the review of income and wealth

Review of Income and Wealth
Series 67, Number 4, December 2021
DOI: 10.1111/roiw.12499

THE WELFARE IMPLICATIONS OF UNOBSERVED HETEROGENEITY

BY SARANTIS TSIAPLIAS*

The University of Melbourne

New conditions are derived for relating household financial well-being to household utility. In particular, a one-for-one mapping between the equivalent incomes stemming from subsistence-based utility functions and probabilistic models of financial well-being is established. This is unique in the literature and enables estimates from reduced-form models based on a cumulative distribution function (e.g. probit and logit models) to be given a formal welfare interpretation. In so doing, it is possible to use reduced-form models of well-being to evaluate welfare distortions associated with unobserved heterogeneity in subsistence levels and marginal utilities of consumption. An Australian household-level data set is used as a case study for exploring the distortions associated with unobserved heterogeneity. The results are significant for better understanding the welfare implications of income and transfer policies, and indicate that the failure to account for unobserved heterogeneity results in large welfare distortions. Finally, I show that the distortions are primarily attributable to heterogeneity in subsistence requirements rather than heterogeneity in marginal utilities of consumption.

JEL Codes: D10, D12, D60, I30

Keywords: households, well-being, utility, subsistence, consumption, heterogeneity

1. INTRODUCTION

A household's financial well-being is often measured using binary indicators of household financial conditions (such as financial stress, financial satisfaction, or whether they have defaulted in a credit obligation). Given such indicators, it is relatively straightforward to estimate probabilistic models of "well-being" that incorporate individual and household characteristics (e.g. using a probit or logit model to assess the probability of financial stress). However, the estimation of a model of well-being is typically a reduced-form exercise with tenuous links to welfare analysis. This paper shows that it is possible to attach a welfare interpretation to the estimates from the well-being function thereby allowing for welfare analysis. In particular, I show how to construct equivalent income measures from probabilistic models of well-being that yield information on the additional income required by a household to achieve the same level of well-being as the benchmark

Note: I thank two anonymous referees and the editor Conchita D'Ambrosio for helpful comments and suggestions. This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to either DSS or the Melbourne Institute.

*Correspondence to: Sarantis Tsiaplias, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne, Parkville, VIC 3010, Australia (sarantis.tsiaplias@unimelb.edu.au).

household after accounting for heterogeneity in its subsistence requirement and its marginal utility of consumption.

The proposed measures are used to evaluate the key issue of measuring the welfare importance of unobserved heterogeneity. This issue is detailed in Pollak and Wales (1979) and Fisher (1987, 1990), who highlight the importance of observed and unobserved differences for undertaking welfare comparisons across households. Using an Australian data set of household-specific financial stress, the welfare distortion measures derived in this paper are empirically evaluated, and the welfare importance of unobserved heterogeneity is estimated using actual data. The results have significant policy implications, with the evidence strongly indicating that the failure to account for unobserved heterogeneity is likely to result in large welfare distortions. An associated implication is that practices such as the provision of a fixed or common payment to households are neither efficient nor welfare maximizing.

The measure of equivalent income adopted in this paper is based on the theoretical income level that would place a household's well-being at the benchmark household's level of well-being. In the context of the paper's financial-stress application, the well-being function is based on household financial stress, with the household's equivalent income determined as the level of income required to render the household no more (and no less) exposed to financial stress than some reference household. The resulting measure of equivalent income bears some similarity to the notion of an indifference income in that it is effectively the level of income that renders a household indifferent to its idiosyncratic consumption and welfare sensitivity differences (Browning *et al.*, 2013; Chiappori, 2016). However, it is estimated using traditional reduced-form models and does not require the observation of product-specific prices or detailed household expenditure (which are often either unavailable or only partially observed). It also bears weak similarity to the notion of an extended equivalent income described in Fleurbaey (2015) because it extends the notion of equivalent income to non-price aspects of the environmental and personal situations of the household members (for further details, see Fleurbaey and Blanchet, 2013).

To measure welfare distortions, the methodology essentially involves evaluating the minimum level of household income required to compensate a household in a manner that accounts for its unique consumption requirements, and the shocks that the household has experienced (such as the illness or death of a family member). The model described in Lim and Tsiaplias (2019), which estimates the income requirements of Australian households using an augmented probit-type model, is used to provide the income requirements that act as inputs into the welfare distortion measures derived in this paper. In this respect, although the sources of household consumption heterogeneity have been explored (Clarida, 1991; Attanasio and Weber, 1995; Browning and Crossley, 2001; Gourinchas and Parker, 2002; Fernandez-Villaverde and Krueger, 2007), significantly less is known about the distribution of unobserved subsistence levels and spending needs across households. Similarly, relatively little is known about the impact of such heterogeneity on the level of compensation households require to achieve a desired level of well-being. The welfare distortion measures directly address the latter issue.

The presence of unobserved heterogeneity implies a *distribution* of equivalent incomes for each household type. In this sense, the proposed approach also complements the significant body of literature concerned with the estimation of models designed to capture unobserved heterogeneity (Briesch *et al.*, 2002, 2009; Blundell *et al.*, 2014; Lewbel and Pendakur, 2017). The moments associated with these distributions have direct welfare implications and can be used to estimate the average welfare distortion and the typical welfare gain or loss when there is a failure to appropriately account for heterogeneity. In so doing, I show that the average welfare distortion stemming from the failure to account for unobserved heterogeneity tends to increase with family size, rising from only 8 percent for a single-person household to approximately 25 percent for a typical four-person household.

Finally, the paper examines the extent to which the estimated distortions are attributable to unobserved heterogeneity in subsistence levels or to heterogeneity in marginal utilities of consumption (van Praag *et al.*, 1980; Aiyagari, 1994; Calvet and Comon, 2003; Chetty and Szeidl, 2016). This distinction is important because distortions stemming from the former may be addressed through particular forms of insurance (such as income or health insurance to reduce the magnitude of income or expenditure shocks) or the elimination of frictions that produce adjustment costs (Blundell *et al.*, 2008; Lusardi *et al.*, 2011). In contrast, the approach to correcting distortions associated with marginal utilities of consumption is less clear. This paper shows that the estimated welfare distortions stemming from the failure to account for unobserved heterogeneity are primarily attributable to the mistaken assumption of a common subsistence level (or a subsistence level that is common across households belonging to a particular demographic group).

This paper is organized as follows. Section 2 provides a formal relationship between the measures of household utility and well-being, in addition to deriving measures of the distortion stemming from the assumption of homogeneity. Section 3 discusses the model used to quantify the welfare distortion measures. Section 4 examines the welfare impact of both group-wise heterogeneity and deeper household-level heterogeneity. Concluding remarks are in Section 5.

2. Deriving Measures of Welfare Distortion Using Models of Well-Being

This section establishes a mapping between the household's utility function and probabilistic measures of well-being. It is shown that there is a one-to-one mapping between the parameters underpinning the equivalent income derived using probabilistic measures of well-being (e.g. the income that renders two given households equal in terms of their probability of financial stress) and the parameters underpinning the equivalent income derived from certain utility functions. The utility functions that satisfy this mapping belong to the Identical-Shape Hyperbolic-Absolute-Risk-Aversion (ISHARA) family and define a household's utility by reference to its specific subsistence threshold and its marginal utility of consumption. Importantly, these utility functions can be used to identify welfare-maximizing policies.[1]

[1]The ISHARA utility function allows for exact linear aggregation (viz. that aggregate welfare is a linear function of individual welfare levels). Gorman (1953) shows that without exact linear aggregation, it is generally not possible to identify welfare-maximizing policies (for further details, see Gorman, 1961).

The conditions under which the utility- and well-being-based approaches generate equivalent incomes that satisfy a one-to-one mapping are also examined, including the formal correspondence between: (i) the subsistence levels and marginal utilities of consumption defined by the household's utility function and (ii) the parameters of the probabilistic model of well-being. The results provide a basis for determining welfare distortions using parameters obtained from typical estimable models of household well-being (e.g. probit, logit, or generalized extreme value models).

2.1. *A Theorem for Relating Household Well-Being and Utility Functions*

Consider an economy with N households. Each household has an ISHARA instantaneous utility function $u_i(c)$. The households can differ in their subsistence level such that household i's utility function is given by

$$u_i(c) = \exp\left(\zeta_i^*\right)\frac{\left[ac+\beta_i\right]^{1-1/a}}{a(1-1/a)} a>0,\ a\neq 1,\ \beta_i<0, \tag{1}$$

where ζ_i^* is household i's (relative) marginal utility of consumption, $-\beta_i$ is household's i's subsistence level of consumption, and a is a curvature parameter (Mazzocco, 2004, 2007; Koulovatianos *et al.*, 2019).

Pursuant to the sign restrictions on a and β_i, households can differ in their subsistence $-\beta_i$ such that any two households with the same level of consumption c and curvature parameter a will nevertheless attain a different utility level if their subsistence levels differ. However, it is noted that the sign restrictions on a and β_i can be motivated by cross-sectional evidence on savings rates (e.g. Dynan *et al.*, 2004) that support the existence of subsistence levels of consumption in the presence of a common positive curvature parameter a.

Equivalent income is defined as the income that results in the same level of utility for households i, j such that

$$u_i\left(\overline{y}_i\right) = u_j\left(\overline{y}_j\right), \tag{2}$$

where $\overline{y}_i$ and $\overline{y}_j$ are the equivalent incomes of households i and j, respectively. In the adopted setting, household j can be treated as the "reference" household. In the absence of the specific formulation of a savings function, income and consumption are synonymous in the setting specified here. This restriction is reasonable for the examination of financial stress undertaken in this paper but will not be reasonable in all settings.[2]

In line with Proposition 1 in Koulovatianos *et al.* (2019), the solution to the problem (2) yields the following equivalent income for household i:

$$\overline{y}_i = \gamma_{ij} + \delta_{ij}\overline{y}_j, \tag{3}$$

[2]In the data set used here, households are deemed to be financially stressed if, *because of lack of money*, they are unable to meet core expenses. This is discussed further in Section 3.

whereby $\bar{y}_i$ is a linear function of household j's income.[3]

To see that (3) does indeed hold, we can explicitly solve for the income levels equivalizing the utility of households i and j when both households have a utility function given by (1). This yields the equivalent income

$$\bar{y}_i = -a^{-1}\left(\beta_i - c_{ij}\beta_j\right) + c_{ij}\bar{y}_j, \tag{4}$$

where $c_{ij} = \exp\left(-\frac{a}{a-1}\left(\zeta_i^* - \zeta_j^*\right)\right)$. Household i's equivalent income is therefore consistent with the general form (3). It is also clear that the γ_{ij}, δ_{ij} parameters reflect both the underlying subsistence levels of households i and j (β_i and β_j), in addition to differences in their marginal utilities of consumption ($\zeta_i^* - \zeta_j^*$).

2.1.1. Defining a Probability Measure for Well-Being

Consider the existence of a probability measure V_i that measures the household's "well-being" and depends, at least partially, on household income. Note that the well-being function V_i can also depend on other factors and is not limited to income. As noted earlier, well-being is defined generally and can be based on binary indicators of financial and economic events (such as financial stress or credit default) or subjective measures such as financial satisfaction and happiness (e.g. consider the models in Senik (2004), Zaidi and Burchardt (2005), Morciano *et al*. (2015) or Decancq *et al*. (2017)).

The well-being function for household i takes on a very general form that satisfies almost any parametric probability model (including the widely used probit or logit forms):

$$V_i = V\left(y_i, x_i\right) = 1 - F\left(\frac{k_i(x_i) - y_i}{\eta_i}\right), \tag{5}$$

where $F(\cdot)$ is a cumulative distribution function (CDF), $k_i(x_i)$ is a household-specific income requirement or threshold that can depend on the set of characteristics x_i, y_i is household i's income, and η_i is a scale parameter. Because the change in $F(\cdot)$ depends on the size of the scale parameter η_i, I follow van Praag (1968, 1971) in labeling η_i as the household's well-being sensitivity parameter. Technically, however, η_i reflects the household's well-being (in)sensitivity with $1/\eta_i$ representing the well-being sensitivity.

The function $F(\cdot)$ encompasses the implicit well-being function adopted in much of the literature that investigates household (or individual) well-being. In particular, the household's well-being can be compared to a binary discrete choice or random utility problem. Such problems can be represented as a linear probability model in an indicator variable that is explained by a set of characteristics x_i,

[3]The resulting functional form has properties that conform to the well-known concept of Generalized Absolute Equivalence Scale Exactness (GAESE) (Blackorby and Donaldson, 1994; Donaldson and Pendakur, 2006; Cherchye *et al*., 2015).

which include household or individual income (e.g. McFadden, 1984; Gul and Pesendorfer, 2006). An example is a model with a probit response variable defining financial stress that depends on household-level income and a set of (observed and unobserved) characteristics for each household (in this example, the household's random utility increases as its probability of financial stress decreases).

Given the general well-being function specified in (5), the following theorem states the one-to-one mapping between the equivalent incomes that equivalize well-being levels and the equivalent incomes that equivalize utility (i.e. that satisfy equation (2)).

Theorem 1 *If u_i is a utility function given by* (1) *and V_i is a well-being function given by* (5)*, then there is a one-to-one mapping between the equivalent income functions generated by u_i and V_i.*

To obtain this result, note that the welfare equivalence problem based on the well-being function (whereby household i's income y_i is "adjusted" such that its well-being function V_i is equal to some reference household j's well-being V_j) is given by

$$\min_{\tilde{y}_i} \| V\left(\tilde{y}_i, x_i\right) - V\left(\overline{y}_j, x_j\right) \| , \tag{6}$$

where the notation $\overline{y}_i$ is used to distinguish between the hypothetical level of household i's income in equation (6) and the household's actual income y_i.[4] The reference household is associated with income $\overline{y}_j$ and observable characteristics x_j.

The solution to the problem (being the income level $\overline{y}_i$ that solves equation (6)) can be obtained exactly and yields the following "equivalent income" resulting in $V\left(\overline{y}_i, x_i\right) = V\left(\overline{y}_j, x_j\right)$:

$$\overline{y}_i = k_i + \frac{\eta_i}{\eta_j}\left(\overline{y}_j - k_j\right) , \tag{7}$$

where, for notational convenience, $k_i = k_i(x_i)$ and $k_j = k_j(x_j)$.

The equivalent income $\overline{y}_i$ that solves problem (6) for household i is therefore a linear function of the reference household's "excess" income $y_j - k_j$, with an intercept given by its threshold k_i and a slope parameter that depends on the ratio of scale terms $\frac{\eta_i}{\eta_j}$.

Setting $\gamma_{ij} = k_i - \frac{\eta_i}{\eta_j}k_j$, $\delta_{ij} = \frac{\eta_i}{\eta_j}$, where $\left(k_i, \eta_i\right)$ and $\left(k_j, \eta_j\right)$ are the income requirement and scale parameters stemming from the well-being functions for households i and j respectively, we obtain

[4]The problem in equation (6) can be read as the definition of the equivalent income for well-being functions that is adopted in this paper.

$$\overline{y}_i = \left(k_i - \frac{\eta_i}{\eta_j}k_j\right) + \frac{\eta_i}{\eta_j}\overline{y}_j, \tag{8}$$

where it is clear that equation (8) accords with the equivalent income specification in equation (3).

It follows from Theorem 1 that there is a unique, one-to-one correspondence between the four relevant arguments of the utility-based equivalent income function (being $\beta_i, \beta_j, \zeta_i^*, \zeta_j^*$) and the four relevant arguments of the well-being-based equivalent income function (being k_i, k_j, η_i, η_j). The mapping between the two equivalent incomes is given by

$$\frac{\eta_i}{\eta_j} = c_{ij} = \exp\left(-\frac{a}{a-1}\left(\zeta_i^* - \zeta_j^*\right)\right) \tag{9a}$$

$$k_i = -a^{-1}\beta_i \tag{9b}$$

$$k_j = -a^{-1}\beta_j. \tag{9c}$$

Equation (9a) identifies the link between household i's well-being sensitivity η_i and the marginal utility of consumption ζ_i^*. It is clear that in the case $\zeta_i^* = \zeta_j^*$, such that marginal utilities of consumption are the same for households i and j, we also have $c_{ij} = \frac{\eta_i}{\eta_j} = 1$. In this case, the same law of motion $\partial\overline{y}_i = \partial\overline{y}_j$ holds for both a household that maximizes V_i or its utility u_i.

The remaining two equations (9b) and (9c) identify the link between the household's subsistence $-\beta_i$ in (1) and the household's income requirement k_i. The relationships imply that that the difference in income requirements $(k_i - k_j)$ is proportional to the difference in the subsistence parameters $(\beta_i - \beta_j)$, with the proportionality depending on the curvature parameter a. Moreover, the ratio of income requirements yields the ratio of subsistence levels independently of a because $k_i / k_j = \beta_i / \beta_j$.[5]

2.2. *Constructing Measures of Welfare Distortion*

To examine the impact of treating households as homogeneous, assume that the household's income requirement k_i is a linear function of a fixed (or permanent) household-specific component γ_{0i} and observed household characteristics x_i such that

$$\hat{k}_i = \gamma_{0i} + \tilde{c}_i = \gamma_{0i} + x_i'\gamma, \tag{10}$$

[5]A benefit of identifying equivalent incomes using well-being functions is that the curvature parameter need not be set a priori by the policy-maker or explicitly estimated.

where $\widehat{k}_i$ is the expected value of k_i, $\tilde{c}_i \equiv x_i' \gamma$, and γ_{0i}, γ are parameters that are either estimated using data or set by the researcher.[6]

For a given level of household income, households are therefore heterogeneous in terms of γ_{0i}, the scale parameter η_i, and observed household characteristics $\tilde{c}_i$ (with γ_{0i} and η_i being unobserved). Notwithstanding the reduced-form nature of these parameters, from Theorem 1 and the associated equations (9a)–(9c), we can assert (rather than simply conjecture) that heterogeneity in these parameters implies heterogeneity in the household's subsistence level β_i and its marginal utility of consumption ζ_i^*. These properties are used to construct equivalent income measures that are interpretable in terms of subsistence levels and marginal utilities of consumption but can be quantified using estimates from reduced-form models.

Relative to a benchmark household characterized by the triple $\left(y^*, \widehat{k}^*, \eta^*\right)$ (representing benchmark levels of actual income, required income, and well-being sensitivity respectively), household i's welfare-equivalizing income can be constructed by substituting equation (10) into equation (7):

$$\bar{y}_i = \widehat{k}_i + \frac{\eta_i}{\eta^*}\left(y^* - \widehat{k}^*\right) = \gamma_{0i} + \tilde{c}_i + \frac{\eta_i}{\eta^*}\left(y^* - \gamma_0^* - \tilde{c}^*\right). \tag{11}$$

To determine the impact of imposing homogeneity on household consumption preferences, I consider the equivalent incomes that stem from assuming that household i's welfare sensitivity η_i and/or fixed level of subsistence γ_{0i} are the same as that of the benchmark household (i.e. $\eta_i = \eta^*$ and $\gamma_{0i} = \gamma_0^*$, respectively). These restrictions yield a homogeneous form of equivalent income, and the difference between (11) and the homogeneous measure is the additional income required by household i to achieve the same level of well-being as the benchmark household after considering heterogeneity in its subsistence requirement and/or marginal utility of consumption.

Three cases considered are as follows: (1) homogeneity in the marginal utility of consumption, $\eta_i = \eta^*$; (2) homogeneity in the fixed part of the subsistence level, $\gamma_{0i} = \gamma_0^*$; and (3) both forms of homogeneity.

Case 1: Homogeneity in the Marginal Utilities of Consumption

Pursuant to (9a), the assumption of common marginal utilities of consumption across all households implies that $\eta_i = \eta^*$ (for any i) in the well-being model. Subject to this restriction, the equivalent income for household i, denoted by $\bar{y}_i | \left(\eta_i = \eta^*\right)$, is then given by the household income of the benchmark household y^* adjusted by the difference in the household-specific income needs of the two households:

$$\bar{y}_i | \left(\eta_i = \eta^*\right) = y^* + \left(\widehat{k}_i - \widehat{k}^*\right). \tag{12}$$

[6]Note that η_i is also a parameter that must be either estimated or set by the researcher.

It can be shown that the distortion in equivalent income (denoted $D_{\eta,i}$) stemming from the assumption of common marginal utilities of consumption is nonzero unless the true η_i is equal to η^*:

$$(13)\qquad D_{\eta,i} = \bar{y}_i - \bar{y}_i \mid \left(\eta_i = \eta^*\right) = \left(\frac{\eta_i}{\eta^*} - 1\right)\left(y^* - \hat{k}^*\right),$$

where $D_{\eta,i}$ is defined as the difference between the heterogeneous (or unrestricted) equivalent income $\bar{y}_i$ in equation (11) and the restricted equivalent income $\bar{y}_i \mid \left(\eta_i = \eta^*\right)$.

In turn, the distortion or bias in the equivalence scale (the equivalence scale being defined as the ratio of the equivalent income $\bar{y}_i$ to the income of the reference household y^*) stemming from the assumption that $\eta_i = \eta^*$ is a linear function of $\hat{k}^*$ with intercept $\left(\frac{\eta_i}{\eta^*} - 1\right)$ and slope parameter $\left(\frac{1-\eta_i/\eta^*}{y^*}\right)$:

$$(14)\qquad D^*_{\eta,i} = \frac{\bar{y}_i}{y^*} - \frac{\bar{y}_i}{y^*} \mid \left(\eta_i = \eta^*\right) = \left(\frac{\eta_i}{\eta^*} - 1\right) + \left(\frac{1-\eta_i/\eta^*}{y^*}\right)\hat{k}^*.$$

Because η_i is nonnegative, the bias stemming from the adoption of $\eta_i = \eta^*$ lies in the interval $\left(\hat{k}^* - y^*, \infty\right)$. Assuming $y^* > \hat{k}^*$, this implies that $\bar{y}_i \mid \left(\eta_j = \eta^*\right)$ may constitute either an over- or under- estimate of the household income required to equate household i's welfare with that of the benchmark household. As the interval is unbounded on the right, however, there is a greater risk that $\frac{\bar{y}_i}{y^*} \mid \left(\eta_i = \eta^*\right)$ understates the additional compensation required to equivalize welfare for household i.[7]

Case 2: Homogeneity in Subsistence Levels

Pursuant to (9b) and (10), the household's subsistence level is decomposed into the part explained by its observed characteristics $\tilde{c}_i$ and the unobserved part that is unique to the household γ_{0i}. If it is assumed that there is no household-specific component, such that $\gamma_{0i} = \gamma_0^*$, then Theorem 1 implies that households will differ in their subsistence levels only by reference to their observed characteristics (e.g. because of age differences in children). In this case, we obtain

$$(15)\qquad \bar{y}_i \mid \left(\gamma_{0i} = \gamma_0^*\right) = \hat{k}_i + \left(\gamma_0^* - \gamma_{0i}\right) + \frac{\eta_i}{\eta^*}\left(y^* - \hat{k}^*\right).$$

[7]Consider, for example, the bias in the equivalence scale when the true ratio η_i / η^* is equal to 2 for a benchmark household with $y^* = \$75,000$. At $\hat{k}^* = \$50,000$, the assumption $\eta_i = \eta^*$ understates the equivalence scale by 1/3, with household i requiring an additional $1/3 \times y^* = \$25,000$ to equivalize welfare in the presence of differences in the marginal utility of consumption.

Consequently, even if the assumption of common marginal utilities of consumption is relaxed, there may still be distortions stemming from the assumption of a common subsistence level.

It can be shown that the distortion stemming from the assumption $\gamma_{0i} = \gamma_0^*$ is an affine function of the simple difference $\gamma_{0i} - \gamma_0^*$ and will therefore only be zero in the case where the true γ_{0i} is equal to γ_0^*:

$$(16) \qquad D_{\gamma_0,i} = \bar{y}_i - \bar{y}_i | \left(\gamma_{0i} = \gamma_0^*\right) = \gamma_{0i} - \gamma_0^*,$$

where $D_{\gamma_0,i}$ is the additional income that would be required to equivalize household i's well-being with that of the benchmark household.

Case 3: Homogeneity in Both Marginal Utilities of Consumption and Subsistence Levels

Finally, it can be shown that the distortion when both $\eta_i = \eta^*$ and $\gamma_{0i} = \gamma_0^*$ are imposed is

$$(17) \qquad \begin{aligned} D_i &= \bar{y}_i - \bar{y}_i | \left(\eta_i = \eta^*, \gamma_{0i} = \gamma_0^*\right) \\ &= \left(\gamma_{0i} - \frac{\eta_i}{\eta^*}\gamma_0^*\right) + \left(1 - \frac{\eta_i}{\eta^*}\right) \%\bar{c}^* + \left(\frac{\eta_i}{\eta^*} - 1\right) y^*. \end{aligned}$$

Equation (17) implies that the assumption of homogeneity in the η_i and γ_{0i} parameters may over- or under-state equivalent income. Ceteris paribus, however, if the household has a greater subsistence requirement, such that $\gamma_{0i} > \gamma_0^*$, then the assumption of homogeneity will *under*-state the household's equivalent income resulting in a reduced level of well-being. Conversely, $\gamma_{0i} < \gamma_0^*$ will *over*-state equivalent income resulting in a greater level of well-being.

Consider also the case where $\eta_i < \eta^*$, such that household i exhibits a greater level of welfare sensitivity to a change in income than the benchmark household. In this case, the assumption of homogeneity will *over*-state the household's equivalent income (with the converse also holding true). The reason for this result is that $\eta_i < \eta^*$ implies a smaller income shift to equivalize $V\left(y_i, x_i\right)$ with $V(y^*, x^*)$ than does $\eta_i \geq \eta^*$. In particular, when $\eta_i < \eta^*$, the slope of household i's marginal utility of consumption is flatter if η^* is imposed instead of η_i. As such, achieving the absolute welfare change $\left|\Delta V\left(y_i, x_i\right)\right|$ requires a greater change in household income y_i than if the true η_i was adopted. The converse holds if $\eta_i > \eta^*$.

3. Modeling Household Income Requirements

It is clear from the preceding discussion that in the absence of exogenous values for β_i and ζ_i^*, the mapping of parameters from V_i to u_i will depend on the policy-maker's choice of V_i. Given the plausible relationship between the household's capacity to consume up to its subsistence level and its experience of financial stress,

V_i is chosen such that it reflects estimates of expenditure requirements that are based on the avoidance of financial stress.

The specific model used as a case study for obtaining empirical quantities of the welfare distortions derived in Section 2 is described in Lim and Tsiaplias (2019). However, the analyst is free to use other models to underpin V_i. The model adopted for the measurement of welfare distortions can be contrasted with the standard ("micro") approach based on the use of expenditure surveys to estimate demand functions (which are subsequently used to compute equivalent income). This approach tries to optimize the estimation of demand functions by allowing for different types of unobserved heterogeneity or using a nonparametric approach (Hausman and Newey, 2017; Pendakur, 2018).

Although the standard approach has several appealing features, it is less useful for the proposed exercise. A key advantage of the chosen model is that it naturally deals with both durables and non-durables. This is often missing in the micro approach because of the absence of reliable demand data for durables. Therefore, it is difficult to fully measure the welfare distortions associated with unobserved heterogeneity because a critical aspect of the household's subsistence level is either fully or partially missing.

Pursuant to the chosen financial-stress-based well-being function, households are averse to a range of indicators of financial stress (e.g. needing to borrow money from friends or family or being unable to pay utility bills) and prefer a lower probability of financial stress. Under (5) or (1), households seek to maximize the distance between their income and some minimum commitment level.

To formalize the financial-stress condition, define the indicator variable m_{it} as equal to 1 when household i is financially stressed in period t and 0 otherwise. The indicator is constructed using

$$m_{it} = I\left(k_{it} > r_{it}\right), \tag{18}$$

where k_{it} is the household's unobserved income requirement and r_{it} is the household's actual income net of its actual accommodation (or housing) expenditure (which is based on the household's annual mortgage repayments or housing rent expenditure).[8] For convenience, r_{it} is hereafter called the household's residual income. $I(\cdot)$ is a binary indicator taking on the value unity if $k_{it} > r_{it}$.[9]

To estimate k_{it} using the methodology described in Lim and Tsiaplias (2019), it is necessary to obtain the indicator m_{it}. This is constructed based on household responses to a rich set of financial-stress variables in the Household, Income and Labour Dynamics in Australia (HILDA) data set, which is a large-scale longitudinal survey containing detailed information about the economic and subjective well-being, labor market dynamics, and family dynamics of Australian households. The data used in the case study cover 3103 households observed over the period

[8]Some households report both mortgage repayments and rental expenditure. In this situation, the sum of the two is treated as the household's housing expenditure.

[9]Without loss of generality, r_{it} is adopted instead of y_{it}. For example, if the household's accommodation expenditure is unobserved, the decision rule becomes $I\left(k_{it}^{+} > y_{it}\right)$, with k_{it}^{+} now inclusive of housing expenditure.

2002–2012 for a total of just under 27,000 observations (and just over 9600 instances of financial stress).[10] In terms of individual households, approximately one in five households experience financial stress in any given year. Further details are provided in Appendix A. A household is deemed to be financially stressed (viz. m_{it} is set to unity) if, because of lack of money, any of the following stress conditions is satisfied: it cannot pay its utility bills, mortgage, or rent on time, has requested financial help from friends or family, has pawned or sold something to make ends meet, is unable to heat its home, goes without meals, or requests help from a charity or similar organization. If the household does not exhibit any of the financial-stress behaviors, then m_{it} is set to zero.

3.1. *Model Overview*

The model assumes that the household forms an expectation of its spending commitments and therefore knows its expected income requirement $\widehat{k}_{it}$. Because the household's spending needs are a function of the unknown (to the econometrician) consumption commitments and adjustment costs of its individual members, each household has its own household-specific spending requirement γ_{0i} that jointly characterizes the needs of its individual members (the importance of adjustment costs is discussed in Chetty and Szeidl, 2007, 2016).

Household *i*'s estimated income requirement is given by

$$\widehat{k}_{it} = \gamma_{0i} + x_{it}'\gamma, \tag{19}$$

where γ_{0i}, γ are estimable parameters and x_{it} is a set of covariates or instruments used to estimate time-variation in household *i*'s spending needs. These include, for example, household-specific events such as changing residency, major events such as serious illness, and year-specific macro effects that are common to all households.

Because the econometrician is unable to observe k_{it}, an error term u_{it} is introduced such that

$$k_{it} = \widehat{k}_{it} + u_{it} = \gamma_{0i} + x_{it}'\gamma + u_{it}, \tag{20}$$

$$u_{it} \sim N\left(0, \eta_i^2\right). \tag{21}$$

Given u_{it}, a likelihood function can be formed for the resulting model characterized by equations (18)–(21). Details regarding model estimation of the unknown parameters $(\gamma, \gamma_{0i}, \eta_i)$ are provided in Appendix A.

Given (20) and (21), household *i*'s probability of financial stress is a function of its residual income r_{it}, the fixed (or permanent) part of its subsistence level γ_{0i} , its welfare sensitivity (which reflects its marginal utility of consumption) η_i, and

[10]Not all households are observed in every period, and therefore the panel is unbalanced.

the information incorporated in the observed regressors x_{it}. It can be shown that this probability is given by

$$\Phi\left(\frac{\gamma_{0i} + x_{it}'\gamma - r_{it}}{\eta_i}\right) = \Phi\left(\frac{\gamma_{0i} + x_{it}'\gamma + \bar{h}_{it} - y_{it}}{\eta_i}\right), \tag{22}$$

where $\Phi(\cdot)$ is the standard normal distribution function and $\bar{h}_{it}$ is the household's housing accommodation expenditure (which, as described earlier, is based on the household's annual mortgage repayments or housing rent expenditure). It is clear from (22) that a greater η_i renders the household less sensitive to a change in income y_{it}, whereas low values of η_i will produce sharp changes in the household's well-being function after a change in income.

The resulting well-being function depends on the extent to which household i's income is sufficient to meet its expenditure requirements (which, in turn, reflect the household's subsistence level):

$$V_{it} = 1 - \Phi\left(\frac{\hat{k}_{it} - r_{it}}{\eta_i}\right) = 1 - \Phi\left(\frac{\gamma_{0i} + x_{it}'\gamma + \bar{h}_{it} - y_{it}}{\eta_i}\right), \tag{23}$$

with V_{it} clearly according with (5) such that the welfare interpretations for k_{it} and η_i stemming from Theorem 1 hold.

4. Measuring the Welfare Distortions Stemming from Heterogeneity

The next two subsections examine the distortions derived in Section 2. In Section 4.1, the paper evaluates whether there are meaningful *group-level* differences in income requirements and welfare sensitivities, with a focus on demographic groups. Although this evaluation provides information about demographic-level differences, it does not consider the welfare impact of household-level heterogeneity. Consequently, Section 4.2 examines *household-level* differences in equivalence scales. Importantly, it examines the shape of the distribution of equivalence scales for different household groups. It also reports on the centrality, variability, and skewness of each household group's distribution of equivalence scales, thereby providing a means for assessing how the assumption of a common group-level equivalent income deviates from the results observed in the data.

4.1. *Group-Level Welfare Distortions*

It is clear from the parameter estimates in Table 1 that the distribution of household-specific income requirements and welfare sensitivities is not uniform across the various household types. The fixed component of household income requirements ($\hat{\gamma}_{0l}$, where the l subscript pertains to the particular household types and family sizes listed in Table 1) is greatest for single-person households or

TABLE 1
Income Requirement and Welfare Sensitivity Parameters, $\hat{\gamma}_{0l}$ and $\hat{\eta}_l$, by Household Type and Size

	Obs.	$\hat{\gamma}_{0l}$	$\hat{\eta}_l$
Household type			
Couple family w/out children	4237	3.717**	1.425**
Couple family with children <15	4244	3.274	1.390
Couple family (dependent children)	513	3.449**	1.397
Couple family (non-dependent children)	641	3.716**	1.418**
Lone parent with children <15	1432	3.090**	1.379**
Lone parent (dependent children)	284	3.539**	1.385
Lone parent (non-dependent children)	551	3.611**	1.408**
Lone person	5527	3.880**	1.428**
Group household	353	3.829**	1.422**
Multi-family household	213	2.986**	1.376
Family size			
1	5527	3.979**	1.428**
2	5855	3.700**	1.419**
3	2904	3.546**	1.404**
4	2785	3.289	1.390
5	1127	2.868**	1.371**
6	370	2.808**	1.369**

Note: Values marked with ** indicate that the sample for the particular group is statistically different (at the 0.05 level) to the sample associated with the benchmark household (being a couple family with children <15 when grouping by household type or a family of size 4 when grouping by family size) using a Kolmogorov–Smirnov test.

households without children, and falls significantly when children are present in the household. Moreover, the value of $\hat{\gamma}_{0l}$ decreases with each additional family member.[11] Welfare sensitivity (in other words, the sensitivity of the household's probability of financial stress to a change in the household's residual income) also appears to differ by household type and size. In particular, $\hat{\eta}_l$ falls, and therefore welfare sensitivity increases, with larger family size. This implies that although larger households have smaller (per-capita) fixed income requirements $\hat{\gamma}_{0l}$, they also have a smaller capacity to adjust consumption thereby resulting in greater sensitivity to income changes.

To examine the welfare distortions associated with these parameters, I adopt the distortion measures discussed in Section 2.2. Essentially, I undertake an exercise based on a policy-maker who compensates household *groups* with the objective of equivalizing group l's welfare V_l with some benchmark $\overline{V}$. The policy-maker is required to make a choice regarding the extent to which group-specific requirements are accounted for in determining the level of compensation. The welfare distortions considered are based on the distortions $D_{\eta,i}$, $D_{\gamma_0,i}$ and D_i associated with the three cases described in Section 2.2, which are described as follows.

Case 1. Heterogeneity is allowed only for the fixed level of subsistence, with group l's marginal utility of consumption forced to be equal to the benchmark η^* (welfare distortions measured by $D_{\eta,i}$).

[11]Note that $\hat{\gamma}_{0l}$ is a per-family member measure of fixed income requirements. Consequently decreasing estimates of $\hat{\gamma}_{0l}$ as family size increases provide evidence of economies of scale.

TABLE 2
EQUIVALENCE SCALES UNDER ALTERNATIVE ASSUMPTIONS REGARDING HETEROGENEITY

	Unrestricted	Case 1	Case 2	Case 3	Distortion
Grouped by household type					
Couple family w/out children	0.519	0.517	0.472	0.470	0.049**
Couple family with children <15	1.000	1.000	1.000	1.000	N/A
Couple family (dependent children)	0.927	0.927	0.893	0.892	0.035**
Couple family (non-dependent children)	0.856	0.854	0.779	0.777	0.079**
Lone parent with children <15	0.688	0.689	0.717	0.717	−0.029**
Lone parent (dependent children)	0.646	0.647	0.611	0.612	0.034**
Lone parent (non-dependent children)	0.549	0.548	0.511	0.510	0.039**
Lone person	0.270	0.269	0.238	0.237	0.033**
Group household	0.618	0.617	0.553	0.552	0.066**
Multi-family household	1.165	1.166	1.244	1.246	−0.081**
Grouped by family size					
1	0.285	0.285	0.248	0.247	0.038**
2	0.547	0.545	0.496	0.495	0.052**
3	0.806	0.806	0.759	0.758	0.048**
4	1.000	1.000	1.000	1.000	N/A
5	1.121	1.123	1.243	1.245	−0.124**
6	1.329	1.332	1.502	1.505	−0.176**

Note: Unrestricted equivalence scales use group l's "true" values $\hat{\eta}_l$ and $\hat{\gamma}_{0l}$ calculated as the sample averages of the parameters for households in group l. Case 1 imposes the restriction that welfare sensitivities are homogeneous ($\hat{\eta}_l = \eta^*$). Case 2 imposes the restriction that the fixed part of the subsistence level is homogeneous ($\hat{\gamma}_{0l} = \gamma_0^*$). Case 3 imposes both restrictions. The estimated distortion is the difference between the equivalence scales obtained in the unrestricted case (which assumes full heterogeneity) and Case 3 (which assumes full homogeneity).

**Significance at the 0.01 level.

Case 2. Heterogeneity is allowed only for the marginal utility of consumption, with group l's fixed level of subsistence forced to be equal to the benchmark γ_0^* (welfare distortions measured by $D_{\gamma_0,i}$).

Case 3. Homogeneity is forced for both the marginal utility of consumption and the fixed level of subsistence (whereby welfare distortions are measured by D_i).

Which Household Groups Benefit from the Assumption of Homogeneity?

For ease of comparison across groups, Table 2 presents the distortions in the equivalence scales (e.g. as per equation (14) for Case 1) rather than the distortions in the equivalent incomes. This is simply the distortion scaled by the benchmark income y^*. The difference between the unrestricted equivalence scales and the equivalence scales for Case 3 (which assume no unobserved heterogeneity) reflects the total distortion stemming from the failure to account for unobserved group-level heterogeneity. The benchmark household is assumed to be the average family of size 4 (in other words, the benchmark values $y^*, \eta^*, \gamma_0^*, \tilde{c}^*$ are sample averages across *all* households of size 4, including periods where these households were stressed), but choosing another type of reference household does not affect the general conclusion. A positive distortion indicates that the failure to account for group-level differences in subsistence levels and marginal utilities of consumption results in *under*-compensation with the household requiring additional income to achieve the well-being of the benchmark household group (the converse holds for a negative distortion).

It is clear from Table 2 that the assumption of homogeneity will substantially over-compensate larger households and under-compensate smaller (up to three persons) households. In all cases, the distortions are statistically significant at the 0.01 level. A household with five members is typically over-compensated by about 12 percent if homogeneity is assumed, rising to nearly 18 percent for a household with six members. In general, the relationship between distortion and family size implies that larger households typically require less to achieve the same welfare level; this suggests that larger households, perhaps like older households (Aguiar and Hurst, 2005), engage in some form of substitution to increase their welfare.

The reason for this over-compensation is that although larger households are less able to adjust their consumption bundles in response to an income change (as they have larger levels of welfare sensitivity $1/\widehat{\eta}_l$, and therefore, by Theorem 1, a greater marginal utility of consumption), they also have smaller subsistence levels (on a per-person basis) and these *more* than offset their higher marginal utility of consumption.

A comparison of the equivalence scales obtained when allowing for heterogeneity with those obtained when imposing the restrictions of a common marginal utility of consumption (Case 1) or a common fixed level of subsistence (Case 2) indicates that *almost all* of the welfare distortions are attributable to the restrictions in subsistence levels. This establishes the first key empirical finding based on the distortion measures constructed in Section 2: a policy-maker who equivalizes income under the assumption that per-capita subsistence levels (or "needs") are common across different household groups is likely to yield significant welfare distortions.

4.2. *Importance of Unobserved Household-Level Differences*

The calculations in the preceding subsection indicate that differences in subsistence levels, rather than marginal utilities of consumption, are responsible for most of the demographic *group-level* distortion resulting from the assumption of homogeneity. This has important consequences for the design of effective income support policies. At the group level, the results provide evidence in favor of the welfare importance of heterogeneity driven by adjustment costs (Chetty and Szeidl, 2016), and against the importance of habit-driven heterogeneity (see, e.g. Calvet and Comon, 2003). However, measurement of the unobserved characteristics that Pollak and Wales (1979) and Fisher (1987) highlight as relevant to any welfare comparison also requires an assessment of *intra-group* differences between households. It is these differences that provide information on what van Praag *et al*. (1980) calls the "contour" of income requirements that stem from household heterogeneity.

To see why intra-group unobserved differences are important, consider the scenario where the policy-maker chooses to ignore unobserved household-level heterogeneity. In this scenario, the equivalence scale for group l is determined by the appropriate transformation of (17):

$$\frac{\overline{y}_l}{y^*} \mid \left(\eta_l = \eta^*, \gamma_{0l} = \gamma_0^*\right) = 1 + \frac{\tilde{c}_l - \tilde{c}^*}{y^*}, \tag{24}$$

where $\tilde{c}_l = x_l' \gamma$ characterizes the observed heterogeneity of group l (with $\tilde{c}^*$ characterizing the observed heterogeneity of the benchmark group). It is clear that all households belonging to group l will have the same equivalence scale. This is the approach that is generally adopted in the literature.[12]

Conversely, if the policy-maker accounts for unobserved household-level heterogeneity, then they will adopt the equivalence scale

$$\frac{\bar{y}_{il}}{y^*} = \frac{\eta_{il}}{\eta^*} + \frac{\hat{k}_{il} - \frac{\eta_{il}}{\eta^*}\hat{k}^*}{y^*}, \tag{25}$$

where $\hat{k}_{il} = \gamma_{0il} + \tilde{c}_l$ is the income requirement of household i belonging to group l (made up of a household-specific component γ_{0il} and a group component $\tilde{c}_l$). The parameters $\hat{k}^*$ and η^* reflect the income requirement and welfare sensitivity of the average four-member household (which is used as the benchmark household for the analysis). In this case, the equivalence scale is able to account for heterogeneity in the marginal utility of consumption (through $\frac{\eta_{il}}{\eta^*}$) and in subsistence levels (through $\hat{k}_{il} - \frac{\eta_{il}}{\eta^*}\hat{k}^*$). If there is no such heterogeneity, then the difference between $\frac{\bar{y}_{il}}{y^*}$ and $\frac{\bar{y}_{il}}{y^*} \mid (\eta_l = \eta^*, \gamma_{0l} = \gamma_0^*)$ will be zero.

Intra-group variation in η_{il}, γ_{0il} will produce a contour of equivalence scales for group l (reflecting the equivalent incomes of the households belonging to the group) that is used to measure the welfare distortion of unobserved differences. To evaluate these distortions, it is instructive to consider the general appropriateness of grouping households by reference to family size as this grouping is ubiquitously adopted for the purpose of reporting equivalence scales (see, e.g. Jorgenson and Slesnik, 1984, 1987; Phipps and Garner, 1994).

In general, households of a given size will differ in terms of both observed and unobserved features. In terms of the former, variation in $\tilde{c}_{il}$ (being the part of the household's income requirement that is explained by observed characteristics) for each household in group l represents variation in observed heterogeneity. To restrict dispersion to that attributable to *unobserved* heterogeneity, a common value $\tilde{c}_l$ is required across all households in group l. Accordingly, without loss of generality, $\tilde{c}_l$ is set equal to $\tilde{c}^*$.

Figure 1 shows the entire distribution (or contour) of welfare distortions that are attributable to unobserved heterogeneity.[13] These are calculated as the difference between equations (25) and (24).[14] For each family size, the level of distortion because of unobserved heterogeneity is substantial. Although differences in both marginal utilities and subsistence levels contribute to the distortion, the latter is

[12]If group l is characterized by a subset of x_l, then equivalence scales will only differ by reference to observed differences in the compliment of that subset with respect to x_l.

[13]To some extent, the figure will also be reflective of factors other than unobserved heterogeneity (such as measurement error). However, it is unlikely that differences between the distributions are attributable to measurement error which should be similar across different household groups.

[14]Because $\tilde{c}_l$ is restricted to be equal to $\tilde{c}^*$, the equivalence scale (24) is, by definition, always equal to unity such that the distributions in Figure 1 are also the distributions of the heterogeneous equivalence scale (25) after subtracting unity.

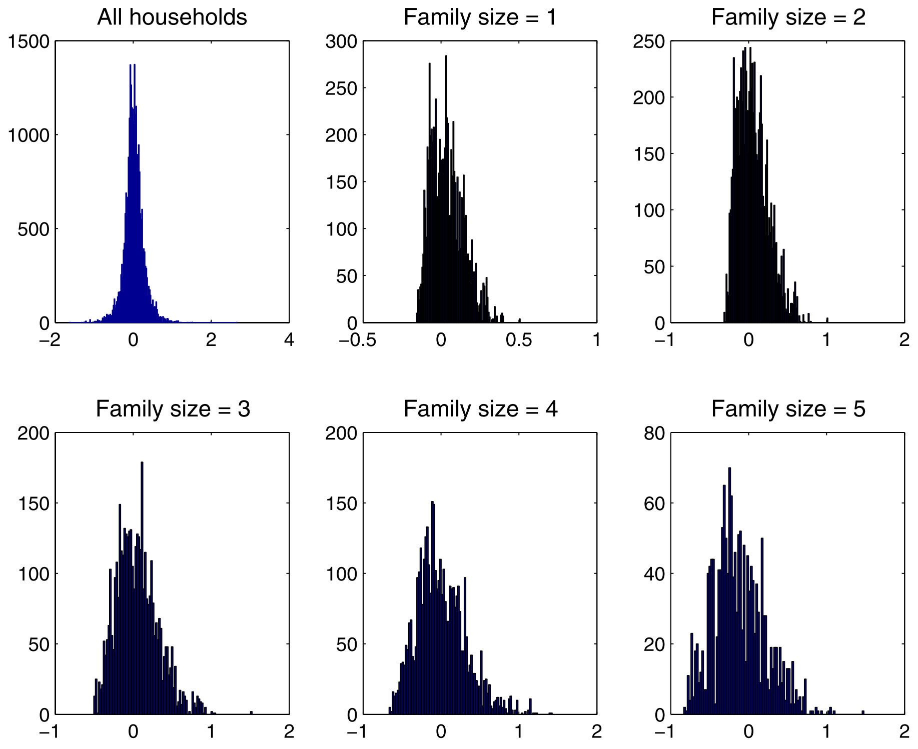

Figure 1. Distribution of Welfare Distortions D_l^* (For All Households and by Groups Based on Household Size) Stemming from Unobserved Heterogeneity. Positive Values Imply That the Relevant Households in the Given Group Require Additional Income to Attain the Well-Being Level of the Benchmark Four-Person Household [Colour figure can be viewed at wileyonlinelibrary.com]

responsible for the majority of the dispersion. However, the distributions clearly differ across the various family sizes, suggesting that the welfare distortion associated with the failure to account for unobserved heterogeneity will be greater for some family sizes relative to others.

This result has significance for welfare analysis. It is a standard, and often implicit, assumption in welfare analysis that the "distribution of unconditional preferences is independent of the distribution of demographic characteristics" (Pollak and Wales, 1979). The distributional differences across family sizes in Figure 1 render it relatively clear that this assumption is untenable. Indeed, the null hypothesis that the distributions are from the same underlying distribution is rejected at the 0.01 level for *every* family size.[15]

The Mean, Variance, and Skewness of the Welfare Distortions

Table 3 specifies the centrality, variance, and skewness of the distribution of heterogeneous equivalence scales. The variance and skewness have direct welfare

[15]The tests are based on the Kolmogorov–Smirnov test, and the resulting p-values are provided in Table 3.

TABLE 3

Summary Statistics for the Distribution of Welfare Distortions Stemming from (i) Unobserved Heterogeneity and (ii) Both Observed and Unobserved Heterogeneity

Family Size	(i) Unobserved Heterogeneity					(ii) Observed and Unobserved Heterogeneity				
	Mean	Std	MAD	Skew	Sig.	Mean	Std	MAD	Skew	Sig.
1	0.287	0.104	0.084	0.638	<0.01	0.285	0.120	0.098	0.491	<0.01
2	0.551	0.196	0.157	0.712	<0.01	0.547	0.225	0.182	0.547	<0.01
3	0.798	0.268	0.212	0.654	<0.01	0.806	0.294	0.231	0.522	<0.01
4	1.000	0.315	0.249	0.707	–	1.000	0.337	0.266	0.594	–
5	1.126	0.336	0.268	0.519	<0.01	1.121	0.377	0.300	0.422	<0.01
6	1.325	0.455	0.348	0.920	<0.01	1.329	0.493	0.375	0.798	<0.01

Note: The summary statistics for unobserved heterogeneity omit the effects of observed heterogeneity, whereas the summary statistics for "observed and unobserved heterogeneity" account for both forms of heterogeneity. MAD is the mean absolute deviation and represents the average welfare distortion or social welfare loss. Significance tests are based on a Kolmogorov–Smirnov test that the sample of equivalence scales for family size j is from the same underlying distribution as the distribution for four-person households.

interpretations if it is assumed that the policy-maker equivalizes welfare by adopting either a homogeneous equivalence scale (24) or the sample mean of the heterogeneous equivalence scale (25). In particular, the standard deviation represents the dispersion of welfare distortions stemming from unobserved heterogeneity, whereas the skewness provides information regarding the typical direction of the distortion. Furthermore, the mean absolute deviation (MAD) represents the average welfare distortion. This is important as it is effectively a measure of the average *social welfare loss* when the policy-maker assumes a transfer policy that does not account for household-level heterogeneity.

The impact of unobserved heterogeneity on the equivalence scales is substantial and increases by family size. For families of size 3, 4, or 5, the average welfare distortion stemming from unobserved heterogeneity is approximately 25 percent. This value rises to 35 percent for families of size 6, but is only about 8 percent for a lone person. In all cases, significant positive skewness implies that the welfare distortions associated with the failure to account for unobserved heterogeneity are *asymmetric* and typically result in under-compensation.

The Welfare Importance of Unobserved Heterogeneity Relative to Observed Heterogeneity

It follows naturally to consider the magnitude of unobserved heterogeneity relative to observed heterogeneity. If the impact of unobserved heterogeneity is small once observed heterogeneity is also accounted for, the results (although still informative) become substantially less important. To assess the magnitude of observed heterogeneity, the equivalence scales are recomputed after allowing $\tilde{c}_l$ to vary across households (i.e. using the actual $\tilde{c}_{il}$ for each household). The resulting distributional statistics are presented as the second set of results in Table 3 and reflect the distortions from both observed and unobserved heterogeneity.

The change in the average distortion after allowing for observed heterogeneity (measured by the difference in the two sets of MAD values in Table 3) is relatively minor. In this respect, because the variance of the observed component of heterogeneity $\tilde{c}_{il}$ must be greater than or equal to zero, it follows that the standard deviation of the equivalence scales when allowing for observed and unobserved heterogeneity will be greater than or equal to the standard deviation obtained when allowing only for unobserved heterogeneity, and therefore $\sigma_{\text{total}} \geq \sigma_{\text{unobserved}}$. The proportion of the variability in the welfare distortions that can be attributed to unobserved heterogeneity can therefore be determined as the ratio of the standard deviations:

$$\frac{\sigma_{\text{unobserved}}}{\sigma_{\text{total}}}, \tag{26}$$

where $\sigma_{\text{unobserved}}$ and σ_{total} are obtained from the standard deviations in Table 3.

Pursuant to equation (26), just under 90 percent of the heterogeneity in the equivalence scale is attributable to unobserved characteristics; the weighted average percentage of heterogeneity attributed to unobserved characteristics is 88.8 percent. The values exhibit relatively little variation across household sizes, ranging from a lowest of 86.7 percent for single-person households to a highest of 93.4 percent for four-person households. Therefore, the distortion stemming from unobserved heterogeneity appears to be substantial, both in its own right and relative to the impact of observed heterogeneity. Similar results are obtained even if the panel's inclusion criteria are modified from households that have been interviewed at least six times to households that have been interviewed eight or ten times. The observed result is also maintained when initially restricting the household lifespan to a minimum of six periods for all households and then repeating the estimation at each time point by adding an extra observation (if available).

Irrespective of the precision of the estimate provided by equation (26), however, the magnitude of the heterogeneity attributed to unobserved factors leads to the second key empirical finding that a large proportion of the variation in equivalence scales is attributable to the unobserved characteristics of households. The failure to account for this unobserved heterogeneity is likely to produce significant welfare distortions, primarily in the form of under-compensation.

5. Conclusion

This paper shows that a formal welfare interpretation can be given to the output obtained from reduced-form models of well-being. This general result is obtained by showing that the equivalent income derived from models of well-being (e.g. reduced-form probit or logit models) can be mapped to the equivalent income stemming from utility functions that allow for household-specific subsistence levels and marginal utilities of consumption. The paper uses this general result to derive measures of the welfare distortion attributable to unobserved heterogeneity in subsistence levels and marginal utilities of consumption, which can be computed using standard reduced-form well-being models. These serve as a basis for examining the critical issue raised in Pollak and Wales (1979) and Fisher (1987) regarding the distortions stemming from a failure to account for unobserved heterogeneity in household expenditure requirements.

The welfare distortion measures are applied to an Australian data set that is used as a case study for examining the distribution of the welfare distortions. The paper finds that unobserved heterogeneity produces welfare distortions in two distinct ways. First, a policy-maker who equivalizes income under the assumption that per-capita subsistence levels (or "needs") are common across different household groups is likely to yield significant welfare distortions. The assumption of a common per-capita subsistence level is also associated with the typical over-compensation of larger households, even though the evidence suggests that larger households have a greater marginal utility of consumption. The primary reason for this is that the greater marginal utility of consumption of larger households is more than offset by their lower per-capita subsistence levels.

Second, the failure to account for unobserved household-level heterogeneity results in substantial and asymmetric welfare distortions, with the asymmetry typically resulting in the under-compensation of households. This is primarily because of the failure to capture differences in subsistence levels, rather than heterogeneity in marginal utilities of consumption.

The results have significant ramifications for understanding the impact of income and transfer policies, and highlight the importance of accounting for heterogeneity in subsistence levels for the formulation of effective income support measures.

References

Aguiar, M. and E. Hurst, "Consumption Versus Expenditure," *Journal of Political Economy*, 113, 919–48, 2005.

Aiyagari, S. R., "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109, 659–84, 1994.

Attanasio, O. and G. Weber, "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 103, 1121–57, 1995.

Australian Bureau of Statistics, *Household Expenditure Survey 2009-10*, Cat. No. 6530.0, Canberra, 2011.

Blackorby, C. and D. Donaldson, "Measuring the Cost of Children: A Theoretical Framework," in R. Blundell, I. Preston, and I. Walker (eds), *The Measurement of Household Welfare*, Cambridge University Press, Cambridge, 51–69, 1994.

Blundell, R., D. Kristensen, and R. L. Matzkin, "Bounding Quantile Demand Functions Using Revealed Preference Inequalities," *Journal of Econometrics*, 117, 112–27, 2014.

Blundell, R., L. Pistaferri, and I. Preston, "Consumption Inequality and Partial Insurance," *American Economic Review*, 98, 1887–921, 2008.

Briesch, R. A., P. K. Chintagunta, and R. L. Matzkin, "Semiparametric Estimation of Choice Brand Behavior," *Journal of the American Statistical Association*, 97, 973–82, 2002.

______, "Nonparametric Discrete Choice Models with Unobserved Heterogeneity," *Journal of Business and Economic Statistics*, 28, 291–307, 2009.

Browning, M., P.-A. Chiappori, and A. Lewbel, "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power," *Review of Economic Studies*, 80, 1267–303, 2013.

Browning, M. and T. F. Crossley, "The Life-Cycle Model of Consumption and Saving," *Journal of Economic Perspectives*, 15, 3–22, 2001.

Buera, F. J., A. Monge-Naranjo, and G. E. Primiceri, "Learning the Wealth of Nations," *Econometrica*, 79, 1–45, 2011.

Calvet, L. and E. Comon, "Behavioural Heterogeneity and the Income Effect," *Review of Economics and Statistics*, 85, 653–69, 2003.

Cherchye, L., I. Crawford, B. De Rock, and F. Vermeulen, "Revealed Preference and Aggregation," *mimeo*, 2015.

Chetty, R. and A. Szeidl, "Consumption Commitments and Risk Preferences," *Quarterly Journal of Economics*, 122, 831–77, 2007.

______, "Consumption Commitments and Habit Formation," *Econometrica*, 84(2), 855–90, 2016.

Chiappori, P. A., "Equivalence Versus Indifference Scales," *Economic Journal*, 126, 523–45, 2016.

Clarida, R. H., "Aggregate Stochastic Implications of the Life Cycle Hypothesis," *Quarterly Journal of Economics*, 106, 851–67, 1991.

Decancq, K., M. Fleurbaey, and E. Schokkaert, "Wellbeing Inequality and Preference Heterogeneity," *Economica*, 84, 210–38, 2017.

Donaldson, D. and K. Pendakur, "The Identification of Fixed Costs From Consumer Behavior," *Journal of Business and Economic Statistics*, 24, 255–65, 2006.

Dynan, K. E., J. Skinner, and S. P. Zeldes, "Do the Rich Save More?," *Journal of Political Economy*, 112, 397–444, 2004.

Fernandez-Villaverde, J. and D. Krueger, "Consumption Over the Life Cycle: Some Facts From Consumer Expenditure Survey Data," *Review of Economics and Statistics*, 89, 552–65, 2007.

Fisher, F. M., "Household Equivalence Scales and Interpersonal Comparisons," *Review of Economic Studies*, 54, 519–24, 1987.

______, "Household Equivalence Scales: Reply," *Review of Economic Studies*, 57, 329–30, 1990.

Fleurbaey, M., "Beyond Income and Wealth," *Review of Income and Wealth*, 61, 199–219, 2015.

Fleurbaey, M. and D. Blanchet, *Beyond GDP: Measuring Welfare and Assessing Sustainability*, Oxford University Press, Oxford, 2013.

Gorman, W., "Community Preference Fields," *Econometrica*, 21, 63–80, 1953.

______, “On a Class of Preference Fields,” *Metroeconomica*, 13, 53–6, 1961.
Gourinchas, P. E. and J. Parker, “Consumption Over the Life Cycle,” *Econometrica*, 70, 47–89, 2002.
Gul, F. and W. Pesendorfer, “Random Expected Utility,” *Econometrica*, 74, 121–46, 2006.
Hausman, J. and W. Newey, “Nonparametric Welfare Analysis,” *Annual Review of Economics*, 9, 521–46, 2017.
Jorgenson, D. W. and D. T. Slesnick, “Aggregate Consumer Behaviour and the Measurement of Inequality,” *Review of Economic Studies*, 51, 369–92, 1984.
______, “Aggregate Consumer Behaviour and Household Equivalence Scales,” *Journal of Business and Economic Statistics*, 5, 219–32, 1987.
Koulovatianos, C., C. Schroder, and U. Schmidt, “Do Demographics Prevent Consumption Aggregates From Reflecting Micro-Level Preferences?,” *European Economic Review*, 111, 166–90, 2019.
Lewbel, A. and K. Pendakur, “Unobserved Preference Heterogeneity in Demand Using Generalized Random Coefficients,” *Journal of Political Economy*, 125, 1100–48, 2017.
Lim, G. C. and S. Tsiaplias, “Household Income Requirements and Financial Conditions,” *Empirical Economics*, 57, 1705–30, 2019.
Lusardi, A., D. Schneider, and P. Tufano, “Financially Fragile Households: Evidence and Implications,” *Brookings Papers on Economic Activity*, 42, 83–150, 2011.
Mazzocco, M., “Saving, Risk Sharing, and Preferences for Risk,” *American Economic Review*, 94, 1169–82, 2004.
______, “Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment,” *Review of Economic Studies*, 74, 857–95, 2007.
McFadden, D., “Econometric Analysis of Qualitative Response Models, ” in Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics Volume II*, North-Holland, Amsterdam, 1395–457, 1984.
Morciano, M., R. Hancock, and S. Pudney, “Disability Costs and Equivalence Scales in the Older Population in Great Britain,” *Review of Income and Wealth*, 61, 494–514, 2015.
Pendakur, K., “Welfare Analysis When People are Different,” *Canadian Journal of Economics*, 5, 321–60, 2018.
Phipps, S. and T. I. Garner, “Are Equivalence Scales the Same for the United States and Canada?,” *Review of Income and Wealth*, 40, 1–17, 1994.
Pollak, R. A. and T. J. Wales, “Welfare Comparisons and Equivalence Scales,” *American Economic Review*, 69, 216–21, 1979.
Senik, C., “When Information Dominates Comparison: Learning from Russian Subjective Panel Data,” *Journal of Public Economics*, 88, 2099–123, 2004.
van Praag, B. M., *Individual Welfare Functions and Consumer Behavior*, North Holland Publishing Co., Amsterdam, 1968.
______, “The Welfare Function of Income in Belgium: An Empirical Investigation,” *European Economic Review*, 2, 337–69, 1971.
van Praag, B. M., T. Goedhart, and A. Kapteyn, “The Poverty Line—A Pilot Survey in Europe,” *Review of Economics and Statistics*, 62, 461–5, 1980.
Zaidi, A. and T. Burchardt, “Comparing Incomes When Needs Differ: Equivalization for the Extra Costs of Disability in the U.K,” *Review of Income and Wealth*, 51, 89–114, 2005.