The Gender Implications of Tax Policy and Inequality Measurement

Patricia Apps
University of Sydney Law School and IZA

Ray Rees
University of Munich, University of Sydney Law School and CESifo

March 18, 2020

Abstract

This paper provides a critique of the standard methodology which bases welfare comparisons between households on deflating household income and consumption values by equivalence scales. We argue that this leads to support for tax/transfer policies that significantly disdvantage low to middle income households as well as women as second earners. We base the critique both on a theoretical model of the family household and a numerical analysis of Australian income and employment data.

Key words: gender; equivalence scales; tax/transfer policy.

1 Introduction

The use of equivalisation indices to deflate measures of household income and consumption in empirical studies of the family household has become widespread and routine, despite the fact that some economists have argued strongly against adopting it.[1] In practice a number of widely used equivalence scales exist. Their aim is based on an assumed need to take account of the size and composition of households and the economies of scale in household consumption in making across-household welfare comparisons.

For example the widely used OECD "square root" scale[2] deflates household aggregates such as gross and disposable incomes and total consumption by the square root of the number of individuals in the household. These are then used to construct indices of inequality across entire economies. Another typical example is the "OECD modified" scale used by the Australian Productivity Commission.[3] A scale of 1 point is used for the first adult, 0.5 for each additional

[1] For example see Pollak and Wales (1979) who completely rejected the procedure, and Atkinson (1970) who criticised the "needs based" approach to their construction.

[2] See, for example, OECD (2000) aand Sila and Dugan (2019).

[3] Report of the Australian Productivity Commission (2018).

person aged 15 years or more, and 0.3 points for each child under 15 years. A number of other scales that use a similar procedure, but with different numbers, have been proposed over the years.

The idea is to take a reference point, typically a single adult household, and on the basis of some calculation of the "needs" of individuals of different ages, to associate with each household in the sample a number of "adult equivalents". The idea of economies of scale in household production, usually expressed in the old adage "two can live as cheaply as one", is reflected in weights of less than one for adults beyond the first. Thus a family of two adults with two children under 15 years is considered under the OECD modified scale to be equivalent to 2.1 single adults, and so if its income or consumption is 2.1 times that of the (average) single person the two *households* are considered equally well off.

The assumption of simple "economies of scale" however does not do justice to the complexity of realistic household production processes - see for example the critique of the Becker model of household production in Pollak and Wachter (1975), who emphasise the importance of multi-activity production functions characterised by significant joint production.[4]

In the more theoretical literature attempts have been made to put these procedures on a less *ad hoc* basis.[5] Angus Deaton and John Muellbauer,[6] at the end of a brilliantly clear exposition of the theoretical fundamentals of the approach (which they basically favour on the grounds of its econometric advantages), write:[7]

> We note, finally, a source of some difficulty in the treatment of children [in the economic analysis of the household]. *So far we have cavalierly ignored the distinction between households and individuals, treating the two terms more or less interchangeably.** Our preference in analysing behaviour is to treat the *household*** as the basic decision-making unit, modeling the behavioral impact of family composition through the equivalence scale [...]. But this is not entirely satisfactory. Social welfare is formed over individual welfares so that society is not likely to be unconcerned about how members of families are treated [...]. The social welfare function should thus have a "slot" for each individual, and *if each family member has the same welfare level,** the family per capita equivalent real income can be used as the welfare indicator for each individual. [This approach is] *not entirely satisfactory without a theory of (or at least some assumption about) allocation within the household.**
>
> *Our italics. **Italics in the original

We should first point out that the authors are ignoring not only the "treat-

[4] The fact for example that one can be minding a small child as well as carrying out other household tasks is often ignored in time use studies which then produce misleadingly low estimates of the time a household spends in child care.

[5] See Appendix 1 for an outline of the Deaton/Muellbauer analysis.

[6] Deaton and Muellbauer (1980).

[7] Deaton and Muellbauer (1980), p 226.

ment of children" but also gender differences in couple households, which of course are closely related to the presence of the children. This is reinforced by the absence of any consideration of household production. It would be missing the point to base an analysis involving gender equity on an assumption that the household's sharing rule involves equal utilities for all its members.

Our aim in this paper is to show that household equivalence scales are both unnecessary and misleading. They give support to tax/transfer policies that make low-to-middle income households, and women in those households as second earners, significantly worse off than they would be under alternative policies. As cases in point, we take the withdrawal of child support payments on the basis of a household's joint income, and the choice of the tax base for personal income taxation as joint rather than individual income, as important examples of this. Following the presentation of our model giving the conceptual framework for the discussion, we present numerical examples based on Australian data to illustrate our argument.

2 A critique of equivalisation indices

2.1 The conceptual framework

This critique is based on a model of household decision-taking behaviour.[8] We use this model[9] to argue that there are two fundamental weaknesses of the equivalisation index approach that undermine its usefulness in applications to policy.

First, the procedure of using indices to deflate household joint income embodies the idea that household income and wellbeing are co-monotonic - one necessarily increases with the other as we move through the equilibria of the given set of households.[10] In fact, given the marked heterogeneity in second earner labour supply, which has been firmly established empirically, joint labour market income is an inaccurate and misleading indicator of household wellbeing, as we show below. The root of the misconception lies in ignoring the existence of household production in a multi-person household.

Secondly, it assumes that the components of the household type vector[11] are fully observable, and therefore rules out consideration of the implications of the

[8] We draw here on the model of the household as a small economy engaged in intra-household production and exchange in Apps (1982). In the original formulation this was a general equilibrium model in which market wage rates were endogenous. The gender wage gap was driven by the "crowding" of women into "female" occupations, which leads to a higher male wage and lower implicit price for the household good. Here we simplify by assuming that the terms of this exchange are determined exogenously.

[9] Set out fully in Appendix 2.

[10] This is true not only of the sophisticated measures such as that in Deaton and Muellbauer, but also of the needs-based counting measures that compute the index by attaching a number to each individual and then adding them up.

[11] The list of household characteristics that can cause two households with the same incomes and preferences, and facing the same prices, to choose different consumption bundles and time allocations. See Appendix 1 for more formal definitions.

fact that important components of this vector are not observable, or at least not available in existing datasets. This again has to do with the omission of household production, as we also show below.

In the model, a household consists of two adults, each of whom works in the labour market and also supplies time to production of a household good, which could, but need not, be thought of as child care.[12] They also consume some of their own time as leisure. They have standard individual utility functions

$$u_{ih} = u(x_{ih}, z_{ih}, l_{ih}) \quad i = 1, 2 \qquad h = 1, 2, ..., H \tag{1}$$

where 1 is the primary and 2 the second earner.[13] Consumptions of a composite market good, the domestic good and leisure are respectively x, z and l, all strictly positive.[14] By definition,1 has the higher labour market earnings.

The household production function is

$$\sum_{i=1}^{2} z_{ih} = z_h = f(k_{1h}a_{1h}, k_{2h}a_{2h}, q_h b_h) \qquad h = 1, 2, ..., H \tag{2}$$

where a_{ih} is i's time input into household production and b_h is a market input bought at price p_h. Note that this price may vary across households.[15] Moreover, the k_{ih}, q_h are productivities/qualities of the inputs into household production that may also vary across households.[16]

The individual time constraints are

$$a_{ih} + l_{ih} + L_{ih} = T \quad i = 1, 2 \tag{3}$$

where T is total time available and $L_{ih} \geq 0$ are market labour supplies.

A key issue in formulating a household model is the budget constraint. Underlying the assumption that utilities of household members are equalised are two conditions:

1. transfer payments are feasible - utility is transferable

[12] In fact the model here is not fully applicable to households in the life cycle phase in which children under school age are present, since in that case additional constraints are placed on time use - someone always has to look after the kids. For further discussion and analysis see Apps and Rees (2018).

[13] Note that these are defined on role rather than gender, but empirically in OECD countries typically around 80% of second earners are women.

[14] For convenience of notation we assume that these 3 consumption goods are scalars rather than vectors, which in turn implies that, since x is the numeraire with price 1, all other prices and wage rates have been deflated by a consumption price index. This is without real loss of generality.

[15] For example, if the input b_h were non-parental child care, there could be a wide range of sources, such as family members and friends, child-minders, childcare centres and posh nannies.

[16] For example it is well-established in the literature that the quality of childcare - in the broad sense of creating human capital as well as simple physical child-minding - depends on parental human capital as well as other resources possessed by the household.

2. only the *pooled household budget constraint* must hold:

$$\sum_{i=1}^{2} x_{ih} + p_h b_h \leq \sum_{i=1}^{2} w_{ih} L_{ih} \quad h = 1, 2, ..., H \tag{4}$$

where w_{ih} are the market wage rates.

In this paper we want to move away from the case in which individual welfares are equalised and so we make the alternative assumption that within-household lump sum transfers are ruled out. We assume that individuals maximise utility subject to budget constraints determined by their own full incomes, given by the value of their time endowments at their own market wage rates, $w_{ih}T$.[17] This gives us a sharing rule that determines the within-household distribution of utility. The values of the individual total consumptions of the market and the household goods as well as leisure are determined by their full income. We then have the basis for the analysis of inequality and its measurement.

The household is fully rational in that it values each earner's time consistently at their outside market wage. In particular it prices individual leisures at their corresponding market wage and applies the implicit price π_h as the marginal opportunity cost of each individual's consumption of the household good.[18] This implies that we can represent the individual choice problems in terms of their *full income budget constraints*:

$$\max_{x_{ih}, z_{ih}, l_{ih}} u(x_{ih}, z_{ih}, l_{ih}) \;\; \text{s.t.} \;\; x_{ih} + \pi_h z_{ih} + w_{ih} l_{ih} \leq w_{ih} T \quad i = 1, 2, \quad h = 1, 2, ..., H \tag{5}$$

This emphasises that the wage rates w_{ih} and the prices π_h jointly determine the utility possibilities of all individuals in all households. From this we could define the type vector in this model as $[w_{1h}, w_{2h}, \pi_h]$ since these are the exogenous variables that determine the household equilibrium. However, the implicit price of the household good is itself a function given by[19]

$$\pi_h = c(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h) \tag{6}$$

So let us now consider the indirect utility and expenditure functions that result from this model, and how they compare to those delivered by the model of Deaton and Muellbauer. We can write them as, respectively:

$$v_{ih} = v(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h) \tag{7}$$

$$e_{ih} = e(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h) \tag{8}$$

First, the relevant functions belong to individuals rather than households. This is then consistent with an approach that says social welfare functions should be defined on individuals rather than on collectives such as households. This

[17] To solve the classical "adding up" problem we have to assume that the household production function has constant returns to scale.

[18] For details of the derivation of this price see the Appendix 1.

[19] See the Appendix 2.

is important in that it *completely removes the need for the ad hoc kinds of calculations that try to adjust for differences in household size and composition across an entire population.* Utilities are already individualised. The number and ages of children are taken into account in the variable z. Underlying this is the idea that children are a source of utility in the household, otherwise why have them?[20] The issue of "needs" created by children is dealt with in the price π_h, as well as by consumption costs, which will be higher when children are present. Underlying this whole discussion is the issue of life cycle phases, which we take up more fully in the next subsection.

Second, it removes household income, which is an endogenous variable, from the index measure, since the relevant functions are defined entirely on exogenous variables.

Third, it takes into account household production and the heterogeneity of second earner labour supplies that result from that. As we see from the analysis in the Appendix, labour supplies are given by the functions

$$L_{ih} = L_i(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h) \quad i = 1, 2, \quad h = 1, 2, ..., H \tag{9}$$

and so the model explains this heterogeneity simultaneously and consistently with its explanation of utility variation.

Therefore we could propose either one of the household's indirect utility function $v(.)$ or, if a monetary metric is required, its expenditure function $e(.)$ as the measure of the household's wellbeing or standard of living, and the basis of a welfare ranking across households. These functions have been extensively used in analyses of optimal tax/transfer policies.

2.2 The importance of the life cycle

It is often acknowledged that welfare comparisons of a population of households that are at different phases of their life cycles are of limited usefulness: what does it mean to compare the wellbeing of a couple with 2 children under school age and a retired couple? The use of standardised equivalisation indices regardless of issues like that essentially just sweeps them under the carpet. It is also often argued that the ideal would be the comparison of households' wellbeing over their entire life cycles, but this seems to be an impossible ideal to achieve.

We would argue instead that when the principal aim of welfare comparisons is to design tax/transfer policies a different approach is required. Quite obviously, there has to be differentiation between the policy issues around families with pre-school age children and retired couples. Households should be assigned to life cycle phases according to the commonality of the preference structures and constraint sets that confront them, since these are the policy-relevant determinants of their wellbeing. For this reason, we have argued for a specification

[20] An alternative approach could specify children as individuals with their own utility functions in the model, but since household decisions are typically taken by adults with the perceived utility of children as arguments of their own utility functions the approach given here may be an acceptable reduced form.

of the life cycle based not on the age of the "head of the household", as is the usual case, but rather upon the the phases through which family households typically go.[21]

For example, perhaps the broadest characterisation would be of a life cycle defined on 5 phases: the first phase in which the couple household has been formed but no children are present; the second, in which there are children of pre-school age in the household; the third, in which the children are older but still present in the household; the fourth, in which children have left home but the parents are still of working age: and the last phase in which both adults have retired. Clearly, across these phases preferences and constraints can be expected to vary significantly, as would the vectors of characteristics exemplified by $(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h)$ in the above analysis. We would need to model the precise form of this vector for each phase, and restrict welfare comparisons based on them to within-phase households.

However, we have to avoid the Nirvana fallacy: defenders of equivalisation scales would point to the non-observability of the type variables p_h, k_{1h}, k_{2h}, q_h as creating the need for some kind of second best measurement procedure. But what if, at any given wage pair (w_{1h}, w_{2h}), although these variables may be randomly distributed across the population with some given joint distribution, they are increasing functions of the wage pairs, which are potentially observable?[22] Then, rather than relying on joint household income which, because of the importance of heterogeneity in time spent in household production, is an unreliable index of household wellbeing, why not take primary earner income? Given the relatively small variation in primary earner working hours, this is a good proxy for the primary wage. In the case of second earners, there is an inverse relationship between household production and market income. If we sum the second earner income and the value of her household production, then it is reasonable to assume that across households this is also an increasing function of the primary wage or income.

In the following section we show numerically the effects on household welfare rankings of moving between joint and primary earner income as the measure of household wellbeing, for samples of households selected on the basis of observable characteristics. The aim of this analysis is to show how the use of equivalisation methods in a setting that implicitly sets household production to zero produces highly misleading implications for the evaluation of alternative tax policies.

[21] For a fuller discussion see Apps and Rees (2009), Ch 5. The approach suggested here could not be adopted if fertility were treated endogenously.

[22] Two points can be made here. In tax analysis of the Mirrlees (1971) type "wage rates" are said to be non-observable, but this is a semantic issue. What is meant there is the non-observability of the *innate ability* of a worker, which in perfect markets is reflected in the wage. In the present model we are concerned with actual wage rates, however determined. Secondly, we could stress the difference between "non-observable" and "unobserved". The importance of these type variables has simply been suppressed in the tendency to fit ideas into the standard framework, which underlies the use of equivalence scales.

3 Numerical analysis

We draw on household survey data to illustrate numerically the implications of switching from primary to household income as the basis for welfare comparisons across couples. The analysis is based on data for two samples of couple income unit records selected from the Australian Bureau of Statistics (ABS) 2015-16 Surveys of Income and Housing (SIH15-16). The first, labelled Data Set 1, is selected on the criteria that the primary earner is aged from 20 to 59 years, earns a minimum of $12/hour and works for a minimum of 30 hours per week. The aim in selecting on these criteria is to control for the observed high degree of heterogeneity in second earner labour supply at a given primary income across an otherwise relatively homogeneous subset of the population. The sample contains 5481 records.

We then select a second sample, labelled Data Set 2, on the further criteria that there are two dependent children present and both are aged under 15 years. The aim in this case is to control for demographic variation, in addition to second earner labour supply heterogeneity, at a given primary income. The sample contains 1055 records.

3.1 Data Set 1: 5842 couple income unit records

We first use Data Set 1 to illustrate the high degree of reordering of households when we switch from primary to household income as the welfare ranking variable. Table 1 presents the ranking by primary income and Table 2, by household income. Row 1 of each table reports decile data means for primary income and row 2, the decile means of second income within each quintile of primary income. Second earner participation rates in each decile of primary income are reported in the subsequent three rows of each table. The results indicate the very high degree of heterogeneity in second earner labour supply at each primary income level. The overall rates are 25%, 36% and 39% for non-participants (Non-P), part-time employed (PT) and full time employed (FT), respectively.

Tables 1 and 2 about here

In the primary income ranking, participation rates within each decile tend to be closely matching. Table 2 indicates the dramatic change in the decile distribution of participation rates when we switch to household income as the welfare ranking variable. An implication of the re-ranking in Table 2 is that two families with the same demographic characteristics are equally well off irrespective of whether the income is earned by both partners in FT work or by only one partner in FT work with the second a non-participant specialising in home production.

The Gini coefficient for the primary income ranking is slightly higher than that for the household income ranking, at 0.32 and 0.31, respectively. This difference reflects the shift of two-earner couples with lower individual incomes towards upper percentiles (see also Tables 5 and 6) and thereby the potential for an underestimate of the degree of inequality.

The profile of the number of dependent children across the primary income ranking tends to be U-shaped, falling from 1.12 in decile 1 to 0.95 in decile 5, and then rising to 1.25 in decile 10. The overall average is 1.05. The averages for Non-P, PT and FT are 1.22, 1.21 and 0.81, respectively. These numbers suggest that little of the heterogeneity in second earner labour supply can be explained by demographic variation. We investigate the issue further using Data Set 2.

3.2 Data Set 2: 1055 couple income unit records

Given the smaller sample size, results are presented for quintile distributions of primary and equivalised incomes. The latter are calculated for the "Oxford modified" scale: 1 point for first adult, 0.5 points for each additional person aged 15 years or older and 0.3 points for each child under 15 years. Give that the sample is limited to couple income units with two children under 15, we have: Equivalised income = Household income/2.1.

Primary income ranking

Table 3, row 1, reports the data means of primary income in each quintile and row 2, the data mean of second incomes within each quintile of primary income. The next three rows report second earner participation rates. Again, as in Table 1, the results reflect the high degree of heterogeneity in second hours within each quintile, with almost half of second earners, 46%, working part-time while 26% are either non-participants, and 28% are in FT work. Again, the quintile profiles of participation rates tend to change very little across primary incomes. Given that the sample is restricted to two-parent families with two dependent children under 15, variation in second earner labour supplies at a given primary income cannot be attributed to demographic characteristics or scale economies. The Gini coefficient is 0.33.

Equivalised income ranking

Consistent with the results in Tables 1 and 2, the quintile distribution of participation rates changes dramatically when we switch from a ranking defined on primary income to one defined on equivalised incomes. Table 4, row 1, reports the quintile data means of equivalised incomes and row 2, the data means of second incomes in each quintile of equivalised income. The last three rows report second earner participation rates. The Gini coefficient is 0.31.

Tables 3 and 4 about here.

The concentration of couple income units with a non-participating partner or part-time second earner towards the lower quintiles in Table 4 highlights the effect of implicitly setting the value of home production to zero in a ranking defined on equivalised income or household income. The percentage of single-earner households in quintile 1 rises from the 29%, as reported in Table 3, to 54%. In quintile 5 the non-participation rate falls from 25% to 12%.

3.3 Household subsets defined by median second hours

To investigate further the effects of switching from primary to equivalised income, we split the records in each quintile of primary income in Table 3 into two subsamples, labelled H1 and H2, according to median second hours within each quintile:

H1: households with 2nd earner working below median 2nd hours

H2: households with 2nd hours working above median 2nd hours.

Next, we split the the records in each quintile of equivalised incomes in Table 4 into two subsamples, labelled E1 and E2, according to median second hours. Thus we have:

E1: households with 2nd earner working at or below median 2nd hours

E2: households with 2nd hours working above median 2nd hours.

We now show that the ranking of households changes dramatically when we switch from primary to equivalised income as the ranking variable.

Primary income ranking

Table 5 compares the quintile data means of primary and second incomes across the two subsamples, H1 and H2. While the data means for H1 and H2 primary incomes are closely matching, the overall average 2nd income of H2 households is over 4 times that of H1 households, with 2nd income gaps across the distribution being of a similar magnitude.

Equivalised income ranking

The high degree of re-ranking associated with switching from primary to equivalised income as the ranking variable becomes evident when we compare the profiles of H1 and H2 primary and second income across a ranking defined on equivalised income, as in Table 6. H2 households with lower primary incomes and longer total hours of work are re-ranked towards the upper quintiles.

Tables 5 and 6 about here

Table 7 reports the percentages of H1 and H2 households across the distribution of equivalised income. The percentage of H1 households in quintile 1 rises from 50% to 76% and in quintile 5, falls from 50% to 28%, while the percentage of H2 households in quintile 1 falls for 50% to 24% and rises from 50% to 72% in quintile 5. This outcome is shown graphically in Figure 1.

Table 7 and Figure 1 about here.

3.4 Implications for tax policy

We now compare the distributional effects of individual vs. joint taxation when the ranking variable is first primary income and then equivalised income. The results illustrate the way in which an equivalised income measure of inequality not only lends support for joint taxation, with the effect of widening the net-of-tax gender pay gap, as is well recognised, but can also of shift the tax burden towards lower and middle income families.

The analysis is based on the Australian family income tax system as a case study. This system combines an individual based income tax, the Personal Income Tax (PIT), with a partly joint income tax system introduced by the withdrawal family payments on joint income, labelled Family Tax Benefit Part A (FTB-A). The following is an outline of the two systems.

The 2015-16 PIT scale has 5 taxable income brackets:

Bracket 1. $0 - $18,200

Bracket 2. $18,201 - $37,000
Bracket 3. $30,001 - $80,000
Bracket 4. $87,001 – $180,000
Bracket 5. $180,000 +

The following progressive rate scale applies to each bracket in turn: 0.0, 0.19, 0.325, 0.37 and 0.45.

The family tax benefit system (FTB-A) has a "maximum rate" which is withdrawn on joint income up to a ""Base Rate + Supplement per child" as follows:

Maximum Rate per dependent child per year:

Child under 13 years	$5,504.20
Child aged 13-14 years	$6,938.65

These maximum payments are withdrawn at 20 cents in the dollar on a family income above $52,706 up to the "Base Rate + Supplement per child" of $2,266.65, which is withdrawn at 30 cents in the dollar on a family income above $94,316.

For the two-child family, with one child aged under 13 and the other aged 13-14 years, the total of the Maximum Rate payment is therefore $12,442.85 per year.

PIT by primary income ranking

Table 8 compares the distribution of PIT burdens on H1 and H2 primary and second incomes across quintiles of primary income. The results are compared graphically in Figure 2. Because H1 and H2 primary incomes within each quintile are closely matching until the 5th quintile (see Table 5) tax burdens are close to equal. In the 5th quintile, the tax burden is significantly higher for H1 households due to their higher primary incomes, a reflection of steeply rising top incomes combined with the tendency for top income earners to have a partner who works fewer hours.

The overall gap between H1 and H2 PIT tax burdens reflect the higher H2 second incomes associated with working longer hours in PT or FT work. Nevertheless, given that second incomes are well below primary incomes, the overall total income tax gaps across the distribution are relatively small. This is due to the progressive marginal tax rate scale of the PIT – lower income earners pay proportionally less tax under a progressive rate scale when the tax base is individual incomes. As a result, an individual based PIT with a progressive rate scale reduces the net-of-tax gender pay gap.

PIT by equivalised income ranking

Table 9 reports income tax burdens on primary and second incomes across households ranked by equivalised income. Figure 3 plots the results graphically. Consistent with Table 7 and Figure 1, we see that switching from primary to equivalised or household income as the ranking variable results in a dramatic reordering of households.

The results differ widely from those reported for the primary income ranking. Partners within H2 households are now misrepresented as paying a lower level of income tax across the entire distribution. This outcome reflects the shift of low to middle income families in quintiles 1 to 3 of primary income towards higher quintiles. The results lend support for joint taxation and, thereby indirectly, for a wider net-of-tax gender pay gap.

Tables 8 and 9 about here

Figures 2 and 3 about here

Income tax adjusted for FTB-A, by primary income

When we introduce the family payment system, with its withdrawal of family payments on joint income, the overall tax gap between H1 and H2 rises. Moreover, the increase in H2 burdens are concentrated in the lower and middle income quintiles of primary income, as illustrated in Table 10 and Figure 4. This outcome reflects the way in which a tax system based on joint income not only widens the net-of-tax gender pay gap, as is well recognised, but also shifts the tax burden towards households in the lower and middle ranges of primary income. In quintile 1 the income tax gap of $2434 widens to $5604 due to an average FTB payment of $8,428 for H1 households but of only $5406 for H2 households. Similarly, the tax gap in quintile 2 widens from $6,193 to $10,371 and in quintile 3, from $9,289 to $11,053 due to the withdrawal of family payments on joint income.

Income tax adjusted for FTB-A, by equivalised income

The impact of FTB-A on the distribution of E1 and E2 tax burdens across equivalised income is illustrated in Table 11 and Figure 5. We can see that the effect is to reduce gains for E2 low and middle income earners that are due to the progressive rate scale of the PIT, a loss that will mistakenly be viewed as a gain in equity. In quintile 1 the tax burden on E2 rises even though they may be working twice the hours of the primary earner in the E1 household with the same household income.

Tables 10 and 11 about here

Figures 4 and 5 about here

4 Conclusions

In this paper we have presented a critique of the standard methods of using equivalence scales to make inter-household welfare comparisons. The basis of this critique is the argument that they produce misleading results in support of tax/transfer policies which disadvantage low to middle income households and the women in those households as second earners. In particular they supply arguments in favour of joint taxation and the withdrawal of child benefits on the basis of joint income, which make working mothers significantly worse off and so create disincentives to female labour force participation.

It appears to be the case that the gender wage gap in terms of pre-tax and transfer wage rates has improved somewhat over the past few decades, but we would argue that inappropriate policies supported by an inadequate economic methodology have had at least to some extent a countervailing effect on net wage rates, which are after all the ultimate determinants of household wellbeing. To the extent that the position of women within the household is influenced by their outside options, which seems to be a characteristic of almost all recent work on models of the family household, it cannot have been improved by policies that weaken them.

Appendix 1

Equivalised income measures: the theory

Here we set out the theoretical derivation of the equivalisation procedure, following Deaton and Muellbauer.[23] Assume that every household $h = 1, ..., H$ possesses a utility function of the form

$$u_h = u(x^h, a^h) \tag{10}$$

where $u(.)$ is the same function for all households and has the properties of the individual utility function in the standard analysis of consumer demand, and x^h is a vector of consumption goods bought on markets at a price vector p, assumed to be the same for all households.[24] The key assumption here is that the vector of household characteristics a^h captures everything that would make two households with the same income[25] y_h (and facing the same price vector) choose different consumption vectors and, moreover, that these are observable to the analyst and so can be controlled for. We will refer to a^h as the household's *type*. These characteristics may take any form but the elements of a^h most used in practice are demographic variables such as the number of household members and their ages.

In the usual way, we can derive the household's indirect utility function $v(p, y_h, a^h) = u(\tilde{x}^h, a^h)$, where $\tilde{x}^h$ is the optimal consumption vector at p, y_h, a^h, and its expenditure function $e(p, a^h, u^h)$, giving the minimum expenditure required to achieve the utility u_h. Again the functions $v(.)$ and $e(.)$ are the same for all households. This implies that households with the same income and type have identical utilities and, holding income constant, the only thing that causes variation in demands and utilities is variation in type. This naturally suggests construction of an index number for household utility along the following lines.

We choose arbitrarily, but without loss of generality, a reference household type, denoted by a^0, for example a household with a single individual. Recall that given the usual assumptions of consumer theory we can always define a *money metric at constant prices* for utility, that is, we can take as our utility measure the minimised expenditure required by type a^h at the given price vector p to achieve a given utility level. So for any p, a^h we can write $u_h \equiv e(p, a^h, v(p, y_h, a^h))$.

We now define the equivalence scale index number as

$$\mu_h = \frac{e(p, a^h, v(p, y_h, a^h))}{e(p, a^0, v(p, y_h, a^h))} \tag{11}$$

[23] The notation is ours.

[24] Deaton and Muellbauer (1980) show that price vectors can differ across households, as long as they are fully observable by the analyst. In that case a price index number can be calculated that becomes part of the equivalisation index.

[25] Note that since the model has no saving, income and expenditure are equal. Deaton and Muellbauer show how this can be generalised.

The denominator is the amount of expenditure required by the type a^0 to achieve the utility level $v(p, y_h, a^h)$, the numerator is the amount of expenditure required by type a^h to reach the same utility level. Thus $\mu_h \gtreqless 1$ according as $e(p, a^h, v(p, y_h, a^h)) \gtreqless e(p, a^0, v(p, y_h, a^h))$, so that for example a household that requires twice the expenditure of the reference household to achieve the same given level of utility has an index value of 2.

Note that:

$$y_h \equiv e(p, a^h, v(p, y_h, a^h)) \tag{12}$$

by definition of the expenditure function, while setting $a^h = a^0$ and the required utility level at $v(p, y_h, a^h)$ gives

$$u_h = e(p, a^0, v(p, y_h, a^h)) \tag{13}$$

at the reference type. Thus, using (2) we can write utility as

$$u_h = \frac{y_h}{\mu_h} \tag{14}$$

which provides the theoretical basis for the equivalisation procedure.

Appendix 2
The Household Exchange Model

Households consist of a primary and a second earner. Both divide their time between market labour supply L, leisure l, and time a spent in production of a household good z, that they both consume.[26] They have market wage rates of w_{ih}, $i = 1, 2$, with $w_{1h} \geq w_{2h}$, and $h = 1, 2, ..., H$ denotes the household.

On the production side, the household chooses its time allocation efficiently by solving the problem[27]

$$\min_{a_{ih}, b_h} \sum_{i=1}^{2} w_{ih} a_{ih} + p_h b_h \quad \text{s.t.} \;\; z_h \geq f(k_{1h} a_{1h}, k_{2h} a_{2h}, q_h b_h) \quad h = 1, 2, ..., H \tag{15}$$

The production function $f(.)$, is identical across households, linear homogeneous in effective labour supplies and strictly quasiconcave.

By setting $z_h = 1$ the solution to this problem yields unit demand functions

$$a_{ih}^{(1)} = a(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h); \;\; b_h^{(1)} = b(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h) \tag{16}$$

and a unit cost function $c(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h)$ independent of the level of output. This defines an implicit price of the domestic good, denoted by

$$\pi_h = c(w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h) = \sum_i w_{ih} a_{ih}^{(1)} + p_h b_h^{(1)} \tag{17}$$

[26] We treat z as a private good rather than as a household publlc good. It is straightforward to extend the model to incorporate this, using Paul Samuelson's analysis of optimal public good choices, but that does not add anything of interest given the focus of this paper.

[27] Details of this solution are given in the appendix.

The Envelope Theorem gives, for $i = 1, 2,\ h = 1, ..., H$:

$$\frac{\partial \pi_h}{\partial w_{ih}} = a_{ih}^{(1)}; \frac{\partial \pi_h}{\partial p_h} = b_h^{(1)}; \frac{\partial \pi_h}{\partial k_{ih}} = -\alpha_h f_i a_{ih}^{(1)}; \frac{\partial \pi_h}{\partial q_h} = -\alpha_h f_3 b_h^{(1)} \tag{18}$$

where $\alpha_h > 0$ are Lagrange multipliers. Because of the linear homogeneity assumption we can write the input demand functions as

$$a_{ij} = a_{ih}^{(1)} z_h;\ b_h = b_h^{(1)} z_h \tag{19}$$

while $a_{ih}^{(1)} z_{jh}$ denotes the amount of time i spends in producing the amount of z consumed by individual j, $i, j = 1, 2$ and $\sum_i z_{ih} = z_h$.

The individual time constraints are:

$$a_{ih} + l_{ih} + L_{ih} = T \tag{20}$$

The constant returns to scale assumption also implies that there is a separation between production and consumption, which greatly simplifies the analysis.

Turning to the consumption side, the individual utility functions are:

$$u_{ih} = u(x_{ih}, z_{ih}, l_{ih})\quad i = 1, 2 \tag{21}$$

where x_{ih} denotes consumption of a composite market good and preferences are identical across all individuals and households, which allows interpersonal comparisons of utility.

We assume that the household is fully rational in that it values each earner's time consistently at their outside market wage. In particular it prices individual leisures at their corresponding market wage and applies the implicit price π_h as the opportunity cost of each individual's consumption of the household good. This implies that we can represent the individual choice problems in terms of their *full income budget constraints*:

$$\max_{x_{ih}, z_{ih}, L_{ih}} u(x_{ih}, z_{ih}, L_{ih}) \text{ s.t. } x_{ih} + \pi_h z_{ih} + w_{ih} l_{ih} \leq w_{ih} T \quad i = 1, 2, \quad h = 1, 2, ..., H \tag{22}$$

Equivalently, we could write the individual budget constraints in terms of income and expenditure.

$$x_{ih} + (w_{jh} h_{jh}^{(1)} + p_h b_h^{(1)}) z_{ih} \leq w_{ih}(l_{ih} + h_{ih}^{(1)} z_{jh})\quad i, j = 1, 2,\ i \neq j \tag{23}$$

We derive these expressions by substituting for $T = l_{ih} + L_{ih} + a_{ih}$ and using $\pi_h = \sum_i w_{ih} h_{ih}^{(1)} + p_h b_h^{(1)}$. The left hand sides of these equations give the expenditure on i's consumption of the market good and the cost of the inputs required for producing i's consumption of the household good, other than its own input, and the right hand sides give the sums of i's market wage earnings and the implicit payment from the partner j for the time i spends in producing j's consumption of the household good, $i, j = 1, 2,\ i \neq n$. This recognises that each individual has not one but two sources of income: as well as labour market earnings there is the implicit income from participating in producing the partner's share of the household good.

Summing these two constraints gives:

$$\sum_i x_{ih} + p_h b_h = \sum_i w_{ih} l_{ih} \tag{24}$$

which is the household's "balance of payments" constraint: the cost of its "imports" in the form of the market consumption good and market input into household production must be covered by the value of its "exports", its market labour supplies. The within-household transactions with respect to the non-traded good z_h of course cancel out in the aggregate. This emphasises the view of the household as a small economy and highlights the limitations of models that ignore its non-traded good and essentially set the value of its GDP equal to its exports.

That the second earner may have a cost of consumption of market goods greater then her market income should not be interpreted as implying that she receives a *lump sum* transfer from the primary earner. To prove the existence of a lump sum transfer in this case, it is necessary to show that her consumption is greater than the amount of her market income plus the value of her contribution to 1's consumption of the market good, net of the cost of the bought in market input required to produce her consumption of the domestic good i.e.:

$$x_{2h} > w_{2h}(l_{2h} + h_{2h}^{(1)} z_{1h}) - p_h b_h^{(1)} z_{2h} > 0 \tag{25}$$

For example, a high wage second earner who supplies little labour time to the market and provides a relatively large amount of the household good to her partner, with relatively little consumption of it herself, could have in a no-transfer equilibrium a large excess of market consumption over her own market labour income. Essentially, in this case the second earner is trading her time in producing the household good for her partner's time in earning the market good. This again demonstrates the limitations of household models that ignore household production.

Solving the individual choice problems yields individual demand functions $x_{ih}(w_{ih}\pi_h)$, $z_{ih}(w_{ih},\pi_h)$, $l_{ih}(w_{ih},\pi_h)$ with all the standard properties, as well as individual indirect utility functions $v_{ih}(w_{ih},\pi_h)$ and expenditure functions $e_{ih}(w_{ih},\pi_h,u_{ih})$. These functions and their properties form the basis for the discussion of inequality in the text of the paper. Since our focus there will be on the indirect utility function, we present here its derivatives with respect to its exogenous determinants.

Because of the two-stage form of our analysis the derivatives of $v_{ih}(w_{ih},\pi_h)$ derived from the household consumption decisions are deceptively simple:

$$\frac{\partial v_{ih}}{\partial w_{ih}} = \lambda_{ih}(T - l_{ih}); \quad \frac{\partial v_{ih}}{\partial \pi_h} = -\lambda_{ih} z_{ih} \quad i = 1, 2 \tag{26}$$

where λ_{ih} is the marginal utility of i's full income. However, this suppresses the fact that π_h is a function of all the variables in the vector $[w_{1h}, w_{2h}, p_h, k_{1h}, k_{2h}, q_h]$. Thus the full derivatives of the indirect utility function with respect to the truly exogenous determinants of utility are both more complex and more interesting:

$$\frac{\partial v_{ih}}{\partial w_{ih}} = \lambda_{ih}[(T - (l_{ih} + z_{ih}h^{(1)}_{ih})] = \lambda_{ih}[l_{ih} + z_{jh}h^{(1)}_{ih}] > 0 \quad i,j = 1,2, \; i \neq j \quad (27)$$

This derivative shows that an increase in the wage changes utility proportionally to not just market labour supply but also to the time given both to this and to the production of the household good for consumption of the other individual in the household, because of the implicit trade relationship. For example an increase in her wage makes the second earner better off even if she works a negligible amount in the market because it raises the implicit return to her household labour supply, given that the household rationally values her time at her market wage rate.

$$\frac{\partial v_{ih}}{\partial w_{jh}} = -\lambda_{ih}z_{ih}h^{(1)}_{jh} < 0 \quad i,j = 1,2, \; i \neq j \quad (28)$$

An increase in one individual's wage however has a negative effect on the utility of the other, because it raises the price of the household good. This utility effect is therefore proportional to the amount of the household good consumed by the individual whose wage has not risen, (a standard Roy's Identity effect), as well as to the time input per unit of output of the individual whose wage has risen (the household production effect).

$$\frac{\partial v_{ih}}{\partial p_h} = -\lambda_{ih}z_{ih}b^{(1)}_h < 0 \quad i = 1,2 \quad (29)$$

Of course an increase in price of the bought in input increases the implicit price of the household good and so makes both individuals worse off, to an extent dependent on their consumption of the household good (Roy's Identity again) and the amount of the good used per unit of output of the household good (household production effect).

$$\frac{\partial v_{ih}}{\partial k_{ih}} = \lambda_{ih}\alpha_h f_i z_{ih}a^{(1)}_{ih} > 0 \quad i = 1,2 \quad (30)$$

$$\frac{\partial v_{ih}}{\partial k_{jh}} = \lambda_{ih}z_{ih}\alpha_h f_j a^{(1)}_{jh} > 0 \quad i,j = 1,2, \; i \neq j \quad (31)$$

$$\frac{\partial v_{ih}}{\partial q_h} = \lambda_{ih}\alpha_h f_3 z_{ih}b^{(1)}_h > 0 \quad i = 1,2 \quad (32)$$

Increases in each of the productivity variables lowers the price of the household good and so must make each individual in the household better off. The size of the effect varies positively with the marginal product of the input concerned and the amount of it used per unit of the household good, as well as with the amount of the good the individual consumes (yet again Roy's Identity).

These derivatives are all perfectly intuitive. Their main aim is not to consider comparative statics effects on any one household in isolation, but rather to consider *what happens to the standard of living of households as we move through the joint population distribution of this vector of variables and observe how this relates to the measurement of household inequality.*

References

[1] Alesina, A., A. Ichino and L. Karabarbounis (2011). Gender based taxation and the division of family chores. *American Economic Journal: Economic Policy,* 3 (2), 1-40.

[2] Andrienko, Y., P. Apps and R. Rees (2016). Optimal taxation and top incomes. *International Tax and Public Finance* 23 (6), 981-1003.

[3] Apps, P. (1982). Institutional inequality and tax incidence. *Journal of Public Economics.* 18(2), 217-42.

[4] Apps, P. (2017). Gender equity in the tax-transfer system for fiscal sustainability. Ch 3 in M Stewart (ed) *Tax, Social Policy and Gender.* 69-98, ANU Press, Canberra.

[5] Apps, P. and R. Rees (1988). Taxation and the household. *Journal of Public Economics* 35 (3), 355–69.

[6] Apps, P. and R. Rees (2009). *Public Economics and the Household.* Cambridge: Cambridge University Press.

[7] Apps, P. and R. Rees (2018). Optimal family taxation and income inequality. *International Tax and Public Finance.* 25 (5), 1093-1128.

[8] Atkinson, A. (1970). On the measurement of inequality. *Journal of Economic Theory, 2, 244-63.*

[9] Basu, K. (2006). Gender and say: A model of household behaviour with endogenously determined balance of power. *The Economic Journal.* 116, 558-580.

[10] Bergmann, B (1971). The effects on white incomes of discrimination in employment. *Journal of Political Economy.* 79, 294-313.

[11] Bick, A., B. Brueggemann, N. Fuchs-Schuendeln and H. Paule-Paludkiewcz (2018). Long-term changes in married couples' labor supply and taxes: Evidence from the US and Europe since the 1980's. *CESifo WP* 7267.

[12] Deaton, A. and J. Muellbauer (1980). Economics of consumer behavior. Cambridge University Press, Camridge.

[13] Konrad K. and K. Lommerud (2000). The bargaining family revisited. *Canadian Journal of Economics.* 33 (2), 471-87.

[14] LaLumia, S. (2008). The effects of joint taxation of married couples on labor Supply and non-wage Income. *Journal of Public Economics.* 92 (7), 1698-1791.

[15] Lundberg, S., R. Pollak and J. Stearns (2016). Family Inequality: Diverging patterns in marriage, cohabitation and childbearing. *Journal of Economic Perspectives.* 30 (2), 79-102.

[16] Lundberg, S., R. Pollak and T. Wales (1997). Do husbands and wives pool their resources? Evidence from the UK child beneft. *Journal of Human Resourcs.* 232(3), 463-80.

[17] Mirrlees J. (1971). An exploration in the theory of optimum Income taxation, *Review of Economic Studies*, 38, 175-208.

[18] OECD (2019), Under pressure: The squeezed middle class, OECD Publishing, Paris.

[19] Sila, U. and V. Dugain (2019). Income, wealth and earnings inequality in Australia: Evidence from the Hilda survey. WP No. 1538. Economics Department, OECD, Paris.

[20] Productivity Commission (2018), Rising inequality? A stocktake of the evidence, Productivity Commission Report, Productivity Commision, Australia.

[21] Pollak, R. and T. Wales (1979). Welfare comparisons and equivalence scales. *American Economic Review*, 69, 316-21.

[22] Pollak, R. and M. Wachter (1975). The relevance of the household production function and its implications for the allocation of time. *Journal of Political Economy.* 68, 349-59.

Table 1: Primary income deciles: 2015-16

Decile	1	2	3	4	5	6	7	8	9	10	All
Prim $pa	41696	54211	63226	72729	82368	92951	105370	121597	151072	300292	108575
2nd inc $pa	17041	26951	30328	34897	37409	40891	43937	46806	49464	62717	39048
Second earner participation rates											
Non-P%	0.39	0.28	0.26	0.24	0.21	0.21	0.22	0.23	0.24	0.27	0.25
PT%	0.36	0.37	0.32	0.32	0.35	0.34	0.34	0.37	0.39	0.38	0.36
FT%	0.25	0.36	0.44	0.44	0.44	0.45	0.44	0.40	0.37	0.35	0.39

Number of records: 5481

Table 2: Household income deciles: 2015-16

Decile	1	2	3	4	5	6	7	8	9	10	All
Prim $pa	46469	60520	67845	73709	84342	93702	104052	119893	145866	289320	108575
2nd inc $pa	3459	13770	21981	29753	32693	39013	46369	53163	64407	85863	39048
Second earner participation rates											
Non-P%	0.75	0.41	0.30	0.22	0.18	0.18	0.11	0.11	0.10	0.14	0.25
PT%	0.17	0.45	0.46	0.36	0.40	0.35	0.37	0.36	0.31	0.35	0.36
FT%	0.08	0.14	0.24	0.42	0.42	0.47	0.52	0.53	0.58	0.51	0.39

Number of records: 5481

Table 3 Quintile distribution of primary income and participation rates

Quintile	1	2	3	4	5	All
Prim. income $pa	50025	74666	95880	122961	247040	118330
2nd income $pa	21641	27999	37660	40657	57105	37034
2nd earner participation rates						
Non-participants %	0.29	0.27	0.23	0.27	0.25	0.26
Part time %	0.43	0.42	0.45	0.49	0.49	0.46
Full time %	0.28	0.31	0.32	0.24	0.26	0.28

Number of records: 1055

Table 4 Quintile distribution of equivalised income and participation rates

Quintile	1	2	3	4	5	All
Prim. income $pa	56470	75722	98689	121625	239056	118330
2nd income $pa	8630	22007	30761	47071	76639	37034
Second earner participation rates						
Non-participant %	0.54	0.30	0.26	0.11	0.12	0.26
Part time %	0.34	0.48	0.46	0.51	0.47	0.46
Full time %	0.12	0.22	0.28	0.38	0.41	0.28

Number of records: 1055

Table 5 Primary and 2nd incomes by primary income

Quintile		1	2	3	4	5	All
H1	Prim. income $pa	49788	74556	95141	123334	268942	121356
	2nd income $pa	10897	9288	13746	16950	19957	14139
	H'hold income $pa	**60685**	**83844**	**108865**	**140280**	**288899**	**135495**
H2	Prim income $pa	50247	74770	96551	122582	227290	115431
	2nd income $pa	32905	45679	58585	65527	90604	59036
	H'hold income $pa	**83179**	**120456**	**155140**	**185682**	**317904**	**174467**

Table 6 Primary and 2nd incomes by equivalised income

Quintile		1	2	3	4	5	All
E1	Prim. income $pa	60601	88105	112942	138710	281570	121356
	2nd income $pa	412	9165	15377	29816	51557	14139
	H'hold income $pa	**61013**	**97270**	**128319**	**168520**	**333130**	**135495**
E2	Prim income $pa	51565	63338	84301	103712	191867	115431
	2nd income $pa	18389	34848	46292	65165	104481	59036
	H'hold income $pa	**69954**	**98187**	**130593**	**168877**	**296348**	**174467**

Table 7 Distribution of H1 and H2 households by equivalised income

Quintile	1	2	3	4	5	All
H1 quintile split %	76	57	46	36	28	50
H2 quintile split %	24	43	54	64	72	50

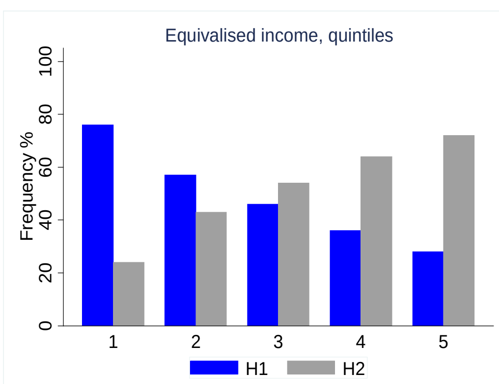

Figure 1 Quintile distribution of H1 and H2 households

Table 8 Income tax payments by primary income

Quintile	1	2	3	4	5	All
H1						
Prim. income tax $pa	7487	15595	22780	32767	93866	34108
2nd income tax $pa	706	384	999	2076	2580	1347
Total income tax $pa	**8192**	**15979**	**23779**	**34843**	**96446**	**35455**
H2						
Prim. income tax $pa	7526	15601	22873	32421	75507	31229
2nd income tax $pa	3100	6570	10195	12905	22270	11130
Total income tax $pa	**10626**	**22172**	**33068**	**45326**	**97777**	**42359**

Table 9 Income tax payments by equivalised income

Quintile	1	2	3	4	5	All
E1						
Prim. income tax $pa	11087	20400	29166	38127	99932	34108
2nd income tax $pa	0	333	1158	3411	10388	1347
Total income tax $pa	**11087**	**20733**	**30324**	**41538**	**110329**	**35455**
E2						
Prim. income tax $pa	8054	11795	18852	25240	61047	31229
2nd income tax $pa	848	3109	6409	12181	26662	11130
Total income tax $pa	**8902**	**14904**	**25260**	**37421**	**87709**	**42359**

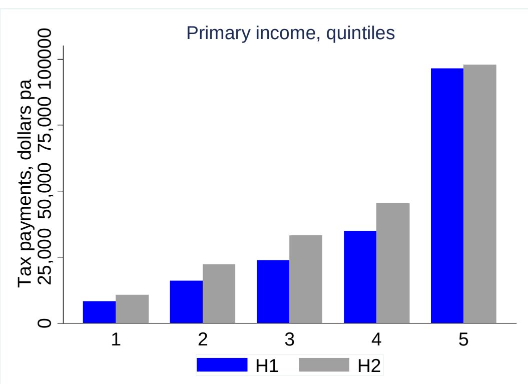

Figure 2 H1&H2 income tax payments

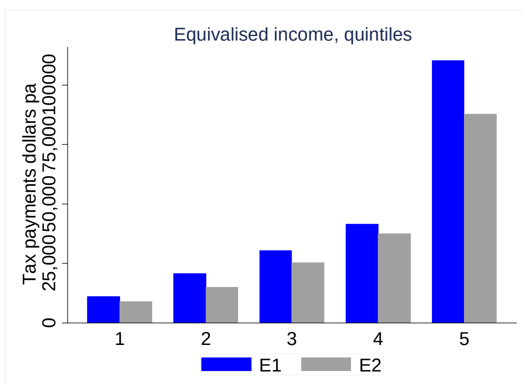

Figure 3 E1&E2 income tax payments

Table 10 Income tax – FTB-A payments: H1 and H2 by primary income

Quintile	1	2	3	4	5	All
H1						
FTB-A $pa	-8577	-5056	-2001	-43	0	-4672
Total Tax $pa	**-385**	**10923**	**21778**	**34800**	**96446**	**30783**
H2						
FTB-A $pa	-5406	-888	-99	0	0	-1405
Total Tax $pa	**5220**	**21284**	**32969**	**45326**	**97777**	**40954**

Table 11 Income tax – FTB-A payments, $pa: E1 and E2 by equivalised income

Quintile	1	2	3	4	5	All
E1						
FTB-A $pa	-8428	-3027	-46	0	0	-2300
Total Tax $pa	**2659**	**17706**	**30278**	**41538**	**110329**	**33155**
E2						
FTB-A $pa	-7087	-3167	-67	30	0	-2070
Total Tax $pa	**1825**	**11737**	**25193**	**37391**	**87709**	**40289**

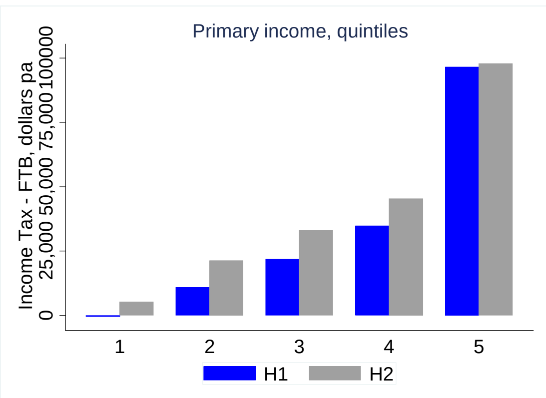

Figure 4 H1&H2 Income tax – FTB-A

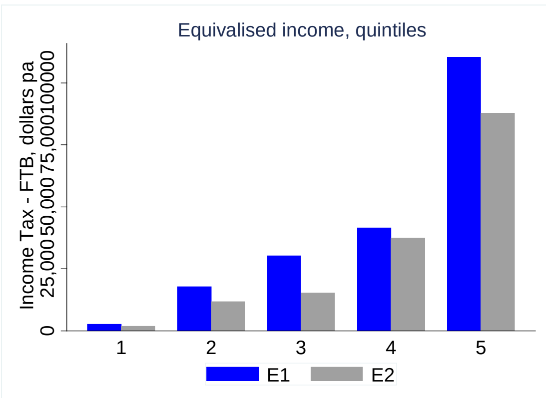

Figure 5 E1&E2 Income tax – FTB-A