differences of means for every two consecutive means for a given household type.

Because all values are reported by the same group of individuals, they are not independent.

¹⁴We index reference incomes by 1, 2.5, 4, 5.5, 7 in order to show how many German single-childless-adult poverty lines each reference income is. Since we have adjusted both countries’ income categories to these German income levels, we avoid any reference to country-specific currency units and nominal values.

¹⁵So, for example, “ACCC” means a household with one adult and three children.

Therefore, the tests we perform are t-tests of differences of pairs of observations.

In Germany, for any given family type, average scales between any two subsequent reference-income levels decline significantly at the 99% level, except from two cases: the case of two childless adults and the case of two adults with one child. For these two cases, as we move from the fourth reference-income class to the fifth reference income class, the change in equivalence scales is statistically insignificant.

In France, in the case of two childless adults as we move from the fourth reference-income class to the fifth reference income class, the change in equivalence scales is statistically insignificant. In the case of two adults and one child, the difference between the fourth and the fifth income class is significant at the 95% level. In all other cases the decline of scales with income is significant at the 99% level.¹⁶

A conclusion that is evident from Table 2 and Figure 1 is that equivalence scales drop with rising reference income. The drop is more intense as we start from the lowest income level close to the poverty line and move upwards and it becomes smoother at the highest income levels.

In Table 3 we present another test for IB. We run the regression of the form

$$E_{i,j} = b_0 + b_1 \text{Ref. Income Dummies} + b_2 \text{PERSONAL_}Y_i + b_3 \text{OTHER_PERSONAL}_i + e_{i,j} ,$$

for each household type and present F -tests on exclusion of reference income dummies for each country.

By $E_{i,j}$ we denote the equivalence scale corresponding to the hypothetical family type “ j ” stated by respondent “ i ”. We include four income dummies, starting from income level equal to 2.5 up to the level 7. The constant b_0 of this regression captures the estimated equivalence scale of the lowest reference income, 1.

The variable $\text{PERSONAL_}Y_i$ is the respondent i 's net household income. This is *not* the

¹⁶We reported these t-tests in a previous version of this paper. All tests are available from the authors upon request.

adjusted after-tax income, but the stated family income. Since the adjusted income is derived by dividing family income by the stated equivalence scale of the respondent's demographic situation and income class, there would be a built-in endogeneity between the endogenous variable (equivalence scales) and the explanatory variable (adjusted income). Therefore, we use only family income even though it does not capture perfectly the variation in living standards across households.

OTHER_PERSONAL_i is a set of conditioning variables that comprise other personal characteristics of each respondent i . We include in the regression all the personal respondent variables appearing in Table 1: whether respondents live with an adult partner, whether they have children in their household, whether respondents had siblings during their childhood, the respondents' gender, their educational level and their occupational characteristics. Finally, $e_{i,j}$ is the error term.

We call the regressions including the income dummies as “unrestricted,” presented in columns having the symbol “U” in Table 3. The regressions under the IB restriction, $b_1 = 0$, are presented in columns named “R” in Table 3 (restricted).

In all cases, none of the personal characteristics of our respondents appeared as significant or robust. Therefore, we only report the estimators of parameters b_0 and b_1 . Underneath each coefficient estimate we provide its t-statistic in parenthesis. At the bottom of each household type regression, for each country, we report the F -test statistic on exclusion of reference income dummies. As we can see, in both countries IB is rejected. The implications of our findings are very strong for the estimation and explanation of equivalence scales by using consumer-expenditure data.

The methodology of using objective observations and eliciting equivalence scales after implementing a consumer-equilibrium concept is the widely accepted way of understanding household economies of scale. The method we suggest in this study is very different. Yet, we do not claim that our method should be a substitute to the usual way scholars think.

It is simply evident to us that preferences and within-household consumption externalities are “too many” unobserved factors to handle both theoretically and empirically. So, we believe, our method can be a useful *complement* to econometric approaches. As Donaldson and Pendakur (2003) mention (p. 177):

*“Confidence in an estimated equivalent-expenditure function [...] depends, in part, on a commitment to the way in which GESE (Generalized Equivalence-Scale Exactness) structures interhousehold comparisons of well-being. Such a commitment may be based on intuition or on independent research [...].”*¹⁷

In other words, Donaldson and Pendakur (2003) stress the fact that some guidance is needed in order to construct a parametric econometric model that may capture both preferences and within-household externalities. In this paper we suggest a simple survey method that can serve as a guide.

As our estimation method is non-parametric, it can serve as a basis for testing particular hypotheses about the construction of demand systems. For example, Figure 1 suggests a direct relationship between equivalence scales and the reciprocal of income.¹⁸ In Koulovatianos et. al. (2001), we show that this direct relationship is not rejected from data obtained by our method. But a more interesting hypothesis to test is a crucial implication of demand systems generating equivalent expenditure functions satisfying “Generalized Equivalence-Scale Exactness” (GESE): that the log of equivalence scales has a direct linear relationship with the log of reference income.

In Table 4 we present regressions of the log of our respondents’ stated equivalence scales against the log of reference income, separately for each household type. We report regressions

¹⁷With the term “Generalized Equivalence-Scale Exactness (GESE)” Donaldson and Pendakur (2003) refer to testable restrictions on the way expenditure functions of different households may be related. Our approach allows the testing of such restrictions. In particular, as we explain below, we find that GESE is supported by our sample.

¹⁸This direct relationship would be in accordance with “Generalized Absolute Equivalence-Scale Exactness” (GAESE) holding in a consumer demand system, as this is stressed in a previous version of Donaldson and Pendakur (2003).

using the specification,

$$\log(E_{i,j}) = a_0 + a_1 \log(\text{Ref. Income}) + a_2 \text{PERSONAL}_i + a_3 \text{OTHER_PERSONAL}_i + e_{i,j} .$$

The variable ‘‘Ref. Income’’ takes the values 1, 2.5, 4, 5.5, 7, for both countries. Again, none of the personal characteristics of our respondents appeared as significant or robust.¹⁹ Thus, again, we only report the estimators of parameters a_0 and a_1 . Underneath each parameter estimate we provide its t-statistic in parenthesis. Table 4 gives affirmative evidence that setting up parametric demand systems that imply equivalent expenditure functions complying with GESE is reasonable, as parameter a_1 is always negative and statistically significant. Moreover, Figures 2.a and 2.b depict scatter plots and the regression lines of this linear relationship between the log of equivalence scales and the log of reference income.

An interesting finding is that respondent characteristics do not play an important role for our subjective estimates of equivalence scales, and especially for establishing the negative relationship between equivalence scales and reference income. We have also estimated a SUR-type 7-equation system of the seven non-reference household types, regressing the log of scales against the log of reference income, which allows for several error correlations that could stem from systematic errors due to personal characteristics. For example, if some respondents think that children are cheaper than the average, could possibly report low children costs for all household types with children and exhibit significant negative deviations for such household types. The 7-equation SUR regression could uncover such biases originated by respondent characteristics. Yet, the estimators for parameter a_1 are almost the same as the ones reported in Table 4, so we do not report them in a new table.²⁰

¹⁹In a previous version of our paper we also reported how much \bar{R}^2 decreases by excluding all personal-characteristics variables and the difference is very small. In contrast to other studies estimating subjective scales, as this of Kapteyn and van Praag (1976), the personal characteristics of respondents do not play an important role.

²⁰The regression outputs of the 7-equation system can be provided by the authors upon request.

4. Investigating the robustness of the results

The results of our survey method may be biased due to two specific characteristics of our questionnaire, namely, (i) that the reference household, for which we also pre-specified a reference income, is always a single adult household instead of a larger household, and (ii) that we presented the single reference incomes in an increasing order starting from the lowest reference income. Ideally, both of these questionnaire characteristics should not influence the responses in our sample. However, alternative means of representing equivalent choice problems may lead to systematic biases in the responses. In other words, questionnaires with different structure may “frame” respondents’ answers towards certain directions (framing effects), even though the questionnaires may pose the same choice problem.²¹ In this section we argue that our qualitative results are not generated by such framing effects.

Concerning our questionnaire characteristic (i) above, especially for the lowest reference income level, it may be that respondents are unwilling to state income amounts that yield very low welfare levels of the single households. Respondents may feel sympathetic towards households with low living standards, and try to compensate them by stating higher increments as the family size rises. If such a framing effect is present, it contributes to finding decreasing equivalence scales in reference income. In order to rule out the possibility that this framing effect is generating our findings, we ran an additional survey in Germany (we refer to it as “new survey” in what follows) in which we pre-specified the income of the largest household (i.e. two adults with three children) instead of the single adult household. If the framing effect we explained above is present, respondents should now be unwilling to subtract too much income at low reference incomes. This would result in higher welfare levels of small households at low reference incomes and, thus, in equivalence scales increasing with reference income. Consequently, if the qualitative results in our original survey are correct

²¹Framing effects in research conducted through questionnaires is a subject of formal research in the fields of experimental economics and psychology. For example, see the study by Tversky and Kahnemann (1974).

and not caused by such a framing effect, equivalence scales should be also decreasing in the new survey.²²

Concerning characteristic (ii) of our questionnaire that we stress above, it may be possible that the order of presenting reference incomes has influenced our results. In the original survey, respondents started the questionnaire by thinking about the costs for additional household members at the lowest reference income. Then, the respondents had to consider all the other reference incomes in increasing order. In order to test for a possible order effect, we performed the new survey in two groups, L and H. In group L, we presented reference incomes in increasing order, as in the original survey. In group H, reference incomes were presented in decreasing order, starting from the highest reference income. If the order effect plays a dominant role, the qualitative results should differ in both groups. However, if the qualitative results in both groups are identical, we can conclude that our method is robust with respect to order effects.

The new survey was conducted in December 2003 in Germany, Kiel and Hannover, with 184 respondents, 84 in group L, and 100 in group H. Since the original survey revealed that the influence of personal characteristics can be neglected, we did not aim at a well-balanced sample and recruited solely students as respondents. Each respondent was rewarded with about \$5 (5 Euros) for participating. Further details about the new survey can be found in Appendix A2.

The results of the new survey are presented in Table 5. For both subsamples, L and H, as well as for the pooled sample (L&H), we run linear regressions identical to those presented in Section 3, apart from omitting personal characteristics. For the pooled sample, we include a dummy in the regression (called “Quest. Type”), which equals one for questionnaires of group L and zero for group H.

In order to make the results comparable to those of the original survey, we take again

²²We thank an anonymous referee for suggesting this alternative questionnaire structure to us.

the single adult household as reference, instead of the household consisting of two adults and three children. Therefore, we divide the incomes of all other household types by the income of the single adult household that each different respondent stated. This means that the reference incomes differ for the single respondents. Consequently, scale values of different respondents cannot be directly compared.

Table 5 shows that equivalence scales are significantly decreasing with reference income for all household types and all three samples. Moreover, the slopes are rather similar to those obtained in the original survey (see Table 4). We can therefore conclude that the qualitative results obtained in the original survey are robust in the sense that they have not been caused mainly by the particular characteristics (i) and (ii) of the original survey questionnaires discussed above. In particular, a framing effect of the original survey questionnaire coming from the possibility that respondents feel sympathetic towards the poor and they are framed by a tendency to increase their living standard by stating higher equivalent incomes, does not seem to generate the negative dependence of equivalence scales on reference income.

However, we cannot claim that framing effects are completely irrelevant in our survey method. First, equivalence scales in group H are always decreasing to a slightly higher extent than in group L. Second, the dummy in the pooled sample shows that there is a slight but significant difference between groups L and H, since the scale values in group H are usually higher. This means that in group H higher income amounts are subtracted.

In summary, we can conclude that framing effects have no influence on our qualitative results. In particular, framing effects do not generate the negative relationship between equivalence scales and reference income. Framing effects related to the order of the given questions do, however, alter the precise scale values slightly. Nevertheless, the small influence of question order is not a general drawback of our survey method, as it can easily be avoided in future studies, for instance, by asking each respondent only about one reference income.

The thought experiment that our respondents perform is similar to this of experts in “ex-

pert approaches” of calculating scales. Experts use insights from data on needs for households of different income levels, they form insights about these households’ needs and they suggest equivalence scales.²³

In our study, a large number of respondents adds more living-standards experiences and more preference profiles over income compared to expert approaches. Do respondents with specific levels of welfare understand household economies of scale in the same way as the rest of the population?

In order to test this question we restrict our sample by taking into account only the stated equivalence scales for which the reference income is closest to the respondent’s adjusted personal income. We therefore consider only 7 stated scales for each respondent (one scale per family type). We call this sample “Welfare Restricted,” denoting it as “WR.” We call the rest of the sample as “Unrestricted excluding Welfare Restricted” and we denote it as UR\WR. We test whether the responses of people concerning equivalence scales corresponding to their own living standard differ from the responses of people whose living standard is different from this living standard.

In Tables 6.a and 6.b we report the means and standard deviations of the two subsamples of respondents, UR\WR and WR, for each household type and for each reference income level. We perform a t-test of difference of means and we report its significance underneath the two stated means in each cell. Three stars (“***”) denote that the two means are different at 99% level of significance, two stars (“**”) at the 95%, one star (“*”) at the 90%, whereas the symbol “X” means that the null hypothesis of equality of the sample means cannot be rejected.

From Tables 6.a and 6.b one can draw the conclusion that, in general, respondents state similar equivalence scales, independently from whether their own welfare level is the same or different from the given hypothetical welfare level. Yet, some biases are present. In

²³See, for example Bradbury (1989) for a review of the “expert” or “budget approach.”

particular, the means of WR (the respondents' own welfare level) are higher at low reference income levels and lower at high income levels, compared to the rest of the respondents. This bias, however, reflects partly the fact that respondents have been sorted to the WR subsample, according to their adjusted personal income, and adjusted personal income has been determined by their own stated equivalence scale values. This endogeneity problem is most likely to be behind this bias. Nevertheless, we can conclude that respondents understand satisfactorily well the needs of households with different living standards.

5. Comparison with previous studies and suggested extensions

How do our results of Table 2 compare to previous studies obtaining objective or subjective scales? Equivalence scales obtained via econometric estimation in Germany appear in the book by Faik (1995). All scales by Faik (1995) that we report are based on the 1983 income and expenditure survey for West Germany, distinguishing households only by the number of household members (number of persons: 1-6). Table 7.a presents the results of Faik (1995), using different demand-system approaches. Our numbers for Germany are not so far from scales presented in Faik (1995).

On the contrary, previous subjective scales have not been close to ones obtained via consumer-expenditure data. Table 7.b presents results from such approaches for AACCC in Germany and France. Except from Riffault and Rabier (1977) who report a scale 2.23, closer to our findings, all other studies follow the Leyden school approach. The column "income level" states the reference income of the scales. For example, minimal is the poverty line and insufficient is below the poverty line.

If we weigh our study's equivalence scales according to each country's income- and household-type distribution, our average scales (not distinguishing among different reference incomes) are close to the equivalence scales stated in studies using consumer-expenditure data and estimating demand systems. Our numbers are also very close to the OECD ones that do

not distinguish among different reference incomes ($AC=1.3$, $ACC=1.6$, $ACCC=1.9$, $AA=1.5$, $AAC=1.8$, $AACC=2.1$ and $AACCC=2.4$). Thus, our scales are plausible. However, we add the important dimension of reference-income dependence. Our income-dependent scales are also close to these of Donaldson and Pendakur (2003) who also report income-dependent scales using Canadian consumer data. The dimension of scale dependence on income is crucial for evaluating horizontal equity: since taxes and transfers change the net income of a household, the horizontal-equity comparison through scales that are independent of base will be erroneous.

The value of our method and results is that they can serve as a guide to building applied models, a stepping stone for dealing with the statistically unobserved factors of household consumption behavior. We do not suggest that our methodology could be the final step before policy evaluation (taxation or redistribution). On the contrary, we believe that our subjective method can be a very important preliminary step within a more general methodology that uses applied dynamic models. Dynamic models that take explicitly into account households' expectations and savings decisions are appropriate for studying the responsiveness of income distributions to taxation schemes and redistributive policies. Models like these of Aiyagari et. al. (2000), Greenwood et. al. (2000) and (2003) are examples of this orientation. However, in these models household production parameters are calibrated to match income-distribution data in the presence of more "free" taste parameters that are jointly calibrated. The presence of too many parameters capturing unobserved factors, such as tastes and economies of scale in household production/consumption economies of scale, may cast doubt on the predictive accuracy of calibrated models. Therefore, estimating household production/consumption economies of scale separately can be of key interest, since a calibrating degree of freedom could be eliminated.

Estimating parameters capturing household-production/consumption economies of scale is a task of well-known difficulty (see, for example Bradbury (1995) and Pendakur (1999)).

One can use a database of subjective scales derived by our method and assume that these are the “true” scales. In a first step, using a plausible household-production parametric form, one can estimate parameters that capture household economies of scale, by regressing subjective scales on household income. Then, in a second step, using consumer-expenditure data and through data mining one can derive objective equivalence scales. If a particular functional-parametric form for household production performs poorly in reproducing the subjective scales, alternative ones can be tried and tested.

These two steps can be a useful iterative procedure that may uncover structural unobserved features of household production/consumption economies of scale. Moreover, fitting objective scales to subjective ones, allows to cross-check the validity of the two approaches.²⁴ Even though we are actively interested in this extension, in this study we confine ourselves to providing compelling cross-country evidence that our new subjective method improves upon existing ones and that it is reliable.

Another point made by Blundell and Lewbel (1991) is that the consumer demand approach can help in identifying the price dependence of equivalence scales across time and price regimes. Yet, the consumer demand approach cannot identify the values of scales in a “base period.” Our method serves as a means for identifying scales in the base period.²⁵

6. Conclusion

We have designed a subjective method for evaluating equivalence scales. Our target was to test and measure the dependence of equivalence scales on income. We directly asked our subjects to state welfare-equivalent incomes for hypothetical households with different demographic composition. Our respondents repeated this task for different income levels.

We implemented our method in two countries: Germany and France. We found, in both

²⁴Blundell and Lewbel (1991) also conclude that it would be fruitful to combine demand data with experimental (or, as they say, “psychometric”) data.

²⁵We thank an anonymous referee for suggesting this point.

countries, that the usual independence of base hypothesis (or equivalence scale exactness) is rejected. In particular, we found that equivalence scales depend negatively on reference income. This is an indication that economies of scale in household consumption increase as living standards go up. Moreover, in both countries, we found a strong linear relationship between the log of equivalence scales and the log of reference income, a key implication of the generalized equivalence-scale exactness hypothesis.

We also found that our method's subjective equivalence scales were not influenced quantitatively by our respondents' personal characteristics. Using an alternative survey in Germany, giving questions in a different order, we concluded that framing effects are not behind our key finding of a negative dependence of equivalence scales on reference income.

We suggested ways of combining our method with existing objective methodologies in future work. Such a synthesis can be useful for a future robust evaluation of household production/consumption economies of scale, equivalence scales and horizontal equity.

REFERENCES

- Aiyagari, S. R., J. Greenwood and N. Guner (2000): "On the State of the Union," *Journal of Political Economy*, 108, pp. 213-244.
- Barten, A. P. (1964): "Family Composition, Prices and Expenditure Patterns." In *Econometric Analysis for National Economic Planning*, edited by Peter E. Hart, Gordon Mills, and John K. Whitaker. London, Butterworths.
- Blackorby C. and D. Donaldson (1993): "Adult-equivalence Scales and the Economic Implementation of Interpersonal Comparisons of Well-being," *Social Choice and Welfare*, 10, pp. 335-361.
- Blundell, R. and A. Lewbel (1991): "The Information Content of Equivalence Scales," *Journal of Econometrics*, 50, pp. 49-68.
- Bradbury, B. (1989): "Family Size Equivalence Scales and Survey Evaluations of Income and Well-Being," *Journal of Social Policy*, 18, pp. 383-408.
- Bradbury, B. (1995): "Household Semi-Public Goods and the Estimation of Consumer Equivalence Scales: Some First Steps," *SPRC Discussion Paper No. 59*, May.
- Browning, M. (1992): "Children and Household Economic Behavior," *Journal of Economic Literature*, 30, pp. 1434-1475.
- Deaton, A. and C. Paxson (1998): "Economies of Scale, Household Size, and the Demand for Food," *Journal of Political Economy*, 106, pp. 897-930.
- Donaldson, D. and K. Pendakur (2003): "Equivalent Income Functions and Income-Dependent Equivalence Scales," *Journal of Public Economics*, 88, pp. 175-208.
- Faik, J. (1995): "Äquivalenzskalen: Theoretische Erörterung, Empirische Ermittlung und Verteilungsbezogene Anwendung für die Bundesrepublik Deutschland," *Volkswirtschaftliche Schriften*, Heft 451, Duncker und Humblot, Berlin.
- Gan, Li and V. Vernon (2003): "Testing the Barten Models of Economies of Scale in Household Consumption: Toward Resolving a Paradox of Deaton and Paxson," *Journal of Political Economy*, 111, pp. 1361-1377.
- Greenwood, J., N. Güner and J. Knowles (2000): "Women on Welfare: A Macroeconomic Analysis," *American Economic Review*, 90, *Papers and Proceedings*, pp. 383-388.
- Greenwood, J., N. Güner and J. Knowles (2003): "More on Marriage, Fertility, and the Distribution of Income," *International Economic Review*, 44, pp. 827-862.
- Hagenaars, A. (1985): "The Perception of Poverty," *Academic Thesis*, University of Leyden, Published in 1986, Amsterdam: North Holland.

- Hagenaars, A. and B.M.S. van Praag (1985): "A Synthesis of Poverty Line Definitions," *The Review of Income and Wealth*, 31, pp. 139–154.
- Kapteyn, A. and B.M.S. van Praag (1976): "A New Approach to the Construction of Family Equivalence Scales." *European Economic Review*, 7, pp. 313-335.
- Knetsch, J. L., and J. A. Sinden (1984): "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value," *Quarterly Journal of Economics*, 99, pp. 507-521.
- Koulovatianos, C., C. Schröder and U. Schmidt (2001): "On the Income Dependence of Equivalence Scales," Mimeo, Univ. of Cyprus.
- Pendakur, K. (1999): "Estimates and Tests of Base-Independent Equivalence Scales," *Journal of Econometrics*, 88, pp. 1-40.
- Pollak, R. A. and T. J. Wales (1979): "Welfare Comparisons and Equivalence Scales." *American Economic Review*, (papers and proceedings), 69, pp. 216-221.
- Riffault, H. and J. Rabier (1977), "La Perception de la Misère en Europe," Brussels: Commission des Communautés Européennes.
- Schröder, C. and U. Schmidt (1999): "A New Subjective Approach to Equivalence Scales: An Empirical Investigation," Mimeo, Univ. of Kiel.
- Schröder, C. and U. Schmidt (2000): "Personal Characteristics and Subjective Evaluation of Equivalence Scales." Mimeo, Univ. of Kiel.
- Tversky, A. and D. Kahnemann, (1974): "Judgments under Uncertainty: Heuristics and Biases," *Science*, 185, pp. 1124-1131.
- Übersicht über das Sozialrecht (1998), Annual Publication of the Bundesministerium für Arbeit und Sozialordnung (Federal Ministry of Labor and Social Order).
- van den Bosch, K. (1999): "Identifying the Poor, Using Subjective and Consensual Measures," Doctoral Thesis, Antwerpen.
- van Praag, B.M.S., T. Goedhart, and A. Kapteyn, (1980): "The poverty line - a pilot survey in Europe," *The Review of Economics and Statistics*, 62, pp. 461-465.
- van Praag, B.M.S., A. Hagenaars, and J. van Weeren (1982), "Poverty in Europe," *The Review of Income and Wealth*, 28, pp. 345-359.
- van Praag, B.M.S., and N.L. Van Der Sar (1988), "Household Cost Functions and Equivalence Scales," *The Journal of Human Resources* 23, pp. 193-210.
- van Praag, B.M.S., and R. Flik (1992), "Poverty Lines and Equivalence Scales: A Theoretical and Empirical Evaluation," Paper Presented at the Multidisciplinary Research Conference on Poverty and Distribution, Oslo, November 16–17.
- Waldfoegel, J. (1993): "The Deadweight Loss of Christmas," *American Economic Review*, 83, pp. 1328-1336.

	Germany			France		
	Sample: 167 obs.		Pop.	Sample: 223 obs.		Pop.
<i>Gender</i>	N	%	%	N	%	%
Female	71	42.5	51.1	106	47.5	51.3
Male	96	57.5	49.9	117	52.5	48.7
<i>Partner in the Household</i>						
Yes	97	58.1	58.0	154	69.1	72.4
No	70	41.9	42.0	69	30.9	37.6
<i>Number of Children in the Household</i>						
None	123	73.7	67.7	102	45.7	57.7
One	18	10.8	15.2	45	20.2	19.0
Two	15	8.9	13.1	46	20.6	15.8
More than two	11	6.6	4.0	30	13.5	7.5
<i>Family After-tax Income Class</i>						
1 ($Y < 1.75P$)	32	19.2	18.1	18	8.1	11.2
2 ($1.75P \leq Y < 1.75P + 1.5P$)	44	26.3	32.6	30	13.5	23.3
3 ($1.75P + 1.5P \leq Y < 1.75P + 3P$)	37	22.2	22.6	41	18.4	19.3
4 ($1.75P + 3P \leq Y < 1.75P + 4.5P$)	37	22.2	12.9	49	22.0	16.2
5 ($1.75P + 4.5P \leq Y$)	17	10.2	13.8	85	38.1	30.0
<i>Adjusted After-tax Income Class</i>						
1 ($Y < 1.75P$)	50	29.9	(32.6)	24	10.8	(28.9)
2 ($1.75P \leq Y < 1.75P + 1.5P$)	64	38.3	(45.6)	92	41.3	(40.8)
3 ($1.75P + 1.5P \leq Y < 1.75P + 3P$)	33	19.8	(16.8)	76	34.1	(17.5)
4 ($1.75P + 3P \leq Y < 1.75P + 4.5P$)	16	9.6	(4.1)	22	9.9	(6.4)
5 ($1.75P + 4.5P \leq Y$)	4	2.4	(0.9)	9	4.0	(6.4)
<i>Occupational Group</i>						
Welfare Recipient	2	1.1	3.3	1	0.4	
Unemployed	5	3.0	5.7	6	2.7	
Blue-collar Worker	10	6.0	20.7	6	2.7	
White-collar Worker	83	49.7	29.9	48	21.5	
Pupil, Student, Trainee	34	20.4	---	102	45.7	
Civil Servant	13	7.8	5.0	29	13.0	
Self-employed	7	4.2	5.8	13	5.8	
Pensioner	10	6.0	32.7	6	2.7	
Housewife, Houseman	3	1.8	---	12	5.4	
<i>Education</i>						
Below 9 years of Education	1	0.6		0	0.0	8.8
Completed Extended Elementary School	21	12.6	5.1	13	5.8	10.6
Completed Special Secondary School	39	23.4	18.6	43	19.3	46.9
Completed Secondary School	65	38.9	26.9	37	16.6	8.8
Technical School and University Degree	41	24.6	9.4	130	58.3	24.9
<i>Number of Siblings during Childhood</i>						
None	31	18.6		37	16.6	
One	55	32.9		72	32.3	
Two	47	28.1		59	26.5	
More than two	34	20.4		55	24.7	

Table 1 Breakdown of the Sample

Notes: For both countries we used our database of equivalence scales for calculating the adjusted income distribution in the way we explain in the text. For both countries we used the OECD equivalence scales for calculating numbers appearing in parentheses. Data in the columns "Population" refer to larger official surveys. Data on the non-adjusted German income distribution come from the 1999 German Microcensus. Data on the German gender distribution and data on education are taken from the 2000 official statistics of the German Statistisches Bundesamt. All other data for Germany are taken from the German 1998 Income and Expenditure Survey (EVS98) that is conducted every fifth year. Data on the French income distribution refer to the whole French population (23.3 million households). French education data refer to a sample of 7602 heads of French households. All French population data were provided to us by Francois Bourguignon.

Germany

Reference Income	AC Scale	ACC Scale	ACCC Scale	AA Scale	AAC Scale	AACC Scale	AACCC Scale
1	1.570 (0.230)	2.020 (0.398)	2.473 (0.612)	1.753 (0.205)	2.269 (0.325)	2.725 (0.498)	3.174 (0.749)
2.5	1.241 (0.114)	1.436 (0.195)	1.629 (0.283)	1.495 (0.266)	1.718 (0.319)	1.919 (0.394)	2.115 (0.474)
4	1.174 (0.110)	1.315 (0.181)	1.451 (0.254)	1.460 (0.279)	1.612 (0.329)	1.755 (0.373)	1.887 (0.435)
5.5	1.128 (0.089)	1.233 (0.150)	1.339 (0.210)	1.387 (0.265)	1.508 (0.311)	1.615 (0.359)	1.726 (0.416)
7	1.112 (0.088)	1.205 (0.146)	1.295 (0.201)	1.389 (0.272)	1.493 (0.317)	1.587 (0.365)	1.677 (0.413)

France

Reference Income	AC Scale	ACC Scale	ACCC Scale	AA Scale	AAC Scale	AACC Scale	AACCC Scale
1	1.579 (0.266)	2.055 (0.468)	2.487 (0.667)	1.734 (0.277)	2.224 (0.416)	2.670 (0.631)	3.092 (0.876)
2.5	1.300 (0.158)	1.539 (0.262)	1.756 (0.369)	1.505 (0.234)	1.763 (0.327)	1.981 (0.431)	2.188 (0.539)
4	1.253 (0.162)	1.444 (0.268)	1.614 (0.383)	1.441 (0.241)	1.636 (0.330)	1.806 (0.433)	1.966 (0.534)
5.5	1.211 (0.160)	1.370 (0.260)	1.511 (0.358)	1.403 (0.260)	1.569 (0.338)	1.714 (0.431)	1.847 (0.529)
7	1.196 (0.158)	1.341 (0.257)	1.473 (0.366)	1.403 (0.266)	1.553 (0.344)	1.683 (0.437)	1.808 (0.545)

Table 2 Average equivalence scales (standard deviations in parentheses)

Table 3 F-tests for exclusion of income dummies, Germany (1999) and France (2002)

Number of adults		Number of Children											
		0			1			2			3		
		A		AC		ACC		AAC		AACC		AAACC	
	Germany U	France R	Germany U	France R	Germany U	France R	Germany U	France R	Germany U	France R	Germany U	France R	
1	Constant	1.53 (62.32)	1.20 (39.40)	1.47 (47.53)	1.20 (36.85)	1.98 (44.83)	1.40 (25.78)	1.91 (36.50)	1.40 (24.50)	2.47 (35.22)	1.63 (20.16)	2.26 (30.64)	1.54 (18.70)
	Dummy Ref. Inc.= 2.5	-0.33 (-16.76)	--	-0.28 (-13.66)	--	-0.58 (-17.32)	--	-0.52 (-14.57)	--	-0.84 (-16.37)	--	-0.73 (-14.63)	--
	Dummy Ref. Inc.= 4	-0.40 (-20.28)	--	-0.32 (-15.92)	--	-0.71 (-21.13)	--	-0.61 (-17.14)	--	-1.02 (-20.13)	--	-0.87 (-17.30)	--
	Dummy Ref. Inc.= 5.5	-0.44 (-23.36)	--	-0.37 (-17.99)	--	-0.79 (-24.20)	--	-0.69 (-19.35)	--	-1.13 (-22.85)	--	-0.98 (-19.64)	--
	Dummy Ref. Inc.= 7	-0.46 (-24.19)	--	-0.38 (-18.75)	--	-0.81 (-25.09)	--	-0.71 (-20.25)	--	-1.18 (-23.81)	--	-1.01 (-20.30)	--
	\bar{R}^2	0.61	-0.01	0.38	0.01	0.63	0.00	0.42	0.01	0.62	0.00	0.43	0.02
	F	321.96		162.78		351.76		198.37		330.69		203.65	
2	Constant	1.69 (39.97)	1.43 (31.40)	1.64 (37.78)	1.41 (32.17)	2.17 (40.81)	1.62 (25.11)	2.04 (34.69)	1.57 (24.39)	2.63 (38.29)	1.83 (21.50)	2.73 (33.50)	1.82 (15.95)
	Dummy Ref. Inc.= 2.5	-0.26 (-9.84)	--	-0.23 (-9.40)	--	-0.55 (-15.54)	--	-0.46 (-13.19)	--	-0.81 (-16.39)	--	-1.05 (-15.47)	--
	Dummy Ref. Inc.= 4	-0.29 (-10.94)	--	-0.29 (-11.93)	--	-0.66 (-18.33)	--	-0.59 (-16.82)	--	-0.97 (-20.18)	--	-1.29 (-19.26)	--
	Dummy Ref. Inc.= 5.5	-0.37 (-14.16)	--	-0.33 (-13.04)	--	-0.76 (-21.91)	--	-0.66 (-18.54)	--	-1.11 (-23.43)	--	-1.45 (-21.90)	--
	Dummy Ref. Inc.= 7	-0.36 (-13.89)	--	-0.33 (-12.89)	--	-0.78 (-22.11)	--	-0.67 (-18.83)	--	-1.14 (-23.87)	--	-1.50 (-22.65)	--
	\bar{R}^2	0.24	0.03	0.20	0.01	0.46	0.01	0.35	0.02	0.53	0.00	0.40	0.03
	F	59.05		66.39		171.42		143.69		233.73		244.66	

Regressions for each different family type
 Endogenous variable: equivalence scales stated by respondents
 Number of observations: 835 (Germany), 1115 (France)
 White's Heteroskedasticity correction for covariance matrix
 t-statistics in parentheses

Table 4 Regressions of log scales against log reference income, Germany (1999) and France (2002)

Number of adults		Number of Children											
		0			1			2			3		
		A		AC		ACC		AAC		AACC		AACCC	
	Germany	France	Germany	France	Germany	France	Germany	France	Germany	France	Germany	France	
1	Constant			0.38 (23.52)	0.35 (17.29)	0.61 (25.74)	0.58 (20.67)	0.81 (26.95)	0.73 (21.18)				
	log(Ref. Inc.)			-0.18 (-29.20)	-0.14 (-22.34)	-0.26 (-31.43)	-0.22 (-24.92)	-0.33 (-31.66)	-0.27 (-25.59)				
	\bar{R}^2			0.60	0.37	0.62	0.42	0.62	0.43				
	F			90.53	48.40	98.36	58.50	97.41	61.82				
2		AA		AAC		AACC		AACCC					
		Germany	France	Germany	France	Germany	France	Germany	France	Germany	France		
	Constant	0.48 (16.85)	0.46 (17.18)	0.70 (22.84)	0.66 (21.28)	0.87 (25.98)	0.79 (22.45)	1.02 (27.73)	0.90 (22.57)				
	log(Ref. Inc.)	-0.13 (-16.34)	-0.12 (-15.85)	-0.22 (-25.31)	-0.19 (-22.57)	-0.29 (-28.11)	-0.24 (-24.14)	-0.33 (-28.56)	-0.28 (-24.54)				
	\bar{R}^2	0.24	0.19	0.41	0.33	0.48	0.37	0.51	0.39				
	F	19.50	20.11	42.22	40.27	55.90	48.49	62.24	52.66				

Regressions for each different family type

Endogenous variable: log of equivalence scales stated by respondents

Number of observations: 835 (Germany), 1115 (France)

White's Heteroskedasticity correction for covariance matrix

t-statistics in parentheses

Table 5 Germany New survey, 2003

Number of adults		Number of Children																				
		0					1					2					3					
		A					AC					ACC					AACC					
		L&H	L	H	L&H	L	H	L&H	L	H	L&H	L	H	L&H	L	H	L&H	L	H	L&H	L	H
1	Constant				0.47 (33.46)	0.41 (27.05)	0.48 (25.63)	0.76 (43.31)	0.66 (35.38)	0.79 (34.07)	0.98 (50.60)	0.86 (38.93)	1.01 (40.37)									
	log(Ref. Inc.)				-0.16 (-19.09)	-0.15 (-16.12)	-0.18 (-13.69)	-0.26 (-24.88)	-0.23 (-20.28)	-0.28 (-18.39)	-0.33 (-28.32)	-0.29 (-21.62)	-0.35 (-21.20)									
	Quest. Type				-0.04 (-4.98)	---	---	-0.06 (-5.59)	---	---	-0.07 (-5.25)	---	---									
	\bar{R}^2 F				0.42 330.84	0.42 308.37	0.39 320.38	0.50 465.46	0.50 414.14	0.48 469.90	0.53 511.69	0.50 425.19	0.52 537.76									
2	Constant				0.89 (45.44)	0.78 (36.31)	0.91 (35.24)	1.01 (55.88)	0.97 (42.06)	1.13 (45.00)	1.27 (62.30)	1.13 (44.65)	1.31 (51.07)									
	log(Ref. Inc.)				-0.19 (-14.65)	-0.17 (-12.28)	-0.20 (-10.30)	-0.28 (-23.47)	-0.25 (-18.98)	-0.30 (-17.02)	-0.34 (-29.02)	-0.31 (-21.52)	-0.34 (-22.51)	-0.42 (-24.77)								
	Quest. Type				-0.05 (-4.21)	---	---	-0.07 (-5.49)	---	---	-0.08 (-5.34)	---	---									
	\bar{R}^2 F				0.29 188.77	0.25 138.18	0.29 200.71	0.46 392.74	0.42 306.43	0.45 412.36	0.52 497.11	0.47 366.07	0.53 552.53	0.55 615.28								

Regressions for the German 2003 sample and its subgroups defined by the type of questionnaire
 Endogenous variable: log of equivalence scales stated by respondents
 Number of observations: L&H 920, L 420, H 500
 White's Heteroskedasticity correction for covariance matrix
 t-statistics in parentheses

Reference Income	AC		ACC		ACCC		AA		AAC		AACC		AACCC	
	UR\WR	WR	UR\WR	WR	UR\WR	WR	UR\WR	WR	UR\WR	WR	UR\WR	WR	UR\WR	WR
1	1.57 (0.23)	1.56 (0.22)	2.02 (0.39)	2.03 (0.41)	2.48 (0.60)	2.46 (0.62)	1.73 (0.27)	1.80 (0.20)	2.26 (0.32)	2.30 (0.33)	2.69 (0.49)	2.80 (0.51)	3.14 (0.73)	3.26 (0.78)
	X		X		X		*		X		X		X	
2.5	1.22 (0.10)	1.27 (0.12)	1.41 (0.18)	1.49 (0.20)	1.59 (0.28)	1.69 (0.27)	1.48 (0.27)	1.52 (0.26)	1.69 (0.33)	1.77 (0.30)	1.87 (0.41)	1.99 (0.35)	2.06 (0.50)	2.20 (0.41)
	**		***		**		X		*		*		*	
4	1.17 (0.11)	1.18 (0.11)	1.31 (0.18)	1.33 (0.18)	1.45 (0.25)	1.46 (0.25)	1.48 (0.28)	1.38 (0.25)	1.63 (0.33)	1.54 (0.29)	1.77 (0.38)	1.69 (0.33)	1.90 (0.45)	1.81 (0.37)
	X		X		X		*		X		X		X	
5.5	1.13 (0.09)	1.10 (0.05)	1.24 (0.15)	1.19 (0.11)	1.35 (0.21)	1.27 (0.15)	1.40 (0.26)	1.30 (0.25)	1.52 (0.31)	1.40 (0.29)	1.63 (0.36)	1.49 (0.32)	1.74 (0.42)	1.60 (0.38)
	X		X		X		X		X		X		X	
7	1.11 (0.09)	1.03 (0.05)	1.21 (0.14)	1.06 (0.10)	1.30 (0.20)	1.09 (0.16)	1.40 (0.27)	1.14 (0.25)	1.50 (0.31)	1.17 (0.30)	1.60 (0.36)	1.20 (0.35)	1.69 (0.41)	1.23 (0.40)
	**		**		**		*		*		**		**	

Table 6.a Average equivalence scales and t-test of means difference for Germany (1999)

Reference Income	AC		ACC		ACCC		AA		AAC		AACC		AACCC	
	UR\WR	WR	UR\WR	WR	UR\WR	WR	UR\WR	WR	UR\WR	WR	UR\WR	WR	UR\WR	WR
1	1.58 (0.25)	1.59 (0.34)	2.05 (0.47)	2.10 (0.47)	2.47 (0.66)	2.62 (0.70)	1.73 (0.26)	1.78 (0.40)	2.22 (0.40)	2.27 (0.50)	2.66 (0.62)	2.77 (0.71)	3.07 (0.85)	3.30 (1.03)
	X		X		X		X		X		X		X	
2.5	1.27 (0.13)	1.34 (0.18)	1.49 (0.22)	1.61 (0.30)	1.69 (0.32)	1.84 (0.41)	1.46 (0.22)	1.57 (0.23)	1.70 (0.30)	1.85 (0.34)	1.90 (0.39)	2.10 (0.46)	2.09 (0.48)	2.32 (0.58)
	***		***		***		***		***		***		***	
4	1.27 (0.18)	1.22 (0.12)	1.48 (0.30)	1.37 (0.18)	1.67 (0.43)	1.51 (0.24)	1.45 (0.26)	1.41 (0.18)	1.66 (0.37)	1.58 (0.23)	1.85 (0.49)	1.71 (0.27)	2.03 (0.60)	1.83 (0.33)
	**		***		***		X		*		**		***	
5.5	1.22 (0.16)	1.17 (0.09)	1.38 (0.27)	1.29 (0.15)	1.52 (0.37)	1.40 (0.21)	1.42 (0.26)	1.28 (0.23)	1.59 (0.34)	1.41 (0.27)	1.74 (0.44)	1.52 (0.31)	1.87 (0.54)	1.62 (0.37)
	X		X		X		**		**		**		**	
7	1.20 (0.16)	1.13 (0.12)	1.34 (0.26)	1.25 (0.25)	1.48 (0.36)	1.38 (0.37)	1.41 (0.26)	1.28 (0.31)	1.56 (0.34)	1.40 (0.42)	1.69 (0.43)	1.52 (0.54)	1.81 (0.54)	1.65 (0.64)
	X		X		X		X		X		X		X	

Table 6.b Average equivalence scales and t-test of means difference for France (2002)

Table 7.a Equivalence scales for Germany obtained from consumer data

Number of Persons	Engel ^a	Barten ^b	Translating ^b	Prais and Houthakker ^b
1	1.00	1.00	1.00	1.00
2	1.81	1.48	1.34	1.55
3	2.19	1.73	1.53	1.84
4	2.45	1.89	1.64	2.02
5	2.77	1.98	1.72	2.17

Source: Faik (1995).

^a Equivalence scales for commodity group food.

^b Equivalence scales for arithmetic mean of income.

Table 7.b Subjective equivalence scales for Germany and France

Author	Country	Income Level	Equivalence Scale for AACCC
van Praag et al. (1980)	France	minimal	1.50
	Germany		1.83
van Praag et al. (1982)	France	mean	1.22
	Germany		1.54
van Praag et al. (1988)	France	insufficient	1.51
	Germany		1.83
van Praag and Flik (1992)	France	mean	1.40-1.60 ^a
Hagenaars (1985)	France	mean	1.24
	Germany		1.38
Riffault and Rabier	France	minimal	2.23

Source: Data are taken from van den Bosch (1999)

^a Equivalence scales vary according to model specification.

Subjective equivalence scales presented in Table 7 are derived from responses to questions such as "Under my (our) conditions I would call an after-tax income [...] of:

about very bad

about bad

[.....]

about very good" (see Hagenaars, 1985, p. 44).

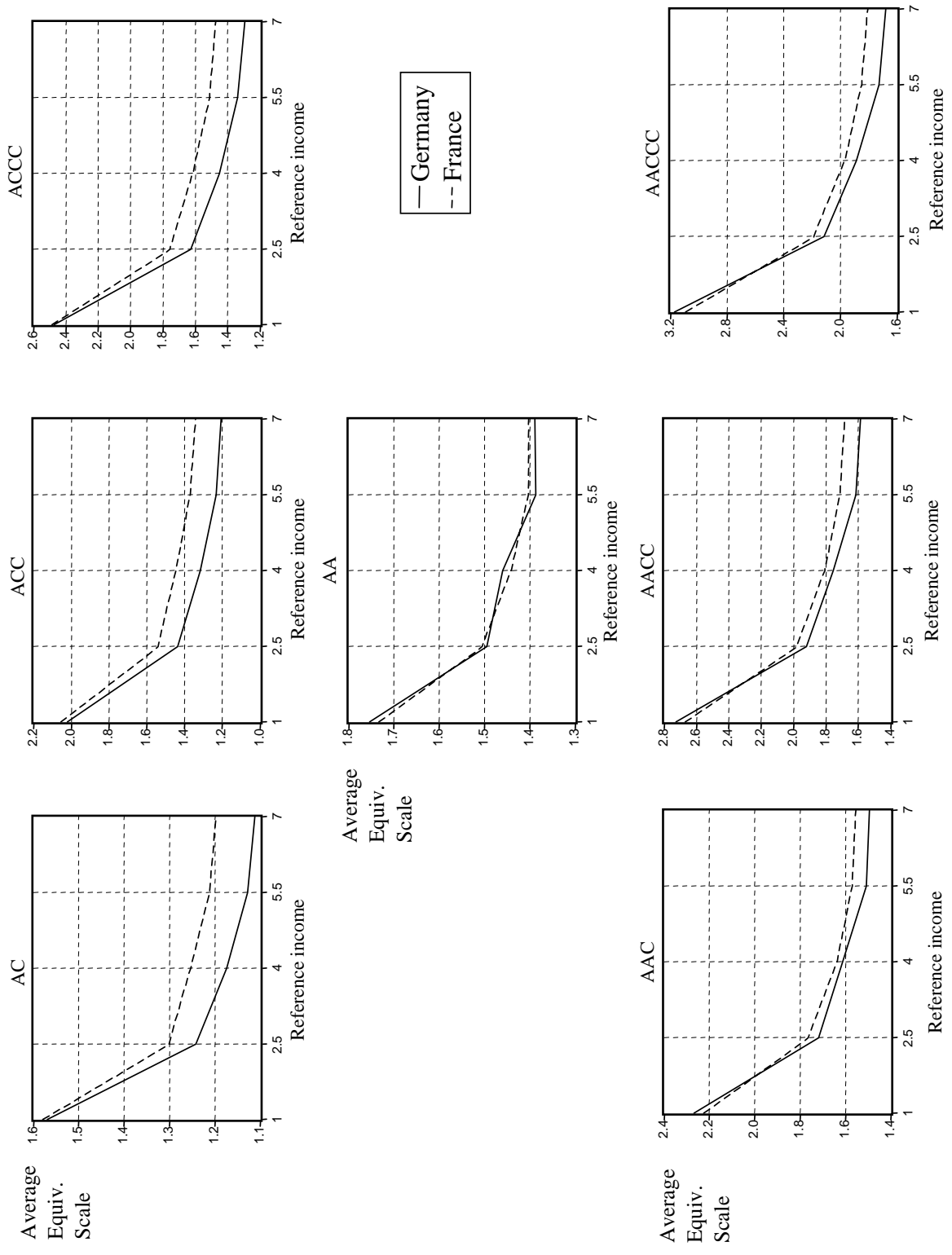


Figure 1 Average Equivalence Scales

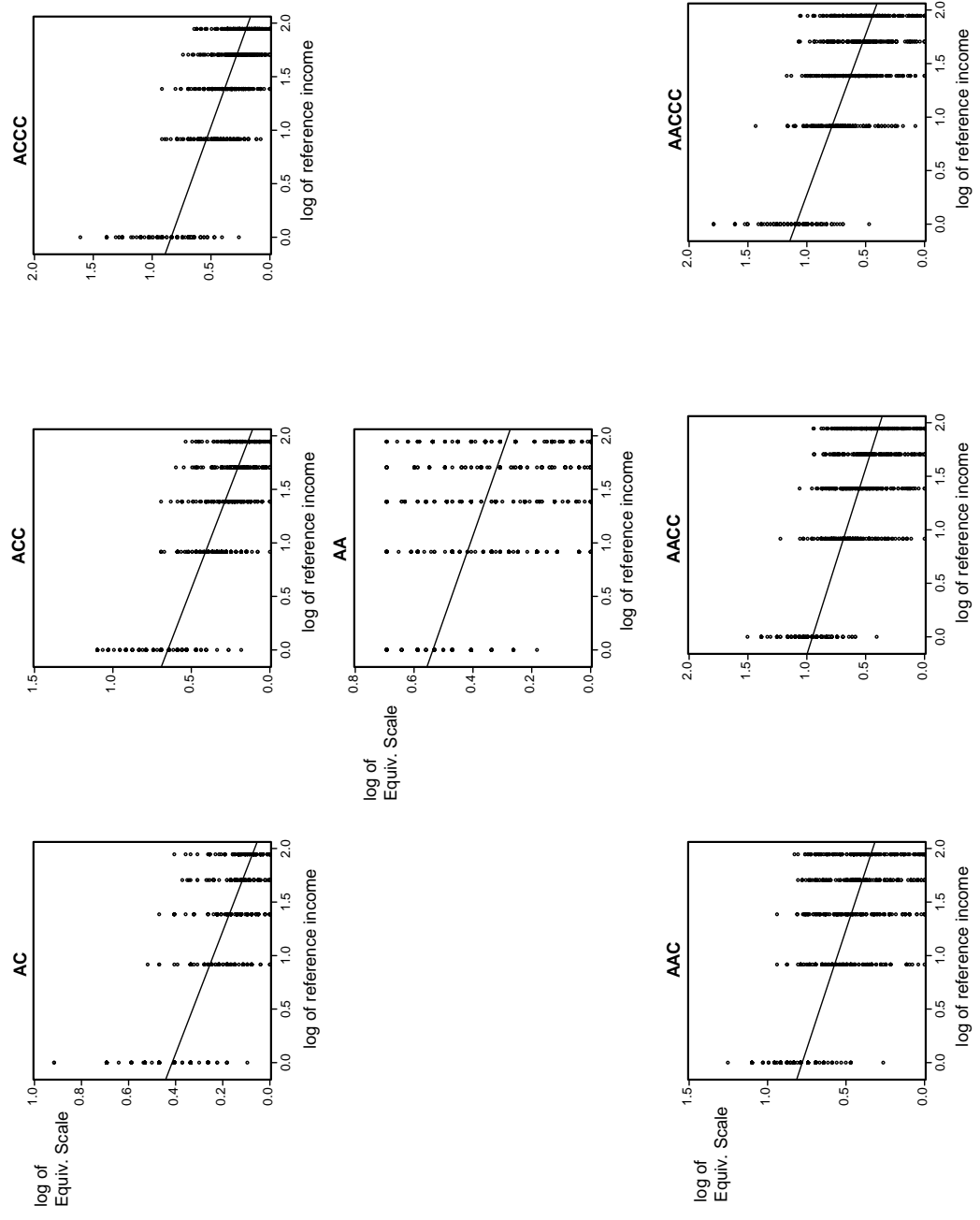


Figure2.a Germany Scatter plots of the log of equivalence scales versus the log of reference income for each family type

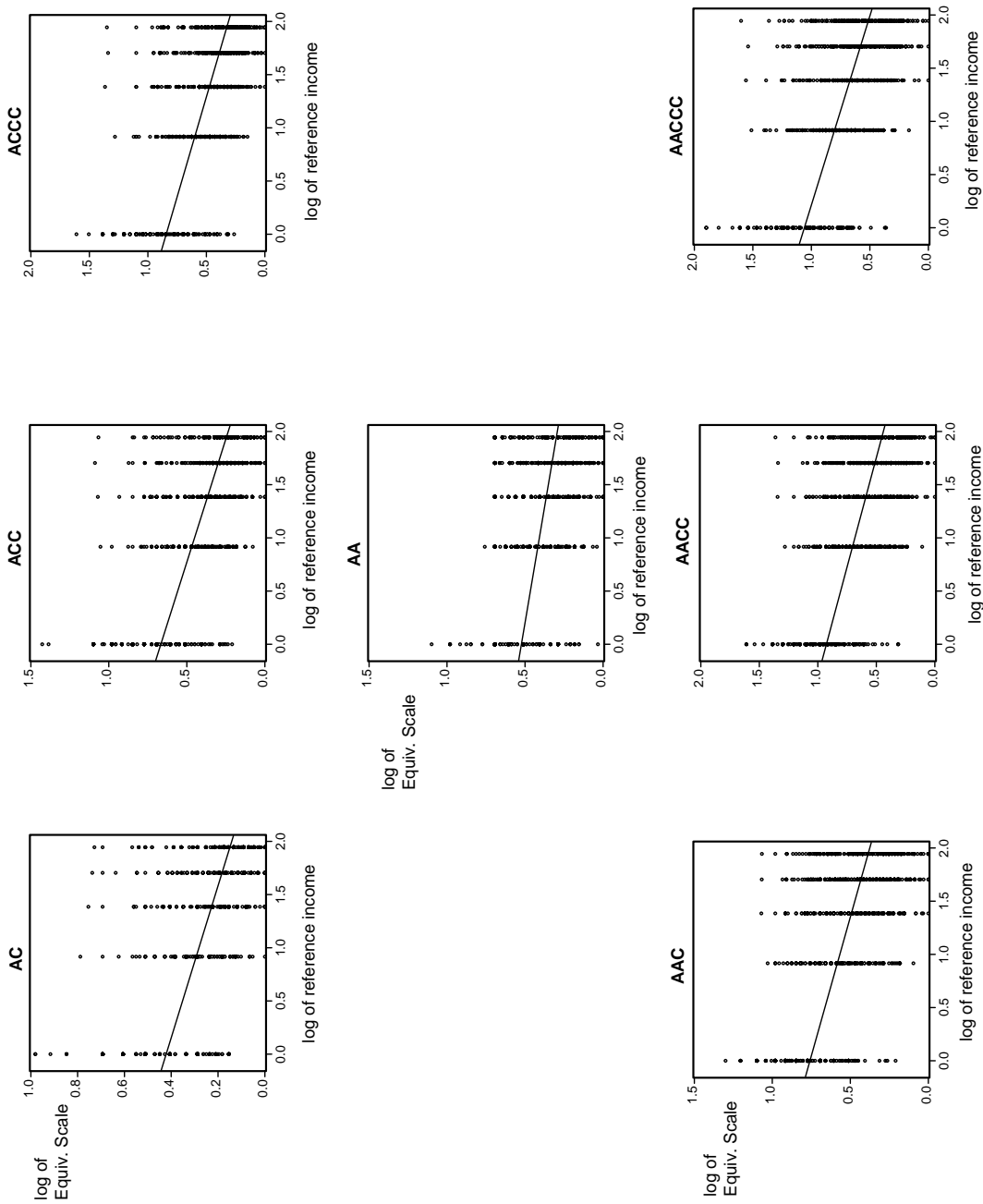


Figure 2.b France Scatter plots of the log of equivalence scales versus the log of reference income for each family type

Appendix

A1 Questionnaire of the original survey

Attention: After filling out this questionnaire, all questionnaires will be immediately put in identical envelopes. This ensures that your responses will be treated anonymously. All collaborators engaged in collecting or analysing data for the present survey are obliged to treat personal data confidentially.

1. Purpose of the survey

In order to determine social assistance levels or tax rates in an equitable way, one has to compare the income needs of households which differ in size or composition. In general, different household types may have different income needs in order to attain a given living standard. Since these income needs are difficult to assess in an objective way, we would like to ask you for your personal evaluation. Please note that in the following questionnaire do not exist objectively “right” or “wrong” answers, which means that your answers should only reflect your personal judgements.

2. Income evaluation questions

In the tables below you shall evaluate five different situations. The situations differ by the pre-specified monthly net income (including all social transfers) of a single adult household. Now consider for each situation separately that the size and composition of the households change according to the table. Which monthly net income would each household type need in order to attain the same living standard as the single adult household with the pre-specified income? You should state precisely this income for each household type in the tables below. Within a given table, all household types should attain an identical living standard. Assume for your assessment that adults are between 35 and 55 and children between 7 and 11 years old.

Single adult household without a child	1000 DM	Two adult household without a child	?
One parent household with 1 child	?	Two parent household with 1 child	?
One parent household with 2 children	?	Two parent household with 2 children	?
One parent household with 3 children	?	Two parent household with 3 children	?

Single adult household without a child	2500 DM	Two adult household without a child	?
One parent household with 1 child	?	Two parent household with 1 child	?
One parent household with 2 children	?	Two parent household with 2 children	?
One parent household with 3 children	?	Two parent household with 3 children	?

Single adult household without a child	4000 DM	Two adult household without a child	?
One parent household with 1 child	?	Two parent household with 1 child	?
One parent household with 2 children	?	Two parent household with 2 children	?
One parent household with 3 children	?	Two parent household with 3 children	?

Single adult household without a child	5500 DM	Two adult household without a child	?
One parent household with 1 child	?	Two parent household with 1 child	?
One parent household with 2 children	?	Two parent household with 2 children	?
One parent household with 3 children	?	Two parent household with 3 children	?

Single adult household without a child	7000 DM	Two adult household without a child	?
One parent household with 1 child	?	Two parent household with 1 child	?
One parent household with 2 children	?	Two parent household with 2 children	?
One parent household with 3 children	?	Two parent household with 3 children	?

3. Questions pertaining the respondent

Please mark the answers that apply to you. Your answers will be treated confidentially.

1) Please state your gender:

- male
 female

2) Do you have a partner living in your household?

- yes
 no

3) How many children live in your household?

- 0
 1
 2
 3 or more

4) In which range is the total net monthly income of your household?

- below 1750 DM
 1750-3249 DM
 3250-4749 DM
 4750-6249 DM
 more than 6249 DM

5) Please state your occupation

- welfare recipient
- unemployed
- blue-collar worker
- white-collar worker
- civil servant

- pupil, student, Trainee
- self-employed
- pensioner
- housewife/-man

5) Please state your education level:

- less than 9 years of education
- completed extended elementary school
- completed special secondary school
- completed secondary school
- technical school or university degree

6) How many siblings did you have during your childhood?

- 0
- 1
- 2
- 3 or more

A2 Questionnaire of the new survey

The only difference between the new questionnaire and the old one is that the five tables now have the following form:

Two parent household with 3 children	reference income	One parent household with 3 children	?
Two parent household with 2 children	?	One parent household with 2 children	?
Two parent household with 1 child	?	One parent household with 1 child	?
Two adult household without a child	?	Single adult household without a child	?

In group L of the new survey, reference incomes were, as in the questionnaire of the original survey, presented in increasing order, i.e. starting from 3400 DM and ending with 12550 DM (the closest rounded numbers of averages of the stated equivalent incomes from the first survey), whereas in group H reference incomes were presented in decreasing order, i.e. starting from 12550 DM and ending with 3400 DM. In order to ensure comparability of the original and the new survey, we have chosen as currency for the stated reference incomes, as well as for the responses, German Marks in the new survey as well. However, for the stated reference incomes we also gave, in parentheses, the corresponding amount in Euros.