ISER Working Paper Series

Do household surveys give a coherent view of disability benefit targeting?

A multi-survey latent variable analysis for the older population in Great Britain

Ruth Hancock
Marcello Morciano
Health Economics Group
University of East Anglia

Stephen Pudney
Institute for Social and Economic Research
University of Essex

Francesca Zantomio
Department of Economics
Ca' Foscari University of Venice

No. 2013-05
May 2013

INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

www.iser.essex.ac.uk

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Non-technical summary

Disability is high on the policy agenda in the UK and many other countries. Given the high and increasing cost of public support for disabled people, it is important to know how well targeted that support is on people in greatest need. Researchers generally address this question by analysing large-scale, nationally representative survey data, containing information at the individual level on the extent of disability, receipt of public support and other household circumstances.

A major concern about this kind of research is the difficulty of measuring disability. Survey interviewers ask questions which invite interviewees to report various kinds of health-related difficulties with everyday activities. Researchers then construct a summary measure of disability from those responses, for use in analysis. The quality of those disability measures naturally depends on the quality of the underlying survey questions. However, surveys vary in the number of disability questions they use and the wording of those questions. Surveys also differ in the way they select their samples and the way they handle cases where the subject is too unwell to answer personally. In practice, researchers often choose to use one survey rather than another quite arbitrarily and it is rare for them to investigate the robustness of their findings with respect to their choice of survey data.

In this study, we explore this issue, using data on older (65+) people from three leading UK surveys: the *Family Resources Survey* (FRS); the *British Household Panel Survey* (BHPS) and the *English Longitudinal Survey of Ageing* (ELSA). We ask whether the three surveys – which ask about disability in quite different ways – nevertheless give a similar statistical picture of the relationship between disability and receipt of the disability benefit called *Attendance Allowance* (AA).

Reassuringly, we find a quite coherent picture of the targeting of AA from the three surveys, particularly after harmonising their sample coverage. It is also important to use the same approach to statistical analysis, which treats disability as an underlying unobservable state indicated – but not directly measured – by responses to the survey questions on disability. The main unresolved cause for concern is in measurement of the cognitive dimension of disability, where the BHPS in particular appears to lack coverage.

Do household surveys give a coherent view of disability benefit targeting?

A multi-survey latent variable analysis for the older population in Great Britain

Ruth Hancock [°*]
Marcello Morciano [°*]
Stephen Pudney [*]
Francesca Zantomio [+*]

[°] *Health Economics Group, University of East Anglia*

**Institute for Social and Economic Research, University of Essex*

[+] *Department of Economics, Ca' Foscari University of Venice*

This version: 2 May 2013

ABSTRACT: We compare three major UK surveys, BHPS, FRS and ELSA, in terms of the picture they give of the relationship between disability and receipt of the *Attendance Allowance* (AA) benefit. Using the different disability indicators available in each survey, we estimate a model in which probabilities of receiving AA depend on latent disability status. Despite major differences in design, once sample composition is standardised through statistical matching, the surveys deliver similar results for the model of disability incidence and AA receipt. Provided surveys offer a sufficiently wide range of disability indicators, the detail of disability measurement appears relatively unimportant.

Keywords: disability indexes, disability benefits, multiple surveys.

JEL codes: C81, I18, I38

Contact details: Marcello Morciano, Health Economics Group, Faculty of Medicine and Health Sciences, University of East Anglia, Norwich Research Park, NR4 7TJ Norwich, UK. `m.morciano@uea.ac.uk`

We are grateful to Monica Hernandez and participants in the 2010 HESG conference at LSE and ECHE conference at Helsinki for helpful comments. This work was supported by the Nuffield Foundation and by the Economic and Social Research Council through the Research Centre on Micro-social Change (MiSoC). The European Research Council (project no. 269874 [DEVHEALTH]) provided support for Pudney's contribution to the research. Data from the Family Resources Survey (FRS) are made available by the UK Department of Work and Pensions through the UK Data Archive. Material from the FRS is Crown Copyright and is used by permission. The BHPS data used were originally collected by MiSoC at the University of Essex (now incorporated within the Institute for Social and Economic Research) and are made available through the UK Data Archive (UKDA). Data from the English Longitudinal Study of Ageing (ELSA) were developed by researchers based at University College London, the Institute for Fiscal Studies and the National Centre for Social Research (NatCen) and are made available through the UKDA. NatCen provided geographical indicators. Neither the collectors of the data nor the UKDA bears any responsibility for the analyses or interpretations presented here. The paper is also available as UEA HEG Working Paper no. 13-03 (http://www.uea.ac.uk/medicine/health-economics-group/working-papers)

1. Introduction

Developed countries like the UK will face severe problems in supporting the projected future growth in the disabled population (McVicar 2008), and in the older disabled population in particular (Karlsson *et al.* 2006, OECD 2005, Pickard *et al.* 2007). In the UK, there has been a long series of policy reviews by a Royal Commission (Sutherland 1999), the independent King's Fund (Wanless 2006), the government (Department of Health 2009), the Commission on Funding of Care and Support (CFCS 2011) and various parliamentary select committees. The current UK government has recently announced changes to some aspects of the long-term care funding system (Department of Health 2013) but debate continues on how best to provide public support to older people with care needs. Such debate and associated policy reform should ideally be evidence-based. This requires a robust and accurate baseline picture of the distribution of support for people with disabilities, allowing the development of statistical models to project changes in this picture as disability levels rise and alternative policy structures are implemented. In turn, this requires good survey data on patterns of disability and receipt of support.

The importance of disability as a policy issue is matched only by the vast range of survey questions that have been used to measure it, and the proliferation of disability indicators across surveys presents difficulties for empirical research. There are many available question designs, supported by limited testing of external validity, internal consistency and test-retest reliability, and some cognitive evaluation of specific question designs (see Sturgis *et al.* 2001 and Jagger *et al.* 2009 for reviews of UK surveys). It is widely recognised that any particular set of disability indicators may give an imperfect description of the concept of disability relevant to the analysis and that bias may result from neglect of the measurement error problem (Bound 1991). However, there has been little cross-survey comparative work which considers the consistency of the empirical 'story' that policy-makers would get from surveys offering different sets of disability indicators. In practice, researchers often use disability indicators that happen to be available in a survey chosen for convenience or to meet other requirements, and the robustness issue is rarely considered systematically. The Green Paper (Department of Health 2009), 'State of the Nation' report (Cabinet Office 2010) and the report of the Commission on Funding Care and Support (CFCS 2011) are examples of policy documents based on research using a mixture of different survey sources for different purposes.

For policy purposes, we are interested not only in the measurement of disability, but also in its relationship with other key variables like receipt of public support. In this study, we focus on a particular form of public support: the disability-linked cash benefits which are available to older people. The main disability benefit for people aged 65 or over in the UK is Attendance Allowance (AA), which is administered by the Department for Work and Pensions (DWP) and designed to help meet the extra costs arising from disability. Besides the age restriction, eligibility for AA requires the claimant to be in need of care in order to perform daily activities. The AA claim form says "you may get Attendance allowance if your disability means that you need help with your personal care or you need someone to supervise you for your own or someone else's safety". It defines help with personal care as "day-to-day help with things like washing (or getting in or out of the bath or shower), dressing, eating, going to and using the toilet, or telling people what you need or making yourself understood"; and supervision as needing "someone to watch over you to help you avoid substantial danger to yourself or other people" (Department for Work and Pensions, 2013) The benefit is not means tested and (in 2012/13) is worth either £51.85 per week, if care is needed during either day or night, or £77.45, if care is needed during both. Eligibility for AA is difficult to assess from survey data. In practice, decisions on claims are made by programme administrators on the basis of claimants' reported health problems and consequent care needs. Once the claim is made, written evidence is examined by administrative assessors, who can require a medical examination of the claimant. An element of judgement is inevitable, so eligibility is uncertain, even with access to the same information as the administrative assessor. A further challenge is that the information on which the award decision is made is not observable directly in survey data. Rather, surveys offer a set of disability-related eligibility indicators, from which inference on the success of disability targeting must be drawn.

Our policy motivation has implications for the appropriate conceptualisation of disability. We are not concerned here with medical concepts of impairment, but rather disability conceived as a set of constraints on functioning which originate from health impairments broadly defined. This corresponds to Sen's (1982, 1985) "capabilities" approach, which sees the individual choosing a consumption vector x from a choice set X and a pattern of commodity utilisation $f(.)$ from a set of possible utilisation functions F. The individual's chosen vector of "functionings" is $b = f(x)$, which is thus constrained by his or her economic entitlements (X) and available ways of using economic resources (F). We view the concept of disability as a health-related limitation on the set F relative to some socially-agreed minimal norm N. The aim of disability policy is to offer support to people for whom $F \subset N$. Support may take the form of cash or services, both of which expand the individual's choice set X, and it may be universal, in which case support is independent of the pre-intervention X, or means-tested, in which case entitlement depends on X. The important point here is that the concept of disability is concerned with constraints on basic functionings, rather than medical conditions themselves. The survey indicators used to measure disability should therefore focus on potential difficulties with everyday activities rather than health or disease.

The contribution of this paper is to investigate whether different indicators of disability, collected in three widely-used household surveys, are consistent with a common set of findings relating to the targeting of disability benefits for older people. If we admit the possibility that underlying disability is multi-dimensional, there are two aspects to this comparability issue: completeness and compatibility. A survey is *complete* in its coverage of disability if its questionnaire content generates disability indicators that are capable of reflecting all the multiple dimensions of disability. Two surveys are mutually *compatible* if their respective indicators of any particular dimension of disability give the same undistorted picture of that underlying concept. For researchers using similar methods but different data sources to be sure of agreeing on their conclusions, both completeness and compatibility are necessary in general. We investigate three British surveys, the Family Resources Survey (FRS), the English Longitudinal Survey of Ageing (ELSA) and the British Household Panel Survey (BHPS), which have been widely used for research on health, disability and related topics. We find that compatibility is not a serious difficulty, although there are some signs that completeness is a problem for the BHPS.

Typically, the statistical analysis of disability benefit receipt employs a single-equation framework, in which a variety of disability indicators (or a count index of them) are used as explanatory covariates, together with other observable socio-economic status (SES) characteristics (see Berthoud and Hancock 2008, Pudney 2009, Forder and Fernandez 2009 and Zantomio 2013 for examples). In this paper, we use a structural equation approach involving a latent concept of disability to study the relationships between disability status, SES characteristics, and receipt of AA in the BHPS, ELSA and the FRS, at (almost) a single time point, 2002/03. While acknowledging that an individual's disability status is not directly observable, we assume it is reflected in varying degrees by members of a set of imperfect but observable survey indicators. The underlying latent disability measure is influenced by a set of SES characteristics and the probability of receiving AA is a function of latent disability and SES characteristics. See Bollen (1989) for a review of this class of latent variable simultaneous equation models.

This methodological approach has two major advantages. First, overcoming the arbitrariness of approaches based on a limited set of disability indicators or a scalar (usually unweighted) count of them, the latent variable framework allows us to develop an index of disability which makes use of all available sample information. This composite index can then be used as a sounder basis for policy analysis focused on the targeting of disability benefit. Second, the latent variable framework reduces the scope for bias arising from the measurement error in observed disability-related indicators and therefore gives more reliable estimates of the relationship between benefit receipt and influences like disability and income – again improving the robustness of an analysis of benefit targeting.

Statistical models are best seen as local approximations, so comparison of evidence derived from different surveys may be influenced by sample composition as well as the design of survey instruments. Weighting is the usual method of compensating for the idiosyncrasies of survey design and response characteristics, but this is problematic for comparative purposes. In practice, weights are often produced using essentially ad hoc calibration methods and the three surveys used here have weights constructed in different ways, using different external information. Consequently, there is no reason to expect that the standard weights issued with the three surveys will overcome comparability problems fully. After estimating statistical models on the full samples in each survey, we make the samples as comparable as possible by using matching techniques to obtain samples which share a (near-)common distribution for the SES covariates, with estimation performed on each set of matched samples. This has the effect of reducing the scope of the comparison slightly (the common support constraint) but it has the advantage of removing differences due to model approximation errors at the periphery of the region covered by the survey samples.

In sections 2 and 3 of the paper, we describe the methodological framework and the three surveys, documenting the distributional characteristics of the variables used. Results from the model fitted to the full (unmatched) samples are discussed in Section 4. The matching procedure and results for matched samples are described in Section 5 and Section 6 examines the sensitivity of our findings to various aspects of our analytical approach.

2. A model of disability status and benefit receipt

In the gerontology literature, Johnson and Wolinsky (1993) conceptualise the dynamics of health status in the older population, viewing functional limitations as outcomes of latent disability. Consistent with this view, we model ‘true’ disability status as an unobservable, possibly multidimensional, phenomenon, which is influenced by socio-economic characteristics and circumstances. We observe a set of survey indicators, each of which provides a ‘noisy’ measure of underlying disability, satisfying the classical measurement error assumption that all correlation with other socio-economic characteristics is explained by latent disability. The main outcome of interest, receipt of AA, depends on latent disability and the set of socioeconomic characteristics which influence an individual’s propensity to claim and be awarded AA.

Analysis is based on independent samples of n^s individuals in surveys s = 1, 2, 3. Each sampled individual i is characterised by: unobserved Q-dimensional 'true' disability $\boldsymbol{\eta}_i = (\eta_{i1} \dots \eta_{iQ})$; socio-economic individual characteristics $\boldsymbol{Z}_i$ observable in all surveys; a set of survey-specific disability-related discrete indicators D_{ij}^s , $j = 1 \dots J^s$; and a binary indicator of benefit receipt (R_i = 1) or non-receipt (R_i = 0). We aim to draw inferences about the conditional distributions $P(\boldsymbol{\eta}|\boldsymbol{Z})$ and $P(R|\boldsymbol{\eta}, \boldsymbol{Z})$ which describe respectively the distribution of disability in the population and the relationship between benefit receipt and the individual's disability and other characteristics. By definition, these population distributions are independent of any survey used to draw inferences about them. An important question is whether the empirical distributions $P^s(R, D_1^s \dots D_{J^s}^s|\boldsymbol{Z})$ produced by the three surveys with their different disability indicators nevertheless give a coherent indication of underlying 'true' disability $\boldsymbol{\eta}$ and its relationship with benefit receipt R.

We use an ordinal quasi-linear structure for disability measurement:

$$\tilde{D}_{ij}^s = \alpha_j^s + \lambda_{j1}^s \eta_{i1} + \dots + \lambda_{jQ}^s \eta_{iQ} + \varepsilon_{ij}^s \quad (1)$$

$$D_{ij}^s = m \quad \text{iff} \quad A_{jm-1}^s \le \tilde{D}_{ij}^s < A_{jm}^s , \quad m = 1, \dots, M_j^s \quad (2)$$

where: the coefficients λ_{jq}^s are factor loadings relating observed indicators in survey s to underlying disability; ε_{ij}^s is a normally-distributed residual term representing random response error, implying an ordered probit link function generating the observable indicator D_{ij}^s from its unobserved continuous form $\tilde{D}_{ij}^s$. M_j^s is the number of response categories for indicator D_{ij}^s and the A_{jm}^s are threshold parameters. The qth disability component η_{iq} is related to $\boldsymbol{Z}_i$ through a linear relationship representing the processes leading to disability:

$$\eta_{iq} = \boldsymbol{\theta}_q \boldsymbol{Z}_i + \upsilon_{iq} \quad (3)$$

where $\boldsymbol{\theta}_q$ is a vector of coefficients. The residual υ_{iq} captures other unobservable factors and satisfies $E(\upsilon_{iq}|\boldsymbol{Z}_i) = 0$). Benefit receipt is modelled by a probit specification:

$$\tilde{R}_i = \boldsymbol{\beta} \boldsymbol{Z}_i + \gamma_1 \eta_{i1} + \dots + \gamma_Q \eta_{iQ} + u_i \quad (4)$$

where the observed benefit receipt status $R_i = 1$ when $\tilde{R}_i > 0$ and $R_i = 0$ otherwise; $\boldsymbol{\beta}$ and the γ_q are coefficients and u_i^s is a stochastic disturbance term. We make the standard assumption underlying probit models like (4) that the stochastic residual u_i is independent of $(\boldsymbol{Z}_i, \boldsymbol{\eta}_i)$ and the residuals in the measurement equations (1). In writing (3) and (4), we allow the same covariates to represent the influences on disability and on benefit claim behaviour. This is not necessary, and there may be exclusion restrictions (which are not necessary for identification) on the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\theta}_q$.

We say that survey s is ***complete*** if the J×Q loadings matrix $\{\lambda_{jq}^s\}$ is of full column rank Q; this requires that, for each dimension of disability q, at least one of the j observed indicators D_{ij}^s has a non-zero loading λ_{jq}^s. Appendix 2 shows that completeness is sufficient to identify the model under our assumptions. The surveys are said to be ***compatible*** if the assumption of common parameters across surveys in equations (3) and (4) is valid.

Several studies have shown that, in the older population, women tend to report significantly higher rates of functional difficulties than comparable men (Rahman and Liu 2000, Crimmins *et al.* 2011). Some researchers have attributed this apparent female functional disadvantage to higher true prevalence of nonfatal but disabling conditions such as arthritis and osteoporosis (Wingard 1984, Verbrugge and Wingard 1987, Kandrack *et al.* 1987). Others have found that, even when controlling for chronic conditions, women still report higher mean levels of functional disability (Waltz and Badura 1984). This could be due to a higher propensity for women to report ill health than men with the same underlying true health status (Verbrugge 1980, Gove 1984, Hibbard and Pope 1983); or to heightened sensitivity to symptoms because of gender-specific social expectations and life experience (Verbrugge and Wingard 1987, Verbrugge 1989, Verbrugge and Balaban 1989); or to task specificity if women are more engaged than men in household tasks that require actions such as bending and lifting. This measurement issue has been termed variously: ‘state-dependent reporting bias’ (Kerkhofs and Lindeboom 1995), ‘scale of reference bias’ (Groot 2000), ‘response category cut-point shift’ (Sadana *et al.* 2000; Murray *et al.* 2001, Lindeboom and van Doorslaer 2004), ‘reporting heterogeneity’ (Shmueli 2002, 2003) and ‘differential item functioning’ (Hays *et al.* 2000). We allow for the possibility of inherent gender differences in survey reporting of disabilities by allowing the parameters of the measurement equations (1)-(2) to be gender-specific. Note that, unless we can specify *a priori* a subset of indicators in each survey for which response behaviour is gender-invariant, it is impossible to distinguish the causal effect of gender on true latent health from its effect on reporting behaviour. We resolve this by excluding gender from the latent disability equation; the results should be interpreted accordingly.

3 Data

The empirical analysis is based on three sample surveys: the first wave of ELSA; the corresponding twelfth wave of BHPS; and the 2002/03 cross section of FRS. All three surveys have been widely used for research on physical health and disability: see, for example, Melzer *et al.* (2005), Banks *et al.* (2006), Mayhew *et al.* (2010), Chan *et al.* (2012), Zaninotto and Falaschetti (2011), Clarke and Smith (2011) for ELSA; Benítez-Silva *et al.*

(2009), Oswald and Powdthavee (2008), Banks *et al.* (2009), Hirst (2004) for BHPS; and Kasparova *et al.* (2007), Hancock and Pudney (2012) and Morciano *et al.* (2012) for FRS.

Although often used for similar research purposes, the three surveys differ in their design and response rates. All are formed of clustered random samples drawn from the Postcode Address File, but at varying time points. The FRS has a sample size of over 25,000 private households. It is an annual cross-section and therefore suffers from non-response but not accumulated attrition. The FRS response rate in 2002/3 was 64% of eligible households (Campbell 2004). The BHPS started in 1991 and has been following a sample of approximately 10,000 households at annual intervals since then. The BHPS wave 1 response rate was 74%; of those original respondents, 67% gave a full interview in wave 12 (Lynn *et al.* 2006). Thus the BHPS sample used here has come through 12 waves of attrition and possible panel conditioning. ELSA is a panel of individuals aged 50+ and their partners in approximately 8,000 private households in England. Panel membership is based on interview in the 1998, 1999 or 2001 Health Surveys for England (HSE). The wave 1 ELSA data are thus potentially affected by non-response in the HSE and a further round of attrition; HSE response rates were 74% (1998), 76% (1999) and 74% (2001) and of those selected for ELSA, around 70% responded to its first wave (Taylor *et al.* 2003). We choose the first wave of ELSA as our common time point to avoid the effects of subsequent attrition. We limit our analysis to people aged 65 years or over, living in England. The former restriction is because only people aged 65 or over can claim AA. The latter is imposed by the ELSA sampling frame. We also exclude respondents receiving Disability Living Allowance (a similar benefit that can be claimed before age 65) because DLA recipients cannot also claim AA.

The surveys differ in content. FRS collects very detailed income and benefit information, used as the basis for most official statistics on welfare and disability program targeting, but a limited set of disability indicators. ELSA provides a richer range of health and disability measures but slightly more limited income data than the FRS (for example, ELSA collects some income components gross of tax and others net). In the BHPS, it is not always possible to distinguish whether a particular income source is gross or net. BHPS information on health and disability is more detailed than the FRS but less so than in ELSA. The surveys differ in the information they collect by proxy for participants who are not able to provide responses themselves, in particular FRS collects information on disability and AA receipt from proxy respondents, whereas BHPS and ELSA do not. We return to treatment of proxy respondents below. Campbell (2004), Taylor *et al.* (2003) and Taylor *et al.* (2006) respectively give detailed descriptions of FRS, ELSA and BHPS sample design and data collection procedures.

In each survey, information about receipt of AA, recorded by the binary variable R_i, is collected through questions following those on health and disability. Thus, none of the three surveys is especially vulnerable to the justification bias in disability measurement that is a concern when the benefits module precedes the health module within the questionnaire (Crossley and Kennedy 2002). There are differences in the reference period for questions on AA receipt: the BHPS covers the year preceding interview; the FRS refers specifically to the time of interview; and ELSA asks separately about different reference points. For ELSA we use receipt of AA at the time of interview, to give comparability with the FRS.

Disability indicators available in one or more of the three surveys cover a wide range. In this study, we use subjective indicators which are the most widely available in social surveys. Appendix Table A1 reports the functional limitation indicators D_j offered by each survey and used in our analysis, with their prevalence rates among AA recipients and non-recipients. Binary indicators in the FRS cover difficulties in eight areas of life. ELSA provides a longer list of indicators including limitations to specific Activities of Daily Living (ADL) (Katz *et al.*, 1963) or Instrumental Activities of Daily Living (IADLs) (Lawton and Brody 1969). The BHPS indicators include binary variables representing activities limited by health and a set of 6-point categorical variables, built from two questions on whether the respondent usually manages to perform a set of mobility and personal care activities alone or only with assistance, and whether he/she finds it very easy, fairly easy, fairly difficult or very difficult. There is a considerably higher sample prevalence of reported functional limitations among AA recipients than non recipients, consistently across surveys and specific indicators.

The choice of other personal characteristics included in $\boldsymbol{Z}$ is governed by previous work on the socio-economic gradient in health or disability (e.g. Goldman 2001) and on older people's benefit claim behaviour (see for example, Pudney 2009 in relation to AA; Hernandez *et al.* (2007) and Pudney *et al.* (2006) for means-tested benefits). We use age (in the form of a spline with a knot at the median age across all samples of 73), gender, being educated beyond the compulsory minimum, housing tenure, and log equivalised pre-benefit income in both equations. Income represents both the socio-economic gradient in health and the basic need for financial support which underlies benefit claim behaviour. It is derived as the sum of income from pensions, earnings, savings and other sources received by any member of the benefit unit, but excludes disability and means tested benefits. Disability benefits must be excluded from the latent disability equation because they are a consequence, and not a cause, of disability, and from the AA equation as it is income in the absence of AA that influences the decision to claim. Means-tested benefits are excluded because their level can also depend on disability through the Severe Disability Premium, an addition to the income thresholds used to assess entitlement to means-tested welfare benefits and applies where the claimant is receiving AA . To account for differences in benefit unit size we apply the modified OECD equivalence scale to income. For this older population, our income measure is dominated by pension income, which is a good indicator of past labour market success, itself strongly related to lifestyle characteristics which have associated health implications. Thus estimates of the impact of income on disability should be interpreted in this wide sense. Log income is entered as a spline with a knot at the median log income level (log of £615.70 per month, 2002 prices). Our definition of housing tenure distinguishes those who own their homes outright from those who rent or are still repaying their mortgage. Outright home-ownership is used to capture an additional long-term socio-economic influence on health. It also allows for the lower financial need (lower housing costs) that outright owners have compared with those who face rent or mortgage costs, to influence their benefit claim behaviour. Additionally, current partnership status (married/cohabiting versus single) is included as a covariate in the AA receipt equation since it has previously been found to affect benefit claim behaviour (Hernandez *et al*. 2007; Pudney *et al*. 2006).

All variables have been derived in a consistent manner as far as possible, although perfect comparability cannot be guaranteed. Sample means and standard deviations for the socio-economic characteristics $\boldsymbol{Z}$ observed in each sample are given in Appendix Table A2. There are some differences between surveys: for example, ELSA sample members are significantly younger and more educated than their BHPS and FRS counterparts; the proportion of outright homeowners is higher in ELSA and the BHPS than in the FRS; and the mean of (log) income is significantly higher in the BHPS than in ELSA and the FRS. FRS reports a higher rate of AA receipt (9.7%) than ELSA or BHPS (7.2%). Comparisons with administrative data suggest that FRS is closest to the population value[1].

Ideally we would use all proxy cases since they are likely to include some of the most severely disabled respondents. This view is supported by an analysis of proxy respondents in the FRS, revealing AA receipt to be about twice as high among proxy respondents as non-proxy respondents (18.1% against 9.1%). However we are forced to exclude proxy responses in ELSA (1.9%) and BHPS (4.1%) as their proxy questionnaires do not collect the respondent's disability (ELSA) or AA receipt (BHPS). We retain the larger proportion of proxy cases (6.5%) in the FRS which does collect this and other relevant information for proxy cases, using a proxy response as an additional disability indicator in the measurement

[1] We estimate that of the over 65 non care home population, excluding those who received DLA, between 12.7 and 13.8% received AA in 2002. This is based on DWP statistics on recipients of AA and DLA (http://statistics.dwp.gov.uk/asd/index.php?page=tabtool) which include, but do not separately distinguish, recipients in care homes, together with estimates from Comas-Herrera *et al.* (2010) on the numbers of over 65s resident in care homes and the proportions of them who receive public support with the care home fees and are therefore not eligible to receive AA. All three surveys therefore seem to under-represent AA recipients but FRS less so than ELSA or BHPS.

model. After these exclusions and dropping cases with missing values for variables used in the analysis, the sample sizes are 1,042, 5,142 and 6,744 individuals from the BHPS, ELSA and FRS respectively. We also assess the sensitivity of the results to the exclusion of FRS proxy cases in which case the FRS sample is reduced to 6,308.

4 Estimation results

4.1 The measurement model

To implement the model, we must specify the dimensionality of latent disability and choose a normalisation to deal with its non-observability and lack of natural units of measurement. Our main results come from a model with a single latent disability factor and a simple normalisation. For the latter, we choose a priori one indicator from each survey that appears to be based on essentially the same question. These are: the FRS question about mobility ('moving about'); the ELSA question about capacity to 'walk 100 yards'; and the BHPS question about 'walking more than 10 minutes'. We then normalise the factor loading for each of these indicators to be unity. In section 6, we explore the sensitivity of the results to our choice of normalisation and number of factors.

The estimates of the measurement model are presented in Table 1: the factor loadings λ_{jq}^{s}, representing the effect of latent disability η on each indicator D_{ij}^{s}, are positive and highly significant in each survey. Although the pattern of estimated factor loadings is similar for male and female respondents in each survey, there are significant differences. In FRS, the loading associated with 'lifting, carrying or moving objects' is significantly higher for women. In ELSA, factor loadings associated with reported difficulties in ADLs like 'bathing or showering', 'eating', 'getting in or out of bed' and 'using the toilet' and IADLs like 'doing work around the house or garden' are significantly lower for women; in BHPS, a significantly lower factor loading for women is also found for difficulties in bed transfers and 'bathing or showering'. The Akaike information criterion suggests that the unrestricted models (which allow the parameters of the measurement equations (3) to be gender-specific) provide slightly better balances of model fit and parsimony. This result is also confirmed by the Satorra-Bentler (2001) test at the 1% level for each of the three surveys.

Table 1: Estimated 1-factor models

Disability Indicator§§	Factor loading	Standard error	Disability Indicator§§	Factor loading	Standard error
		MEN			
	FRS			*ELSA*	
MOBILITY	1	-	WALK100	1	-
LIFTING	1.005†	(0.088)	SITTING	0.386†	(0.031)
DEXTERITY	0.723†	(0.064)	CHAIR	0.581†	(0.039)
CONTINENCE	0.395†	(0.036)	CLIMBSEV	0.724†	(0.049)
COMMUNIC	0.385†	(0.042)	CLIMB1	0.990†	(0.066)
MEMORY	0.420†	(0.042)	STOOP	0.641†	(0.042)
DANGER	0.510†	(0.093)	ARMS	0.503†	(0.042)
OTHER	0.098†	(0.027)	PULL/PUSH	1.008†	(0.077)
PROXY	0.116†	(0.029)	LIFTING	0.934†	(0.066)
	BHPS		COIN	0.379†	(0.047)
HOUSEWORK	0.851†	(0.129)	DRESSING	0.661†	(0.048)
STAIRS	0.959†	(0.131)	WALKING	1.052†	(0.135)
DRESS	0.660†	(0.115)	BATH	0.863†	(0.067)
WALKING	1	-	EATING	0.596†	(0.087)
STAIRS	1.112†	(0.172)	BED	0.879†	(0.085)
MOBILITY	1.358†	(0.265)	TOILET	0.738†	(0.091)
BED	1.346†	(0.247)	CONTINENCE	0.299†	(0.030)
NAILS	0.585†	(0.082)	MAP	0.406†	(0.049)
BATH	1.001†	(0.162)	MEAL	0.806†	(0.101)
ROAD	1.151†	(0.169)	SHOPPING	1.018†	(0.085)
			PHONE	0.358†	(0.046)
			MEDICATION	0.477†	(0.071)
			HOUSEWORK	1.132†	(0.086)
			MONEY	0.453†	(0.057)
		WOMEN			
	FRS			*ELSA*	
MOBILITY	1	-	WALK100	1	-
LIFTING	1.186†	(0.101)	SITTING	0.399†	(0.029)
DEXTERITY	0.643†	(0.047)	CHAIR	0.532†	(0.033)
CONTINENCE	0.431†	(0.035)	CLIMBSEV	0.671†	(0.043)
COMMUNIC	0.365†	(0.037)	CLIMB1	0.899†	(0.053)
MEMORY	0.416†	(0.035)	STOOP	0.653†	(0.040)
DANGER	0.426†	(0.052)	ARMS	0.500†	(0.035)
OTHER	0.060‡	(0.024)	PULL/PUSH	0.899†	(0.056)
PROXY	0.121†	(0.024)	LIFTING	0.900†	(0.058)
	BHPS		COIN	0.433†	(0.036)
HOUSEWORK	0.968†	(0.151)	DRESSING	0.650†	(0.042)
STAIRS	1.201†	(0.171)	WALKING	0.959†	(0.091)
DRESS	0.910†	(0.164)	BATH	0.722†	(0.047)
WALKING	1	-	EATING	0.428†	(0.055)
STAIRS	0.911†	(0.122)	BED	0.686†	(0.054)
MOBILITY	1.066†	(0.154)	TOILET	0.577†	(0.050)
BED	0.965†	(0.141)	CONTINENCE	0.251†	(0.022)
NAILS	0.582†	(0.076)	MAP	0.343†	(0.029)
BATH	0.777†	(0.105)	MEAL	0.811†	(0.074)
ROAD	1.110†	(0.154)	SHOPPING	1.135†	(0.080)
			PHONE	0.327†	(0.045)
			MEDICATION	0.479†	(0.074)
			HOUSEWORK	0.926†	(0.060)
			MONEY	0.479†	(0.048)

Statistical significance: † p < 0.01; ‡ p < 0.05; § p < 0.1. §§ A more detailed description for each D_j^s indicator can be found in appendix Table A1.

4.2 The disability model

Estimates for the model (3) of latent disability status are reported in Table 2, together with t-tests of individual coefficient equality and the overall χ^2 Wald tests for equality of the whole coefficient vector for each pair of surveys. The conditional mean of latent disability η

increases with age: the FRS and ELSA display a nonlinear relation between age and disability, with a higher gradient beyond age 73. In the BHPS we find a strong and near-linear relationship between age and disability. Higher education and pre-benefit income are associated with lower disability, giving evidence of a socio-economic gradient in disability that is consistent across surveys. Being a homeowner decreases the conditional mean of η, particularly in ELSA. The variance of the latent disability factor is greater in the BHPS than in the FRS or ELSA, but we find that the factor variances are quite comparable across surveys (a 10% significant difference is found only for the FRS-ELSA contrast). The estimated coefficients for FRS and ELSA are similar in size and the Wald test cannot reject the hypothesis of equality; when the BHPS is used as the basis for comparison, the null hypothesis of joint equality of coefficients is rejected at the 5% level.

Table 2: Estimates of the latent disability equation

	Coefficients			Tests and coefficient differences		
Covariates	FRS	ELSA	BHPS	FRS-ELSA	FRS-BHPS	ELSA-BHPS
Spline age 65-73	0.038†	0.035†	0.127†	0.003	-0.089†	-0.092†
	(0.013)	(0.012)	(0.035)	(0.018)	(0.038)	(0.037)
Spline from age 73+	0.091†	0.099†	0.128†	-0.008	-0.037§	-0.029
	(0.008)	(0.007)	(0.020)	(0.011)	(0.021)	(0.021)
Post-compulsory education	-0.279†	-0.280†	-0.182	0.001	-0.096	-0.097
	(0.064)	(0.060)	(0.148)	(0.088)	(0.161)	(0.159)
Income spline to median	-0.162†	-0.046	-0.172	-0.116§	0.009	0.125
	(0.045)	(0.051)	(0.107)	(0.068)	(0.116)	(0.118)
Income spline from median	-0.336†	-0.310†	-0.558†	-0.025	0.223	0.248
	(0.084)	(0.069)	(0.191)	(0.109)	(0.209)	(0.203)
Outright owner	-0.382†	-0.487†	-0.185	0.105	-0.197	-0.302§
	(0.062)	(0.061)	(0.143)	(0.087)	(0.155)	(0.155)
Variance (σ_v^2)	3.012†	2.543†	3.298†	0.469§	-0.286	-0.755
	(0.271)	(0.223)	(0.743)	(1.336)	(0.362)	(0.973)
	Sample size			*Coefficient equality χ^2(6)*		
	6,744	5,142	1,042	4.675	13.060‡	15.745‡

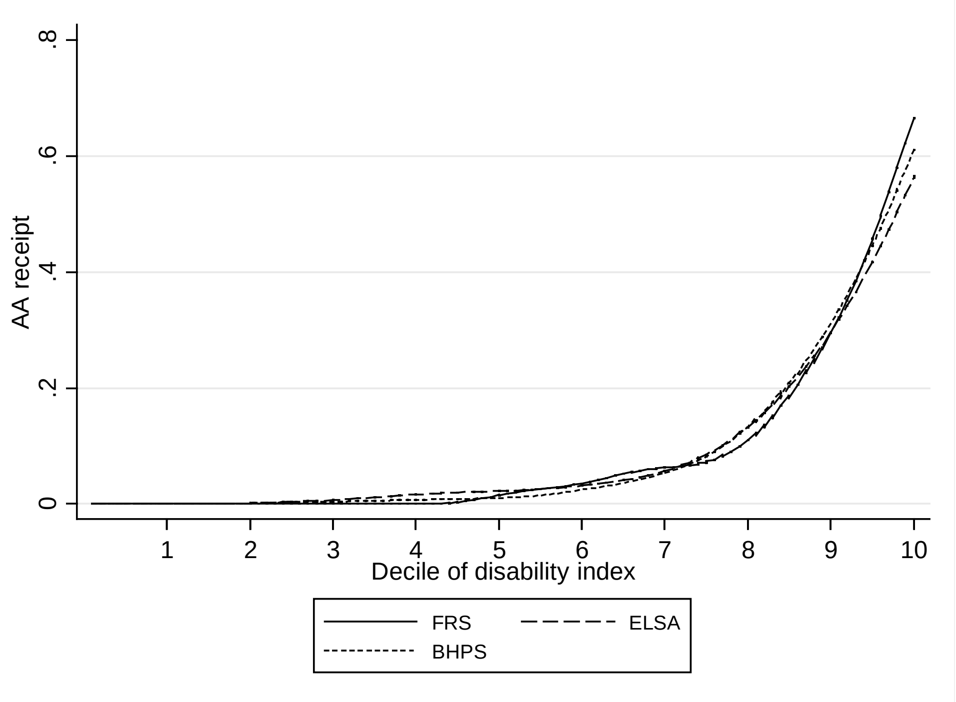
Statistical significance of the coefficient, cross-sample coefficient difference and χ^2 statistic:
† $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parenthesis.

4.3 The benefit receipt model

Estimates for equation (4), describing the relationship of AA receipt with socio-economic characteristics and latent disability, are reported in Table 3. Receipt of AA is clearly disability-related in each of the surveys, and disability consistently emerges as the dominant variable in explaining AA receipt. Although disability might raise barriers to claiming and at the same time reduce individuals' capacity to benefit from additional cash income, the survey evidence suggests there is successful targeting of AA on the disabled older population, irrespective of the source of survey data. This is clear from Figure 1, which shows the mean prevalence of AA receipt within each decile of the distribution of the posterior prediction of latent disability for each individual. The strong disability-targeting of AA emerges very clearly for all three surveys.

The estimated probability of receiving AA declines nonlinearly with income. We find that, below median income, the coefficient is significant at the 10% level only in ELSA, so the income gradient in AA receipt operates primarily among higher-income people. The negative gradient is due both to the low incidence of disability among high-income groups (Pudney 2010) and to the low propensity of these groups to claim benefit (Hernandez *et al.* 2007). Consequently, although AA is not means-tested, patterns of receipt mimic to some degree the effect of means testing for those in the top half of the pensioner income distribution.

Figure 1: Proportion of people in receipt of AA by predicted severity of disability



Note: Smoothed local linear regressions applied on the FRS (solid line), the ELSA (long dashed line) and the BHPS (dotted line) samples. Bandwidth set equal to 0.4.

We find significant evidence of a negative association between the level of education and AA receipt in both ELSA and FRS. This suggests that any advantage that more educated people may have in navigating the benefits system is outweighed by factors such as less contact throughout their lives with the benefit system, or greater perceived stigma from claiming benefits (as also found in Zantomio 2013). Owning one’s home outright reduces significantly the probability of AA receipt in the FRS and the BHPS. This could reflect a lower financial need among homeowners, or the same factors that may be at work for more educated people could play a similar role for outright homeowners.

Receipt of AA appears gender-related in the FRS and the ELSA, where men are less likely to receive AA than women; gender differences are insignificant in the BHPS. In all three surveys, age affects the probability of AA receipt non-linearly, with a convex age profile. There is again a significant difference between the estimated age profile for the BHPS compared with FRS and ELSA, with a less significant upturn at older ages. Finally, none of the surveys suggests that the presence of a partner significantly affects the probability of receiving AA. Inspection of coefficients in this piecemeal way creates a bias in favour of finding significant differences, because of the multiple comparisons involved. However, a joint Wald test finds a significant difference between BHPS and the other two samples (*P*-values 0.086 and 0.078). We do not reject coefficient equality between the FRS and ELSA.

Table 3: Estimates of the equation for receipt of Attendance Allowance

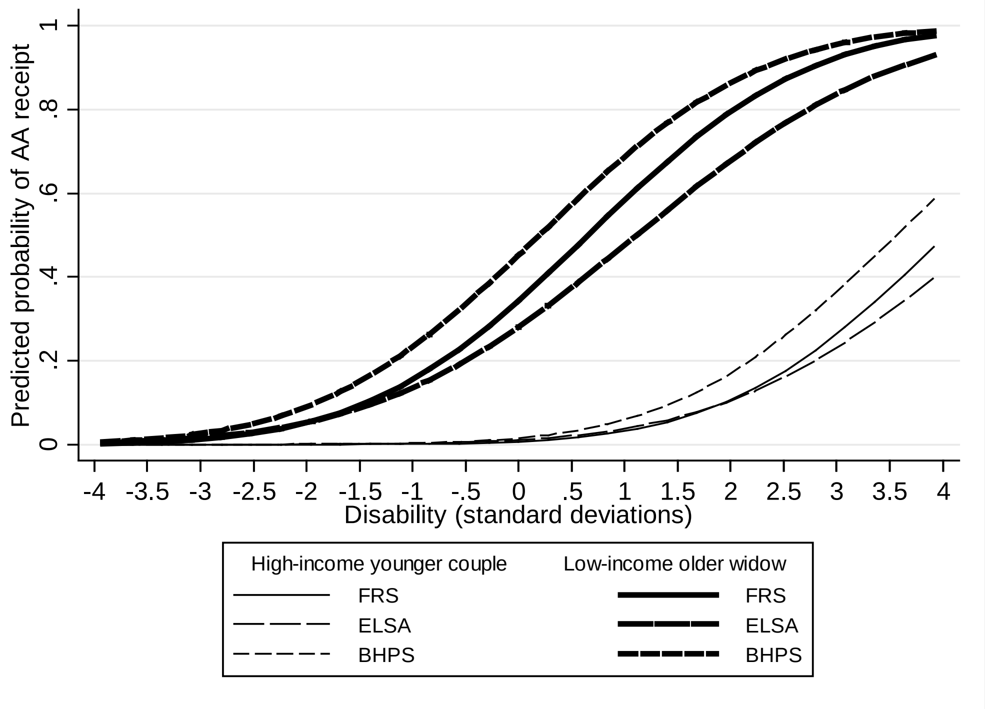
	Coefficients			Coefficient differences		
Covariates	FRS	ELSA	BHPS	FRS-ELSA	FRS-BHPS	ELSA-BHPS
Latent disability η	0.569†	0.477†	0.538†	0.092§	0.031	-0.060
	(0.040)	(0.035)	(0.091)	(0.053)	(0.099)	(0.098)
Female	0.122§	0.251†	-0.068	-0.129	0.190	0.319§
	(0.067)	(0.076)	(0.177)	(0.101)	(0.189)	(0.193)
Spline age 65-73	-0.040†	-0.036†	-0.084†	-0.004	0.043‡	0.048‡
	(0.008)	(0.007)	(0.020)	(0.011)	(0.022)	(0.021)
Spline from age 73+	0.0580†	0.0460†	0.028§	0.012	0.030§	0.017
	(0.006)	(0.006)	(0.015)	(0.009)	(0.016)	(0.016)
Post- compulsory education	-0.161‡	-0.238†	-0.070	0.077	-0.090	-0.167
	(0.064)	(0.071)	(0.160)	(0.096)	(0.172)	(0.175)
(ln) income spline to median	-0.008	-0.092§	-0.041	0.083	0.033	-0.050
	(0.047)	(0.050)	(0.084)	(0.069)	(0.096)	(0.098)
(ln) income spline from median	-0.392†	-0.422†	-0.411§	0.030	0.019	-0.011
	(0.115)	(0.153)	(0.249)	(0.191)	(0.274)	(0.292)
Outright owner	-0.136‡	-0.006	-0.265§	-0.130	0.128	0.259
	(0.061)	(0.070)	(0.160)	(0.093)	(0.171)	(0.174)
Married/cohabiting	-0.076	0.087	-0.171	-0.163§	0.094	0.257
	(0.063)	(0.074)	(0.173)	(0.097)	(0.184)	(0.188)
$\chi^2(9)$ test of coefficient equality				14.460	15.174§	15.483§

Statistical significance of the coefficient, cross-sample coefficient difference and χ^2 statistic:
† p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parenthesis.

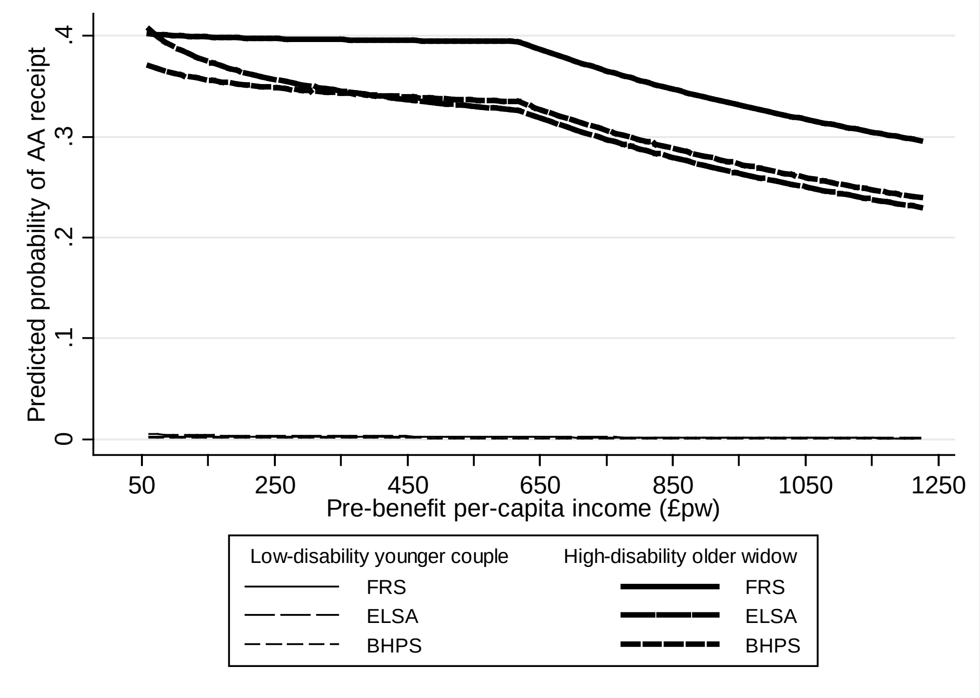
In Figure 2(a), we compare the implications of the estimated models, for two illustrative individuals: a 65-year old man living with his partner as an outright homeowner with income 50% above the median; and an 85-year old non-homeowner widow, with equivalised income 75% of the median. Both have compulsory minimum education. In Figure 2a, the between-survey differences in their AA-disability profiles are modest in comparison with the predicted differences between hypothetical individual types. For example, at a disability level one standard deviation above the mean, the three models predict a 4-7% rate of receipt for the couple compared to a 50-71% rate for the widow. At disability level of 2.5 standard deviations above the mean, the ranges are 16-26% for the couple and 77-92% for the widow.

Figure 2: Predicted probabilities of AA receipt by survey for two benchmark cases

(a) The AA-disability relation



(b) The AA-income relation



In Figure 2(b), we compare the estimated AA-income profiles. Again, the between-survey differences in these profiles are modest in comparison with the predicted differences between hypothetical individual types. The rate of receipt for the low-disability type (at the 25th percentile of the disability index distribution) couple is essentially zero, whereas the rate of receipt for the high-disability type (at the 75th percentile of the disability index distribution) ranges from 31% to 37% in the income interval we consider. The rate of receipt is nonlinear in income: almost flat below median equivalised income and steadily declining thereafter. For example, the rate of receipt for the highly-disabled widow ranges from 34 to 39% at the 25th (£435 per month) and at the 50th percentile of the income distribution, and 27-33% at the 75th percentile (£917 per month).

5 Controlling for sample composition

We now present estimates computed after using matching techniques to define sub-samples from each survey that are as comparable as possible in terms of the set of socio-economic characteristics used as covariates in the AA receipt equation (age, gender, post-compulsory education, partnership, housing tenure and the log of pre-benefit net income). We take each survey in turn as a baseline and construct matched sub-samples from the other two surveys yielding six pairs of matched samples. The matching algorithm (Leuven and Sianesi 2003) minimises the Mahalanobis distance, for the socio-economic variables. Matching is performed without replacement, to avoid repeated use of the same observation from the matched survey, at the cost of possibly reducing the size of successfully matched samples. According to available sample size, in each round of pairwise matching we impose a caliper (ranging from 0.04 to 0.5) to prevent poor matches, equivalent in practice to exact matching of binary variables and very close matching for the continuous income and age variables; *t*-tests for the equality of means between each baseline sample and the corresponding matched samples were used to confirm the success of the algorithm in balancing the conditioning covariates. We also discarded matched pairs of observations whose income difference was in the top 5% when matching BHPS to ELSA and the top 10% when matching ELSA to BHPS. Means of socio-economic variables and AA receipt in the matched samples are given in appendix table A3.

We repeated estimation of the system of equations (1), (3) and (4) on each of the six pairs of matched samples. Results obtained for the measurement equations (1)-(2) broadly confirm the patterns described in Section 4, with mobility indicators playing a dominant role as indicators of latent disability (see appendix table A4-A6). The three panels of Table 4 report estimated regression coefficients for the latent disability equation (3) obtained from samples mimicking the FRS, ELSA and BHPS sample compositions respectively. As in the unmatched samples (Table 2), we obtain significant disability gradients in age (positive) and income (negative), consistently across surveys (although some of the coefficients lose significance in smaller samples); *t*-tests of coefficients' cross-sample stability reject the null hypothesis (at the 10% level) of coefficient equality only for the first spline of income coefficient, when FRS or ELSA are used to mimic the BHPS sample composition. The striking similarity of estimated coefficients is confirmed by the χ^2 tests of coefficients' joint equality: in none of the six paired survey comparisons is the null hypothesis rejected.

Table 4: Estimates of the disability equation in matched samples

Covariate	Coefficient Estimates (standard errors)			
FRS sample composition	***ELSA matched to FRS***		***BHPS matched to FRS***	
	FRS	ELSA	FRS	BHPS
Spline age 65-73	0.047 (0.016)	0.036 (0.013)	0.073 (0.084)	0.142 (0.037)
Spline age 73+	0.090 (0.010)	0.098 (0.008)	0.077 (0.034)	0.119 (0.019)
Post- compulsory education	-0.182 (0.081)	-0.231 (0.065)	-0.001 (0.264)	-0.090 (0.160)
Income spline to median	-0.258 (0.097)	-0.113 (0.093)	-0.662 (0.662)	-0.925 (0.370)
Income spline from median	-0.314 (0.119)	-0.391 (0.085)	-0.308 (0.320)	-0.469 (0.238)
outright owner	-0.447 (0.081)	-0.491 (0.066)	-0.146 (0.166)	-0.226 (0.157)
χ^2 (6) *for coefficient equality*	1.958		6.236	
N	4,587		973	
ELSA sample composition	***FRS matched to ELSA***		***BHPS matched to ELSA***	
	FRS	ELSA	ELSA	BHPS
Spline age 65-73	0.033 (0.016)	0.037 (0.013)	0.061 (0.031)	0.072 (0.035)
Spline age 73+	0.096 (0.010)	0.098 (0.008)	0.082† (0.016)	0.128† (0.021)
Post- compulsory education	-0.205 (0.079)	-0.271 (0.067)	-0.043 (0.142)	-0.257 (0.169)
Income spline to median	-0.125 (0.083)	-0.093 (0.095)	-0.284 (0.189)	-0.608 (0.378)
Income spline from median	-0.340 (0.115)	-0.362 (0.086)	-0.245 (0.194)	-0.512 (0.253)
outright owner	-0.437 (0.077)	-0.524 (0.068)	-0.442 (0.146)	-0.230 (0.157)
χ^2 (6) *for coefficient equality*	1.573		7.049	
N	4,596		850	
BHPS sample composition	***FRS matched to BHPS***		***ELSA matched to BHPS***	
	FRS	BHPS	ELSA	BHPS
Spline age 65-73	0.040† (0.038)	0.143† (0.036)	0.044† (0.034)	0.133† (0.039)
Spline age 73+	0.089 (0.021)	0.116 (0.019)	0.089 (0.019)	0.112 (0.021)
Post- compulsory education	-0.075 (0.169)	-0.053 (0.156)	0.112 (0.156)	-0.091 (0.176)
Income spline to median	-0.444 (0.418)	-0.941 (0.356)	0.138 (0.296)	-0.296 (0.261)
Income spline from median	-0.403 (0.256)	-0.423 (0.236)	-0.606 (0.271)	-0.551 (0.292)
outright owner	-0.457 (0.180)	-0.209 (0.152)	-0.648 (0.171)	-0.318 (0.174)
χ^2 (6) *for coefficient equality*	8.265		10.571	
N	966		791	

Note: Significance of cross-sample coefficient difference and χ^2 statistic: † p < 0.01; ‡ p < 0.05; § p < 0.1.

Estimated coefficients for the AA receipt equation (4) are reported in Table 5. The positive disability gradient in AA receipt found in the unmatched samples (Table 3) is evident also in the matched samples: estimates for the disability coefficient γ are positive, significant and remarkably similar in size. The negative income gradient is also confirmed, except for an insignificant positive coefficient when ELSA mimics the BHPS sample composition. The negative association between homeownership and receipt of AA is again found whenever the coefficient on homeownership is significant. For age, coefficient equality is rejected at the 5% level only for the second spline when BHPS observations are used to mimic the ELSA sample composition; but such isolated rejections are likely to arise from sampling error when large numbers of individual *t*-tests are used. None of them would be significant if a Bonferroni correction were used, and joint Wald χ^2 tests of coefficient equality fail to reject the hypothesis of joint coefficient equality in any of the six pairwise comparisons.

Table 5: Estimates of the AA receipt equation in matched samples

Covariate	Coefficient Estimates (standard errors)			
FRS sample composition	***ELSA matched to FRS***		***BHPS matched to FRS***	
	FRS	ELSA	FRS	BHPS
Latent disability η	0.550 (0.047)	0.498 (0.038)	0.622 (0.118)	0.517 (0.091)
Female	0.031 (0.084)	0.179 (0.083)	-0.037 (0.182)	-0.128 (0.191)
Spline age 65-73	-0.031 (0.010)	-0.025 (0.009)	-0.004 (0.057)	0.001 (0.035)
Spline age 73+	0.062 (0.008)	0.050 (0.007)	0.023 (0.021)	0.025 (0.016)
Post- compulsory education	-0.119 (0.080)	-0.209 (0.081)	-0.107 (0.180)	0.146 (0.173)
Income spline to median	-0.125 (0.080)	-0.203 (0.086)	-0.349 (0.433)	-0.688 (0.357)
Income spline from median	-0.398 (0.160)	-0.492 (0.199)	-0.644 (0.339)	-0.304 (0.272)
outright owner	-0.113 (0.077)	0.010 (0.077)	-0.223 (0.173)	-0.297 (0.168)
Married/Cohabiting	-0.010 (0.082)	0.079 (0.082)	0.110 (0.177)	-0.047 (0.185)
χ^2 (9) *for coefficient equality*	6.509		2.953	
N	4,587		973	
ELSA sample composition	***FRS matched to ELSA***		***BHPS matched to ELSA***	
	FRS	ELSA	ELSA	BHPS
Latent disability η	0.581 (0.050)	0.480 (0.038)	0.658 (0.119)	0.508 (0.098)
Female	0.084 (0.084)	0.172 (0.082)	0.420 (0.224)	0.025 (0.198)
Spline age 65-73	-0.028 (0.010)	-0.027 (0.009)	-0.037 (0.026)	-0.003 (0.033)
Spline age 73+	0.057 (0.008)	0.050 (0.007)	0.057‡ (0.019)	0.021‡ (0.017)
Post- compulsory education	-0.139 (0.082)	-0.207 (0.080)	-0.542 (0.207)	0.075 (0.181)
Income spline to median	-0.154 (0.079)	-0.184 (0.084)	-0.241 (0.206)	-0.388 (0.197)
Income spline from median	-0.415 (0.166)	-0.530 (0.200)	-0.525 (0.447)	-0.232 (0.314)
outright owner	-0.089 (0.078)	0.027 (0.077)	-0.017 (0.189)	-0.251 (0.179)
Married/Cohabiting	-0.066 (0.081)	0.084 (0.081)	0.023 (0.219)	-0.275 (0.200)
χ^2 (9) *for coefficient equality*	7.745		12.104	
N	4,596		850	
BHPS sample composition	***FRS matched to BHPS***		***ELSA matched to BHPS***	
	FRS	BHPS	ELSA	BHPS
Latent disability η	0.519 (0.096)	0.530 (0.093)	0.566 (0.100)	0.510 (0.101)
Female	-0.115 (0.178)	-0.131 (0.189)	0.0590 (0.203)	-0.128 (0.191)
Spline age 65-73	-0.005 (0.031)	-0.001 (0.035)	-0.038 (0.023)	-0.047 (0.029)
Spline age 73+	0.048 (0.017)	0.026 (0.016)	0.057 (0.016)	0.032 (0.017)
Post- compulsory education	-0.076 (0.172)	0.147 (0.174)	-0.388 (0.210)	0.050 (0.181)
Income spline to median	-0.223 (0.329)	-0.692 (0.356)	-0.265 (0.206)	-0.381 (0.206)
Income spline from median	-0.524 (0.371)	-0.334 (0.276)	0.131 (0.382)	-0.318 (0.308)
outright owner	-0.259 (0.168)	-0.302 (0.168)	0.011 (0.202)	-0.289 (0.178)
Married/Cohabiting	-0.021 (0.190)	-0.031 (0.184)	-0.095 (0.206)	-0.103 (0.185)
χ^2 (9) *for coefficient equality*	3.444		6.581	
N	966		791	

Note: Significance of cross-sample coefficient difference and χ^2 statistic: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

6 Robustness

6.1 The number of factors

In the estimated 1-factor measurement models of Table 1, there is a strikingly low correlation between the latent disability index and those indicators which might be thought to represent cognitive rather than physical disability. To allow for a distinction between physical and cognitive disability, we have also estimated a 2-factor model for each sample, following an exploratory factor analysis of the disability indicators. The attempt failed for the BHPS, where only a single factor could be detected, arguably because the BHPS disability questions lack completeness and have poor sensitivity to the cognitive dimension of disability. For the FRS and ELSA 2-factor models can be estimated (see Tables A7-A9 in the Appendix). The second factor appears to distinguish satisfactorily the cognitive aspect of disability for the FRS where difficulties in communication, in memory/concentration/learning/understanding and in recognising physical danger are fairly obviously related to cognitive functioning. Since incontinence could stem from physical and/or cognitive problems, we allow for a cross-loading between the 2 factors for difficulties with continence. In ELSA, the second factor is determined from four cognitively-demanding IADLs (using a map, telephone use, self-medication, and handling finances) and, as for the FRS, we allow a cross-loading for continence. It is well known that there are limitations in the extent to which IADLs capture difficulties in cognitive functioning (Cromwell *et al.* 2003). We find the two factors to be strongly correlated (a similar result for the US is reported by Wallace and Herzog 1995). In the 2-factor latent disability equations (Table A8) the estimated coefficients for the first factor are close to those found in the 1-factor model for ELSA but are generally lower for the FRS, particularly for age and home-ownership. Using unmatched samples, we can reject the hypothesis of equal coefficients in the FRS and ELSA models for latent disability factor 1 but not factor 2 (Table A8). Results in Table A9 suggest a larger role for physical than cognitive influences on AA receipt with statistically insignificant differences between the estimated coefficients in the two surveys (*P*-values 0.133 and 0.185, respectively). The 2-factor specification confirms our previous findings on the relationship of AA receipt to socio-economic characteristics, since tests of coefficient equality do not reject the null hypothesis that coefficients ($\boldsymbol{\beta}$) of the observed covariates in the 2-factor models are equal to those obtained with the 1-factor specification in both surveys. The estimated coefficients of the 2-factor models are similar in size for FRS and ELSA. Based on a Wald-test, we reject the hypothesis of equality for the full AA coefficient vector ($\boldsymbol{\beta}$, $\boldsymbol{\gamma}$) (*P*-value = 0.011) but we do not reject for $\boldsymbol{\beta}$ alone (Wald *P*-value = 0.235). Cross-survey differences in the magnitude of the coefficients are not large and, for practical research purposes, one would draw essentially the same conclusions from the FRS and ELSA results.

6.2 Alternative normalisations

The 1-factor models set out above were estimated under the normalisation to unity of the factor loading associated with difficulties in mobility in each survey. Here we discuss the robustness of those findings to two alternative normalizations of η: in the first, we constrain an alternative factor loading; in the second, we set the residual variance of η equal to 1.

The comparability of estimates of the disability and AA equations can be improved by normalising the loadings of more similar questionnaire items. For instance, the FRS and ELSA have questions on the capacity to lift weights (variable LIFTING) which are arguably more similar than those on general mobility. When the factor loading for LIFTING is normalised to unity, the concordance between the FRS and ELSA disability equation and AA coefficients does indeed improve, with the Wald χ^2 *P*-values rising to 0.237 and 0.284 respectively (1-factor specification, unmatched samples). Details of the estimates are in appendix tables A10-A12. However, the scope of this exercise is reduced by the lack of a similar indicator in the BHPS.

6.3 Proxy cases in the FRS

Since we are forced to exclude proxy cases from the analysis of ELSA and BHPS, we investigate the consequences of also excluding them from the FRS and dropping the proxy indicator from the disability measurement equations (see Appendix Tables A13-A15). This has the effect of changing slightly the factor loadings on the other indicators. Nevertheless, all factor loadings remain positive and highly significant. The largest changes in loadings are for men, where the factor loading on lifting increases from 1.005 to 1.039, while those for memory problems and recognising when in danger fall from 0.420 to 0.356 and from 0.510 to 0.355 respectively. The estimated latent disability and AA receipt equations are not changed substantially. However, there are some small effects on the statistical significance of differences between the surveys in the estimated coefficients. In both the disability and the receipt of AA equations, after dropping proxy cases, the differences between the FRS and ELSA become smaller but increase slightly when FRS is contrasted with BHPS.

7 Conclusions

Our aim in this study is to contribute to the current policy debate over reform prospects for the social care system by investigating the robustness of survey-based evidence on the targeting of public support for older people with disabilities. We have examined the three UK surveys (FRS, ELSA and BHPS) which have been the basis for much of the empirical analysis underpinning the debate on policy on disability in the pensioner population. Despite substantial differences between the three surveys in terms of their design, non-response characteristics and questionnaire content, we have found that they have a coherent story to tell about the targeting of one form of public support in relation to disability, income and other personal and household characteristics.

We also claim to offer some advance in terms of the statistical modeling methodology typically used in the disability research literature. Adopting a latent variable approach, we are able to exploit the existence of multiple – but largely arbitrary and individually unreliable – survey indicators, whilst avoiding the common practice of using ad hoc count indexes as disability measures. Results confirm that the probability of receiving AA increases strongly with the severity of disability and decreases with income – especially for those in the top half of the income distribution – after allowing for the socio-economic gradient in health that associates higher living standards with lower disability. Contrary to some suggestions, we can say there is no evidence of people receiving AA without any disability revealed by their survey interview. In allowing for two latent disability factors we find evidence from the FRS and ELSA that physical disability has a larger influence on AA receipt than cognitive disability. Limitations in the BHPS survey instrument meant that we were unable to confirm this in the BHPS. Our use of Mahalanobis matching to improve comparability by removing differences in sample composition also provides a valuable reminder of the need to consider sample coverage as a factor when reviewing a range of research findings.

References

Banks, J., Kapteyn, A., Smith, J. P. and van Soest, A. (2009) Work Disability is a Pain in the ****, Especially in England, the Netherlands, and the United States. In: *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly* (eds D. M. Cutler and D.A. Wise), pp. 251-293. National Bureau of Economic Research.

Banks, J., Marmot, M., Oldfield, Z. and J.P. and Smith, J.P. (2006) Disease and disadvantage in the United States and in England. *Journal of the American Medical Association*, **295**, 2037–2045.

Benítez-Silva, H., Disney, R. and Jiménez-Martín, S. (2009) Disability, capacity for work and the business cycle: an international perspective, *Fundacin de Estudios de Economa Aplicada Working Papers*, 2009-28.

Berthoud, R. and Hancock, R. M. (2008) Disability benefits and paying for care, in *Advancing Opportunity: Social Care* (ed N. Churchill). London: The Smith Institute.

Bollen, K. (1989) *Structural Equations with Latent Variables*, New York: Wiley.

Bound, J. (1991) Self-Reported versus Objective Measures of Health in Retirement Models. *Journal of Human Resources* **26**(1), 106-138.

Cabinet Office (2010) *State of the Nation Report: Poverty, Worklessness and Welfare Dependency in the UK*. London: Cabinet Office.

Campbell, A. (2004) *Family Resources Survey: Annual Technical Report: 2002/03*, London: Office for National Statistics. Available at: http://www.statistics.gov.uk/downloads/theme_social/FRS_Tech0203.pdf

Chan, K.S., Kasper, J.D. and Pezzin, L.E. (2012) Measurement equivalence in ADL and IADL difficulty across international surveys of aging: findings from the HRS, SHARE, and ELSA. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, **67**(1), 121-132.

Clarke, P. and Smith, J.P. (2011) Aging in a cultural context: cross-national differences in disability and the moderating role of personal control among older adults in the United States and England. *Journal of Gerontology: Medical Sciences*, **66**(4), 457-67.

Comas-Herrera, A., Wittenberg, R and Pickard, L. (2010) The long road to universalism? Recent Developments in the Financing of Long-term Care in England. *Social Policy and Administration*, **44** (4), 375-391

Commission on Funding Care and Support (2011) *Fairer care funding: the report of the commission on funding care and support*. Available at: https://www.wp.dh.gov.uk/carecommission/files/2011/07/Fairer-Care-Funding-Report.pdf

Crimmins, E. M., Ki Kim J. and Solé-Auró A. (2011) Gender differences in health: results from SHARE, ELSA and HRS. *European Journal of Public Health*, **21**(1), 81-91.

Cromwell, D. A., Eagar K. and Poulos R. G. (2003) The performance of instrumental activities of daily living scale in screening for cognitive impairment in elderly community residents. *Journal of Clinical Epidemiology*, **56**(2), 131-137

Department of Health (2009) *Shaping the Future of Care Together*. Available at: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_102338

Department of Health (2013) *Policy statement on care and support funding reform and legislative requirements*. London: Department of Health. Available at: http://www.dh.gov.uk/health/files/2013/02/Policy-statement-on-funding-reform.pdf

Department for Work and Pensions (2013) *AA claim form*. Available at: http://www.dwp.gov.uk/advisers/claimforms/aa1a_print.pdf

Forder, J. and Fernandez, J-L. (2009) *Analysing the costs and benefits of social care funding arrangements in England: technical report*. London: LSE. PSSRU Discussion Paper 2644. Available at: http://eprints.lse.ac.uk/24977/

Goldman, N. (2001) Social inequalities in health: disentangling the underlying mechanisms. In *Strengthening the Dialogue between Epidemiology and Demography* (eds M. Weinstein and A. Hermalin), Annals of the New York Academy of Sciences.

Gove, W. R. (1984) Gender differences in mental and physical illness; the effect of fixed roles and nurturant roles. *Social Science and Medicine,* **19**, 77-84.

Groot, W. (2000) Adaptation and scale reference bias in self assessments of quality of life. *Journal of Health Economics,* **19**, 403–420.

Hancock, R. M. and Barker, G. (2005) The quality of social security benefit data in the British Family Resources survey: implications for investigating income support take-up by pensioners. *J. R. Statist. Soc. A,* **168**, 63-82.

Hernandez, M., Pudney, S. E. and Hancock, R. M. (2007) The welfare cost of means-testing: pensioner participation in Income Support. *J. Appl. Econometrics,* **22**, 581-598.

Hibbard, J. H. and Pope, C. R. (1983) Gender roles, illness orientation and use of medical services. *Social Science and Medicine,* **17**, 129-137.

Hirst, M. (2004) The British Household Panel Survey: a longitudinal perspective on informal care (eds S. Becker and A. Bryman) *Understanding Research for Social Policy and Practice,* pp.190-193. Bristol: The Policy Press.

Jagger C., Matthews R., King D., Comas-Herrera A., Grundy E., Stuchbury R, Morciano M., Hancock R. and the MAP2030 team (2009) *Calibrating disability measures across UK national surveys*. Report prepared for Department of Work and Pensions. Available at: http://www.esrc.ac.uk/my-esrc/grants/RES-339-25-0002/outputs/read/04254e47-90ff-4f77-af7a-0533490acd3c

Johnson, R.J. and Wolinsky F.D. (1993) The Structure of Health Status Among Older Adults: Disease, Disability, Functional Limitation, and Perceived Health*.* *Journal of Health and Social Behavior*, **34** (2), 105-121.

Kasparova D., Marsh A. and Wilkinson D. (2007) *The take-up rate of Disability Living Allowance and Attendance Allowance: feasibility study*. Research Report No 442. London: Department for Work and Pensions.

Katz S., Ford A.B., Moskowitz R.W., Jackson B.A., Jaffe M.W. and Cleveland M.A. (1963) Studies of Illness in the Aged. The Index of ADL: A Standardized Measure of Biological and Psychosocial Function. *J. Am. Medical Ass.* **185**(12), 914-919.

Karlsson M., Mayhew, L., Plumb, R. and Rickayzen B. (2006) Future cost for long term care: cost projections for long term care for older people in the United Kingdom. *Health Policy*, **75**, 187–213.

Lawton M.P. and Brody E.M. (1969) Assessment of older people: self-maintaining and instrumental activities of daily living. *The Gerontologist,* **9**(3), 179-186.

Leuven, E. and Sianesi B. (2003) *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. Available at: http://ideas.repec.org/c/boc/bocode/s432001.html.

Lindeboom, M. and van Doorslaer, E. (2004) Cut-point shift and index shift in self-reported health. *Journal of Health Economics,* **23**, 1083–1099.

Lynn, P. (ed.), with Buck, N., Burton, J., Laurie, H. and Uhrig S.C.N. (2006) *Quality Profile: British Household Panel Survey: Waves 1 to 13, 1991-2003.* Institute for Social and Economic Research, University of Essex, Colchester. Available at: https://www.iser.essex.ac.uk/files/bhps/quality-profiles/BHPS-QP-01-03-06-v2.pdf.

Mayhew, L., Karlsson, M. and Rickayzen B. (2010) The Role of Private Finance in Paying for Long Term Care. *The Economic Journal*, **120**(548), F478–F504.

McVicar, D. (2008) Why have UK disability benefit rolls grown so much? *Journal of Economic Surveys* **22**, 114-139.

Melzer, D. Gardener, E. and Guralnik J. (2005) Mobility disability in the middle-aged: cross-sectional associations in the English Longitudinal Study of Ageing, *Age and Ageing*, **34** (6), 594-602.

Morciano, M. Hancock, R. and Pudney S. (2013) Disability costs and equivalence scales in the older population, *Review of Income and Wealth*, forthcoming.

OECD (2005) Long Term Care for Older People, Paris: OECD Publishing.

Oswald, A. J. and Powdthavee N., (2008) Does happiness adapt? A longitudinal study of disability with implications for economists and judges. *Journal of Public Economics*, **92**(5-6), 1061-1077.

Pickard, L, A. Comas-Herrera, J. Costa-Font, C. Gori, A. Di Maio, C. Patxot, A. Pozzi, H. and Wittenberg R. (2007) Modelling an Entitlement to Long-Term Care Services for Older People in Europe: Projections for Long-Term Care Expenditure to 2050. *Journal of European Social Policy* **17**, 33–48.

Pudney, S. Hancock, R. and Sutherland H. (2006) Simulating the reform of means-tested benefits with endogenous take-up and claim costs. *Oxford Bulleting of Economics and Statistics* **68** (2) 135-166

Pudney, S. (2009) *Participation in disability benefit programmes. A partial identification analysis of the British Attendance Allowance system*, ISER Working Paper 2009-19. Institute for Social and Economic Research, University of Essex, Colchester.

Pudney, S. E. (2010) *Disability benefits for older people: How does the UK Attendance Allowance system really work?* University of Essex: ISER Working Paper no. 2010-02.

Satorra, A. and Bentler P. M. (2001) A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* **66**, 507−514.

Sen, A. K. (1982) *Choice, Welfare and Measurement*. Oxford: Blackwell.

Sen, A. K. (1985) *Commodities and Capabilities*. Amsterdam: North Holland.

Sturgis, P., Thomas, R., Purdon, S., Bridgwood, A. and Dodd T. (2001) *Comparative review and assessment of key health state measures of the general population*. Research report, London: Department of Health, UK.

Sutherland, S. (ed.) (1999) *With Respect to Old Age: Long-term Care – Rights and Responsibilities. A Report by the Royal Commission on Long Term Care*. London: The Stationery Office CM4192.

Taylor, M.F. (ed.), with Brice, J., Buck, N. and Prentice-Lane E. (2006) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Institute for Social and Economic Research, University of Essex, Colchester. Available at http://www.iser.essex.ac.uk/ulsc/bhps/doc/pdf_versions/ volumes/bhpsvola.pdf.

Taylor, R., Conway, L., Calderwood, L., and Lessof C. (2003) Methodology. In *Health, wealth and lifestyles of the older population in England: The 2002 English Longitudinal Study of Ageing* (eds M. Marmot, J. Banks, R. Blundell, C. Lessof and J. Nazroo), London :The Institute for Fiscal Studies, pp. 357–374.

Verbrugge, L. M. (1980) Sex differences in complaints and diagnoses. *Journal of Behavioural Medicine*, **3**, 327-355.

Verbrugge, L. M. (1985). Gender and health: Update on hypotheses and evidence. *Journal of Health and Social Behaviour*, **26**, 156–182.

Verbrugge, L. M. and Wingard D. L. (1987) Sex differentials in health and mortality. *Women’s Health*, **12**,103-145.

Verbrugge L. M. (1989) Recent, present and future health of American adults. *Annual Review of Public Health*, **10**, 33-61.

Wallace R.B. and Herzog A. R. (1995) Overview of the Health Measures in the Health and Retirement Study. *The Journal of Human Resources*, Special Issue on the Health and Retirement Study: Data Quality and Early Results, **30**:S84-S107

Wanless, D. (2006) Securing Good Care for Older People: Taking a Long-Term View. London: King`s Fund. Available at: http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/securing-good-care-for-older-people-wanless-2006.pdf

Zaninotto, P. and Falaschetti E. (2011) Comparison of methods for modelling a count outcome with excess zeros: application to Activities of Daily Living (ADL-s). *Journal of Epidemiology and Community Health*, **65** (3), 205-210.

Zantomio, F. (2013) Older people's participation in extra-cost disability benefits. *Journal of Health Economics*, **32** (1), 320-330.

Appendix 1: Additional tables

Table A1: Survey specific functional limitations indicators D

Data Source:		not receiving AA		receiving AA		Difference†
		mean	sd	mean	sd	
FRS:						
Has difficulty with:						
MOBILITY	mobility (moving about)	0.251	0.434	0.814	0.389	-0.563
LIFTING	lifting, carrying or moving objects	0.221	0.415	0.745	0.436	-0.524
DEXTERITY	manual dexterity using hands for everyday tasks	0.077	0.267	0.396	0.490	-0.319
CONTINENCE	with continence (bladder control)	0.055	0.228	0.237	0.425	-0.182
COMMUNICATION	communication (speech, hearing or eyesight)	0.039	0.194	0.204	0.404	-0.165
MEMORY	memory/concentration/learning/understanding	0.049	0.217	0.252	0.434	-0.203
KNOWING DANGER	recognising when in physical danger	0.005	0.073	0.068	0.251	-0.062
OTHER	other area of life	0.040	0.195	0.092	0.289	-0.053
PROXY	interviewed by proxy	0.059	0.235	0.121	0.327	-0.063
Observations		*6,093*		*651*		
***ELSA:** Has difficulty with:*						
WALKING 100 YDS	walking 100 yards	0.117	0.321	0.572	0.495	-0.455
SITTING 2 HRS	sitting for about two hours	0.126	0.332	0.285	0.452	-0.158
CHAIR TRANSFERS	getting up from a chair after sitting for long periods	0.282	0.450	0.626	0.485	-0.344
STAIRS (several flights)	climbing several flights of stairs without resting	0.424	0.494	0.821	0.384	-0.397
STAIRS (1 flights)	climbing one flight of stairs without resting	0.161	0.368	0.650	0.477	-0.489
STOOPING	stooping, kneeling, or crouching	0.411	0.492	0.791	0.407	-0.381
REACHING	reaching or extending arms above shoulder level	0.103	0.304	0.344	0.476	-0.241
PULL/PUSHING	pulling or pushing large objects e.g. living room chair	0.183	0.387	0.675	0.469	-0.492
LIFTING	lifting/carrying weights over 10 lbs, e.g. heavy bag	0.281	0.449	0.797	0.403	-0.516
PICKING-UP COIN	picking up a 5p coin from a table	0.049	0.216	0.241	0.428	-0.192
DRESSING	ADL:dressing, including putting on shoes an	0.126	0.332	0.472	0.500	-0.346
WALKING	ADL:walking across a room	0.025	0.157	0.203	0.403	-0.178
BATHING	ADL:bathing or showering	0.128	0.334	0.566	0.496	-0.438
FEEDING	ADL:eating, such as cutting up your food	0.012	0.110	0.092	0.290	-0.08
BED TRANSFERS	ADL:getting in or out of bed	0.044	0.205	0.287	0.453	-0.243
USING TOILET	ADL:using the toilet, including getting up	0.029	0.167	0.179	0.384	-0.15
CONTINENCE	Problem with continence	0.157	0.364	0.336	0.473	-0.179
USING MAP	IADL:using a map to figure out how to get around	0.057	0.231	0.222	0.416	-0.165
PREP HOT MEAL	IADL:preparing a hot meal	0.029	0.167	0.282	0.451	-0.253
SHOPPING	IADL:shopping for groceries	0.083	0.275	0.504	0.501	-0.422
PHONING	IADL:making telephone calls	0.020	0.139	0.095	0.293	-0.075
MEDICATION	IADL:taking medications	0.010	0.102	0.084	0.278	-0.073
HOUSEWORK	IADL:doing work around the house or garden	0.159	0.366	0.650	0.477	-0.491
MANAGING MONEY	IADL: managing money, e.g. paying bills	0.023	0.151	0.154	0.362	-0.131
Observations		*4,773*		*369*		
***BHPS:** Health hinders:*						
HOUSEWORK	doing the housework	0.089	0.285	0.573	0.498	-0.484
CLIMBING STAIRS	climbing the stairs	0.105	0.307	0.600	0.493	-0.495
DRESSING	getting dressed	0.036	0.187	0.173	0.381	-0.137
WALKING > 10 mins	walking more than 10 mins	0.094	0.292	0.520	0.503	-0.426
How manages ..(6-point scale)						
STAIRS	Stairs	1.856	1.132	3.920	1.566	-2.064
AROUND HOUSE	getting around house	1.350	0.735	2.613	1.345	-1.264
BED TRANSFERS	getting in/out bed	1.360	0.721	2.547	1.233	-1.187
CUTTING TOENAILS	cutting toenails	2.555	1.792	4.920	1.333	-2.365
BATHING	bathing/showering	1.572	1.002	3.280	1.564	-1.708
WALKING DOWN ROAD	walking down road	1.678	1.163	3.773	1.729	-2.095
Observations		*967*		*75*		

† All differences are significantly different from 0 at the 1% level.

Table A2: Sample means of SES and AA receipt in FRS, ELSA and BHPS

	FRS		ELSA		BHPS	
	mean	sd	mean	sd	mean	sd
Female	0.559	*0.497*	0.557	*0.497*	0.560	*0.497*
Age	74.537	*6.692*	74.154	*6.643*	74.698	*6.437*
Post-compulsory education	0.505	*0.500*	0.539	*0.499*	0.513	*0.500*
Ln pre-benefit equivalised income†	6.454	*0.806*	6.412	*0.751*	6.551	*0.732*
Outright owner	0.664	*0.472*	0.690	*0.463*	0.701	*0.458*
Married/cohabiting	0.579	*0.494*	0.565	*0.496*	0.553	*0.497*
Receives AA	0.097	*0.295*	0.072	*0.258*	0.072	*0.259*
Observations	6,746		5,142		1,042	

Notes: Based on unweighted selected samples. † Household income excludes disability and means tested benefits and it has been equivalised using the modified OECD equivalence scale.

Table A3: Sample means of SES and AA receipt in in the post-matched samples

FRS sample composition

	FRS		ELSA		FRS		BHPS	
	mean	sd	mean	sd	mean	sd	mean	sd
Female	0.561	*0.496*	0.561	*0.496*	0.566	*0.496*	0.566	*0.496*
Age	73.972	*6.512*	73.973	*6.509*	74.458	*6.283*	74.443	*6.259*
Post-compulsory schooling	0.530	*0.499*	0.530	*0.499*	0.506	*0.500*	0.506	*0.500*
ln pre-benefit equivalised income†	6.457	*0.582*	6.456	*0.582*	6.576	*0.503*	6.600	*0.500*
Accommodation own it outright	0.690	*0.462*	0.690	*0.462*	0.716	*0.451*	0.716	*0.451*
Married/cohabiting	0.572	*0.495*	0.572	*0.495*	0.565	*0.496*	0.565	*0.496*
Receives AA	0.088	*0.283*	0.071	*0.257*	0.094	*0.291*	0.072	*0.259*
Observations	4,587				973			

ELSA sample composition

	FRS		ELSA		ELSA		BHPS	
	mean	sd	mean	sd	mean	sd	mean	sd
Female	0.562	*0.496*	0.562	*0.496*	0.575	*0.495*	0.575	*0.495*
Age	73.961	*6.497*	73.976	*6.500*	74.484	*6.352*	74.502	*6.358*
Post-compulsory schooling	0.531	*0.499*	0.531	*0.499*	0.504	*0.500*	0.504	*0.500*
ln pre-benefit equivalised income†	6.458	*0.578*	6.455	*0.582*	6.563	*0.513*	6.533	*0.527*
accommodation own it outright	0.690	*0.463*	0.690	*0.463*	0.720	*0.449*	0.720	*0.449*
Married/cohabiting	0.574	*0.495*	0.574	*0.495*	0.552	*0.498*	0.552	*0.498*
Receives AA	0.089	*0.284*	0.070	*0.255*	0.072	*0.258*	0.066	*0.248*
Observations	4,596				850			

BHPS sample composition

	FRS		BHPS		ELSA		BHPS	
	mean	sd	mean	sd	mean	sd	mean	sd
Female	0.565	*0.496*	0.565	*0.496*	0.564	*0.496*	0.564	*0.496*
Age	74.404	*6.255*	74.404	*6.247*	74.498	*6.412*	74.469	*6.399*
Post-compulsory schooling	0.505	*0.500*	0.505	*0.500*	0.497	*0.500*	0.497	*0.500*
ln pre-benefit equivalised income†	6.575	*0.499*	6.599	*0.496*	6.488	*0.496*	6.513	*0.500*
accommodation own it outright	0.716	*0.451*	0.716	*0.451*	0.718	*0.450*	0.718	*0.450*
Married/cohabiting	0.566	*0.496*	0.566	*0.496*	0.550	*0.498*	0.550	*0.498*
Receives AA	0.085	*0.279*	0.072	*0.259*	0.068	*0.252*	0.078	*0.269*
Observations	966				791			

Notes: Based on unweighted selected samples. † Household income excludes disability and means tested benefits and it has been equivalised using the modified OECD equivalence scale.

Table A4: Factor loadings 1-factor models, post-matched samples, FRS sample composition

FRS			ELSA		
Disability Indicator	Factor loading (St. err.) Men	Factor loading (St. err.) Women	Disability Indicator	Factor loading (St. err.) Men	Factor loading (St. err.) Women
MOBILITY	1 (-)	1 (-)	WALKING 100 YDS	1 (-)	1 (-)
LIFTING	1.022† (0.111)	1.216† (0.133)	SITTING 2 HRS	0.384† (0.033)	0.387† (0.030)
DEXTERITY	0.666† (0.069)	0.643† (0.057)	CHAIR TRANSFERS	0.587† (0.042)	0.524† (0.035)
CONTINENCE	0.410† (0.046)	0.429† (0.042)	STAIRS (several flights)	0.695† (0.049)	0.670† (0.044)
COMMUNIC	0.407† (0.051)	0.351† (0.044)	STAIRS (1 flight)	1.007† (0.071)	0.880† (0.055)
MEMORY	0.434† (0.050)	0.405† (0.044)	STOOPING	0.641† (0.045)	0.637† (0.041)
DANGER	0.551† (0.119)	0.437† (0.071)	REACHING	0.513† (0.045)	0.507† (0.038)
OTHER	0.092† (0.032)	0.057‡ (0.029)	PULL/PUSHING	1.027† (0.085)	0.897† (0.059)
PROXY	0.112† (0.034)	0.088† (0.029)	LIFTING	0.927† (0.070)	0.876† (0.060)
			PICKING-UP COIN	0.389† (0.049)	0.436† (0.039)
			DRESSING	0.658† (0.051)	0.650† (0.046)
			WALK ACROSS ROOM	1.022† (0.136)	0.934† (0.091)
			BATHING	0.853† (0.071)	0.711† (0.049)
			FEEDING	0.586† (0.091)	0.451† (0.061)
			BED TRANSFERS	0.909† (0.088)	0.682† (0.057)
			USING TOILET	0.767† (0.096)	0.592† (0.055)
			CONTINENCE	0.292† (0.032)	0.240† (0.023)
			USING A MAP	0.422† (0.053)	0.327† (0.030)
			PREP. HOT MEAL	0.817† (0.109)	0.803† (0.079)
			SHOPPING	1.065† (0.094)	1.121† (0.084)
			PHONING	0.345† (0.048)	0.327† (0.049)
			MEDICATION	0.500† (0.076)	0.455† (0.074)
			HOUSEWORK	1.169† (0.096)	0.907† (0.063)
			MANAGING MONEY	0.452† (0.060)	0.486† (0.051)
Sample size: 4,587					

FRS			BHPS		
Disability Indicator	Factor loading (St. err.) Men	Factor loading (St. err.) Women	Disability Indicator	Factor loading (St. err.) Men	Factor loading (St. err.) Women
MOBILITY	1 (-)	1 (-)	HOUSEWORK	0.841† (0.14)	0.937† (0.145)
LIFTING	1.109† (0.287)	1.578† (0.454)	CLIMBING STAIRS	0.976† (0.145)	1.172† (0.167)
DEXTERITY	0.708† (0.150)	0.766† (0.149)	DRESSING	0.621† (0.116)	0.936† (0.179)
CONTINENCE	0.400† (0.092)	0.601† (0.147)	WALKING > 10 mins	1 (-)	1 (-)
COMMUNIC	0.441† (0.114)	0.457† (0.135)	STAIRS	1.087† (0.169)	0.920† (0.124)
MEMORY	0.437† (0.117)	0.479† (0.113)	AROUND HOUSE	1.329† (0.256)	1.050† (0.153)
DANGER	0.701 (0.615)	0.396‡ (0.171)	BED TRANSFERS	1.316† (0.237)	0.959† (0.141)
OTHER	0.067 (0.094)	0.084 (0.066)	CURRING TOENAILS	0.580† (0.081)	0.588† (0.077)
PROXY	0.085 (0.080)	0.048 (0.072)	BATHING	0.991† (0.153)	0.762† (0.104)
			WALK DOWN ROAD	1.143† (0.169)	1.115† (0.157)
Sample size: 973					

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table A5: Factor loadings 1-factor models, post-matched samples, ELSA sample composition

ELSA			FRS		
Disability Indicator	Factor loading (St. err.) Men	Factor loading (St. err.) Women	Disability Indicator	Factor loading (St. err.) Men	Factor loading (St. err.) Women
WALKING 100 YDS	1 (-)	1 (-)	MOBILITY	1 (-)	1 (-)
SITTING 2 HRS	0.384† (0.033)	0.38† (0.029)	LIFTING	1.028† (0.112)	1.118† (0.110)
CHAIR TRANSFERS	0.584† (0.042)	0.511† (0.034)	DEXTERITY	0.736† (0.078)	0.660† (0.060)
STAIRS (several flights)	0.678† (0.048)	0.647† (0.044)	CONTINENCE	0.413† (0.046)	0.449† (0.044)
STAIRS (1 flight)	0.980† (0.069)	0.856† (0.053)	COMMUNIC	0.352† (0.048)	0.384† (0.048)
STOOPING	0.615† (0.043)	0.612† (0.039)	MEMORY	0.453† (0.053)	0.453† (0.048)
REACHING	0.516† (0.046)	0.502† (0.038)	DANGER	0.525† (0.112)	0.428† (0.066)
PULL/PUSHING	1.035† (0.087)	0.890† (0.059)	OTHER	0.095† (0.033)	0.044 (0.03)
LIFTING	0.935† (0.07)	0.853† (0.058)	PROXY	0.118† (0.035)	0.087† (0.03)
PICKING-UP COIN	0.384† (0.049)	0.410† (0.037)			
DRESSING	0.638† (0.049)	0.639† (0.045)			
WALK ACROSS ROOM	0.998† (0.13)	0.880† (0.087)			
BATHING	0.842† (0.07)	0.696† (0.048)			
FEEDING	0.598† (0.093)	0.432† (0.059)			
BED TRANSFERS	0.882† (0.085)	0.677† (0.058)			
USING TOILET	0.754† (0.096)	0.558† (0.051)			
CONTINENCE	0.282† (0.031)	0.236† (0.022)			
USING A MAP	0.398† (0.052)	0.329† (0.03)			
PREP. HOT MEAL	0.820† (0.111)	0.794† (0.078)			
SHOPPING	1.039† (0.092)	1.092† (0.082)			
PHONING	0.353† (0.048)	0.323† (0.049)			
MEDICATION	0.496† (0.076)	0.471† (0.079)			
HOUSEWORK	1.127† (0.092)	0.896† (0.062)			
MANAGING MONEY	0.446† (0.059)	0.481† (0.05)			
Sample size: 4,596					

ELSA			BHPS		
Disability Indicator	Factor loading (St. err.) Men	Factor loading (St. err.) Women	Disability Indicator	Factor loading (St. err.) Men	Factor loading (St. err.) Women
WALKING 100 YDS	1 (-)	1 (-)	HOUSEWORK	0.839† (0.149)	1.014† (0.169)
SITTING 2 HRS	0.522† (0.1)	0.475† (0.081)	CLIMBING STAIRS	1.053† (0.168)	1.182† (0.184)
CHAIR TRANSFERS	0.772† (0.121)	0.512† (0.083)	DRESSING	0.721† (0.145)	0.838† (0.162)
STAIRS (several flights)	0.749† (0.121)	0.734† (0.118)	WALKING > 10 mins	1 (-)	1 (-)
STAIRS (1 flight)	1.130† (0.193)	0.837† (0.119)	STAIRS	1.087† (0.173)	0.995† (0.14)
STOOPING	0.708† (0.112)	0.709† (0.103)	AROUND HOUSE	1.587† (0.356)	1.155† (0.18)
REACHING	0.585† (0.119)	0.670† (0.114)	BED TRANSFERS	1.307† (0.239)	1.043† (0.165)
PULL/PUSHING	1.451† (0.293)	0.889† (0.137)	CURRING TOENAILS	0.637† (0.097)	0.602† (0.083)
LIFTING	1.132† (0.210)	0.863† (0.131)	BATHING	0.980† (0.168)	0.798† (0.115)
PICKING-UP COIN	0.474† (0.154)	0.600† (0.122)	WALK DOWN ROAD	1.182† (0.177)	1.097† (0.158)
DRESSING	0.646† (0.115)	0.830† (0.131)			
WALK ACROSS ROOM	2.197† (0.833)	1.037† (0.210)			
BATHING	0.936† (0.176)	0.792† (0.119)			
FEEDING	1.153† (0.344)	0.555† (0.151)			
BED TRANSFERS	0.990† (0.252)	0.726† (0.132)			
USING TOILET	0.687‡ (0.280)	0.736† (0.147)			
CONTINENCE	0.340† (0.087)	0.311† (0.061)			
USING A MAP	0.754† (0.194)	0.264† (0.067)			
PREP. HOT MEAL	2.084† (0.676)	0.924† (0.179)			
SHOPPING	1.706† (0.436)	1.130† (0.191)			
PHONING	0.219‡ (0.091)	0.414† (0.129)			
MEDICATION	0.635† (0.181)	0.925† (0.327)			
HOUSEWORK	1.404† (0.273)	0.973† (0.145)			
MANAGING MONEY	0.583† (0.161)	0.552† (0.139)			
Sample size: 850					

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table A6: Factor loadings 1-factor models, post-matched samples, BHPS sample composition

BHPS			FRS		
	Factor loading (St. err.)		Disability	Factor loading (St. err.)	
Disability Indicator	Men	Women	Indicator	Men	Women
HOUSEWORK	0.829† (0.129)	0.965† (0.151)	MOBILITY	1 (-)	1 (-)
CLIMBING STAIRS	0.943† (0.129)	1.204† (0.174)	LIFTING	0.824† (0.212)	1.179† (0.267)
DRESSING	0.622† (0.112)	0.966† (0.185)	DEXTERITY	0.696† (0.167)	0.725† (0.149)
WALKING > 10 mins	1 (-)	1 (-)	CONTINENCE	0.494† (0.118)	0.404† (0.098)
STAIRS	1.146† (0.183)	0.964† (0.131)	COMMUNIC	0.419† (0.122)	0.350† (0.097)
AROUND HOUSE	1.415† (0.287)	1.089† (0.159)	MEMORY	0.374† (0.099)	0.499† (0.113)
BED TRANSFERS	1.404† (0.268)	0.994† (0.147)	DANGER	1.205§ (0.701)	0.391‡ (0.163)
CURRING TOENAILS	0.616† (0.088)	0.606† (0.08)	OTHER	0.139§ (0.075)	0.024 (0.067)
BATHING	1.047† (0.169)	0.798† (0.11)	PROXY	0.133§ (0.072)	0.061 (0.068)
WALK DOWN ROAD	1.201† (0.183)	1.155† (0.164)			
Sample size: 966					

BHPS			ELSA		
Disability	Factor loading (St. err.)		Disability Indicator	Factor loading (St. err.)	
Indicator	Men	Men	Women	Men	Men
HOUSEWORK	0.978† (0.172)	1.169† (0.176)	WALKING 100 YDS	1 (-)	1 (-)
CLIMBING STAIRS	1.061† (0.175)	1.268† (0.208)	SITTING 2 HRS	0.406† (0.084)	0.417† (0.08)
DRESSING	0.634† (0.135)	1.034† (0.218)	CHAIR TRANSFERS	0.569† (0.109)	0.610† (0.105)
WALKING > 10 mins	1 (-)	1 (-)	STAIRS (several flights)	0.932† (0.16)	0.712† (0.121)
STAIRS	1.062† (0.182)	0.932† (0.139)	STAIRS (1 flight)	0.932† (0.157)	0.794† (0.12)
AROUND HOUSE	1.378† (0.292)	1.069† (0.172)	STOOPING	0.565† (0.105)	0.640† (0.104)
BED TRANSFERS	1.248† (0.243)	1.022† (0.167)	REACHING	0.403† (0.099)	0.551† (0.111)
CURRING TOENAILS	0.593† (0.092)	0.579† (0.082)	PULL/PUSHING	1.503† (0.342)	0.793† (0.133)
BATHING	0.979† (0.166)	0.781† (0.117)	LIFTING	1.245† (0.246)	0.728† (0.125)
WALK DOWN ROAD	1.153† (0.18)	1.108† (0.172)	PICKING-UP COIN	0.505† (0.152)	0.352† (0.088)
			DRESSING	0.642† (0.127)	0.779† (0.14)
			WALK ACROSS ROOM	0.682† (0.206)	0.806† (0.17)
			BATHING	0.795† (0.151)	0.742† (0.129)
			FEEDING	1.202† (0.362)	0.497† (0.166)
			BED TRANSFERS	0.905† (0.229)	0.711† (0.127)
			USING TOILET	0.553† (0.154)	0.831† (0.184)
			CONTINENCE	0.366† (0.081)	0.341† (0.068)
			USING A MAP	0.439† (0.134)	0.267† (0.068)
			PREP. HOT MEAL	1.307† (0.425)	0.679† (0.14)
			SHOPPING	1.310† (0.326)	1.105† (0.234)
			PHONING	0.314† (0.106)	0.369† (0.112)
			MEDICATION	0.443† (0.125)	0.409† (0.15)
			HOUSEWORK	1.547† (0.347)	0.969† (0.159)
			MANAGING MONEY	0.548† (0.127)	0.556† (0.134)
Sample size: 791					

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table A7: Factor loadings for the FRS and ELSA 2-factor models and squared correlations of disability indicators with latent indexes (*η_q*)

Functional limitation indicator	Male		Female	
	Factor 1 (η_1)	Factor 2 (η_2)	Factor 1 (η_1)	Factor 2 (η_2)
FRS cov(η_1,η_2)	1.172		0.854	
MOBILITY	1		1	
LIFTING	1.586†		2.226†	
DEXTERITY	0.768†		0.736†	
CONTINENCE	0.315†	0.235†	0.363†	0.275†
COMMUNIC		1		1
MEMORY		0.837†		0.987†
DANGER		1.005†		1.078†
OTHER	0.009	0.144‡	-0.064	0.208†
PROXY		0.204†		0.270†
ELSA cov(η_1,η_2)	1.058		0.890	
WALK100	1		1	
SITTING	0.394†		0.409†	
CHAIR	0.593†		0.545†	
CLIMBSEV	0.736†		0.689†	
CLIMB1	1.014†		0.918†	
STOOP	0.657†		0.669†	
ARMS	0.511†		0.511†	
PULL/PUSH	1.025†		0.921†	
LIFTING	0.954†		0.919†	
COIN	0.383†		0.44†	
DRESSING	0.673†		0.665†	
WALKING	1.082†		0.980†	
BATH	0.879†		0.736†	
EATING	0.586†		0.431†	
BED	0.897†		0.705†	
TOILET	0.751†		0.592†	
CONTINENCE	0.196†	0.235‡	0.275†	-0.047
MAP		1.052†		1.031†
MEAL				
SHOPPING	0.999†		1.129†	
PHONE		1		1
MEDICATION		1.231†		1.319†
HOUSEWORK	1.137†		0.938†	
MONEY		1.25†		1.731†

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table A8: Estimates of the latent disability equation for the FRS and ELSA 2-factor models

Covariates	η_1			η_2		
	FRS	ELSA	Tests and coefficient differences	FRS	ELSA	Tests and coefficient differences
Spline age 65-73	0.033†	0.035†	-0.002	0.025	-0.015	0.04§
	(0.003)	(0.011)	(0.011)	(0.016)	(0.013)	(0.021)
Spline from age 73+	0.064†	0.095†	-0.031†	0.079†	0.071†	0.008
	(0.005)	(0.006)	(0.008)	(0.008)	(0.009)	(0.012)
Post-compulsory education	-0.237†	-0.276†	0.039	-0.142§	-0.241†	0.1
	(0.05)	(0.057)	(0.076)	(0.075)	(0.069)	(0.102)
Income spline to median	-0.103†	-0.039	-0.063	-0.175†	-0.119†	-0.056
	(0.036)	(0.049)	(0.061)	(0.037)	(0.046)	(0.059)
Income spline from median	-0.293†	-0.305†	0.013	-0.086	-0.17§	0.084
	(0.071)	(0.068)	(0.098)	(0.1)	(0.089)	(0.134)
Outright owner	-0.334†	-0.484†	0.15§	-0.12§	-0.135‡	0.015
	(0.051)	(0.059)	(0.078)	(0.071)	(0.06)	(0.093)
coefficient equality $\chi^2(6)$			20.553†			7.662
N	6744			5142		

Statistical significance of the coefficient, cross-sample coefficient difference and χ^2 statistic:
† p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parenthesis.

Table A9: Estimates of the AA receipt equation for the FRS and ELSA 2-factor models

Covariates	FRS		ELSA		tests and coefficient differences	
Latent disability η_1	0.508†	(0.039)	0.419†	(0.044)	0.089	(0.059)
Latent disability η_2	0.295†	(0.046)	0.164§	(0.087)	0.131	(0.099)
Female	-0.043†	(0.006)	-0.032†	(0.007)	-0.012	(0.01)
Spline age 65-73	0.055†	(0.006)	0.042†	(0.007)	0.013	(0.009)
Spline from age 73+	-0.166†	(0.064)	-0.222†	(0.073)	0.056	(0.097)
Post- compulsory education	-0.001	(0.047)	-0.078	(0.051)	0.077	(0.069)
(ln) income spline to median e	-0.406†	(0.116)	-0.421†	(0.152)	0.015	(0.191)
(ln) income spline from median	-0.149‡	(0.063)	-0.015	(0.07)	-0.135	(0.094)
Outright owner	-0.079	(0.064)	0.084	(0.074)	-0.163§	(0.098)
Married/cohabiting	0.183‡	(0.073)	0.271†	(0.078)	-0.088	(0.107)
coefficient equality $\chi^2(10)$					22.989‡	
N	6744		5142			

Statistical significance of the coefficient, cross-sample coefficient difference and χ^2 statistic:
† p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parenthesis.

Table A10: Factor loadings for the FRS and ELSA 1-factor models with alternative factor loading constraints

FRS			ELSA		
Disability Indicator	Factor loading (St. err.)		Disability Indicator	Factor loading (St. err.)	
	Men	Women		Men	Women
MOBILITY	$0.849^{\dagger}$ (0.071)	$0.962^{\dagger}$ (0.077)	WALKING 100 YDS	$1.118^{\dagger}$ (0.079)	$1.077^{\dagger}$ (0.039)
LIFTING	1 (-)	1 (-)	SITTING 2 HRS	$0.422^{\dagger}$ (0.034)	$0.436^{\dagger}$ (0.030)
DEXTERITY	$0.663^{\dagger}$ (0.058)	$0.579^{\dagger}$ (0.040)	CHAIR TRANSFERS	$0.635^{\dagger}$ (0.042)	$0.582^{\dagger}$ (0.035)
CONTINENCE	$0.360^{\dagger}$ (0.035)	$0.392^{\dagger}$ (0.033)	STAIRS (several flights)	$0.792^{\dagger}$ (0.050)	$0.735^{\dagger}$ (0.041)
COMMUNIC	$0.351^{\dagger}$ (0.039)	$0.333^{\dagger}$ (0.035)	STAIRS (1 flight)	$1.084^{\dagger}$ (0.070)	$0.984^{\dagger}$ (0.058)
MEMORY	$0.382^{\dagger}$ (0.040)	$0.380^{\dagger}$ (0.035)	STOOPING	$0.701^{\dagger}$ (0.044)	$0.715^{\dagger}$ (0.040)
DANGER	$0.461^{\dagger}$ (0.086)	$0.388^{\dagger}$ (0.049)	REACHING	$0.550^{\dagger}$ (0.044)	$0.547^{\dagger}$ (0.037)
OTHER	$0.089^{\dagger}$ (0.025)	$0.055^{\ddagger}$ (0.022)	PULL/PUSHING	$1.100^{\dagger}$ (0.071)	$0.987^{\dagger}$ (0.050)
PROXY	$0.105^{\dagger}$ (0.027)	$0.110^{\dagger}$ (0.022)	LIFTING	1 (-)	1 (-)
			PICKING-UP COIN	$0.415^{\dagger}$ (0.051)	$0.474^{\dagger}$ (0.039)
			DRESSING	$0.723^{\dagger}$ (0.051)	$0.711^{\dagger}$ (0.045)
			WALK ACROSS ROOM	$1.154^{\dagger}$ (0.151)	$1.048^{\dagger}$ (0.100)
			BATHING	$0.944^{\dagger}$ (0.072)	$0.790^{\dagger}$ (0.050)
			FEEDING	$0.652^{\dagger}$ (0.093)	$0.468^{\dagger}$ (0.060)
			BED TRANSFERS	$0.962^{\dagger}$ (0.093)	$0.751^{\dagger}$ (0.058)
			USING TOILET	$0.808^{\dagger}$ (0.098)	$0.631^{\dagger}$ (0.054)
			CONTINENCE	$0.327^{\dagger}$ (0.032)	$0.275^{\dagger}$ (0.023)
			USING A MAP	$0.445^{\dagger}$ (0.052)	$0.375^{\dagger}$ (0.031)
			PREP. HOT MEAL	$0.883^{\dagger}$ (0.109)	$0.886^{\dagger}$ (0.082)
			SHOPPING	$1.115^{\dagger}$ (0.091)	$1.241^{\dagger}$ (0.087)
			PHONING	$0.392^{\dagger}$ (0.049)	$0.357^{\dagger}$ (0.049)
			MEDICATION	$0.523^{\dagger}$ (0.077)	$0.524^{\dagger}$ (0.081)
			HOUSEWORK	$1.239^{\dagger}$ (0.092)	$1.014^{\dagger}$ (0.063)
			MANAGING MONEY	$0.496^{\dagger}$ (0.061)	$0.524^{\dagger}$ (0.052)
Sample size	6,744			5,142	

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table A11: Estimates of the latent disability equation for the FRS and ELSA 1-factor models with alternative factor loading constraints

Covariates	Coefficients and Standard Errors		Tests and coefficient differences	
	FRS	ELSA		
Spline age 65-73	0.042† (0.014)	0.032† (0.011)	0.010	(0.018)
Spline from age 73+	0.100† (0.008)	0.090† (0.006)	0.010	(0.011)
Post-compulsory education	-0.307† (0.073)	-0.255† (0.054)	-0.052	(0.091)
Income spline to median	-0.180† (0.049)	-0.042 (0.046)	-0.137‡	(0.068)
Income spline from median	-0.369† (0.094)	-0.284† (0.063)	-0.085	(0.113)
Outright owner	-0.416† (0.069)	-0.444† (0.055)	0.028	(0.089)
	Sample size		*Coefficient equality* χ^2 (6)	
	6,744	5,142	8.014	

Statistical significance of the coefficient, cross-sample coefficient difference and χ^2 statistic:
† p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parenthesis.

Table A12: Estimates of the AA receipt equation for the FRS and ELSA 1-factor models with alternative factor loading constraints

Covariates	Coefficients and Standard Errors		Tests of coefficient equality	
	FRS	ELSA	FRS-ELSA	
Latent disability η	0.516† (0.041)	0.522† (0.038)	-0.006	(0.055)
Female	0.118§ (0.067)	0.252† (0.076)	-0.134	(0.101)
Spline age 65-73	-0.040† (0.008)	-0.036† (0.007)	-0.004	(0.011)
Spline from age 73+	0.058† (0.006)	0.046† (0.006)	0.012	(0.009)
Post- compulsory education	-0.161‡ (0.064)	-0.238† (0.071)	0.077	(0.096)
(ln) income spline to median	-0.007 (0.047)	-0.092§ (0.050)	0.085	(0.069)
(ln) income spline from median	-0.39† (0.115)	-0.422† (0.153)	0.032	(0.191)
Outright owner	-0.138‡ (0.061)	-0.006 (0.070)	-0.132	(0.093)
Married/cohabiting	-0.077 (0.063)	0.087 (0.074)	-0.164§	(0.097)
	Sample size		*χ^2(9) test of coefficient equality*	
	6,744	5,142	*10.874*	

Statistical significance of the coefficient, cross-sample coefficient difference and χ^2 statistic:
† p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parenthesis.

Table A13: Factor loadings for the FRS 1-factor model obtained by dropping proxy cases from the FRS sample (and the proxy indicator from the measurement model)

	FRS			
	Factor loading (St. err.)			
Disability Indicator	Men		Women	
MOBILITY	1	(-)	1	(-)
LIFTING	1.039†	(0.103)	1.203†	(0.122)
DEXTERITY	0.683†	(0.065)	0.602†	(0.048)
CONTINENCE	0.343†	(0.035)	0.426†	(0.037)
COMMUNIC	0.338†	(0.040)	0.317†	(0.036)
MEMORY	0.356†	(0.039)	0.382†	(0.036)
DANGER	0.355†	(0.091)	0.408†	(0.063)
OTHER	0.101†	(0.029)	0.068†	(0.026)
Sample size	6,308			

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table A14: Estimates of the latent disability equations obtained by dropping proxy cases from the FRS sample (and the proxy indicator from the measurement model)

	Coefficients and Standard Errors			Tests and coefficient differences		
Covariates	FRS	ELSA§§	BHPS§§	FRS-ELSA	FRS-BHPS	ELSA-BHPS§§
Spline age 65-73	0.039†	0.035†	0.127†	0.003	-0.089†	-0.092†
	(0.014)	(0.012)	(0.035)	(0.018)	(0.038)	(0.037)
Spline from age 73+	0.084†	0.099†	0.128†	-0.015	-0.044‡	-0.029
	(0.008)	(0.007)	(0.020)	(0.011)	(0.021)	(0.021)
Post-compulsory education	-0.301†	-0.280†	-0.182	-0.021	-0.119	-0.097
	(0.067)	(0.06)	(0.148)	(0.089)	(0.162)	(0.159)
Income spline to median	-0.114‡	-0.046	-0.172	-0.068	0.057	0.125
	(0.052)	(0.051)	(0.107)	(0.073)	(0.119)	(0.118)
Income spline from median	-0.317†	-0.310†	-0.558†	-0.007	0.241	0.248
	(0.086)	(0.069)	(0.191)	(0.111)	(0.210)	(0.203)
Outright owner	-0.389†	-0.487†	-0.185	0.098	-0.204	-0.302§
	(0.065)	(0.061)	(0.143)	(0.089)	(0.157)	(0.155)
	Sample size			*Coefficient equality* χ^2 (6)		
	6,308	5,142	1,042	3.574	14.613‡	15.745‡

Statistical significance of the coefficient, cross-sample coefficient difference and χ^2 statistic:
† $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parenthesis. §§ Estimates are the same reported in Table 2.

Table A15: Estimates of the AA receipt equations obtained by dropping PROXY cases from the FRS sample (and the proxy indicator from the measurement model)

	Coefficients and Standard Errors			Tests of coefficient equality		
Covariates	FRS	ELSA§§	BHPS§§	FRS-ELSA	FRS-BHPS	ELSA-BHPS§§
Latent disability η	0.573†	0.477†	0.538†	0.096§	0.036	-0.060
	(0.042)	(0.035)	(0.091)	(0.055)	(0.100)	(0.098)
Female	0.156‡	0.251†	-0.068	-0.095	0.224	0.319§
	(0.071)	(0.076)	(0.177)	(0.104)	(0.191)	(0.193)
Spline age 65-73	-0.041†	-0.036†	-0.084†	-0.004	0.0430‡	0.048‡
	(0.009)	(0.007)	(0.020)	(0.011)	(0.022)	(0.021)
Spline from age 73+	0.059†	0.046†	0.028§	0.014	0.031§	0.017
	(0.006)	(0.006)	(0.015)	(0.009)	(0.016)	(0.016)
Post- compulsory education	-0.153‡	-0.238†	-0.070	0.085	-0.083	-0.167
	(0.068)	(0.071)	(0.160)	(0.099)	(0.174)	(0.175)
(ln) income spline to median	-0.044	-0.092§	-0.041	0.048	-0.002	-0.050
	(0.063)	(0.05)	(0.084)	(0.08)	(0.105)	(0.098)
(ln) income spline from median	-0.493†	-0.422†	-0.411§	-0.071	-0.082	-0.011
	(0.129)	(0.153)	(0.249)	(0.2)	(0.280)	(0.292)
Outright owner	-0.137‡	-0.006	-0.265§	-0.131	0.128	0.259
	(0.065)	(0.070)	(0.160)	(0.095)	(0.172)	(0.174)
Married/cohabiting	-0.058	0.087	-0.171	-0.145	0.112	0.257
	(0.067)	(0.074)	(0.173)	(0.099)	(0.185)	(0.188)
	Sample size			*Coefficient equality* χ^2 (9)		
	6,308	5,142	1,042	13.452	15.61§	15.483§

Statistical significance of the coefficient, cross-sample coefficient difference and χ^2 statistic:
† p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parenthesis. §§ Estimates are the same reported in Table 3.

Appendix 2: Identification

After using equation (3) to solve out the latent disability variables η_{iq} from the model, the structure can be written in matrix form as:

$$\tilde{\mathbf{D}} = \mathbf{\Lambda\Theta z} + \mathbf{\Lambda\upsilon} + \boldsymbol{\varepsilon} \quad \text{(A1)}$$

$$\tilde{R} = (\boldsymbol{\beta} + \boldsymbol{\gamma\Theta})\mathbf{z} + \boldsymbol{\gamma\upsilon} + u \quad \text{(A2)}$$

where **Λ**, **Θ**, **β** and **γ** are respectively $K_s \times Q$, $Q \times p$, $1 \times p$ and $1 \times Q$ dimensional coefficient matrices and we have omitted the individual *i* suffix from the covariates **z**, the latent variables $\tilde{\mathbf{D}}$, and $\tilde{R}$ underlying the observed ordinal variables **D** and *R*, and the unobservable random terms **υ**, **ε** and *u*. Equations (A1)-(A2) together comprise a system of correlated reduced form (ordered) probit equations, from which we can identify the following coefficient matrices and residual covariances:

$$\mathbf{B}_1 = \mathbf{\Lambda\Theta} \quad \text{(A3)}$$

$$\mathbf{B}_2 = \boldsymbol{\beta} + \boldsymbol{\gamma\Theta} \quad \text{(A4)}$$

$$\mathbf{C}_{11} = \mathbf{\Lambda\Omega\Lambda}' + \mathbf{\Sigma} \quad \text{(A5)}$$

$$\mathbf{C}_{22} = \boldsymbol{\gamma\Omega\gamma}' + \boldsymbol{\gamma\delta} + \sigma^2 \quad \text{(A6)}$$

$$\mathbf{C}_{12} = \mathbf{\Lambda\Omega}\boldsymbol{\gamma}' + \mathbf{\Lambda}\boldsymbol{\delta} \quad \text{(A7)}$$

where **Ω** is the covariance matrix of **υ**, **Σ** is the diagonal covariance matrix of **ε**, **δ** is the vector of covariances between **υ** and *u*, and σ^2 is the variance of *u*.

Some normalisations are necessary, because the observed variables **D** and *R* do not reveal the scale of $\tilde{\mathbf{D}}$ and $\tilde{R}$ and because the latent **η** can be replaced by arbitrary linear combinations with the loadings **Θ** and **γ** transformed accordingly. Without loss of generality, we resolve these indeterminacies by setting C_{22} and the diagonal elements of C_{11} to unity and by imposing the restrictions:

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{I} \\ \mathbf{\Lambda_2} \end{pmatrix} \quad \text{(A8)}$$

Given these normalisations, the first *Q* rows of $\mathbf{B}_1$ identify **Θ**. Provided the rank of **Θ** is *Q*, $\mathbf{\Lambda}_2$ can then be found by solving the last K_s -*Q* equations in (A3). This rank condition implies that the *Q* latent factors in the measurement equations (1) cannot be replaced by a smaller number of linear combinations of the factors.

Now consider identification of **Ω**. Write the vector of *Q* diagonal elements of **Ω** as $\boldsymbol{\omega}_d$ and the vector of (*Q*-1)/2 sub-diagonal elements as $\boldsymbol{\omega}_s$. We can construct an identity: vec(**Ω**) = $\mathbf{S}_d\, \boldsymbol{\omega}_d + \mathbf{S}_s\, \boldsymbol{\omega}_s$ where $\mathbf{S} = (\mathbf{S}_d\ \mathbf{S}_s)$ is a $Q^2 \times Q(Q+1)/2$ permutation matrix containing 1s and 0s and vec(.) is the operation of stacking the rows of a matrix into a column vector. Let $\mathbf{C}_{11}^{1,1}$ be the leading $Q \times Q$ block of $\mathbf{C}_{11}$ and note that $\mathbf{\Sigma}$ is diagonal so that $\mathbf{S}_s' \text{vec}\left(\mathbf{C}_{11}^{1,1}\right) = \boldsymbol{\omega}_s$. This determines the off-diagonal elements of $\boldsymbol{\omega}$. Now let $\mathbf{C}_{11}^{1,2}$ be the submatrix of $\mathbf{C}_{11}$ containing elements from the first Q rows and last K_s-Q columns: then $\mathbf{C}_{11}^{1,2} = \mathbf{\Omega \Lambda}_2'$ and, if c_{qj} is the typical element of $\mathbf{C}_{11}^{1,2}$, each of the ω_{qq} can be deduced as $\omega_{qq} = \left(c_{qj} - \sum_{r \neq q} \omega_{qr} \lambda_{jr}^s \right) / \lambda_{jq}^s$, provided there exists at least one non-zero element in the qth column of $\mathbf{\Lambda}_2$, for each q = 1...Q. With $\mathbf{\Omega}$ determined, $\mathbf{\Sigma}$ is immediately given by (A5).

Without further restrictions, this is as far as we can go. Once $\mathbf{\Theta}$, $\mathbf{\Lambda}$, $\mathbf{\Omega}$ and $\mathbf{\Sigma}$ are known, this still leaves $p + 2Q + 1$ parameters $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$ and σ^2 to be determined by the $p + Q + 1$ equations in (A4), (A6) and (A7). At least Q further restrictions are necessary. Natural possibilities are $\boldsymbol{\delta} = \text{cov}(\boldsymbol{\upsilon}, u) = \mathbf{0}$ or exclusion restrictions on the vector $\boldsymbol{\beta}$. The latter requires the existence of covariates that can be assumed a priori to influence disability status (relevance) but have no causal role in determining benefit receipt (validity).