



Australian Government

Department of Family and Community Services



HILDA PROJECT DISCUSSION PAPER SERIES NO. 2/03, OCTOBER 2003

How Best To Impute Taxes And Measure Public Transfers?

Bruce Headey

Acknowledgements

Many thanks to John Freebairn of the Department of Economics at the University of Melbourne, and to Guyonne Kalb, Rosanna Scutella, Nicole Watson and Mark Wooden of the Melbourne Institute for discussions and advice on these issues. The usual disclaimer applies.

**The HILDA Project was initiated, and is funded, by the
Australian Government Department of Family and Community Services**



Introduction

The Household, Income and Labour Dynamics in Australia (HILDA) survey began in 2001, and wave 3 is currently in the field. In the HILDA Survey, as in other surveys that measure income, respondents are asked detailed questions about their income from market sources, from private transfers and also from public transfers, but they are not asked about the taxes they pay. Such questions would be futile because most people would be unable to report accurately. So taxes have to be imputed in order to estimate disposable incomes.

The main aim of this paper is to make recommendations to improve our tax imputations in HILDA. Subsequently, and especially given the ABS's difficulties in measuring public transfer income in its own surveys (see Siminski, Saunders and Bradbury 2003), it was suggested that it might also be helpful to make some recommendations about improving HILDA's measures of transfer income.

The current state of play with *tax imputation* in HILDA is that the program used for the wave 1 data release implemented just two aspects of the Tax Code: the four marginal rates of Commonwealth income tax and the Medicare levy rules (which have different taxable income bands from the income tax code).

The current state of play with *public transfer income* – pensions, benefits and allowances – is discussed in more detail below, but basically we appear to be in fairly good shape and just need to make a few amendments to our questionnaire.

How to Improve – Main Recommendations

In regard to both taxes and public transfer income we need a gold standard of data against which we can judge how good our measures and calculations are, and then keep improving them.

- For taxes see the annual *Taxation Statistics* volume issued by the Australian Taxation Office (ATO), and especially the CD-rom of supplementary data. A practical problem is that these publications come out over two years after the end of the financial year in question. For example, the issue for the 2000-2001 financial year is due out in September 2003.
- For transfers, see occasional papers on *Income Support and Related Statistics*, issued by Department of Family and Community Services (FaCS).

These volumes give aggregate national statistics (with some additional State and demographic breakdowns relating to sex, age and marital status) on the total taxes the ATO says it collected each year, and the total public transfers FaCS says it paid out.

So we can use these official Government sources as a basis of comparison to check both the total national tax bite implied by the HILDA data, and the national transfer pay-out which HILDA predicts. Making these checks, it transpires that the simple method of tax imputation described above (Commonwealth income tax bands plus Medicare levy) has the effect of over-estimating taxes by a large margin. It follows that we are substantially under-estimating the disposable incomes of Australian individuals and families; not errors we should cheerfully accept. We are also somewhat under-estimating public transfers; but apparently no worse than ABS surveys (see Siminski et al. 2003).

Our tax estimation problems appear to be mainly due to three compounding errors.

- i More or less equating taxable income with gross income (gross income comprises all sources of income, including private and public transfers). In fact, ATO reports that taxable income averages about 95% of gross income, because tax-payers can claim deductions which make up this difference. The largest deduction is for work-related expenses, but many public transfers are also deductible.
- ii Not allowing enough for rebates/offsets (about 3% of gross income on average). The largest rebate is for dividend imputation.
- iii The current method of imputing taxes greatly over-taxes retired people, who rarely have an average tax rate over 15% (ATO, 1999-2000).

Specific Recommendations About Tax Imputation

There are several sensible approaches to improving our tax imputations. These approaches are not so much alternatives as modes that could be combined. The approaches are first outlined and then a combination is recommended.

- A basic approach would be to rely very heavily on the data given in *Taxation Statistics*. This would ensure that our figures matched the national aggregates; a big plus. We could simply attribute to each HILDA respondent a net tax payment equal to the average paid out by a person of his/her sex, age, marital status and wage/salary income. Note that some additional information is also given in *Taxation Statistics* about taxation of pension income and business income.

The main defect of this approach is that much interpersonal variance in taxes would be lost. However, note that the published data (including the supplementary CD-rom) give information about the average tax rates of people on a very wide range of taxable incomes, and also give sex, age and marital status breakdowns. For example, data are given for the ten richest and ten poorest postcodes in each State and Territory (80 postcodes in all). Combining this information and treating it as providing a picture for the whole of Australia gives a good picture of the mean tax rates of taxpayers at all levels of income.

- A second approach is to write code for HILDA's tax imputation that directly follows the actual tax return form used by individual taxpayers for the financial year in question (or, much the same, the summary version – ready reckoner version – of the form given towards the end of the Tax Pack each year). Looking at the tax form, which is really quite short, it seems that it would not be difficult to write a program to follow a large part of it, including a few of the deductions and many offsets (rebates).

In writing the code we would begin with the marginal rates and the Medicare Levy rates and then add code for deductions and rebates, where we had the necessary information from HILDA respondents.

The strength of this approach might appear to be that we would take account of the particular circumstances of individuals and not just apply ATO averages. However, this advantage is far from certain. *Taxation Statistics* (1999-2000) comments that a major factor affecting the amount taxpayers

claim in deductions and tax offsets is whether or not they employ an accountant, and we do not have this information from our respondents. Other defects of this approach are that (1) it involves a lot more staff time than the first approach and (2) our predicted national tax bite numbers would not be guaranteed to match the ATO's. Indeed, they would presumably, at first, miss by a fair margin. My guess is that they would miss on the high side, since we would assume that most people pay the taxes they appear to owe. In reality many people, with the assistance of upstanding accountants, pay less.

- A third approach would simply be to borrow someone else's imputations; the Melbourne Institute Tax and Transfer Simulator (MITTS) and the Reserve Bank currently do the job and both have expressed a willingness to confer and offer assistance.

Combining the Approaches: Some Detail

Step 1: We need to obtain a second measure of gross income that is in line with the ATO's measure. Our current measure of gross income is fine – it sums income from all sources – but it is not the same as the ATO's measure. The ATO does not include child support payments (assumed to be part of the payer's income for tax purposes, and hence not the payee's income) and does not include foreign pensions.

Step 2: Derive taxable income from gross income, using both information from HILDA respondents and, if this does not give big enough deductions (which it almost certainly won't), then add extra amounts, so that total deductions equal the amounts given for different income groups in *Taxation Statistics*. The standard tax form for the year in question should be used to guide us through listing and programming deductions for each person. The main deductions include work expenses and some public transfers.

Step 3: Apply the Commonwealth income tax (i.e., the four marginal rates).

Step 4: Apply the Medicare Levy marginal rates, which are different from the income tax rates.

Step 5: Make adjustments for the lower taxation of business income relative to labour income. Information is available in *Taxation Statistics*.

Step 6: Special treatment is required for retired people. Most retired people pay little tax and so need to be treated separately in our imputations. Information about the special treatment of pension income is given in *Taxation Statistics*. Among other things, we need to take account of the rules governing tax-free earnings and the very low income taxation and zero capital gains taxation of money in allocated pension funds.

Step 7: Most people who receive income support payments – pensions, benefits and allowances – pay no tax on incomes up to \$20,000. This needs detailed inquiries and programming, but clearly income support recipients in HILDA need to be programmed so that they are taxed at low rates.

Step 8: As far as possible, work out the offsets (rebates) which HILDA respondents could claim. The main offsets include dividend imputation (about which we have no information in HILDA) and low income offset. There is also an offset for living in the tropics. If we under-estimate offsets (almost certain in view of our lacuna regarding dividend imputation), then we should add in extra amounts to bring respondents up to

the figures given in *Taxation Statistics* for people of their sex, age, marital status and income band.

Step 9: Check to see if the total number of taxpayers (i.e., those paying more than zero) and the total national tax take implied by HILDA matches the results given in *Taxation Statistics*. If not, find reasons for the discrepancies.

Step 10: Having obtained estimates of each person's tax payment, add a variable 'imputed tax paid' to the HILDA file. Also, calculate a household level tax variable and add to file.

Results – Comparing Our Imputations with ATO Data

The latest ATO data are for the financial year 1999-2000, whereas our HILDA first wave data, analysed here, are for 2000-01. So one would expect estimates in current dollars to be about 6 per cent higher for the later year.

The estimates for HILDA given here come directly from following the steps listed above. No adjustments were made for discrepancies between our estimates and ATO data, because no substantial discrepancies were found.

In 1999-2000 the ATO reports that 8.5 million people paid some taxes – income tax and/or Medicare Levy. Our estimate for FY2000-01 is 8.75 million. The ATO estimate of average tax paid was \$9100, giving a total national tax take of \$76.7 billion. For 2000-01 we estimate average tax paid of \$9275 for a national tax take of \$81.1 billion.

I recommend adding the following new variables to the file: individual income tax, individual Medicare Levy, individual total tax (sum of first two), household income tax, household Medicare Levy and household total tax.

Recommendations relating to benefits

Now some notes relating to public transfers. Three public transfers are omitted from HILDA's current measure of total benefits. These are family tax benefit, child care benefit and rent assistance. We now have code from FaCS to impute the first two (see above). There is nothing to be done about rent assistance in the first three waves, because no question was asked. However, some respondents will have included it under 'other payments' in the list of questions on pensions and benefits. Foreign pensioner income has been included in our current measure of total benefits. It should probably be ejected on the grounds that users of HILDA will tend to see the measure as relating to Australian Government payments only; for example in doing research on the redistributive impact of public transfers.

So here are steps we need to take to impute benefits not at present recorded in the HILDA file:

Step 1: Impute Family Tax Benefits (Parts A and B). FACS has supplied code to do this task.

Step 2: Impute Child Care benefit (CCB) for those respondents who state that they received it. FaCS has also supplied code for this.

Step 3: Add the new benefit variables to the HILDA file: FTA-A, FTB-B, and child care benefit (CCB). These are household level income variables. We may want to add them to the current measure of benefits (_bnf), which at present is only given as an individual level variable, but could also be calculated for households.

Results – Comparing Our Imputations with FaCS Data

FaCS's latest data are for June 1999. The HILDA Survey estimates given here are for FY 2000-01, which is approximately two years later. So one might expect current dollar amounts for the later time point to be about 12 per cent higher.

As with taxes, the HILDA estimates given here follow directly from the listed steps and involve no adjustments to bring them into line with FaCS's national aggregate data.

For June 1999 FaCS reports that 43.3% of the population received some benefits. Our HILDA figure for FY 2000-01 is 42.8%. FaCS reports a national total pay-out in benefits for 1999 of \$44.57 billion. The HILDA estimate for 2000-01 is \$49.27 billion. As a further check on our FTB-A and FTB-B imputations we ran about 50 of our respondents claims through the FaCS ready-reckoners supplied to their clients. The estimates given by the ready-reckoner and by the HILDA imputations were in all cases within 5% of each other.

I recommend adding the following benefit variables to HILDA files: FTB-A, FTB-B and CCB. These are household (strictly income unit) variables and cannot really be treated as individual level variables. A variable called 'total household benefits' should also be added. This would be the list of benefits summed in our current _bnf variable, aggregated to the household level, and added to FTA-A, FTB-B and CCB.

Recommendation regarding disposable income

New corrected measures of disposable income become available at both the individual and household levels. These measures involve inclusion of the new imputed measures of taxes and benefits. In order to provide an individual level measure it would be necessary to divide FTB and CCB by the number of persons in the household and attribute an equal share to each person.

Conclusion

It is not difficult to make substantial improvements on the tax imputations and measures of disposable income provided in the Wave 1 release of HILDA. We should regard the tax imputation task as one requiring continuous improvements over future years. The 'gold standard' data from the ATO and FaCS make it possible to keep improving our estimates.

References

ATO (2002), *Taxation Statistics 1999-2000*, Canberra, ATO.

ATO (forthcoming), *Taxation Statistics 2000-2001*, Canberra, ATO.

Department of Family and Community Services (2001), *Income Support and Related Statistics: A Ten Year Compendium*, Occasional Paper no. 1, Canberra, Department of Family and Community Services.

Siminiski, P., Saunders, P. and Bradbury, B. (2003), 'Reviewing the Intertemporal Consistency of ABS Household Income Data Through Comparisons with External Aggregates', *The Australian Economic Review* 36, September, 333-349.