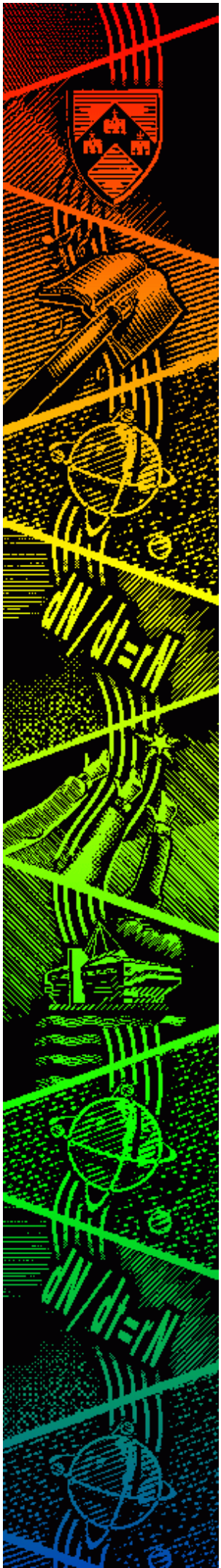

THE UNIVERSITY of York

Discussion Papers in Economics

No. 12/23

Two can live as cheaply as one... But three's a crowd

Christopher R. Bollinger, Cheti Nicoletti and Stephen Pudney

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Two can live as cheaply as one... But three's a crowd

Christopher R. Bollinger

University of Kentucky and ISER, Essex University

Cheti Nicoletti

DERS, University of York and ISER, University of Essex

Stephen Pudney

ISER, University of Essex

This version: September 12, 2012

Keywords: Equivalence scales; subjective wellbeing

JEL codes: I32, J12, J13

We are grateful to the Leverhulme Trust for financial support for Bollinger's Visiting Professorship at ISER. Nicoletti and Pudney's involvement was supported by the ESRC Research Centre on Micro-Social Change (award no. RES-518-285-001).

Abstract

To measure poverty, incomes must be equivalized across households with different structures. In this paper, we use a very flexible ordered response model to analyze the relationship between income, demographic structure and subjective assessments of financial wellbeing drawn from the 1991-2008 British Household Panel Survey. Our results suggest the existence of large scale economies within marital/cohabiting couples, but substantial diseconomies from the addition of children or further adults. This pattern contrasts sharply with commonly-used equivalence scales, and is consistent with explanations in terms of the capital requirements associated with additions to the core couple.

1 Introduction

The analysis of poverty and inequality requires an ability to compare households or individuals with each other and with any choice of poverty line. This requires us to take account of differences in household size and structure: a five-person household clearly needs more resources than a one-person household to achieve the same standard of living for its members. The necessary demographic equivalence relation will reflect the number of household members, any differences that exist in the needs of different types of individual (particularly adults and children) and the extent of returns to scale in the production of "wellbeing" within the household. This is unavoidable: demographic equivalence relations are implicit in any complete specification of the criterion determining who is classified as being in poverty.

Various methods have been used to establish demographic equivalence relations, usually in the form of adult equivalence scales which 'deflate' household income or consumption to express it as resources per equivalent adult. There are four principal approaches used in practice.[1] The *cost of basics* approach attempts to determine for each household type a consumption bundle interpretable as providing a minimally acceptable standard of living. Valuing these basic consumption plans at current market prices then gives a poverty line for each household type and the relationship between those poverty lines defines an equivalence scale. This approach was implemented by Orshansky (1963) and has played a dominant role in US poverty analysis since that time (Fisher 1997). In the US, the basic consumption plan is specified as an 'economy' food bundle (see Ziliak 2006), but it has been criticized (Ruggles 1990, Citro and Michael 1995) as being an emergency plan, rather than a basic needs plan.

A second method uses household demand theory, deriving equivalence scales by observation of *consumer behavior* (see for example Prais and Houtaker 1955, Deaton and Muellbauer 1980, Blundell and Lewbel 1991 and Lewbel 1997). Households are assumed to choose their consumption by maximizing utility given prices and their income and demographic characteristics.The equivalence scale is then derived as the ratio of incomes needed by two households with different demographic characteristics to attain the same level of utility. The main issue with this method is that the identification of the equivalence scales cannot be based exclusively on observations of household budgeting and generally requires additional assumptions,

[1]For a detailed review of approaches used to derive equivalence scales see Coulter et al (1992).

which may be questionable (see Blundell and Lewbel 1991, and Lewbel 1997). One of the most influential early attempts to construct equivalence scales from observations of household budgeting was by McClements (1977), who modified the iterative algorithm developed by Singh and Nagar (1973) and used the UK Family Expenditure Survey to produce an equivalence scale which was quite close to the scale implicit in the design of the UK welfare benefit system at the time.[2] As pointed out by Muellbauer (1979), McClements' method is seriously flawed, because it fails to address the fundamental identification problem. More recent work has sought to overcome this problem using a priori information, for example by attributing specific expenditure categories (such as children's clothes or alcoholic drink) to specific age groups within the household – an idea that dates back to Rothbarth (1943) and has been extended and implemented in more recent work (for example, Bargain and Donni 2009). Despite the flaws in the McClements scale, it remained in favor for poverty analysis for many years in official UK poverty statistics (DWP 2011).

A third approach is to use *arbitrary scales*, not formally derived from any analysis of basic needs or consumer behavior. The simplest of these is a constant returns to scale (CRTS) assumption, implying household income per capita as the adjusted measure of resources. Much more widely used is the OECD equivalence scale which, in its current form (Haagenars et al. 1994) counts the first adult, each additional adult and each child (aged under 15) as 1, 0.5 and 0.3 equivalent adults respectively. Another arbitrary scale in common use is the square root of household size (the SQRT scale) which has recently been adopted by the OECD (2009) and which can be generalized to the class of equivalence scales with constant elasticity with respect to household size; the current OECD scale usually produces results close to those of the SQRT scale.[3] These arbitrary scales have one or more of three potentially restrictive features built into their design: (*i*) a relatively low degree of returns to scale for adults within a marital or cohabiting couple; (*ii*) a similar low degree of returns to scale for further adults additional to the core couple; and (*iii*) a high degree of returns to scale

[2]Renormalizing on an individual-equivalent basis, the McClements scale counts the household head as 1 unit; a spouse as 0.64; a second (non-spouse) adult as 0.75; a third adult as 0.69; each subsequent adult as 0.59; and each child in the age ranges 0-1, 2-4, 5-7, 8-10, 11-12, 13-15 and 16-18 as 0.15, 0.30, 0.34, 0.38, 0.41, 0.44, 0.62 respectively.

[3]Buhmann et al (1988) are the first economists to introduce the class of equivalence scales with constant elasticity and show that most of the scales used in practice can be approximated by such scales. Coulter et al (1992) and Jenkins and Cowell (1994) consider this class of equivalence scales and provide together with Buhmann et al (1988) extensive evidence on the sensitivity of poverty and inequality measures to the choice of the equivalence scale.

for children. Our findings clearly reject all three features and suggest the need for radical reconsideration of the construction of equivalence scales.

A fourth approach uses subjective scales i.e. scales which are derived by asking people to report the minimum income they deem to be needed to make ends meet (see the subjective poverty line approach), or the levels of income that they deem to be very bad, bad, insufficient, sufficient, good and very good (see Leyden poverty line approach).[4] These methods have been criticized because the respondents might find difficult to make judgements on situations which could be very far from their actual situation and may be confused about the precise definition of income they have to consider (see Kapteyn and van Praag 1976 and Tummers 1994). We avoid completely these issues by considering questions which ask people to judge their actual financial wellbeing by reporting whether they are living comfortably, doing alright, just about getting by, finding it quite difficult, or finding it very difficult. This question is similar in spirit to that used by the Centre for Social Policy (CPS) for poverty analysis, based on a question asking respondents whether they are able of make ends meet with their actual income with great difficulty, with difficulty, with some difficulty, rather easily, easily and very easily (see Flik and van Praag 1991). Nevertheless, our approach to derive the equivalence scales is different from the CPS approach and it is closer to methods adopted by Melenberg and van Soest (1996), Pradhan and Ravallion (2000) and Charlier (2002), who use self-reported perception of consumption adequacy and satisfaction with life and work.

The method we pursue here for constructing demographic equivalence assumes that subjective assessments of financial wellbeing are monotone transformations of individual utility plus a perception error. We estimate how utility relates to household demographic characteristics and income and use the results to compute equivalence scales as the ratio of household incomes required by two households with different demographic characteristics to attain the same level of individual utility. Subjective assessments of this kind are now widely accepted. See Diener et al (1999), Kahneman and Krueger (2006) and Kahneman and Deaton (2010) for excellent reviews of various uses of subjective wellbeing measures, and Berthoud et al (2004) who review their empirical relationship to poverty statistics using our data source, the British Household Panel Survey (BHPS).

[4]For more details on these subjective scales approach see Goedhart et al (1977), Kapteyn and van Praag (1976), van Praag and van der Sar (1988), Kapteyn et al (1988), Flik and van Praag (1991).

Nevertheless, we have to recognize that equivalence scales based on subjective assessments have not been accepted with universal enthusiasm and economists are still sceptical about the reliability of subjective measures (for a review of advantages and limits of subjective measures of wellbeing, see Ravallion and Lokshin 2001 and Ravallion 2012). There are concerns that different people might understand subjective questions differently or might differ in their use of the scale provided to rank their situation. This heterogeneity in response style across individuals can lead to biases. In our empirical application we explicitly take account of this issue by controlling for differences in the response style. We assume that individuals adopt a set of thresholds to divide the range of possible values of their financial wellbeing into contiguous intervals each one corresponding to a different level of the 5-point scale used to rate their financial wellbeing. The response style of each person can be summarized by his/her choice of thresholds.

We initially allow the thresholds to be individual-specific by considering an additive individual effect which can shift all the thresholds up or down by the same amount, leading to ordered models with individual fixed or random effects for the relationship between subjective financial wellbeing and household income and composition.[5] Unlike earlier papers, we also consider other types of heterogeneity in response style by allowing each threshold to depend on personal characteristics, such as psychological wellbeing, personality traits, education and age.[6] The inclusion of personal characteristics allows each threshold to shift by a different amount. This implies that we can capture differences in response style that cause a stretching or a squeezing of the scale. As an example, suppose that neuroticism causes people to report extreme levels of financial well-being more frequently. Fixed effect estimation would not be able to capture such differences in response style, whereas allowing each threshold to depend on a measure of neuroticism allows the thresholds to capture this effect.

For our primary analysis, we use individual data on financial wellbeing from the BHPS that reflects household members' perceptions of the adequacy of their monetary resources. We explain this indicator in section 3 below; Pudney (2011) discusses it in relation to income and various concepts of wellbeing. We use a range of ordinal modeling techniques for panel data, including fixed effects ordered logit and random effects probit, both extended to allow

[5]For previous papers modeling subjective wellbeing using individual effects see Winkelmann and Winkelmann (1998), Ravallion and Lokshin (2001) and Charlier (2002).

[6]The only other paper on subjective wellbeing taking account of different types of heterogeneity in the response style is Beegle et al (2012), who use anchoring vignettes.

for variation across respondents in the way that their interpretation of the 5-point response scale relates to the underlying perception of their financial wellbeing. The statistical model underlying these methods is outlined in section 2, starting from a theoretical foundation of co-operative decision making over private and public good consumption within the household. Section 4 sets out our main findings, which reveal an equivalence scale significantly different than any currently in use. Our estimates suggest high returns to scale for cohabiting couples and diseconomies of scale for non-cohabiting adults. Our results also suggest that all scales significantly understate the high diseconomies of scale for children, which we argue reflect capital costs which are ignored or understated by other approaches. This suggests the need for a revision of equivalence scales. Section 5 examines the robustness of our results in several dimensions, including the assumed separability and base independence of the equivalence relation, the characteristics assumed to influence response behavior, the choice of subjective indicator and the use of a panel rather than a cross-section as the basis for the analysis. The essential character of our findings proves to be remarkably robust to these variations in analytical approach. Section 6 examines the implications for poverty measurement and shows that, while current scales may not significantly alter the overall poverty rate, rates of child poverty are understated while poverty rates for childless couples are overstated.

2 The model

In practice, household equivalence scales are largely arbitrary constructs which often have no clear basis in the theory of family consumption and production behavior. This is understandable, since the theory of family decision-making is itself uncertain, with a range of possibilities including the classical unitary model and alternative resource-sharing models based on bargaining theories. As a consequence, simplicity and the need for comparability of results have been the dominant considerations in the choice of an equivalence scale. To illustrate the concepts involved in construction of an equivalence scale, we use a simple model of co-operative decision-making in which public and private consumption coexist within the family. Consider a household containing n individuals, with household income Y and demographic characteristics $\boldsymbol{d} = (d_1, \ldots, d_n)$, and suppose that household consumption decisions are made by maximizing family utility:

$$U\left(u\left(C_1, H; d_1\right), \ldots, u\left(C_n, H; d_n\right) ; \boldsymbol{d} \right) \tag{1}$$

where $u(C_i, H; d_i)$ is the utility of the ith household member, C_i is his or her private consumption of a vector of goods purchased at prices p and H is the consumption of a within-household public good (housing, say) purchased by the household at a rental r. The household's budget constraint is:

$$\sum_{i=1}^{n} pC_i + rH = Y \tag{2}$$

where prices and income are expressed in real terms. Note that the model (1)-(2) contains as a special case the Pareto-efficient household bargaining model where $U = \sum_i \omega_i(\boldsymbol{d}) u(C_i, H; d_i)$ and the $\omega_i(\boldsymbol{d})$ are weights reflecting the bargaining positions of household members (see Bourguignon and Chiappori 1994).

Let $v_i(Y, p, r; \boldsymbol{d})$ be member i's achieved level of utility at the family optimum. If family member i were a 1-person household with income Y^*, his or her maximized utility would be given by a different indirect utility function $v_i^*(Y^*, p, r; d_i)$. We then define an equivalence index[7] $I_i(Y, p, r; \boldsymbol{d})$ to be the level of 1-person household income Y^* at which he or she would be as well off as in the multi-person household with income Y

$$v_i(Y, p, r; \boldsymbol{d}) = v_i^*(I, p, r; d_i). \tag{3}$$

In this general setting, each individual may have a different equivalence index depending on his or her characteristics d_i, and the index would also depend on prices p, r. In contrast, the poverty research literature generally uses equivalence indeces that depend on household size and composition but have three special properties: (1) uniformity across family members; (2) invariance to relative price variation; and (3) proportionality in income. In that case, the index I_i for each household member can be written:

$$I(Y, p, r; \boldsymbol{d}) = g(\boldsymbol{d})Y \tag{4}$$

where $g(\boldsymbol{d})$ is the reciprocal of the equivalence scale. This index is normalized so that it coincides with real income for a benchmark 1-adult household, so that $g(\boldsymbol{d}_0) = 1$ where $\boldsymbol{d} = \boldsymbol{d}_0$ indicates the benchmark household type.

Common equivalence scales of this form include the McClements and OECD scales, the latter specifying:

$$g(\boldsymbol{d}) = \left[1 + 0.5(n_a - 1) + 0.3n_c\right]^{-1} \tag{5}$$

[7] Essentially equivalent to the *indifference scale* of Browning et al (2006).

where n_a and n_c are the numbers of adults and children respectively.

The model can also be used as a basis for the iso-elastic scale. Suppose the household welfare function is multiplicative and individual preferences are Cobb-Douglas:

$$U = \prod_{i=1}^{n} \left\{ C_i^{\alpha} H^{1-\alpha} \right\} \tag{6}$$

This implies $v_i^* = Y^* \alpha^{\alpha} (1-\alpha)^{1-\alpha} / [p^{\alpha} r^{1-\alpha}]$ and $v_i = v_i^* n^{-\alpha}$, so that $g(\boldsymbol{d}) = n^{-\alpha}$ and the equivalence scale is proportional to family size raised to the power α.[8] If $\alpha = 1/2$, so that half the family's spending goes on the public good, then the square root formula gives the appropriate equivalence scale. A further special case is the use of per capita income based on an equivalence scale equal to family size n, which corresponds to $\alpha = 1$, where there is no public consumption good within the household and there is constant returns to scale (CRTS).

In the primary model, we adopt the following specific functional form for $g(\boldsymbol{d})$

$$g(\boldsymbol{d}) = \exp\left[\sum_{j=2}^{J} \beta_j A_j + \sum_{k=1}^{K} \gamma_k C_k \right]. \tag{7}$$

where A_j is a binary indicator for households with exactly j adults and C_k is an indicator for households with k children; J and K are the maximum numbers of adults and children respectively. The exponential function ensures that g is always positive and the restriction $\beta_1 = 0$ embodies the normalization $g(\boldsymbol{d}_0) = 1$ for a 1-adult household. This specification separates adults from children and allows for considerable flexibility in economies of scale. In section 5 we consider further extended specifications involving the ages and genders of children and ages of some adults.

The index I is a measure of the individual's access to resources, expressed in a form comparable with the benchmark case of a lone individual. If this is so, it is reasonable to assume that individuals' own perceptions of their economic circumstances are related to the same measure. We assume that an individual i from household h, sampled at time t has perceived (log) financial wellbeing generated as follows:

$$F_{iht}^* = \sum_{j=2}^{J} \beta_j A_{j,ht} + \sum_{k=1}^{K} \gamma_k C_{k,ht} + \ln Y_{ht} + \eta_i + v_{iht} \tag{8}$$

[8] This class of iso-elastic scales was first introduced as a class of arbitrary scales by Buhmann et al (1988) with the aim to use them as approximation of a large range of scales used in empirical applications.

where η_i is an unobserved individual effect and v_{iht} is a random deviation distributed independently across individuals and time periods. Since F^*_{iht} is not directly observed, its origin and scale are arbitrary and we normalize it by setting the intercept and income coefficient to 0 and 1 respectively. Notice that by adopting model (8), we are implicitly assuming that F^*_{iht} is a monotonic transformation of the perceived utility of an individual i in household h at time t and that $[F^*_{iht} - \eta_i - v_{iht}]$ is a monotonic transformation of his/her actual level of utility once we have corrected for his/her perception error η_i and the idiosyncratic error v_{iht}. Given model (8), the utility of a reference (1-person) household with income Y_0 is a monotonic transformation of Y_0, say $m(Y_0)$, while the utility of an individual living in a multi-person household with demographic characteristics $[A_2, A_3, A_4, C_1, C_2, C_3]$ and household income Y_1 is equal to $m\left(exp\left\{\sum_{j=1}^{4}\beta_j A_j + \sum_{k=1}^{3}\gamma_k C_k\right\} Y_1\right)$. The equivalence index to compare these two types of households can be derived by finding the value, $Y_0 = exp\left\{\sum_{j=1}^{4}\beta_j A_j + \sum_{k=1}^{3}\gamma_k C_k\right\} Y_1$, which equates the utilities of the two households, implying an equivalence scale $\exp\left\{\sum_{j=2}^{J}\beta_j A_j + \sum_{k=1}^{K}\gamma_k C_k\right\}^{-1}$. This scale is invariant to the value of income Y_1 as well as to the value of utility $u = m\left(\sum_{j=1}^{4}\beta_j A_j + \sum_{k=1}^{3}\gamma_k C_k + lnY_1\right)$ at which the two households are compared. Thus our equivalence index satisfies multiplicative separability and the base independence property (Dickens et al 1993).

The BHPS contains a question asking each respondent to give a subjective assessment of their financial wellbeing on a 5-point scale. The resulting observed ordinal variable is related to F^*_{iht} as follows:

$$F_{it} = \begin{cases} 1 & \text{if} \quad F^*_{iht} < \tau^1_{iht} \\ j & \text{if} \quad \tau^{j-1}_{iht} \le F^*_{iht} < \tau^j_{iht}\,, \qquad j = 2, \ldots, 4 \\ 5 & \text{if} \quad F^*_{iht} \ge \tau^4_{iht} \end{cases} \tag{9}$$

where $\tau^j_{iht} = \tau^{j-1}_{iht} + \lambda^j_0 + \lambda^j_1 X_{iht}$ and $\tau^0_{iht} \equiv 0$. The covariates X_{iht} appearing in these response thresholds include any variables which are thought to influence individuals' interpretation of the response scale specified by the question designer. Some of these variables, like gender, are time invariant, while others, such as age, change with time. Perhaps the most interesting are measures of personality traits and cognitive and non-cognitive ability, which we explore below.

The empirical model (8)-(9) is also consistent with a more general theoretical structure. Suppose the index (4) is not uniform across household members, but takes the more general

form $I(Y, p, r; \boldsymbol{d}) = g(\boldsymbol{d})Yh(d_i)$, where $\ln h(d_i)$ is expressible as a linear function of covariates X_{iht}, which would then appear as additional individual-level covariates in (8). It would be impossible to identify the effects of these variables on the equivalence index $F^{\star}_{iht}$, since they cannot be distinguished from their effects (via the coefficients λ^{j}_{1}) on the respondent's interpretation of the response scale, so the equivalence scales we estimate should be seen as capturing only the common household component of the equivalence relation.

Estimation of this model requires distributional and other assumptions. Our base results use fixed-effects ordered logit using the approach of Baetschmann et al (2011), and we explore other approaches, including maximum likelihood random-effects with generalized forms of unobserved heterogeneity.

3 Data

The BHPS is an annual longitudinal survey collecting data on socioeconomic characteristics at both individual and household levels, during the period 1991-2008. The target population of the original sample consists of all individuals aged 16 or over, resident in private households in England, Wales or Scotland (south of the Caledonian Canal). All original sample members are retained in the panel as long as possible, even when moving to new households. Those who join the household of a sample member are also included in the survey for as long as they remain in the same household as a sample member. Additional samples for Northern Ireland and the rest of Scotland were introduced later but are excluded from our analysis. In its first wave (1991), the BHPS covered more than 5,000 households and about 10,000 individuals and the sample has remained broadly representative of the population of Britain across time.

The subsample we use in our main analysis consists of all individuals aged 18 to 80 who are household heads or the partner of the head of the household in any of the waves from 1991 to 2008. The household head is defined as the owner or renter of the property and, if more than one, the oldest of them. Partners are those who report being married or in a cohabiting partnership. We further restrict the sample by removing households with unrelated individuals and other relationships such as adult siblings who are living together. We do not require the panel to be balanced, nor do we require continuous presence in the

panel. This allows for the creation and dissolution of households and it allows for individuals who play different roles in different households. Overall, our sample covers 12,372 individuals of whom 1,895 are present in all eighteen waves; 1,567 are present in only one wave and over half the sample is present for at least eight waves. There are 7,391 individuals in wave 1 and 5,318 in wave 18. Generally, sample sizes decline over time. The variables we use in our analysis are summarized in Appendix Table A1, which gives three sets of means and sample sizes, calculated for all individuals across all waves, and for waves 1 and 18.

The self-assessed financial situation is measured by asking each individual:

> *"How well would you say you yourself are managing financially these days? Would you say you are (1) Living comfortably, (2) Doing alright, (3) Just about getting by, (4) Finding it quite difficult, (5) Finding it very difficult?"*

The advantage of the BHPS for our purposes is that it is the only nationally representative survey observing financial wellbeing over a long period of time. The US Health and Retirement Survey carries a financial wellbeing measure comparable to that in the BHPS, but only for three waves and only for individuals over 50 years of age. No other panel survey in the US has a financial well being question.

Appendix Table A2 summarizes the distribution of responses to this question and their relationship with real family income. The the modal category is the second highest ("doing alright"), which accounts for 35% of responses. The lowest two categories ("finding it very/quite difficult") account for only 8% of responses overall, but there is variation with macroeconomic conditions. In wave 1 (1991, a recession period) there is a clear negative shift relative to other years, with the mode shifting to the third category ("just about getting by") and the bottom two categories accounting for over 13% of the sample. There is a tendency for self-assessments to improve over the length of the panel, as the initial population ages; however, the rise in measured income is much greater than the improvement in financial wellbeing: the overall average real pre-tax household income is £2,524 per month in 2005 prices, rising over the life of the panel from £2,069 to £3,073 (see Appendix Table A1). These average income figures are quite consistent with other measures of household income in Britain.

The average household structure is 2 adults and 0.6 children. Almost two-thirds of the observations are from households with two adults and 78% are from households headed by a married couple. The next largest category (19%) covers single adult households and almost two-thirds of observations are from childless households. Families with one or two children are each about 15% of the observations. The average age of respondents is 47; while there is some aging of the sample over the panel, the inflow of new entrants to the sample produces remarkable stability. No demographic group displays dramatic changes through the sample.

In addition to standard demographic covariates, we also use in some variants of the model measures of subjective wellbeing and personality traits, to control for differences in individual interview response behavior. We use a time-varying psychological wellbeing measure based on the 12-item version of the general health questionnaire, which generates a scale ranging from 0 to 36, with high scores indicating extreme mental strain, psychological disorder and mental illness. Validation studies have found that this psychological score is highly predictive of psychiatric conditions identified using standard clinical evaluations (Goldberg and William 1988 and Bowling 1997).

For part of the analysis we also make use of measures of the "big five" personality traits: openness to experience, conscientiousness, extroversion, agreeableness and neuroticism. These indicators were only collected in wave 15 so the sample is limited to individuals who were present (but not necessarily head of household) in that wave. We measure personality traits using the 15-item personality inventory and three questions for each of the big five personality traits (see John and Srivastava 1999). Although they are usually assessed with a longer set of questions, there is empirical evidence supporting the reliability of measures based on concise inventory (see Benet-Martinez and John 1998 and Gosling et al 2003). Respondents are asked to rate a set of claims on how they see themselves on a 7-point scale, from 1 "does not apply" to 7 "applies perfectly". We measure each personality trait as the average score of the responses to the three questions. The standardized Cronbach alpha reliability index is 0.680, 0.520, 0.577, 0.514 and 0.664 for openness, conscientiousness, extroversion, agreeableness and neuroticism respectively. These are comparable to reliability indexes cited by Heineck and Anger (2009) and Heineck (2011), but slightly worse than those found in studies using a larger set of questions. To make panel analysis possible, we treat

these personality traits as time-invariant: an assumption that is contentious (see Almlund et al 2011) but receives some support from Cobb-Clark and Schurer (2011).

Appendix Table A3 gives summary statistics broken down by gender. Women have a slightly higher percentage in the two lowest categories of financial wellbeing and slightly lower percentage in the highest category. Women's self reported financial situation is correspondingly slightly worse than men's and they have significantly lower average educational attainment than men. Differences in the wellbeing measure and personality types between men and women are small but all are statistically significant. Women have lower subjective wellbeing than men, they appear slightly more agreeable and extroverted and slightly more neurotic. Men appear to be more open to experience and less conscientious.

4 Baseline estimates

Our starting point is a set of estimates from a fixed effect ordered logit specification with time-invariant individual-specific thresholds. Formally, we assume that the idiosyncratic errors v_{it} are independent draws from a logistic distribution. We estimate this fixed effect ordered logit model using the "blow-up and cluster" (BUC) method of Baetschmann et al (2011). BUC estimation is based on maximization of a quasi-log-likelihood function constructed from log conditional likelihoods for binary fixed effect logit models, corresponding to the four possible choices for the division of the 5-point wellbeing indicator into a binary variable. Each constituent likelihood is conditioned on a sufficient statistic for the individual effect, η_i (see Chamberlain 1980), allowing estimation of the model parameters without the need for any assumptions about the distribution of η_i or its correlation with covariates. To take account of the within-individual stochastic dependence of the constituent likelihood terms, the variance of the BUC estimator is computed using cluster robust estimation.

Threshold crossing models such as the fixed effect ordered logit model are only identified up to scale, and usual practice is to normalize the variance of the error term, v_{it}. We report estimates using a more natural normalization (in view of (8)) of the coefficient of the log income covariate to unity. The scale-invariance entailed by this normalization allows direct comparison of coefficients estimated for separate subgroups of the population: in this case,

Table 1: Fixed effects ordered logit model for financial wellbeing (model (8)).

	Men		Women		Test diff, $\chi^2(1)$	$P-$value
	Coeff	SE	Coeff	SE		
Log Monthly Income per person	1.000		1.000			
Two Adults	0.552	0.096	0.425	0.068	2.161	0.142
Three Adults	0.324	0.108	0.303	0.083	0.044	0.834
Four or More Adults	0.182	0.110	0.264	0.091	0.775	0.379
One Child	-0.276	0.074	-0.231	0.061	0.490	0.484
Two Children	-0.251	0.093	-0.097	0.073	4.167	0.041
Three or More Children	0.007	0.099	0.202	0.075	6.323	0.012
$Var(v_{iht})$	1.684	0.184	1.309	0.122		
Joint test equality coefficients between gender, $\chi^2(6)$					12.445	.053

women and men. The baseline estimates are presented in Table 1.[9] Wald tests indicate that statistically significant differences between gender apply only to the coefficients on two children or three or more children.

The estimated coefficients in Table 1 are difficult to interpret directly, and we move to Table 2 where the coefficients are used to derive the adjustment factors needed to equivalize incomes between households with different demographic composition. We focus our discussion on households with one or two adults and zero, one or two children. The estimated coefficients from Table 1 are used to derive the equivalence scale:

$$g^*(\boldsymbol{d})^{-1} = n \exp\left[\sum_{j=1}^{4} \beta_j A_j + \sum_{k=1}^{3} \gamma_k C_k\right]^{-1}, \tag{10}$$

where n is the household size. We use $g^*(\boldsymbol{d})$ rather than $g(\boldsymbol{d})$ as defined in (7) because we estimate our model by using per capita income (household income divided by the household size) rather than household income, which we then use to compute the proportional increase in household income that a household with demographic characteristics $\boldsymbol{d}_1$ would need to compensate for a change in the household size to $\boldsymbol{d}_2$

$$A(\boldsymbol{d}_1, \boldsymbol{d}_2) = \frac{g^*(\boldsymbol{d}_1)}{g^*(\boldsymbol{d}_2)} - 1 \tag{11}$$

[9]The variable "4 or more adults" is constructed as zero for cases with fewer than 4 adults and as the number of adults minus 4 otherwise. The "3 or more children" variable is constructed analogously. This specification implies a constant cost for additional adults and children after the third and second respectively. This affects fewer than 1% of cases.

Table 2: Percentage increase in required income for a change in the household composition. Model (8) estimated for self-reported financial wellbeing.

					Our estimates		Test	
		CRTS	OECD	SQRT	Men	Women	gender diff	$P-$value
From	To						$\chi^2(1)$	
1-adult	2-adult	100.0	50.0	41.4	15.1	30.8	2.161	0.142
					(11.0)	(8.9)		
2-adult	3-adult	50.0	20.0	22.5	88.5	69.3	3.808	0.051
					(13.0)	(9.5)		
1-adult	1-child, 1-adult	100.0	30.0	41.4	163.6	152.1	0.508	0.476
					(19.5)	(15.5)		
2-adult	1-child, 2-adult	50.0	20.0	22.5	97.7	89.1	3.315	0.069
					(14.6)	(11.6)		
1-child, 1-adult	2-child, 1-adult	50.0	23.0	22.5	46.2	31.1	0.508	0.476
					(10.9)	(7.5)		
1-child, 2-adult	2-child, 2-adult	33.0	16.0	15.5	30.0	16.6	3.315	0.069
					(9.7)	(6.7)		
Joint test equality coefficients between gender, $\chi^2(6)$							8.193	0.224

Note : Standard errors in parenthesis.

Finally we compare these adjustment coefficients derived from our estimated ordered logit models for men and women (columns 4-5 of Table 2) with the corresponding adjustment coefficients implied by the CRTS, OECD and SQRT equivalence scales (columns 1-3).

Estimates from both the male and female samples lead to qualitatively and statistically similar family size adjustments. Given the sample sizes and demographics, one of the most interesting comparisons is a 2-adult household compared to a 1-adult household. For both the male and female samples, the estimated adjustment coefficient indicates economies of scale in marriage and cohabitation, with two adults in a couple estimated to require 15% (male) or 31% (female) more income to achieve the same level of financial wellbeing. These large within-couple economies of scale are strikingly different from those implied by conventional equivalence scales, but they are perhaps not too surprising given the conventional wisdom of "two can live as cheaply as one". Differences with the CRTS, OECD and SQRT scales are statistically significant, except for the single insignificant difference from the SQRT ratio in the female sample.

Other papers which have used self-reported wellbeing measures to derive equivalence

scales (see van Praag and van der Saar 1988, Melenberg and van Soest 1996, Ravallion 2012, and Charlier 2002) find big economies of scale when increasing the household size. But in our case this does not extend to additional adults. The estimated adjustment coefficients comparing a 3-adult to a 2-adult household are higher than the corresponding OECD, SQRT and CRTS ratios, indicating that large economies of scale only accrue within a marital/cohabiting couple. When a third adult joins a 2-adult household, our estimates suggest a perceived need for 89% (men) or 69% (women) more income. In comparison, conventional equivalence scales suggest a need for 50% (CRTS) or around 20% (OECD and SQRT) more income. We can reject statistically the hypothesis of equality between our estimated scale and the CRTS, OECD and SQRT coefficients.

There are further striking differences from conventional equivalence scales when we introduce children as household members. For a single-adult household, adding a child requires 164% (men) or 152% (women) more income to achieve the same level of financial wellbeing. Adding a first child to a 2-adult household requires 98% (men) or 89% (women) more income. The comparable ratios would be 100% and 50% respectively for the CRTS scale, 30% and 20% for the OECD scale and 41% and 22% for the SQRT scale. The differences between these ratios and our estimates are statistically significant in every case. This result suggests that adding a child is much more expensive than adding a partner to a single adult, but quite comparable with the cost of adding a third adult. We speculate that these findings relate to capital requirements, primarily housing. When two adults live as a couple, they typically share a bedroom and bathroom as well as other facilities. Housing is a major portion of the household budget, averaging around 21% of total expenditure in 2008[10]. When the household includes a child or a third adult, additional bedrooms, bathrooms and other facilities may be needed. Unlike adults, children outgrow clothing rapidly, and this too may add to the expense of a child as compared to an adult.

After the first child, addition of a second does involve some returns to scale. For a 1-adult, 1-child baseline household, a further child increase income need by 46% (men) or 31% (women). For a couple with a single child, a second child increases the required level of income by 30% (men) or 17% (women). As estimated from the female sample, the adjustment coefficients for the addition of a second child are close to, and not statistically different from,

[10]Office for National Statistics: *Family Spending* 2009 edition.

those suggested by the OECD and SQRT equivalence scales, whereas they are significantly lower than the CRTS adjustment coefficients. This is a clear sign of substantial returns to scale for a second child. The results are somewhat different for the adjustment coefficients estimated from the male sample, where there is evidence of significant economies of scale for the addition of second child to a two-adult household but not to a one-adult household.

5 Robustness of the baseline estimates

The baseline results are striking and plausible. However, the simple model on which they are based has some potentially restrictive features, which we group into five broad issues. Section 5.1 varies the degree of detail on the household's demographic structure by distinguishing age groups of children and separating adult children and other adults who are not marital/cohabiting partners from the core singleton or couple. In section 5.2, we generalize our model by allowing the thresholds to be functions of covariates and we also consider random effects ordered probit specifications. Section 5.3 relaxes the base independence assumption built into the equivalence relation (4). In section 5.4, we explore the use of an alternative subjective wellbeing measure based on satisfaction with income rather than financial wellbeing. Section 5.5 looks at the possibility of carrying out this type of analysis on a cross-section basis, using single waves of the BHPS and section 5.6 investigates alternative income definitions.

5.1 Demographic structure

Different types of adult and children of different ages may have different costs. Few equivalence scales take this into account, although the McClements scale is an exception. We use a similar approach to McClements by allowing the equivalence scale to depend on the number of children in different age categories.

$$\begin{aligned} F^*_{iht} = \sum_{j=2}^{4} \beta_j AEC_{j,ht} + \beta_{oa} O2Aht + \delta_1 C_{0-1,ht} + \delta_2 C_{2-4,ht} + \delta_3 C_{5-7,ht} + \delta_4 C_{8-10,ht} \\ + \delta_5 C_{11-12,ht} + \delta_6 C_{13-15,ht} + \delta_7 C_{16-18,ht} + \beta_{AC} AC_{18p,ht}) + \ln Y_{ht} + \eta_i + v_{iht} \end{aligned} \tag{12}$$

where $AEC_{j,ht}$ is a binary indicator for households with exactly j adults[11] for $j = 1, 2, 3$; $AEC_{4,ht}$ is constructed as the number of adults minus 3 for households with 4 or more adult members and zero otherwise; $O2A_{ht}$ is a binary indicator for the presence of a second adult in the household who is not a spouse or partner; the variable $C_{a-b,ht}$ is the number of children in the age range (a,b); and the variable $AC_{18p,ht}$ counts the number of adult children living at home. The model (12) differs from our main specification by imposing additivity of children within but not between age groups. This new specification also distinguishes three types of adults: a partner or spouse of the household head, an adult child of the household head or his/her partner or spouse, and any other adult. It also allows us to investigate the effect of the arrival of a new child. A large coefficient on C_{0-1} and smaller coefficients on other C_{a-b} could be interpreted as a transient shock from the arrival of a new child which does not reflect long term costs.

Table 3 presents our estimated scale together with those for the McClements, CRTS, OECD and SQRT scales.[12] We first consider the effect of adding different types of additional adult to a single-adult baseline household. If the two adults constitute a couple, we see still larger returns to scale than in the baseline model of Table 5. Indeed, there is no statistically significant evidence of any increased income need for a couple compared to a single adult. A remarkably different story emerges when the additional adult is not a spouse or partner. It is not possible to draw reliable inferences about the case of an unrelated adult introduced into a single-adult household, because the paucity of data leads to very wide confidence intervals. However, the addition of an adult child living at home with a single parent or a parental couple gives a result that is qualitatively similar to the case of the third adult in Table 2: a cost increase of 120-140% (single parent) or 65-75% (parental couple), compared with the 70-90% estimated cost increase in Table 5 for inclusion of a third adult in a 2-adult household. Clearly, large economies of scale for additional adults only derive from cohabitation, and this finding represents a large and statistically significant difference from the CRTS, OECD, SQRT and McClements scales.

An interesting U-shaped pattern emerges from the age-specific costs in Table 3, with the lowest additional costs associated with 8-12 year-olds, for whom CRTS is approximately satisfied. Infants and toddlers, aged zero to four have the highest diseconomies of scale,

[11] Henceforth, we define the term "adult" to refer to all over-18s, but excluding adult children.

[12] Full estimates of the model parameters are available from the authors on request.

Table 3: Percentage increase in required income for a change in the household composition. Model (12) estimated for the self-reported financial wellbeing.

	CRTS	OECD	SQRT	McClements	Our estimates for Men	Our estimates for Women
Adding to an 1-adult household the following additional member						
a partner or spouse	100.0	50.0	41.4	64.0	7.5	-5.4
					(11.4)	(8.0)
an adult (no child, no spouse)	100.0	50.0	41.4	75.0	84.1	-0.7
					(247.3)	(103.1)
a child 0-1	100.0	30.0	41.4	15.0	166.1	171.7
					(19.1)	(17.5)
a child 2-4	100.0	30.0	41.4	30.0	154.5	148.6
					(16.2)	(13.3)
a child 5-7	100.0	30.0	41.4	34.0	106.8	94.8
					(12.0)	(9.2)
a child 8-10	100.0	30.0	41.4	38.0	93.5	83.0
					(10.7)	(8.6)
a child 11-12	100.0	30.0	41.4	41.0	93.4	84.1
					(12.3)	(9.6)
a child 13-15	100.0	30.0	41.4	26.8	121.1	98.8
					(12.7)	(9.6)
a child 16-18	100.0	50.0	41.4	62.0	138.8	125.0
					(14.7)	(11.1)
a child 18 or more	100.0	50.0	41.4	69.0	133.6	124.6
					(13.9)	(11.0)
Adding to a 2-adult household the following additional member						
a child 0-1	50.0	20.0	22.5	9.1	99.5	103.7
					(14.3)	(13.1)
a child 2-4	50.0	20.0	22.5	18.3	90.9	86.4
					(12.2)	(10.0)
a child 5-7	50.0	20.0	22.5	20.7	55.1	46.1
					(9.0)	(6.9)
a child 8-10	50.0	20.0	22.5	23.2	45.2	37.2
					(8.0)	(6.5)
a child 11-12	50.0	20.0	22.5	25.0	45.1	38.1
					(9.2)	(7.2)
a child 13-15	50.0	20.0	22.5	26.8	65.8	49.1
					(9.5)	(7.2)
a child 16-18	50.0	33.3	22.5	37.8	79.1	68.8
					(11.0)	(8.3)
a child 18 or more	50.0	33.3	22.5	42.1	75.2	68.4
					(10.4)	(8.3)

Note : Standard errors in parenthesis.

quite comparable to the results in Table 2 and much higher than those implied by the CRTS, OECD, SQRT and McClements scales. Our estimated age pattern is markedly different from the CRTS, OECD and SQRT scales which are flat with respect to age, and the McClements scale which assumes diseconomies of scale increasing monotonically with the age of the child. We interpret the economically and statistically significant drop between pre-school and school-age children as the consequence of child care costs and the significant capital expenditures often associated with a new child. The rising scale diseconomies in the teenage years may also represent a capital story: couples can share bedrooms, while others in the house often do not. This is a feature that conventional equivalence scales fail to capture.

5.2 Econometric specification

We investigate two aspects of econometric specification. The first is to use alternative random effects logit and probit models to assess the impact of assumptions about the distribution of v_{iht} and the covariation of η_i with the covariates. The second is to allow the threshold parameters in (9) to vary with individual level characteristics. Four specifications of the covariates are used: first, an indicator for existence of a child under age one and the GHQ measure of well being; the second adds education and age dummies; the third adds the "big five" measures of personality (which cannot be included in the fixed effect model because they are only measured once and do not vary over the panel); the fourth also includes averages of all time-varying covariates, as an approach to controlling for fixed effects (see Chamberlain 1980). The estimated scale parameters appear in Tables 4 and 5 for the male and female samples respectively.[13] We focus on the scale for couples compared to singles, adding a third adult to a couple and adding a child to a couple and a single parent. Other results are available upon request and are qualitatively similar.

Estimates from the baseline model are reproduced in the first row of each panel of the tables for comparison. The following rows show the effects of introducing progressively more covariates into the thresholds. Perhaps the most striking aspect of these generalized results are the qualitative stability of the estimates. The fixed effects and random effects differences are statistically quite small and qualitatively irrelevant. The addition of covariates causes some increase in the scale estimates for couples, implying that our baseline estimates may

[13]Full parameter estimates are available from the authors on request.

Table 4: Percentage increase in required income for a change in the household composition. Different model specifications for men.

	From 1-adult to 2-adult	From 2-adult to 3-adult	From 1-adult to 1-child 1-adult	From 2-adult to 1-child 2-adult
Fixed Effect Ordered Logit				
Primary Model	15.1	88.5	163.6	97.7
	(11.0)	(13.0)	(19.5)	(14.6)
Linear Threshold 0	23.1	82.1	156.3	92.2
	(12.5)	(12.7)	(20.7)	(15.5)
Linear Threshold 1	20.1	79.6	142.5	81.8
	(13.0)	(13.5)	(21.0)	(15.8)
Random Effect Ordered Logit				
Basic Specification	16	106.1	166.8	100.1
	(5.3)	(7.7)	(10.3)	(7.7)
Linear Threshold 0	28.6	95.2	155.5	91.6
	(6.4)	(7.5)	(11.0)	(8.2)
Linear Threshold 1	35.7	85.7	139.8	79.8
	(6.9)	(8.0)	(11.5)	(8.6)
Linear Threshold 2	32	84.3	158.9	94.2
	(8.1)	(9.2)	(13.6)	(10.2)
Chamberlain Approach	30.6	76.2	131.3	73.5
	(9.0)	(9.0)	(13.1)	(9.9)
Random Effect Ordered Probit				
Basic Specification	14.3	101.4	175.3	106.5
	(5.6)	(8.2)	(11.7)	(8.8)
Linear Threshold 0	25	92.4	169.9	102.5
	(6.4)	(8.2)	(12.7)	(9.5)
Linear Threshold 1	24.3	89.4	145.2	83.9
	(6.7)	(8.7)	(12.7)	(9.5)
Linear Threshold 2	24	86.2	161.4	96
	(8.1)	(10.1)	(15.9)	(11.9)
Chamberlain Approach	28.2	76.7	136.8	77.6
	(9.4)	(9.8)	(14.6)	(10.9)

Note : Standard errors in parenthesis. The basic and primary specifications do not consider variables in the thresholds; while Linear Threshold 0 considers GHQ and an indicator for a first child aged between 0 and 1, Linear Threshold 1 includes also education and age, Linear Threshold 2 adds the big five personality traits too. Finally, the Chamberlain Approach, beside considering all the above variables in the thresholds, includes the average across wave of each explanatory variable in the main index.

Table 5: Percentage increase in required income for a change in the household composition. Different model specifications for women.

	From 1-adult to 2-adult	From 2-adult to 3-adult	From 1-adult to 1-child 1-adult	From 2-adult to 1-child 2-adult
Fixed Effect Ordered Logit				
Primary Model	30.8	69.3	152.1	89.1
	(8.9)	(9.5)	(15.5)	(11.6)
Linear Threshold 0	33.6	68	145.1	83.8
	(9.2)	(9.4)	(15.8)	(11.8)
Linear Threshold 1	24.2	65.5	136.1	77
	(9.1)	(9.8)	(15.9)	(11.9)
Random Effect Ordered Logit				
Basic Specification	26.1	78.4	170.3	102.7
	(4.2)	(6.0)	(8.9)	(6.7)
Linear Threshold 0	29.7	73.9	161.6	96.2
	(4.4)	(5.8)	(9.5)	(7.1)
Linear Threshold 1	18.6	67.7	139.1	79.4
	(4.4)	(5.9)	(9.3)	(7.0)
Linear Threshold 2	31.5	66.2	143.5	86.2
	(6.1)	(6.8)	(10.3)	(7.7)
Chamberlain Approach	42.5	61.2	114.2	60.7
	(7.0)	(7.1)	(10.2)	(7.6)
Random Effect Ordered Probit				
Basic Specification	21.2	80.4	176.4	107.3
	(4.4)	(6.6)	(10.0)	(7.5)
Linear Threshold 0	24.6	77.5	162.8	97.1
	(4.6)	(6.4)	(10.3)	(7.7)
Linear Threshold 1	13.8	69.6	138	78.5
	(4.5)	(6.5)	(10.0)	(7.5)
Linear Threshold 2	26.6	68.9	133.5	75.1
	(6.2)	(7.7)	(11.4)	(8.6)
Chamberlain Approach	0.7	63.2	220	140
	(8.1)	(16.3)	(27.2)	(20.4)

Note : Standard errors in parenthesis. The basic and primary specifications do not consider variables in the thresholds; while Linear Threshold 0 considers GHQ and an indicator for a first child aged between 0 and 1, Linear Threshold 1 includes also education and age, Linear Threshold 2 adds the big five personality traits too. Finally, the Chamberlain Approach, beside considering all the above variables in the thresholds, includes the average across wave of each explanatory variable in the main index.

slightly overstate the returns to scale within couples, but it is important to note they remain uniformly lower than even the OECD scale parameters, and in all but one case (random effects with the Chamberlain approach) the difference is statistically significant. Hence our finding that conventional equivalence scales greatly understate the scale economies household formation appears robust to econometric specification.

Adding a child to a single-adult household or adding an additional adult or child to a couple again shows the diseconomies of scale found using the baseline specification. The results are quite robust across ordered logit and ordered probit specifications. Including covariates in the specification lowers the estimated adjustment coefficient in most cases, although not uniformly. The estimate of larger than constant returns to scale continues to be supported by all model specifications.

5.3 Base independence and separability assumptions

The equivalence scale implied by our model imposes multiplicative separability and base independence assumptions. Given that utility is monotonically increasing in household income, if one of these properties fails, then so does the other. Consequently, we can test both properties by re-estimating the fixed effects ordered logit model allowing the coefficients to differ across thresholds. If the equivalence scales change significantly across thresholds (and thus across different levels of utility and income), we infer that the two properties can be rejected. In effect, we estimate four fixed-effects logit models, one for each adjacent pair of categories in the dependent variable, imposing a common fixed effect across all four models. The scale estimates produced by this specification are presented in Table 6. The baseline model is reproduced in the first column of the table, the next four columns showing the adjustment coefficients based on successive thresholds separating the five response categories for the measure of financial well being. The lowest threshold separates "finding it very difficult" from "finding it quite difficult", while the highest separates "doing alright" from "living comfortably". The estimated diseconomies of scale are least pronounced for women with a partner and one child at the first threshold while, for men at the first threshold comparing two adults to three, we find economies of scale even higher than those implied by the OECD scale. In neither of these cases can CRTS be rejected statistically, because of the large standard errors. We can accept the null hypothesis of uniformity of adjustment coefficients across

Table 6: Percentage increase in required income for a change in the household composition. Fixed effect ordered model allowing the equivalence scale to change at different level of financial wellbeing.

From	To	Primary	Financial Wellbeing Level			
		Model	First level	Second level	Third level	Fourth level
		Men estimates				
1-adult	2-adult	15.1	24.9	33.1	12.7	9.4
		(11.0)	(29.2)	(19.7)	(12.7)	(14.3)
2-adult	3-adult	88.5	24.5	53.2	110.3	98.2
		(13.0)	(25.0)	(19.1)	(17.1)	(19.6)
1-adult	1-child and 1-adult	163.6	113.6	115.4	155.3	225.9
		(19.5)	(44.9)	(26.4)	(22.3)	(38.5)
2-adult	1-child and 2-adult	97.7	60.2	61.5	91.5	144.4
		(14.6)	(33.6)	(19.8)	(16.7)	(28.9)
		Women estimates				
1-adult	2-adult	30.8	10.4	23.8	44.1	23.1
		(8.9)	(16.4)	(13.4)	(10.7)	(12.2)
2-adult	3-adult	69.3	58.2	78.2	76.9	59.3
		(9.5)	(24.7)	(19.7)	(11.5)	(12.0)
1-adult	1-child and 1-adult	152.1	105.9	144.6	144.5	176.7
		(15.5)	(33.1)	(27.5)	(16.7)	(25.3)
2-adult	1-child and 2-adult	89.1	54.4	83.4	83.4	107.6
		(11.6)	(24.8)	(20.6)	(12.5)	(18.9)

Note : Standard errors in parenthesis.

thresholds in all cases except the 2-adult to 3-adult comparison in the male sample. Thus, although there is evidence of a little difference across the thresholds the separability and the base independence assumptions do not appear to be unreasonably strong. In qualitative terms, our conclusions continue to be supported: large economies of scale for couples, but diseconomies of scale for other adults and children.

5.4 An alternative subjective measure: satisfaction with income

We would argue that our choice of the BHPS financial wellbeing variable as the family welfare measure is matched more directly to the requirements of an equivalence scale than are other measures like satisfaction with income or life in general. However, the BHPS also asks about the respondent's satisfaction with household income, which we would expect to be more coherent across different respondents within the household than the commonly

used life satisfaction measure. The income satisfaction question is only asked after wave six of the BHPS, so sample sizes are smaller and standard errors consequently larger. Income satisfaction is measured on a 1-7 numerical scale where 1 represents "completely dissatisfied" and 7 represents "completely satisfied"; other scale points bear no textual labels.[14]

Table 7 compares our baseline estimates with the adjustment coefficients derived from the same fixed effect ordered logit model estimated from income satisfaction data. There are two striking features of the comparison. First, there is much closer agreement of the results derived from the samples of male and female respondents for financial wellbeing than for income satisfaction. The financial wellbeing measure gives a much more coherent picture of family wellbeing and is clearly preferable for our purposes. Second, the income satisfaction and financial wellbeing results are qualitatively similar in terms of their strong contrast with conventional equivalence scales. There are only two exceptions, both involving male respondents' reaction to an additional child, where there is a lower adjustment coefficient than CRTS for the income satisfaction measure but higher for financial wellbeing. There are no qualitative contradictions with respect to the comparison with OECD and SQRT scales. Consequently, although we have strong grounds for preferring the financial wellbeing measure, the general character of our conclusions is robust to the choice of a subjective welfare indicator.

Table 7: Percentage increase in required income for a change in the household composition. Model (8) estimated for satisfaction with household income.

					Our estimates for	
From	To	CRTS	OECD	SQRT	Men	Women
1-adult	2-adult	100.0	50.0	41.4	49.1	28.1
					(20.5)	(12.7)
2-adult	3-adult	50.0	20.0	22.5	130.7	81.0
					(22.2)	(14.2)
1-adult	1-child and 1-adult	100.0	30.0	41.4	86.3	127.8
					(18.0)	(19.7)
2-adult	1-child and 2-adult	50.0	20.0	22.5	39.8	70.8
					(13.5)	(14.8)

Note : Standard errors in parenthesis.

[14]See Pudney (2011) for a discussion of the different characteristics of the financial wellbeing and income satisfaction variables.

5.5 Cross-section estimation

There are few longitudinal surveys that contain direct measures of financial wellbeing, but a much wider range of subjective variables is available in cross-section surveys. Despite their many drawbacks, cross-section surveys also have the advantage of avoiding the response problems caused by the cumulative effect of sample attrition over time. To assess the feasibility of our approach for cross-section data, we estimate ordered logit models using waves 1, 10 and 18, which are presented in Table 8. The standard errors are somewhat larger than those in Table 2. We first note that the two main findings: high returns to scale for couples and high diseconomies of scale for children are still very clear, and even exaggerated here. While we prefer to use the panel because of the more robust estimation techniques available, it is possible to obtain estimates using only a cross section. Together with the finding of robustness to choice of subjective measure, this implies that our approach can be implemented in many settings. Second, this could allow further investigation of the way these demographic adjustment coefficients change over time. This in turn may be important in understanding the evolution of poverty.

Table 8: Percentage increase in required income for a change in the household composition. Model (8) estimated for self-reported financial wellbeing separately across waves.

		Our estimates for men			Our estimates for women		
		wave 1	wave 10	wave 18	wave 1	wave 10	wave 18
1-adult	2-adult	14.4	8.0	2.4	31.0	24.6	26.6
		(9.6)	(12.6)	(13.9)	(8.5)	(11.3)	(13.4)
2-adult	3-adult	73.1	140.4	135.3	62.9	154.1	101.6
		(14.0)	(28.6)	(32.5)	(12.2)	(27.2)	(24.9)
1-adult	1-child and 1-adult	145.1	187.7	182.3	193.6	198.0	206.3
		(19.8)	(32.3)	(37.3)	(21.4)	(28.9)	(34.5)
2-adult	1-child and 2-adult	83.9	115.7	111.7	120.2	123.5	129.7
		(14.8)	(24.2)	(28.0)	(16.1)	(21.7)	(25.9)

Note : Standard errors in parenthesis.

5.6 Income definition

The results presented so far use a gross (pre-tax) definition of household income. This income variable is available for all 18 waves of the BHPS, but a constructed estimate of post-tax

net income is also available for a subset of households for waves 1-16. Table 9 compares the results of an analysis of the restricted sample using net income with our baseline analysis. The net income analysis produces slightly smaller adjustment coefficients, but qualitatively, the results are again similar, with large significant differences from the conventional equivalence scales.

Table 9: Percentage increase in required income for a change in the household composition. Model (8) estimated using gross and net monthly income.

		Net income		Gross income	
From	To	Men	Women	Men	Women
1-adult	2-adult	21.1	31.3	15.1	30.8
		(9.6)	(7.7)	(11.0)	(8.9)
2-adult	3-adult	69.6	55.7	88.5	69.3
		(10.0)	(7.9)	(13.0)	(9.5)
1-adult	1-child and 1-adult	138.6	133.3	163.6	152.1
		(14.7)	(12.3)	(19.5	(15.5)
2-adult	1-child and 2-adult	79.0	75.0	97.7	89.1
		(11.0)	(9.2)	(14.6)	(11.6)
1-child and 1-adult	2-child and 1-adult	38.5	25.1	46.2	31.1
		(8.8)	(6.3)	(10.9)	(7.5)
1-child and 2-adult	2-child and 2-adult	23.1	11.2	30.0	16.6
		(7.8)	(5.6)	(9.7)	(6.7)

Note : Standard errors in parenthesis.

6 Implications for poverty measurement

The most important use of equivalence scales is in poverty analysis. Given our results, we would expect poverty rates calculated for incomes equivalized with our new scale to indicate higher poverty among families with children, but lower poverty among couples. We investigate this using illustrative BHPS wave 18 gross income data relating to 2008. To estimate a poverty measure, we calculate the adjustment coefficient (11) for each household in the sample and adjust household income accordingly. We set the poverty threshold at 60% of the overall median of this equivalized income, weighting the household observations by household size, to give a poverty headcount in individual, rather than household terms. Poverty rates for children are constructed in the same way, weighting the headcount by the

number of children rather than household size.[15] Poverty rates are estimated using alternative parameter values estimated from the male and female samples to give an indication of robustness (note that calculation is made using the full sample of individuals in each case, irrespective of the sample used for parameter estimation).

The results are shown in Table 10, together with results calculated using the conventional CRTS and OECD scales. The overall poverty rate in our sample is 21.0% or 21.8%, calculated using the latter scales. Our new scale yields a similar overall poverty rate of 21.7% (male) or 20.6% (female). Part of these differences are due to the difference in median equivalized income: the OECD median is the highest and the CRTS the lowest. Dispersion also differs across methods of equivalization: the CRTS scale gives least dispersion, while the remaining three measures are broadly comparable.

Table 10: Poverty measures using different equivalence scales.

	CRTS	OECD	Estimated Scale Male Sample	Estimated Scale Female sample
Mean Scaled Income	1251	1892	1628	1637
Median Scaled Income	1011	1602	1270	1295
Poverty Threshold	607	961	762	777
Standard Deviation of Scaled Income	1004	1369	1336	1304
IQR of Scaled Income	880	1337	1169	1173
Overall Rate	21.0	21.8	21.7	20.6
Children	35.0	26.1	36.6	33.5
Couples only	12.8	20.9	7.6	8.6

The proportion of children in poverty is an important target for government policy, and the two estimates of our new measure are dramatically higher than the OECD-based rate of 26.1% for child poverty. We estimate rates of 33.5% and 36.6% using parameter estimates based respectively on female and male subjective assessments; these bracket the CRTS rate of 35.0%. We also report estimated poverty rates for adults in households formed of married or cohabiting couples only. Our measures give the lowest poverty rates: 7.6% and 8.6% for estimates based on the male and female estimation samples respectively, which are less than half the 20.9% poverty rate calculated using the OECD scale. Our approach therefore gives

[15] But we keep the poverty threshold fixed at 60% of the household size-weighted median.

a quite different picture of the pattern of UK poverty and suggest the need for greater focus on child poverty.

In conclusion, our results provide evidence that income poverty measures may be affected by the choice of the equivalence scale as found in previous papers (e.g. Buhamnn 1998, Coulter et al 1992, and Jenkins and Cowell 1994) and this is especially evident when considering subgroups of the population such as children and when comparing equivalence scales derived by self-reported financial wellbeing with other commonly used equivalence scales.

7 Conclusions

There is a vast empirical research literature on patterns of poverty and income inequality, which use arbitrary conventional scales to compare households with different demographic structures. There is a lot to be said for agreement on a common standard for income analyses, but it is important to remember that the standard which researchers adopt may have a big influence on the nature of the conclusions that are generally reached.

Our findings give grounds for serious concern about conventional equivalence scales like the OECD and square root of family size scales. We have used an easily-implemented survey measure of financial wellbeing to analyze the relationship between income, demographic structure and family welfare, in a panel data setting. Unlike the more widely-used income satisfaction question, this financial wellbeing question elicits the same coherent picture of family circumstances from male and female respondents and generates remarkably robust conclusions, which contrast sharply with the assumptions built into conventional equivalence scales. We find very strong economies of scale within marital/cohabiting couples: "two can live [nearly] as cheaply as one", and we find also that an additional child or adult brings strong diseconomies of scale. We attribute this to the additional capital requirements that the introduction of non-core family members bring.

References

[1] Almlund, M., Duckworth, A.L., Heckman, J.J., Kautz, T. (2011). Personality psychology and economics, in Eric A. Hanushek, Stephen J. Machin, Ludger Woessmann. (Eds) *Handbook of the Economics of Education*, Volume 4, Chapter 1.

[2] Baetschmann, G., Staub, K. E. and Winkelmann, R. (2011). Consistent estimation of the fixed effects ordered logit model. Bonn: IZA Discussion Paper no. 5443.

[3] Bargain, O. and Donni, O. (2009). The measurement of child costs: a Rothbarth-type method consistent with scale economies. Bonn: IZA Discussion Paper no. 4654.

[4] Beegle, K., Himelein, K. and Ravallion, M. (2012). Frame-of-reference bias in subjective welfare, *Journal of Economic Behavior Organization*, Elsevier, 81, 2, pages 556-570.

[5] Benet-Martinez, V. and John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology* **75**, 729-750.

[6] Berthoud, R. M. Bryan and E. Bardasi (2004). The dynamics of deprivation: the relationship between income and material deprivation over time. Leeds: Department for Work and Pensions, Research Report no. 219.

[7] Blank, R. (1997). *It Takes a Nation: A New Agenda for Fighting Poverty*, Princeton, NJ: Princeton University Press.

[8] Blundell, R. and Lewbel, A. (1991). The information content of equivalence scales. *Journal of Econometrics*, **50** , 49-68.

[9] Bourguignon, F. and Chiappori, P.-A. (1994). The collective approach to household behavior, in *The Measurement of Household Behavior*, Blundell, R., Preston, I. and Walker, I. (eds.), 70-85, Cambridge: Cambridge University Press

[10] Bowling, A. (1997). Measuring health. A review of quality of life measurement scales. Buckingham: Open University Press, 2nd edition.

[11] Browning, M., Chiappori, P.-A. and Lewbel, A. (2006). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. University of Oxford: Department of Economics Discussion Paper no. 289.

[12] Buhmann, B., Rainwater, L., Schumauss, G. and Smeeding, L. (1988). Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database, *Review of Income and Wealth*, 34, 115-142.

[13] Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *Review of Economic Studies*, 47, 225-238.

[14] Charlier, E. (2002). Equivalence scales in an intertemporal setting with an application to the former West Germany, *Review of Income and Wealth* **48**, 99-126.

[15] Citro, C. and R. Michael (1995). *Measuring Poverty: A New Approach*, Washington, D.C.: National Academy Press.

[16] Cobb-Clark, D. and Schurer, S. (2011). The stability of big-five personality traits. Bonn: IZA Discussion Paper no. 5943.

[17] Coulter F.A.E., Cowell, F.A., Jenkins S.P. (1992). Differences in needs and assessment of income distributions. *Bulletin of Economic Research*, 44, 2, 77-124.

[18] Deaton, A. and Muellbauer, J. (1980), Economics and Consumer Behavior, Cambridge: Cambridge University Press.

[19] Department for Work and Pensions (2011). *Households Below Average Income*. London: The Stationery Office.

[20] Dickens, R., Fry, V., and Pashardes, P. (1993). Non-linearities and equivalence scales, *T*he Economic Journal, **103**, 359-368.

[21] Diener, E. E.M. Suh, R.E. Lucas and H.L. Smith, (1999). Subjective well-being: three decades of progress. *Psychological Bulletin* **125**, 276-302.

[22] DWP (2011). *Households Below Average Income. An Analysis of the Income Distribution 1994/95-2009/2010*. London: Department for Work and Pensions.

[23] Eurostat (1998). *Recommendations on Social Exclusion and Poverty Statistics*. Luxembourg: 31st Meeting of the Statistical Programme Committee, Income and Living Conditions.

[24] Fisher, G. M. (1997). The development and history of the u.s. poverty thresholds - a brief overview, Newsletter of the Government Statistics Section and the Social Statistics Section of the American Statistical Association, Winter 1997.

[25] Flik, R.J., and van Praag, B.M.S. (1991). Subjective poverty line defintions. *De Economist*, 139, 3, 311-330.

[26] Goedhart, Th., Halberstad, V., Kapteyn, A., van Praag B.M.S. (1977). The poverty line: Concepts and measurement. *The Journal of Human Resources*, 12, 503-520.

[27] Goldberg, D., and Williams, P. (1988). *A User's Guide to the General Health Questionnaire*. Windsor: NFER-NELSON.

[28] Gosling, S. D., Rentfrow, P. J., and Swann, W. B. (2003). A very brief measure of the Big-Five personality domains, *Journal of Research in Personality*, **37**, 504-528.

[29] Hagenaars, A., de Vos, K. and Zaidi M. A. (1994). Poverty statistics in the late 1980s: research based on micro-data, Luxembourg: Office for Official Publications of the European Communities.

[30] Heineck, G. (2011). Does it pay to be nice? Personality and earnings in the UK. *Industrial & Labor Relations Review* **64**, 5.

[31] Heineck, G. and Anger, S. (2009). The Returns to Cognitive Abilities and Personality Traits in Germany, *Labour Economics*, **17**, 3, 535-46.

[32] Jenkins, S. P. and Cowell, F. A. (1994). Parametric equivalence scales and scale relativities. *Economic Journal*, 104, 891-900.

[33] John, O. P., and Srivastava, S. (1999). *The Big Five trait taxonomy: History, measurement, and theoretical perspectives.* In O. P. John & L. A. Pervin (Eds.), Handbook of Personality: Theory and Research. New York: Guilford Press.

[34] Kahneman, D. and Krueger, A. B. (2006). Developments in the measurement of subjective well-being, *Journal of Economic Perspectives* **20**, 3-24.

[35] Kahneman, D. and Deaton, A. S. (2010). High income improves evaluation of life but not emotional well-being, *Proceedings of the National Academy of Sciences* **107**, 16489-16493.

[36] Kapteyn A., Kooreman, P., and Willemse, R. (1988). Some methodological issues in the implementation of subjective poverty definitions, *Journal of Human Resources*, 23, 2, 222-242.

[37] Kapteyn, A., van de Geer, S. and van de Stadt, H. (1984). The impact of changes in income and family composition on subjective measures of well-being, in *Horizontal Equity, Uncertainty and Economic Well-Being*, M. David and T. Smeeding, eds. Chicago, IL: University of Chicago Press.

[38] Kapteyn. A., and van Praag, B. (1976). A new approach to the construction of family equivalence scales. *European Economic Review*, 7, 313-335.

[39] Kapteyn. A.,Kooreman, P., and Willemse, R. (1988). Some methodological issues in the implementation of subjective poverty definitions. *The Journal of Human Resources*, 23, 2, 222-242.

[40] Lewbel, A. (1997). Consumer demand systems and household equivalence scales, in *Handbook of Applied Econometrics*, Volume II: Microeconomics, M. H. Pesaran and P. Schmidt (eds.), Oxford: Blackwell Publishers Ltd.

[41] Melenberg, B., and van Soest, A. H. 0. (1996). Measuring the costs of children: parametric and semi parametric estimators. *Statistica Neerlandica*, 50, 1, 171-192.

[42] McClements, L. D. (1977). Equivalence scales for children, *Journal of Public Economics* **8**, 191-210.

[43] Muellbauer, J. (1979). McClements on equivalence scales for children. *Journal of Public Economics* **12**, 221-231.

[44] OECD (2009) What are equivalence scales? Paris: OECD Project on Income Distribution and Poverty (http://www.oecd.org/dataoecd/61/52/35411111.pdf)

[45] Orshanskey, M. (1963). Children of the Poor. Social Security Bulletin 26(7): 3-13.

[46] Pradhan, M. and Ravallion, M. (2000). Measuring Poverty Using Qualitative Perceptions of Consumption Adequacy. *Review of Economics and Statistics* **82**, 462-71.

[47] Pudney, S. E. (2011). Perception and retrospection: The dynamic consistency of responses to survey questions on wellbeing. *Journal of Public Economics* **95**, 300-310.

[48] Prais, S. J. and Houthakker, D H. S. (1955). The Analysis of Family Budgets. Cambridge: Cambridge University Press.

[49] Ravallion, M., and Lokshin, M. (2001). Identifying welfare effects from subjective questions. *Economica* 68, 271, 335357.

[50] Ravallion, M. (2012). Poor, or Just Feeling Poor? On Using Subjective Data in Measuring Poverty. *The World Bank*, Policy Research Working Paper 5968.

[51] Rothbarth, E. (1943). Note on a method of determining equivalent income for families of different composition, in *War-time Pattern of Saving and Spending*, C. Madge (ed.), Appendix 4, 123-130. Cambridge: Cambridge University Press.

[52] Ruggles, P. (1990). Drawing the Line: Alternative Poverty Measures and Their Implications for Public Policy. Washing, D.C.: Urban Institute.

[53] Singh, B. and Nagar, A.L. (1973). Determination of consumer unit scales, *Econometrica*, **41**, 347-35s.

[54] Tummers, M. (1994). The Effect of Systematic Misperception of Income on the Subjective Poverty Line, in Richard Blundell, Ian Preston and Ian Walker (eds) *The Measurement of Household Welfare*, Cambridge, England: Cambridge University Press.

[55] van Praag, B.M.S. and van der Sar, N.L. (1988). Empirical uses of subjective measures of well-being. Household cost gunctions and equivalence scales. *The Journal of Human Resources*, 23, 2, 193-210.

[56] Winkelmann, L. and R. Winkelmann (1998). Why are the unemployed so unhappy? Evidence from panel data, *Economica*, 65, 1-15.

[57] Ziliak, J. P.(2006). Understanding poverty rates and gaps: concepts, trends and challenges. University of Kentucky: UKCP Discussion Paper Series #2006-06.

Appendix: Additional tables

Table A1 Sample means and sample numbers

	Full Panel Mean	Wave 1 Mean	Wave 18 Mean
Self Reported Financial Situation			
Finding it very difficult	0.0245	0.0467	0.0231
Finding it quite difficult	0.0567	0.0874	0.0555
Just about getting by	0.2602	0.3203	0.2625
Doing Alright	0.3532	0.2703	0.3793
Living Comfortably	0.3045	0.2744	0.2796
Main Index Variables			
Real Monthly Income	£2,524	£2,069	£3,073
Real Monthly Income/person	£1,033	£839	£1,254
Log-real Monthly Income	7.565	7.378	7.777
Log-real Monthly Income/person	6.697	6.505	6.908
Household Size	2.673	2.692	2.665
Adults	2.038	2.036	2.063
Children	0.635	0.656	0.602
Threshold Variables			
Female	0.538	0.54	0.538
Age	47.1	46.8	48.9
Age 18-30	0.165	0.184	0.127
Age 31-50	0.44	0.432	0.434
Age 51-65	0.232	0.213	0.266
Age 65-80	0.164	0.172	0.172
Education Level 1	0.23	0.35	0.14
Education Level 2	0.091	0.11	0.066
Education Level 3	0.178	0.192	0.145
Education Level 4	0.375	0.271	0.472
Education Level 5	0.126	0.077	0.178
GHQ-12 wellbeing	11.2	10.9	11.4
Agreeable	5.41	5.419	5.41
Extrovert	4.44	4.373	4.486
Neurotic	3.61	3.573	3.654
Openness to experience	4.45	4.357	4.509
Conscientiousness	5.33	5.307	5.349
No. observations	112,489	7,391	5,318

Table A2: Relationship between Self Reported Financial Situation and Household Income

	Mean	Median	S.D.	No. obs.
Self Reported Financial Situation		Income		
Finding it very difficult	1205.482	890.1219	1032.866	2989
Finding it quite difficult	1577.866	1255.183	1300.606	6860
Just about getting by	1803.813	1504.683	1384.736	31176
Doing Alright	2464.064	2179.183	1704.517	41760
Living Comfortably	3018.987	2520.705	2413.046	35692
Total	2374.435	1984.15	1921.575	118477

Table A3: Gender differences

	Men	Women
Self Reported Financial Situation		
Finding it very difficult	0.0217	0.0269
Finding it quite difficult	0.0527	0.0602
Just about getting by	0.2608	0.2596
Doing Alright	0.3508	0.3554
Living Comfortably	0.3141	0.2979
Main Index Variables		
Real Monthly Income	2,649	2,417
Real Monthly Income/person	1,086	989
Log-real Monthly Income	7.634	7.506
Log-real Monthly Income/person	6.745	6.656
Household Size	2.719	2.634
Adults	2.09	1.994
Children	0.629	0.64
Threshold Variables		
Age	47.3	47
Age 18-30	0.148	0.179
Age 31-50	0.456	0.427
Age 51-65	0.241	0.224
Age 65-80	0.155	0.171
Education Level 1	0.201	0.225
Education Level 2	0.083	0.098
Education Level 3	0.155	0.198
Education Level 4	0.419	0.336
Education Level 5	0.141	0.113
Wellbeing	10.5	11.8
Agreeable	5.209	5.575
Extrovert	4.309	4.558
Neurotic	3.246	3.925
Open to experience	4.528	4.375
Conscientiousness	5.287	5.36
Sample size	51,984	60,505
Sample sizes for personality traits	37,464	44,327