The Economic and Social Review, Vol. 48, No. 2, Summer, 2017, pp. 171-205

Integrated Modelling of the Impact of Direct and Indirect Taxes Using Complementary Datasets

Michael Savage
Economic and Social Research Institute, Dublin

Abstract: Comprehensive modelling of the impact of taxes and tax policy options requires data on the impact at micro-level of both direct and indirect taxes. With some exceptions, most national expenditure surveys are not suitable for use in detailed modelling of the direct tax and welfare system. This makes approaches that impute expenditure data into detailed income surveys of considerable interest. Using a commonly applied method of imputing expenditure data into an income survey, this paper examines the degree of sensitivity in the distribution of household expenditure, and following that the distributional impact of indirect taxes, as the analysis moves from being based on actual expenditure data as recorded in the Irish Household Budget Survey (HBS), to being based on expenditure data imputed to the Irish Survey on Income and Living Conditions (SILC). Most sensitivity is found in the estimation and imputation of expenditure data for the bottom income decile. Nonetheless, the imputation process is shown to produce highly comparable distributional results when compared with actual expenditure data. An application of an indirect tax microsimulation model integrated with SWITCH, the ESRI tax-benefit microsimulation model, examines the distributional effects of simultaneous reforms of the direct tax, indirect tax, and social welfare systems, illustrating the analytical capability of the approach tested here.

I INTRODUCTION

Comprehensive modelling of the impact of taxes and of tax policy options requires data on the impact at micro-level of both direct and indirect taxes. To be able to accurately simulate the distributional effects of the direct and indirect tax systems, as well as the social welfare system, detailed income information is needed for each individual in the household, as well as household expenditure

Acknowledgements: I am grateful to two anonymous referees and to the editor for helpful comments. I thank participants at the ESRI Seminar Series, as well as David Madden, Tim Callan, Emer Smyth, Frances Ruane and Alan Barrett for comments. I thank the CSO and ISSDA for providing access to the SILC data and HBS data respectively for Ireland.

information. From a research point of view, a single dataset combining detailed data on incomes, expenditures, labour market participation, and a range of socio-demographic variables would be ideal. There are, however, limits on the amount of data that can be gathered by any one survey. Recording expenditure data imposes a heavy burden on respondents: they are often required to keep expenditure diaries covering all purchases over a reasonably long period of time. This limits the extent to which other data of interest can be gathered. Consequently, as noted in Sutherland *et al*. (2002), most countries do not have a single source of micro-data including high-quality disaggregated information on both incomes and expenditures.[1]

In the absence of a single source of detailed micro-level income and expenditure data, a number of alternative methods for analysing the distributional effects of tax and benefit systems have been devised. Traditionally, distributional analyses focus on either the direct tax and benefit systems (see Avram *et al*.(2014), Callan *et al*. (2010) for example), or the indirect tax system (see Collins (2014), Leahy *et al*. (2011) for example) in isolation, using income surveys or expenditure surveys alone. Distributional results from such analyses based on separate income and expenditure surveys can be combined to provide a more complete picture of the distributional impact of tax and benefit systems (see Callan *et al*. (2009) for example). However, the approach of combining results from separate analyses is limited in the variety of scenarios that can be modelled (limitations of this approach are discussed in Section II).

Methods based on imputing expenditure information into a detailed income survey, so that the indirect tax system can be modelled simultaneously with the direct tax and welfare system, are therefore of considerable interest for policy analysis, and have been applied in a number of countries in recent years.[2] Commonly, a two-stage imputation procedure is used to impute expenditure data to an income survey. In the first stage, expenditure on a range of goods and services is estimated in the expenditure survey based on a set of characteristics observable in both the expenditure survey and the income survey. Second, the coefficients from the estimation stage are used to impute expenditure into the income survey. The result of the imputation procedure is a single dataset with detailed income information, along with imputed expenditure information. Applying microsimula-

[1] The UK Living Costs and Food Survey (LCF) is an exception. The LCF is somewhat unusual in that it contains household expenditure information, as well as detailed income and household composition information which can form the basis of a tax-benefit microsimulation model. TAXBEN, the tax and benefit microsimulation model used in the Institute for Fiscal Studies (IFS), is based on the LCF. TAXBEN can therefore model the distributional impacts of direct taxes and social welfare, as well as indirect taxes based on actual expenditure data.

[2] The imputation approach has been used widely in other fields of economics – Arellano and Meghir (1992) for example, combining the UK Labour Force Survey with the UK Family Expenditure Survey to estimate female labour supply and on-the-job search.

tion techniques to this expanded dataset then allows for simultaneous analysis of the direct tax and welfare system along with the indirect tax system. A number of authors, including Decoster *et al*. (2010), O'Donoghue *et al*. (2004) and Pestel and Sommer (2016), have used this approach to measure the distributional effects of taxes and welfare payments across a range of European countries.

Building on the existing expenditure imputation analyses, in this paper the method of imputing expenditure information into a detailed income survey is examined in detail, with Irish data. Ultimately, the question we wish to address is whether we can combine data from one survey with another via a predictive approach on common variables, in order to broaden our understanding of the distributional effects of the direct tax, indirect tax and welfare systems in Ireland. To examine the suitability of this approach, the first step is to assess the degree of sensitivity in the distribution of household expenditure, and following that the distributional impact of indirect taxes, as the analysis moves from being based on actual expenditure data as recorded in the Irish Household Budget Survey (HBS), to being based on expenditure data imputed to the Irish Survey on Income and Living Conditions (SILC). To examine the sensitivity of results as we move between data sources, the distributional effects of indirect taxes is simulated using actual expenditure data, as recorded in the nationally representative HBS, and used as a benchmark. Second, the distributional analysis is replicated on expenditure data that have been estimated based on a set of observable characteristics in the expenditure survey. These estimated expenditure data are the result of the first stage of the imputation process described above. Finally, having imputed expenditure data from the expenditure survey into the income survey based on a set of common variables, the imputed expenditure data are used as the basis for the distributional analysis. The degree of sensitivity in the distributional results is identified by comparing the results based on the actual expenditure data, the estimated expenditure data and the imputed expenditure data. Using actual, estimated and imputed expenditure data to analyse the sensitivity of results to the imputation process in this way adds a level of detail to this approach that has not been applied before in the literature.

With evidence that the imputation approach provides a promising basis on which to base a combined analysis of the direct tax, indirect tax, and welfare systems, the second contribution of the paper is to illustrate the increased analytical capability that this approach provides. In Section V of this paper, the indirect tax microsimulation techniques developed in Section IV are integrated into SWITCH, the ESRI tax-benefit microsimulation model, and applied to two hypothetical reforms of the the tax and welfare system in Ireland. While these illustrative examples are neither actual policy evaluations nor policy proposals, they serve to highlight the expanded modelling capacity that the analysis here develops. The analysis therefore serves as an updated picture of the distributional effects of the indirect tax system in Ireland, as well as a base for future microsimulation analysis of simultaneous direct and indirect tax, and welfare reform.

While the results of the imputation are encouraging from a policy-analysis point of view, the limitations of the imputation approach must be recognised. Sutherland *et al*. (2002) discussed the imperfections of the imputation procedure and noted that it is particularly difficult to predict expenditure patterns that capture micro-level diversity of expenditure on a range of highly disaggregated goods and services, as exhibited in the actual data. They highlighted that imputed expenditure variables (necessarily aggregated into broader representative expenditure categories) are only ever an adequate second-best substitute for actual data. The results of this paper should be interpreted with this in mind.

II METHODS FOR ANALYSING THE DISTRIBUTIONAL IMPACT OF THE TAX AND WELFARE SYSTEM

Household expenditure data can be used to directly model the impact of changes in indirect taxation. Collins (2014), Leahy *et al*. (2011) and Barrett and Wall (2006) are recent examples of applications of this approach in Ireland. Analyses based on such surveys are commonly limited to examining the indirect tax system, and ignoring reforms to direct taxes and welfare payments. The main limitation of these surveys is the unit of observation in the data. In most cases, the unit of observation is the household, with some information on individuals within that household. For tax and welfare purposes, however, there is an intermediate unit – often called the family unit or tax unit – which is very relevant for policy purposes. Most tax and welfare policies do not operate at the level of the household, though household income and household welfare are of key concern to policy. Instead, tax and welfare policies tend to operate at either individual level (for example, contributions to social insurance) or at a family unit level. Detailed information on family and household composition is needed to ensure that it is possible to group individuals into tax units, usually defined as a married couple or single person, together with all children aged under 15, and children aged less than 18 who are in full-time education (Keane *et al*., 2014). Household expenditure surveys, by their nature, contain information at household level, and therefore omit the necessary detail to group people into family or tax-units, as well as some tax-relevant information on individuals within a household. It is therefore usually not possible to simulate direct tax and welfare reforms using household expenditure surveys.

One method that can be applied to combine direct tax and welfare analysis based on income surveys with indirect tax analysis based on expenditure surveys is to use what can be called a "results-matching" approach. Detailed in Figure 1a, the "results-matching" approach combines the results at an aggregated level (income deciles or family types, for example) from separate analyses based on different datasets. Callan *et al*. (2009) examined the distributional effects of the imposition of a carbon tax, based on the HBS, compensated by an increase in welfare payments and tax credits, modelled with SILC data, using such an approach. By comparing the change in income in each income decile from the separate analyses, they showed that the regressive impact of a carbon tax could be fully offset by increases in welfare payments and tax credits. Similarly, Callan *et al*. (2011) included the impact of indirect tax changes in an analysis of the distributional impact of austerity measures in six European countries during the Great Recession by matching results from separate analyses, again at income decile level. They showed that in all countries where VAT was increased, it had a regressive impact on the distributional effect of austerity measures.

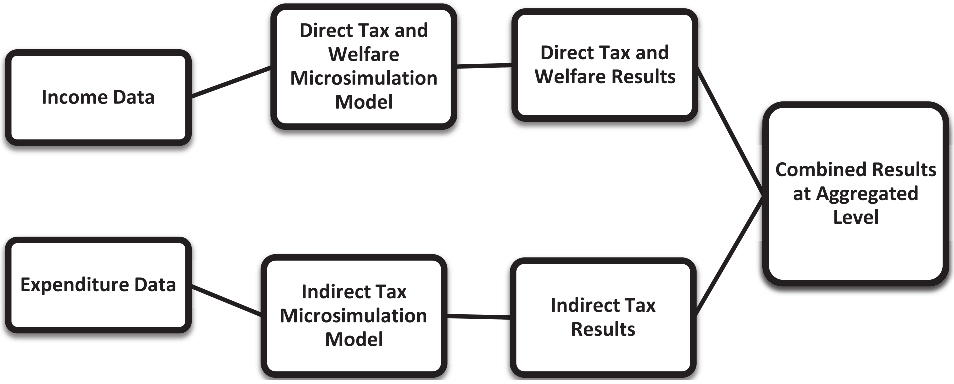
The “results-matching” approach, however, is limited in the scenarios that can be modelled. The primary limitation of the approach is subtle, but is vital from a policy-analysis point of view. One key advantage of tax-benefit microsimulation is the ability to model alternative (counterfactual) scenarios. These counterfactual scenarios can then be used not only as “reform” scenarios in the microsimulation analysis, but also as the “baseline” scenario from which policy reforms can be analysed. For example, suppose the latest available detailed income and expenditure data are from 2010, but we wish to examine the distributional impact of a reform to the 2015 tax system. Microsimulation techniques can be used to simulate the 2015 tax and welfare policy rules, provided the relevant tax-units can be identified in the data. Due to the limitations of expenditure surveys discussed above, counterfactual direct tax and welfare policies (such as the 2015 policy) cannot be simulated on the income information in the expenditure survey. Distributional analyses using the “results-matching” approach, therefore, can only be based on reform of the tax and benefit system in the year of data collection, as the income information in the expenditure survey is representative only of the direct tax and welfare system in that year.

The imputation of expenditure data to an income survey has therefore become a commonly used approach to measure the distributive impact of indirect taxes. Figure 1b illustrates the how the expenditure imputation approach differs from the “results-matching” approach. While Figure 1a shows that “results-matching” is based on two separate analyses, with the imputed expenditure approach the full microsimulation analysis is done on a common set of households with full detailed income information and an estimate of household expenditure. This allows for analysis of reform to tax and benefit systems beyond the year of data collection.

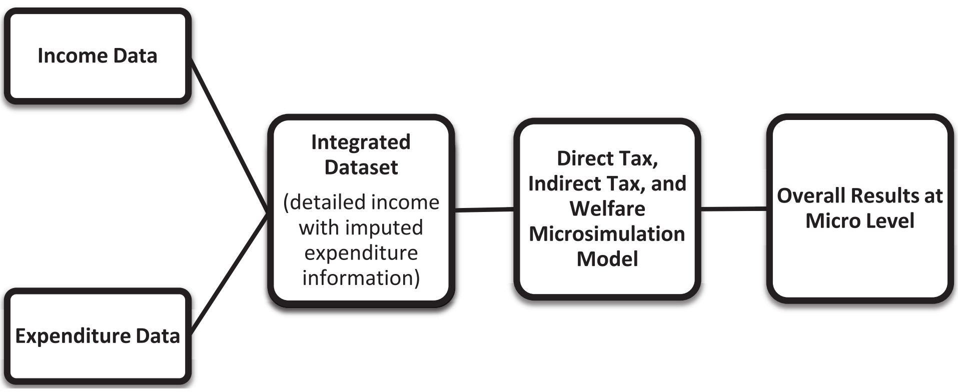
A number of methodologies have been developed to combine expenditure information with income surveys. Pestel and Sommer (2016) categorise the methods into implicit and explicit techniques. Implicit methods, they state, seek to infer expenditures based on observations that are as similar as possible to the target

Figure 1: Combined Analysis of Direct Tax, Indirect Tax and Welfare Systems

(a) "Results-Matching" Method



(b) Expenditure Imputation Method



observation. The main advantage of using implicit methods is that theoretical assumptions regarding the relationship between common variables and expenditure are not required. Explicit methods, however, usually base expenditure imputation on some form of Engel Curve relationship, where a functional form of expenditure is assumed.[3]

[3] In the absence of expenditure data, Browing and Leth-Petersen (2003) derived a measure of total expenditure at the household level from administrative micro-data on income and wealth. They validated their derived expenditure data by matching their administrative data to survey expenditure data and found encouraging results.

Decoster *et al*. (2007) carried out a detailed comparison of four methods to impute expenditure into an income survey. The four methods were parametric and kernel estimation of Engel curves, Mahalanobis distance matching and grade correspondence. The results suggested that parametric estimation provided the most accurate results in terms of the imputed budget shares for a range of goods in the income survey compared to the observed values in the expenditure survey. At least partly based on these results, parametric estimation of expenditure has become the most common method of imputing expenditure to an income survey for indirect tax reform analysis. To investigate the effect of increases in the Value Added Tax in Germany for example, Pestel and Sommer imputed expenditure from the German Sample Survey of Income and Expenditure (EVS) to the German Socio-Economic Panel Study (SOEP) which does not report expenditures. Based on the recommendations of Decoster *et al*. (2009), Pestel and Sommer used parametric estimation of Engel Curves for the imputation of expenditure data. Using similar methods, Decoster *et al*. (2010), Figari and Paulus (2012), and O'Donoghue *et al*. (2004) separately examined the impact of indirect tax changes for range of European countries. Common to each study is the use of a tax-benefit microsimulation model based on an income survey which lacks expenditure data. In each case, parametric estimation of expenditure based on a range of common variables between the income and expenditure surveys was used.

While in each of these analyses the comparison of actual and imputed expenditure was encouraging, the limitations of the imputation approach must be recognised. Sutherland *et al*. (2002) discussed the imperfections of the imputation procedure. They highlighted that imputed expenditure variables are only ever an adequate second-best substitute for actual data, and recommended that the imputed variables are used at a sufficient level of aggregation to mask differences that are not controlled by the imputation procedure. The first approach based on the expenditure survey alone is of interest in itself, they argued, and also provides a benchmark against which to assess approaches based on prediction of expenditure. Sutherland *et al*. conducted an evaluation of the imputation procedure by comparing results of indirect tax simulations using imputed and actual expenditures. They concluded that they could only be confident in the imputations when the dimensions that are important to the end analysis have been controlled for in the imputation procedure. These limitations of the expenditure imputation approach must be considered when interpreting the results of this paper.

In presenting results, in common with many other distributional analyses based on imputed expenditure data, we group households into income deciles. This step is motivated by a number of factors. First, as argued by Sutherland *et al*., we do not expect the imputation procedure to provide perfect estimates of expenditure on a household-by-household basis. Rather, the imputation procedure is used to provide results which are as accurate as possible at the level of tabulations sought: in this case, income deciles. In addition, Decoster *et al*. (2007) address this issue when discussing the level of aggregation at which to present results in their analysis. A problem they identified was the effect observed for most countries that the tails of the disposable income distributions tend to differ in income and expenditure surveys. However, since the number of observations in the "diverging tails" is often observed to be very small,[4] they argued that it is more likely that these values can be considered as outliers rather than as systematic measure differences. Therefore grouping by decile minimises the risk that these outliers would have a significant effect on the outcome. In later sections, we group households by categories other than income deciles to examine if other sources of heterogeneity exist in the results that are not visible by income decile.

The analysis in this paper is intended as a complement to the existing literature on the subject discussed in this section. O'Donoghue *et al*. (2004) included Ireland as one of 12 countries for which they imputed expenditure into an income survey for use in a tax-benefit microsimulation model. This paper builds on O'Donoghue *et al*.'s work in a number of respects. First, this paper is the first to compare in detail how the distribution of expenditure and indirect taxes change as we move from using actual expenditure data to imputed expenditure data in Ireland. Second, the treatment of observed zero-expenditures differs substantially between the two analyses. O'Donoghue *et al*. grouped expenditure items to minimise the occurrence of zero-expenditures, whereas this paper explicitly models the zero-expenditure decision. As the results of the analysis show, explicit treatment of zero-expenditures results in an imputation of expenditure that more closely resembles the observed expenditure data. Third, the analysis here is designed to supplement SWITCH, the ESRI tax-benefit model (see Section V). This model is widely used both in Irish policy debate (see Callan *et al*. (2013), Savage *et al*. (forthcoming), for example) as well as by Irish policymakers (see Department of Social Protection (2013), O'Connor *et al*. (2016), Tax Strategy Group (2015) for example). Finally, the use of more recent data in this paper (O'Donoghue's expenditure data are 15 years older than the data used here) provides a more recent picture of the distribution of expenditure in Ireland.

III DATA

The expenditure survey used in this analysis is the 2009/2010 wave of the Irish Household Budget Survey (HBS).[5] The main purpose of the HBS is to gather detailed information that is used in the construction of the Consumer Price Index.

[4] Less than 1 per thousand in the Belgian Budget Surveys analysed in Decoster *et al*. (2007).
[5] Survey conducted by the Central Statistics Office (CSO) of Ireland.

In a nationally representative sample of 5,891 households,[6] responding households are asked to maintain a detailed diary of household expenditure over a two-week period. Some expenditure items, such as durable expenditure, are collected over a longer timeframe. The survey also contains demographic and socio-economic information of the households and household members, such as age, location and labour force status. Information on direct, gross and disposable income is also contained in the dataset.

The indirect taxes included in the analysis are VAT, excise duties and the carbon tax. To estimate the VAT payments made by each household using the actual expenditure data, we assign the appropriate VAT rate to each of the expenditure items in the HBS data. Using these VAT rates, the VAT contribution of each household can be measured. Given that VAT is applied as a percentage of producer price, the VAT element of expenditure can be calculated as:

$$VAT_i^j = x_i^j * \frac{t^j}{1+t^j} \tag{1}$$

where x_i^j is household i's expenditure on good j, and t^j is the VAT rate on good j. The distributional impact of VAT reforms can then be simulated by changing t^j. In later sections when examining the distributional impact of VAT reforms, we assume demand does not respond to relative price changes and that the full indirect tax is passed on to the consumer, as is common in the literature.[7]

To estimate excise and carbon tax payments, an assumption is required on the price of the relevant goods.[8] SEAI (2013) report average prices for the goods on which carbon taxes apply,[9] while the Revenue Commissioners (2012) report national average retail prices for goods on which excise duties apply.[10] Using these price estimates, we can estimate excise payments as

$$EX_i^j = \frac{x_i^j}{\tilde{p}^j} * ExciseDuty^j \tag{2}$$

[6] In the main analysis of the paper, we keep the entire sample of households. Sensitivity analysis was carried out dropping certain proportions of the sample. The main conclusions of the paper did not change.
[7] See Pestel and Sommer (2016), and Picos-Sanchez (2011) for examples of this approach, and Banks *et al*. (1996) for a study on the bias caused by using these first order approximations to measure welfare effects. The assumption of zero demand responses is discussed further in Section V.
[8] Excises and carbon taxes are charged per quantity. Only total expenditure is observed in the HBS data. Therefore, a representative price is required to estimate quantities purchased, and so excises paid.
[9] Phase 1 of the carbon tax, in respect of petrol and auto diesel, was implemented with effect from 10 December 2009. The carbon tax on kerosene, marked gas oil, fuel oil, LPG (Other), auto LPG and natural gas commenced in May 2010. In 2013, it was extended to include solid fuels.
[10] Various forms of alcohol, tobacco and energy products.

where $\tilde{p}^j$ is the assumed price of good j, and $Excise\ Duty^j$ is the excise duty applied per unit of good j. Carbon tax payments can be estimated in a similar manner. With the estimates of VAT, excise duties and carbon taxes paid by each household, we can use the income data contained in the HBS to assess the distributional impact of the indirect tax system, as well as the effects of reforms to that system. This "actual data" approach is used as the benchmark measure with which we compare our estimated and imputed expenditure data.

To estimate household expenditure on different goods and services, expenditure items in the HBS are aggregated into a smaller number of representative goods. Hicks' Composite Commodity Theorem suggests that commodities among which the relative prices are constant can, in a natural way, be treated as a single commodity (Bradford, 1974). Similarly, the expenditure items we group together should be as homogenous as possible. A trade-off exists between the desire to estimate expenditure for as many different goods and services as possible, with the need to be able to estimate the parameters of the model with precision. A review of the previous literature (Decoster *et al*. (2007), Pestel and Sommer (2016), for example) suggests aggregating expenditure items into ten to 18 representative goods. Here we group the expenditure items in the HBS into 15 aggregated commodities.[11]

The income survey used in the paper is the 2010 wave of the Survey of Income and Living Conditions (SILC). SILC is an annual survey conducted by the Central Statistics Office (CSO) designed to obtain information regarding the income and living conditions of Irish households. It is the Irish component of an EU-wide survey which aims to capture information on poverty and social exclusion across Europe. SILC 2010 is a survey of 4,642 households, consisting of 11,587 individuals, between January 2010 and January 2011, with an income reference period being the 12 months prior to interview. Crucially, it contains information on household composition so that individuals within a household can be grouped into tax-units. SWITCH, the tax-benefit microsimulation model of the ESRI, is based on this survey.

Webber and Tonkin (2013) described the importance of ensuring similar distributions of matching variables between the income and expenditure surveys, while also arguing that having a smaller number of variables increases the risk that any model fitted to the data will be misspecified and the results of the matching will not be reliable. A range of common variables exist between the HBS and

[11] Food, alcohol, tobacco, fuel and light, clothing, housing, household non-durables, health, public transport, private transport, communications, education, recreation and culture, durables, and other goods. A sensitivity analysis was carried out where expenditure items were into ten aggregated goods, rather than 15. Little sensitivity was found in the results.

SILC.[12] For the majority of variables, the variable distributions are very comparable between the two datasets. For example, approximately 23 per cent of households contain at least one person with an honours degree or higher in both surveys, while the regional split of households is also very similar in the surveys. Some reasonably large differences do emerge between the surveys however. Mean disposable income is almost €60 per week higher in the HBS compared to SILC. As explained by the CSO (CSO, 2012), these difference emerge due to different reference periods in the two surveys. In addition, a higher proportion of households outright own their house in the SILC, while more households have a mortgaged property or rent their property in the HBS. The labour force status (LFS) distribution also differs between the datasets. In particular, the HBS contains a higher proportion of households where the household reference person (HRP) is employed, and a lower proportion where the HRP is engaged in home duties, compared to SILC. To minimise differences in the LFS distribution between the surveys, we use an alternative specification of this variable in the estimation equations, which indicates the predominant LFS in a household, rather than the LFS of the HRP. While some differences remain, this alternative specification brings the LFS distribution closer between the surveys.

IV ESTIMATION AND IMPUTATION OF EXPENDITURE AND INDIRECT TAXES

The method of estimating expenditure in the HBS proceeds in two steps. First, we estimate total expenditure for each household. We then estimate budget shares, defined as expenditure on good *i* by household *h* as a percentage of total expenditure by household *h*, for the range of aggregated commodities. To get household level estimated expenditure on each of our aggregated commodities, we simply multiply the estimated total expenditure by the estimated budget shares. Using the coefficients of each of these estimation steps, we impute total expenditure and budget shares from the HBS into SILC based on a set of common variables, and estimate expenditure on each of the goods in the same way as above. The following sections examine in more detail the estimation methods, and present results from these estimations.

4.1 Total Expenditure

Total household expenditure is estimated in the expenditure survey using the following specification:

$$ln(\mathrm{E_i}) = \alpha + \sum_{k=1}^{3} \gamma_k \, ln(y_i)^k + x_i'\beta + \varepsilon_i \tag{3}$$

[12] See Appendix Table A1 for summary statistics of common variables.

for households $i = 1, \ldots, H$. $ln(E_i)$ is the logarithm of total expenditures at the household level. $ln(y_i)$ denotes the logarithm of disposable household income and x_i denotes a vector of household-specific characteristics, as detailed in Section III.[13] The error term ε_i is assumed to be independent and identically distributed.[14] As the variance of the predicted expenditure is lower than actual expenditure, an additional error, μ_i, is added to the predicted expenditure. This error term is mean zero, and is specified so that the variance of actual and estimated expenditure is equal in the HBS data.[15]

The coefficients from the total expenditure equation are used to impute expenditure into the income survey, as in Equation 4.[16] Total expenditure estimations and imputations are corrected for retransformation bias via use of the smearing estimator (see Duan *et al*. (1983) and Manning (1998) for details). Similar to the estimated total expenditure in the HBS data, imputed total expenditure in the income survey ($\tilde{E}_{SILC}$) has lower variance than actual total expenditure in the expenditure survey. Therefore, the additional error term μ_i as before, is added to the imputed expenditure in the SILC data.

$$\tilde{E}_{SILC} = \Phi\left(\hat{\alpha}_{HBS} + \sum \hat{\gamma}k_{\text{HBS}}\, ln(y_{SILC})^k + x'_{SILC}\hat{\beta}_{HBS}\right) + \mu \qquad (4)$$

where $\hat{\alpha}_{HBS}$, $\hat{\gamma}k_{\text{HBS}}$ and $\hat{\beta}_{HBS}$ are the parameters estimated by Equation 3 using the HBS, y_{SILC} and x_{SILC} are the observed values of y_i and x_i in SILC, and Φ is a function transforming estimated log expenditure correcting for retransformation bias.

As in Sutherland *et al*. (2002), our validation of the imputation procedure relies on comparison of the distribution of expenditure (and later indirect taxes) when we use actual, estimated or imputed data. Figure 1 compares the distribution of expenditure by decile of equivalised disposable income.[17] With the actual HBS data, mean expenditure is higher in the bottom decile than either of the second or third deciles. From the second decile onwards, mean expenditure is increasing in income. The pattern is less pronounced in the estimated and imputed data, where mean expenditure remains relatively constant between deciles one, two and three.

[13] Regression results can be found in Appendix Table A2.

[14] O'Donoghue *et al*. (2004) suggest that results of this regression are likely to be affected by heteroscedasticity. However, given that expenditure imputation relies on coefficients rather than standard errors, and that validation of the approach is based largely on comparison of actual and imputed expenditure (see Sutherland *et al*., 2002), the presence of heteroscedasticity in the regression results will not alter the results of the analysis.

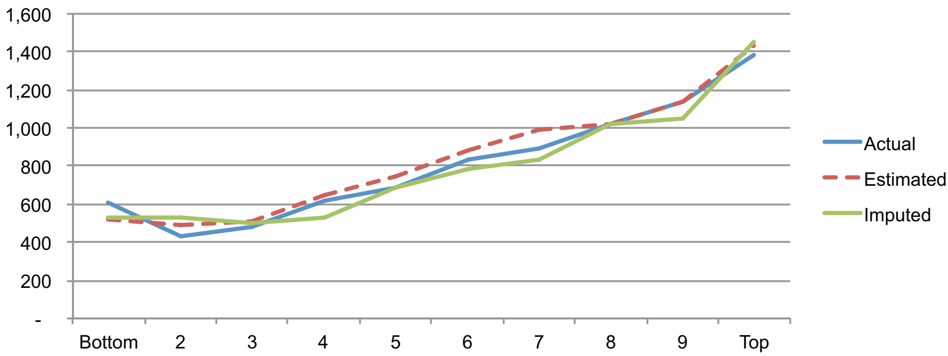
[15] A number of alternative specifications of μ_i were applied, including setting the standard deviation equal to the root mean square error from the regression in Equation 3, with little resulting sensitivity in the distribution of household expenditure.

[16] Household subscripts dropped for clarity.

[17] See Appendix Table A3 for comparison across categories of expenditure rather than income deciles.

From the third decile onwards, the pattern of expenditure by income is very similar in each of the three approaches.

Figure 2: Mean Expenditure by Decile of Equivalised Disposable Income



4.2 Estimating Budget Shares of Aggregated Commodities

Having estimated total expenditure for each household, and decided upon the level of aggregation of expenditure items, the next step in the process is to predict expenditure shares for each of the aggregated goods. Here the aggregated goods are split into two groups. The first group contains goods for which all households record a positive expenditure.[18] Household expenditure for this group of goods is estimated using the following specification, using the full sample of households in the expenditure data.[19]

$$w_i^j = \alpha^j + \gamma_1^j \, ln(E_i) + \gamma_2^j \, ln(E_i)^2 + x_i' \beta^j + \eta_i^j \tag{5}$$

for households $i = 1,...,N$ and commodities $j = 1,..., J$. x_i contains the same set of variables as in the specification for total expenditure. Again the error term, η_i^j, is assumed to be independent and identically distributed.

The second group contains goods for which a significant proportion of households record zero-expenditures. For example, a significant proportion of households will commonly not have any expenditure on goods such as alcohol and tobacco. For this group of goods, we apply a two-stage estimation procedure. Following Decoster *et al*. (2013), in the first step a probit model is estimated for the probability of positive demand for the respective good:

[18] In practice, we include goods for which at least 99 per cent of households have positive expenditure.
[19] Results here estimated via OLS regression. As a sensitivity test, we also estimated budget share equations using the Papke-Wooldridge fractional logit estimator. The results were insensitive to the choice of estimator.

$$Pr(D_i^k = 1) = \phi(\alpha^k + \delta_1^k \, ln\,(y_i) + \delta_2^k \, ln\,(y_i)^2 + x_i' \beta_0^k + v_i^k) \qquad (6)$$

for households $i = 1,...,N$, where D_i^k denotes a dummy variable that is 1 for household i if demand for good k is positive, and zero otherwise. The term α^k denotes a commodity-specific constant, and v_i^k is an error term, assumed to be independent and identically distributed. Equation 6 is estimated separately for each of the goods in the "zero-expenditures" group. In the second step, budget share equations are estimated for each of the goods, conditional on the outcome of the first step. Predicted probabilities are set equal to zero if $Pr(D_i^k = 1)$ is lower than the mean probability for good k, and set equal to one otherwise. For households for which demand is estimated to be positive ($D_i^k = 1$), we estimate on a restricted sample:

$$w_i^k = \alpha^k + \gamma_1^k \, ln\,(E_i) + \gamma_2^k \, ln\,(E_i)^2 + x_i' \beta^k + \eta_i^k \qquad if\ \hat{D}_i^k = 1 \qquad (7)$$

for goods $k = 1,..., K$.

4.3 Imputing Budget Shares into SILC

Using the results of the budget share equations, budget shares are imputed into the income survey and expenditure on each representative good is estimated by multiplying the predicted budget shares by the predicted total expenditure. In some cases, negative budget shares are predicted for some households in the income survey. Negative imputed budget shares are set to zero, and the other budget shares are rescaled, to comply with adding-up conditions.

For the goods with significant proportions of zero-expenditures, imputation is based on the two-stage estimation procedure. To incorporate a degree of randomness in estimating which households have positive expenditure on these goods, we follow the approach suggested in Decoster *et al.* (2014). After imputing the probabilities for positive expenditures, a number is drawn for each commodity from a uniform distribution between 0 and 1. If the number drawn is smaller than the predicted probability, $\hat{D}_i^k$ in the SILC data is coded 1, and 0 otherwise. Positive expenditure is then assigned using the second stage regression output for households with $\hat{D}_i^k = 1$. This approach is also used with the estimated data, and ensures that a portion of households with a relatively low predicted probability of having positive expenditure are estimated to have positive expenditure (as is observed in the HBS data).

Table 1 compares the actual, estimated and imputed budget shares for the 15 aggregated goods in the analysis. The estimated and imputed budget shares are very similar for the majority of goods. For example, the mean and median budget share for food is close to 20 per cent for both. For most commodities, mean and median budget shares are within 0.4 percentage points of each other. The estimated and imputed budget shares generally compare well with the actual budget shares. Comparing the mean budget shares of the actual and imputed expenditure data, most are within 1 percentage point of each other, with the largest differences not exceeding 2.5 percentage points. The proportion of households reporting zero-expenditure on the relevant subset of goods also compares well across the actual, estimated and imputed expenditures.[20]

Table 1: Budget Shares Actual, Estimated and Imputed

	HBS				*SILC*	
	Actual		*Estimated*		*Imputed*	
	Mean (%)	*Median (%)*	*Mean (%)*	*Median (%)*	*Mean (%)*	*Median (%)*
Food	18.4	16.9	20.1	19.4	20.8	20.0
Alcohol	3.2	1.8	2.3	2.6	2.3	2.6
Tobacco	2.3	0.0	1.0	0.0	1.0	0.0
Clothing	4.5	2.5	3.3	3.9	3.3	4.0
Fuel and Light	5.8	4.4	6.4	5.8	6.8	6.2
Housing	18.2	16.3	19.9	21.4	17.6	17.0
Household Non-Dur.	2.2	1.7	2.4	2.4	2.4	2.3
Health	2.2	0.4	1.4	1.0	1.6	1.2
Public Transport	1.1	0.0	0.6	0.0	0.5	0.0
Private Transport	9.2	8.0	9.2	10.0	9.4	10.2
Communications	3.8	3.1	4.1	4.0	4.2	3.9
Rec. and Culture	4.2	3.0	4.5	4.4	4.7	4.7
Education	2.0	0.0	1.1	0.0	0.9	0.0
Durables	6.4	3.5	5.8	6.5	6.0	6.9
Other	16.5	14.5	17.8	18.2	18.5	18.7

Note: Survey weights applied here, and in all other tables and figures

4.4 Estimating Indirect Tax Payments using Estimated and Imputed Data

The final pieces of information we require to assess the distributional effects of indirect taxes, given our estimates of expenditure on each of the aggregated goods, are the VAT rates, excise duties and carbon taxes applicable to each of the aggregated goods. To calculate the proportion of expenditure on each aggregated good that was paid in VAT, we categorise the expenditure items into groups according to whether the item had a zero, reduced, second reduced or standard VAT

[20] See Appendix Table A4 for details. Appendix Table A5 shows that the frequency of estimated and imputed zero-expenditures for the relevant goods drops significantly if the budget shares of these goods are estimated by Equation 5.

rate applied. We then calculate, by decile, the share of each aggregated good that was spent on each VAT category. For example, in the bottom decile 64 per cent of expenditure on food was on food items that had a zero VAT rate applied in 2010; the corresponding figure in the top decile was 51 per cent. With this information we can simply estimate the VAT element of expenditure on each aggregated good as:

$$\widehat{VAT}_i^j = \sum_{v=1}^{V} \hat{x}_i^j * \pi_d^{jv} * \frac{t^v}{1 = t^v} \tag{8}$$

where $v = 1,\ldots,V$ are the different VAT rates in the indirect tax system, and π_d^{jv} is the share of total expenditure on good j in decile d ($d = 1$ to 10) that was spent on expenditure items taxed at VAT rate v. In the 2010 Irish system, goods could be subject to a zero, reduced or standard VAT rate, so in this case $V = 3$. From July 1, 2011, a second reduced rate of VAT of 9 per cent was introduced in Ireland, resulting in four VAT rates in total, so V increased to 4. Excise duty and carbon tax payments can be estimated in a similar manner, again using the assumed prices to estimate quantities purchased.

4.5 Distributional Effects of Indirect Taxes and Indirect Tax Reform

Imputation procedures, such as the one used in this analysis, cannot be expected to accurately predict expenditure levels on a household by household basis. In this section, we analyse the distribution of indirect tax payments in Ireland, comparing the results using the estimated and imputed expenditure data with the results using the actual HBS data. As argued by Sutherland *et al*. (2002), the imputed expenditure should be used at a sufficient level of aggregation to mask differences that are not controlled by the imputation procedure. Therefore, like much distributional analysis (see Callan *et al*. (2013), Decoster *et al*. (2011) among others), the results are shown by income decile.

Table 2 shows the proportion of disposable income paid in indirect taxes, as well as the proportion of income paid in VAT, across the income distribution under the 2010 indirect tax system. The difference between total indirect taxes and VAT is made up of excises and the carbon tax. The results portray the regressive nature of the indirect tax system in Ireland, with the proportion of income paid in indirect taxes falling as income rises.[21] Encouragingly, in eight out of the ten deciles, the estimated and imputed indirect tax payments are within 1.2 percentage points of the actual data. Similarly, the state average indirect tax payment is approximately 10 to 11 per cent in each case. The largest difference emerges in the bottom decile. The estimated data suggest indirect taxes were under 18 per cent of disposable

[21] Measuring indirect taxes as a proportion of expenditure rather than income can result in quite a different impression of the distributional impact of the indirect tax system. See Appendix B for more details.

income, while the imputed data figure of 20.5 per cent is closer to the actual figure of 24.6 per cent.[22]

It is worth bearing in mind that while the actual data are presented as the "benchmark" measure, the income measure in the HBS may be subject to a certain degree of measurement error, particularly in the extremes of the distribution.[23] Differences in the bottom decile may therefore be due to a combination of imperfections in the imputation procedure (affecting the "imputed" figures), and measurement error in the HBS data (affecting the "actual" figures). Therefore, the relationship between income and expenditure in the bottom decile can be particularly difficult to model. This is supported by Meyer and Sullivan (2011) who suggested that that income is likely to be poorly measured for households with low resources and, in particular, likely to be under-reported. Similarly, Brewer and O'Dea (2012) found evidence from UK data of under-reporting of income among

Table 2: Indirect Taxation by Decile of Equivalised Disposable Income, 2010 as a Percentage of Disposable Income

	HBS				*SILC*	
	Actual (%)		*Estimated (%)*		*Imputed (%)*	
Decile	*Total Indirect Taxes*	*VAT*	*Total Indirect Taxes*	*VAT*	*Total Indirect Taxes*	*VAT*
Bottom	24.6	16.4	17.6	13.1	20.5	14.9
2	15.3	9.9	14.4	10.4	15.2	10.9
3	14.6	9.3	12.6	9.1	12.7	9.3
4	12.6	8.4	11.7	8.4	11.4	8.4
5	12.1	8.0	11.4	8.2	11.6	8.4
6	11.7	7.9	10.9	7.9	10.7	7.8
7	10.0	6.8	10.1	7.3	9.7	7.0
8	9.7	6.9	9.0	6.4	9.8	7.0
9	9.0	6.4	8.2	6.0	8.6	6.3
Top	6.4	4.8	6.5	4.9	6.6	5.0
State	10.5	7.2	9.7	7.1	10.0	7.3

[22] Aggregate VAT receipts simulated by the model represent approximately 50 per cent of VAT receipts reported by the Revenue Commissioners (2012, 2013a). Collins (2014) reports that households contributed on average 49 per cent of the total VAT tax take in between 2000 and 2011, suggesting the results of the model capture well the VAT contributions of households in Ireland. Excise and carbon tax receipts are closer to 40 per cent of Revenue reported receipts, which is unsurprising given the nature of the goods subject to these excise duties and carbon tax.

[23] The Central Statistics Office (2012) states that "although income data from the two sources are not directly comparable, calculation of key income distribution and poverty indicators using income data from both surveys resulted in similar values." However, they go on to say that "The SILC is recognised as the primary source of data on income in Ireland", and elsewhere state that HBS income data are "used primarily for categorical purposes (e.g. for analysis of households according to different levels of income) rather than the provision of information on income levels".

households with low resources. They argue that household expenditure may be more likely to be accurately measured. Therefore, given the possibility of some data error in the bottom decile in the HBS data in particular, it is unsurprising to see some element of discrepancy in the results here. The similarity of the distribution of indirect tax payments using the imputed and actual data however are an encouraging basis from which to undertake some indirect tax reform analysis.

Table 3: The Distributive Impact of 2010 to 2014 VAT Reforms, Percentage of Disposable Income

	HBS		*SILC*
Decile	*Actual (%)*	*Estimated (%)*	*Imputed (%)*
Bottom	–0.59	–0.65	–0.75
2	–0.38	–0.46	–0.50
3	–0.35	–0.35	–0.37
4	–0.32	–0.38	–0.37
5	–0.31	–0.39	–0.40
6	–0.29	–0.37	–0.37
7	–0.23	–0.33	–0.32
8	–0.21	–0.29	–0.32
9	–0.20	–0.26	–0.28
Top	–0.13	–0.22	–0.22
State	–0.24	–0.32	–0.33

Between 2010 and 2014 in Ireland, two major VAT reforms occurred. The first was the introduction of a second reduced rate of VAT. A new rate of 9 per cent was applied to a range of goods and services in the tourism sector. The second major VAT reform was an increase in the standard rate of VAT from 21 per cent to 23 per cent. Table 3 compares the distributive impact of these reforms, again using the three sources of expenditure data. The results show that for a marginal indirect tax change, such as those in Ireland between 2010 and 2014, the actual and imputed data produce very similar results. The largest difference between the actual and imputed expenditure again occurs in the bottom decile, yet qualitative results are largely unaffected.

V INTEGRATED MODELLING OF REFORMS TO THE DIRECT TAX, INDIRECT TAX, AND SOCIAL WELFARE SYSTEMS

In this section we undertake an integrated analysis of direct tax and welfare reform in conjunction with an indirect tax reform. Such joint reforms are common in practice. In the UK in 2011, for example, the standard rate of VAT was increased from 17.5 per cent to 20 per cent, while the income tax allowance for those aged under 65 was increased by £1,000. Similarly, in Budget 2012 in Ireland, the lower threshold for the Universal Social Charge[24] was increased at the same time as an increase in the standard rate of VAT. In both cases, the reforms were implemented along with several other tax and welfare reforms.

The reforms considered here are not policy recommendations, nor are they analyses of actual governmental policy. Rather they serve to highlight the capability of the model developed in this analysis to examine the distributional impact of simultaneous direct tax, indirect tax and welfare reform. The ESRI tax-benefit microsimulation model, SWITCH, is used to model the impact of the direct tax and welfare element of the integrated tax and welfare reform. SWITCH is based on the SILC 2010 survey.[25] For the integrated analysis here, SILC is supplemented with the imputed expenditure data described in previous sections. SWITCH is used to simulate the impact of the welfare reforms, while the indirect tax microsimulation model described in Section IV is used to simulate the impact of the indirect tax reforms. The integrated model produces results at micro (individual) level, which are then aggregated by some observable characteristic, such as income group or family type to produce the results presented in this section.

The results here are "first-round" effects, as they do not include any behavioural responses to the tax reforms. These behavioural responses could come in several forms. Labour supply responses to direct tax and welfare reforms are likely, while labour supply may also respond to changes in indirect taxation. Similarly, consumer demand is likely to respond to indirect tax reform, while it may also respond to direct tax and welfare reform (via income effects). On top of these responses, firm supply of goods and services may respond to changes in their consumer price and changes in demand. Firm demand for labour may respond to changes in labour costs following direct tax reform. Finally, further behavioural responses such as changes in take-up of welfare payments and tax avoidance or evasion are possible. Mirroring the approach taken here, Adam *et al*. (2015) omit behavioural responses from their analysis of UK tax and benefit reforms between 2010 and 2015. Ideally, they argue, distributional effects would be assessed on the basis of a utility functions that account for the value of income and leisure. In the absence of such a structural model however, they argue that "measuring changes in household incomes before behavioural responses is preferable to analysing them afterwards", as it is not clear that the change in income after a behavioural response is a better measure of welfare change than the change in income before a behavioural response. For example, suppose a child benefit cut resulted in an individual choosing to work more hours. If measured after the behavioural change, the person's net income might rise, yet

[24] The Universal Social Charge is a progressively-structured tax on income first introduced in 2011.
[25] See Keane *et al*. (2014) for details.

their welfare may have reduced (their income at any hours choice would be the same as or lower than before).

We analyse two hypothetical reform packages. Each scenario is modelled as a reform to the 2015 tax and benefit system. First, we remove the zero VAT-rating on children's clothing and footwear, and redistribute the additional revenue gained from this reform through an increase in Child Benefit. The second reform package includes an increase in the carbon tax, with the revenue redistributed through an increase in the Fuel Allowance payment.[26] In each case, we examine the impact of the reforms by income decile, and also by other relevant categories of the population to identify the households most affected by the reform.

5.1 Child Benefit and VAT on Children's Clothing (Reform 1)

Children's clothing and footwear in Ireland is subject to a zero rate of VAT. This includes clothing for children aged up to 11 years old, and children's footwear up to UK size 5.5. In this section, we examine the distributional effects of subjecting all children's clothing and footwear to the standard rate of VAT,[27] and redistributing the additional VAT revenue raised through an increase in Child Benefit. Child benefit in Ireland is paid universally, so that any household that contains a child under the age of 16, or 18 if the child is still in education, receives the payment. In 2015 the value of Child Benefit was €135 per child per month. After a process of trial and error using SWITCH, we found increasing Child Benefit by €5.50 per month fully exhausted the €75 million per annum revenue raised though the VAT increase. Given the different age classifications of children across the two schemes, Child Benefit is not perfectly targeted at the households most likely to lose out from the VAT increase on children's clothing and footwear. We can use the model developed in previous sections to examine which households are most likely to gain from such a reform, and which households are most likely to lose.

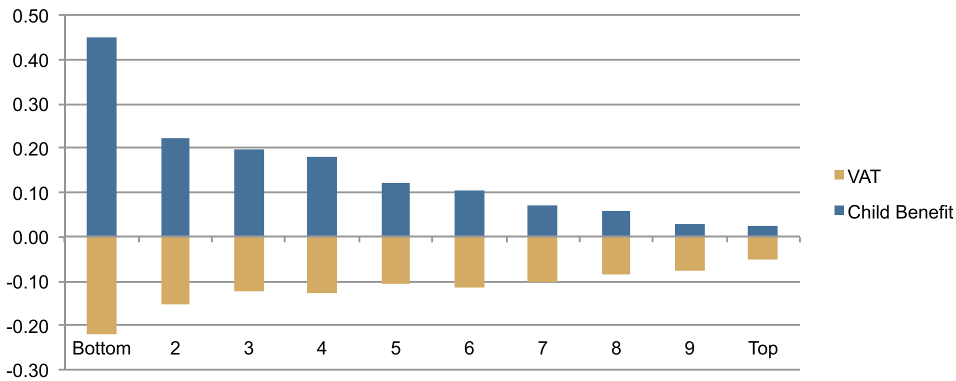
Figure 3 shows the percentage change in household disposable income by income decile. As expected, the VAT increase results in a loss in income in each decile. In percentage terms, the VAT reform is regressive, with the negative impact on income largest in the bottom decile and becoming smaller higher up the income distribution. When examining the impact by income decile, the increase in child benefit performs well in compensating for the increased VAT payments. Income in the bottom decile is increased by 0.4 per cent due to the child benefit increase, resulting in a net increase of 0.2 per cent of income when the VAT increase is also

[26] Both scenarios are joint reforms to the indirect tax system and the welfare system. Welfare reforms, similar to direct tax reforms, can be modelled in this framework due to the ability to identify the relevant tax-units (or welfare-units) within each household.

[27] In the HBS, expenditure on children's clothing is aggregated into expenditure on clothing for children aged five to 15. We assume this expenditure is evenly distributed across the age range, so that €6 of every €11 spent on this category is subject to the zero rate of VAT. Collins (2014) employed a similar approach.

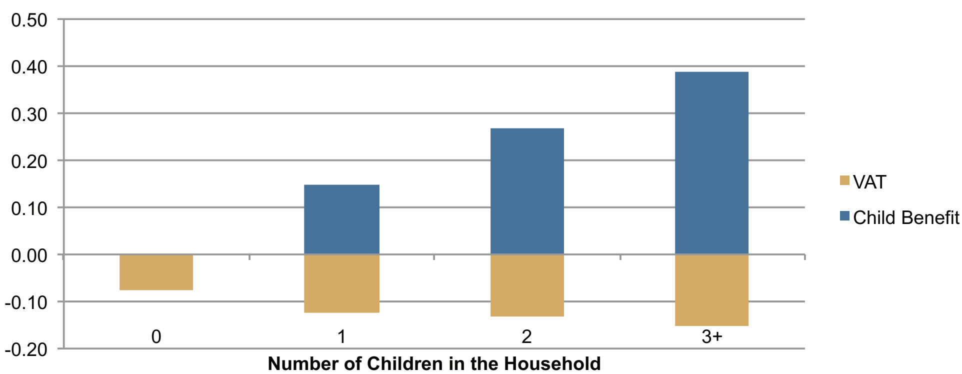
taken into account. The second third and fourth deciles also experience a net gain in their disposable income, while the fifth to seventh deciles are almost perfectly compensated. The top three deciles face a small net loss in disposable income, meaning the overall impact of the reform is progressive.

Figure 3: Percentage Change in Disposable Income by Decile – Reform 1



Of course, not all households within each decile fare equally due to the VAT and Child Benefit reforms. We would expect households with children to be affected to a larger degree by both tax-benefit reforms. Figure 3 categorises households by the number of children present, where children are defined as aged 17 or under. Given that Child Benefit is paid in respect of children aged at a maximum of 17, households with no children aged 17 or under receive no increase in their income as a result of the increase in child benefit. Nevertheless, there is a small decrease in their income due to the increase in VAT on children's clothing and footwear. The more children present in a household, the greater the changes in income due to both

Figure 4: Percentage Change in Disposable Income by Number of Children in Household – Reform 1



the VAT change and the Child Benefit change. The Child Benefit increase almost perfectly compensates households with just one child for the VAT increase, while households with two or more children experience a net increase in their income. It is important to note that, despite three of the four categories of households in Figure 3 maintaining or increasing their income due to the tax-benefit reforms, households with no children are by far the most common household type in Ireland, making up well over 50 per cent of the sample of households in SILC 2010.

5.2 Carbon Tax and the Fuel Allowance (Reform 2)

A carbon tax was introduced in Ireland in December 2009. The carbon tax, which operates much like excise duties, applies to a range of energy related goods, such as natural gas, liquid fuels and solid fuels. Initially, the tax applied to transport fuel (petrol and auto-diesel) only. In May 2010 it was extended to include non-transport fuels, including kerosene, liquid petroleum gas (LPG) and natural gas. Solid fuels, such as coal and commercial peat, became subject to the tax in May 2013. Since May 1, 2014, the carbon tax has been charged at a rate of €20 per tonne. This rate is used to calculate the rate of tax that applies per unit of each of the different fuels. For example, a rate of €4.10 per megawatt hour applies to natural gas. It is calculated by multiplying the emission factor of natural gas, expressed in kilograms of CO_2 per terajoule (57,022), by the number of terajoules per megawatt-hour (0.0036), and multiplying the resultant figure by a rate of 2 cent per kilogram of CO_2 emitted (Revenue Commissioners, 2013b).

In this section, we examine the impact of a 10 per cent net increase[28] of the carbon tax on all fuels, with the exception of transport fuels.[29] Therefore, in this reform, households will predominantly experience an increase in home heating and energy fuels such as natural gas, home heating oil, and solid fuels like coal and turf. We exclude transport fuels from the carbon tax increase for two reasons. First, they are already subject to relatively high excise duties, and in 2013 contributed over half of total carbon tax revenue (Revenue Commissioners, 2013a). Second, a scheme exists within the Irish Social Welfare system that is designed to help low income households with their home heating costs. We can therefore use this scheme to redistribute the additional carbon tax revenue in an equitable manner.

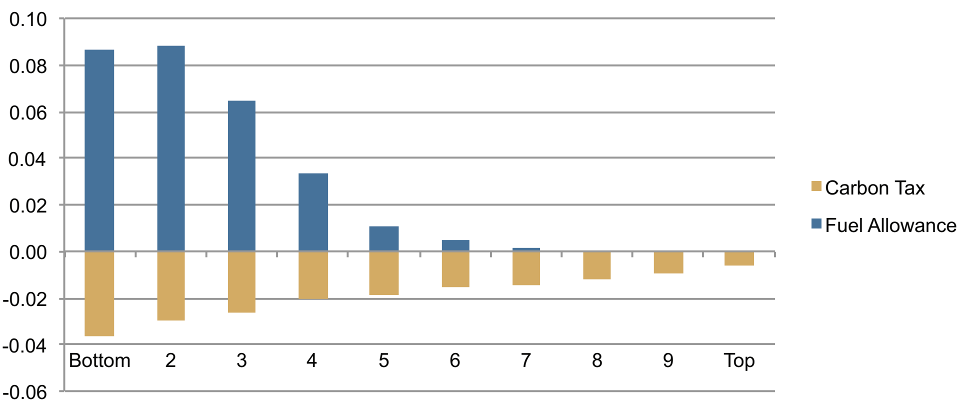
Given that the carbon tax increase modelled here applies largely to fuels used in the home, we redistribute the additional revenue raised through an increase in the Fuel Allowance. The Fuel Allowance is a payment under the National Fuel Scheme to help with the cost of home heating. The benefit is paid for 26 weeks of the year at a rate of €20 per week. It is paid to people who are dependent on long-term social welfare, and is therefore concentrated at the lower end of the income

[28] The 10 per cent increase we model here is inclusive of the increase in VAT that will occur as a result of the carbon tax increase. In Ireland, VAT is applied after the imposition of excise duties and carbon taxes.
[29] The model predicts an increase in revenue of €11.2 million per annum due to this reform.

distribution. As before, we use SWITCH to examine the distributional effects of an increase in this payment. An increase of €1.50 per week in the Fuel Allowance fully redistributes the additional revenue raised from the increase in the carbon tax.

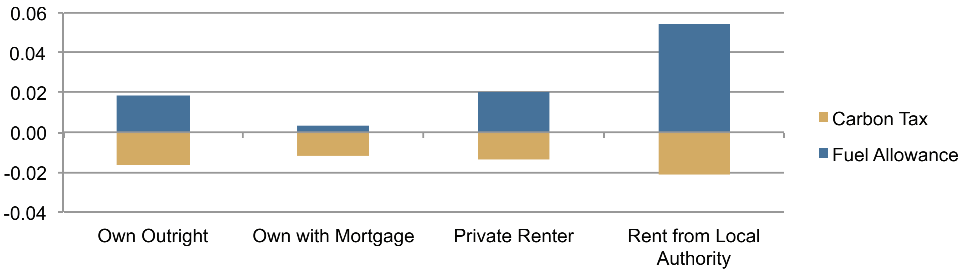
Figure 5 shows the impact by income decile of the reforms to the carbon tax and Fuel Allowance. The increase in the Fuel Allowance is largely progressive, with the largest increases in income occurring in the bottom two deciles. The percentage increases in income fall sharply in the next four deciles, and are essentially zero in the top four deciles. This pattern is due to the means-test element of the scheme, whereby high income households will be ineligible for the payment. The carbon tax increase shows a regressive pattern, with the percentage decreases in income becoming smaller as we move up the income distribution. This pattern is consistent with the findings of Callan *et al*. (2009). Overall, the distributional impact of the reform would be broadly progressive, although the second decile gains most from the reform, closely followed by the first, then third, deciles. Small losses are recorded in each decile in the top half of the distribution, with the seventh and eighth deciles losing the highest proportion of income. Given the relatively small size of the reform, gains and losses in each decile are small, below 0.1 per cent of income in each case.

Figure 5: Percentage Change in Disposable Income by Decile – Reform 2



Households renting from the local authority households are the biggest net gainers from the reform, due to a relatively large gain in income from the Fuel Allowance reform. Both the Fuel Allowance and access to local authority housing is means-tested; low income households are therefore concentrated in this tenure type, and are most likely to gain from the Fuel Allowance increase. Households with a mortgage gain least from the Fuel Allowance increase. This tenure type is dominated by working age adults with relatively high incomes. They are therefore least likely to qualify for the Fuel Allowance.

Figure 6: Percentage Change in Disposable Income by Household Tenure – Reform 2



VI CONCLUSION

Accurate appraisal of public policy should include as much of the relevant information as possible. In measuring the distributional impact of the tax and welfare system, it is therefore important to include not only the impact of the direct tax and welfare system, but also the impact of the indirect tax system. Comprehensive modelling of the impact of both direct and indirect taxes, and of tax policy options, requires data on the impact at micro-level household and tax-unit level. Most countries, however, do not have a single source of micro-data including high-quality disaggregated information on both incomes and expenditures. This makes approaches which impute expenditure data into detailed income surveys of considerable interest.

In this paper, we investigated the feasibility of imputing expenditure information from a household expenditure survey to an income survey, so that the impact of reforms to the direct tax, indirect tax and welfare systems could be simultaneously examined. The distributions of estimated expenditure data in the HBS and the imputed expenditure data in SILC were highly comparable to the distribution of actual HBS expenditure data, despite some sensitivity in the bottom income decile to the choice between the expenditure data. When we used each of the expenditure data sources to estimate indirect tax payments, again the results were highly comparable across datasets.

The empirical findings in the analysis confirm the regressive nature of the indirect tax system when measured relative to income in Ireland, on a snapshot basis at least. It is important to remember, however, that the overall distributional impact of a tax system or tax reform depends not just on the indirect tax system, but also on the direct tax system and the welfare system.[30] An integrated analysis of all three

[30] The distributional impact can also depend on the time period under consideration. A lifecycle analysis, as discussed in a preliminary study by Roantree and Shaw (2014), can substantially alter the degree of progressivity or regressivity found in a tax system. Including such issues is beyond the scope of the current analysis.

components – direct tax, indirect tax and welfare – is needed to provide a comprehensive picture of the overall distributional impact of the tax and benefit system. A number of recent recommendations on the structure and analysis of the effects of taxation are particularly relevant here. In a comprehensive review of the tax system in the UK, Mirrlees *et al*. 2011 argued that a failure of government was not examining the tax system as a whole and recognising that the rate schedule of personal income taxes and benefits is the instrument most suited to achieving redistributive ends. Elsewhere, the World Bank (2003) recommended that regressivity be studied in an overall context of the applicable fiscal policies including direct, indirect taxes, and public expenditures. They also argued that regressivity can be addressed more effectively with sound income taxation and government expenditures. Assessing the distributional impact of the indirect tax system alone, without considering the distributional impact of other parts of the tax and welfare system, therefore omits some of the most important schemes used to target distributional concerns.

With this in mind, the final section of the paper applied the model to two reform scenarios, both of which included an increase in an indirect tax with the additional revenue raised redistributed through an existing social welfare scheme. The analysis showed that it was possible to use the revenue raised through an increase in indirect taxes to compensate households in a manner that was neutral or progressive across the income distribution by increasing a suitable social welfare payment. The reform packages examined were quite straightforward and resulted in small changes in disposable income of households across the income distribution. They are neither actual policy evaluations nor policy proposals. Yet they served to highlight two important points in the analysis of tax and welfare reform. First, taken in isolation, either social welfare reform resulted in a largely progressive effect, while either indirect tax reform resulted in a largely regressive effect. Without the benefit of analysing the impact of both reforms, incorrect conclusions may be drawn about the distributional impact of reform package. A similar argument can be applied when categorising households by characteristics other than income, such as the number of children in the household or the tenure status of the household. Second, the ability to simultaneously model the impact of direct and indirect tax reforms, as well as social welfare reforms, allows policymakers to examine the distributional impact of a broader range of options for the distribution and redistribution of revenue. As a result, the information available to policymakers when evaluating implemented tax reforms, or considering a range of possible tax reforms, is substantially increased by this approach.

Ultimately, the aim of this analysis is to assess the feasibility of using the imputation process to enhance the SILC data with expenditure data. This process allows simultaneous analysis of the direct tax, indirect tax and welfare systems. The methods analysed here are not, of course, a perfect substitute for a single dataset containing detailed income, expenditure, labour market participation and other information. As argued by Sutherland *et al*. (2002), expenditure imputation cannot be expected to reproduce the micro-level diversity captured in actual expenditure data. Nonetheless, taking the advice of Mirrlees and others, it is important to examine the distributional impact of the indirect tax system in the context of the full tax and benefit system. The methods and results in this paper are encouraging in that regard. By showing that the imputation procedure used in this analysis produced results with the SILC data that are highly comparable to results produced with actual expenditure data in the HBS, the analysis here increases the capacity of measuring the distributional impact of the tax and benefit system in Ireland.

BIBLIOGRAPHY

Adam, Stuart, James Browne and William Elming, 2015. "The Effect of the UK Coalition Government's Tax and Benefit Changes on Household Incomes and Work Incentives", *Fiscal Studies*, Vol. 36, pp. 375-402.

Arellano, Manuel and Costas Meghir, 1992. "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets", *Review of Economic Studies*, Vol. 59, pp. 537-59.

Avram, Silvia, Horacio Levy and Holly Sutherland, 2014. "Income redistribution in the European Union", *IZA Journal of European Labor Studies* 3 (December 2014), pp. 1-29.

Banks, James, Richard Blundell and Arthur Lewbel, 1996. "Tax Reform and Welfare Measurement: Do We Need Demand System Estimation?", *Economic Journal*, Vol. 106, pp. 1227-1241.

Barrett, Alan and Caeman Wall, 2006. *The Distributional Impact of Ireland's Indirect Tax System*. Combat Poverty Agency.

Bradford, David F., 1974. "Generalisation and Proof of the Hicks Composite Commodity Theorem without a Utility Function." *Econometric Research Program Princeton University Research Memorandum No.161*, Princeton University.

Brewer, Mike and Cormac O'Dea, 2012. "Measuring living standards with income and consumption: evidence from the UK", IFS Working Papers, Institute for Fiscal Studies.

Browning, Martin and Soren Leth-Petersen, 2003. "Imputing consumption from income and wealth information", *Economic Journal*, Vol. 113, pp. F282-F301.

Callan, Tim, Brian Nolan, Claire Keane and John R. Walsh, 2010. "Inequality and the Crisis: The Distributional Impact of Tax Increases and Welfare and Public Sector Pay Cuts", *The Economic and Social Review*, Vol. 41, pp. 461-471.

Callan, Tim, Brian Nolan, Claire Keane, Michael Savage and John R. Walsh, 2013. "Crisis, Response and Distributional Impact: The Case of Ireland", IZA Discussion Papers, Institute for the Study of Labor (IZA).

Callan, Tim, Chrysa Leventi, Horacio Levy, Manos Matsaganis, Alari Paulus and Holly Sutherland, 2011. "The Distributional Effects of Austerity Measures: A Comparison of Six EU Countries", EUROMOD Working Papers, EUROMOD at the Institute for Social and Economic Research.

Callan, Tim, Sean Lyons, Susan Scott, Richard S. J. Tol and Stefano Verde, 2009. "The Distributional Implications of a Carbon Tax in Ireland", *Energy Policy*, Vol. 37, pp. 407-412.

Central Statistics Office, 2012. "Household Budget Survey 2009/10 " Volume 1", Tech. rep., Central Statistics Office, Dublin, Stationery Office.

Collins, Micheál L., 2014. "Total Direct and Indirect Tax Contributions of Households in Ireland: Estimates and Policy Simulations", Working Paper, Nevin Economic Research Institute.

Cragg, John G., 1971. "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods", *Econometrica*, Vol. 39, pp. 829-44.

Crossley, Thomas F., David Phillips and Matthew Wakefield, 2009. "Value Added Tax" in *The IFS Green Budget: January 2009*, edited by Robert Chote, Carl Emmerson, David Miles and Jonathan Shaw. IFS.

Decoster Andre, Richard Ochmann and Kevin Spiritus, 2013. "Integrating Indirect Taxation into EUROMOD. Documentation and Results for Germany", EUROMOD Working Papers, EUROMOD at the Institute for Social and Economic Research.

Decoster Andre, Bram De Rock, Kris De Swerdt, Jason Loughrey, Cathal O'Donoghue and Dirk Verwerft, 2007. "Comparative Analysis of Different Techniques to Impute Expenditures into an Income Data Set", Accurate Income Measurement for the Assessment of Public Policies Deliverable 3.4, KULeuven.

Decoster Andre, Jason Loughrey, Cathal O'Donoghue and Dirk Verwerft, 2010. "How Regressive Are Indirect Taxes? A Microsimulation Analysis For Five European Countries", *Journal of Policy Analysis and Management*, Vol. 29, pp. 326-350.

Decoster Andre, Jason Loughrey, Cathal O'Donoghue and Dirk Verwerft, 2011. "Microsimulation of Indirect Taxes", *International Journal of Microsimulation*, Vol. 4, pp. 41-56.

Decoster Andre, Kris De Swerdt, Jason Loughrey, Cathal O'Donoghue and Dirk Verwerft, 2009. "Imputation of Expenditures Into the Income Datasets of Five European Countries", Accurate Income Measurement for the Assessment of Public Policies Deliverable, 3.5:, KULeuven.

Decoster Andre, Richard Ochmann and Kevin Spiritus, 2014. "Integrating VAT into EUROMOD. Documentation and Results for Belgium", Tech. rep., EUORMOD Working Paper Series 12/14.

Department of Finance "Energy and Environmental Taxes, 2012", TSG 12/23, Department of Finance, Ireland.

Department of Social Protection, 2013. "Social Impact Assessment of the Main Welfare and Direct Tax Measures in Budget 2013", Tech. rep., Department of Social Protection.

Duan, Naihua, Willard G. Manning, Carl N. Morris and Joseph P. Newhouse, 1983. "A Comparison of Alternative Models for the Demand for Medical Care", *Journal of Business & Economic Statistics*, Vol. 1, pp. 115-126.

Engel, Christoph and Peter G. Moffatt, 2014. "dhreg, xtdhreg and bootdhreg: Commands to implement double-hurdle regression", *Stata Journal*, Vol. 14, pp. 778-797.

Figari, Francesco and Alari Paulus, 2012. "GINI DP 28: The Impact of Indirect Taxes and Imputed Rent on Inequality: A Comparison With Cash Transfers and Direct Taxes in Five EU Countries", GINI Discussion Papers, AIAS, Amsterdam Institute for Advanced Labour Studies.

Heckman, James J., 1979. "Sample Selection Bias as a Specification Error", *Econometrica*, Vol. 47, pp. 153-61.

Herault, Nicolas and Francisco Azpitarte, 2016. "Understanding Changes in the Distribution and Redistribution of Income: A Unifying Decomposition Framework", *Review of Income and Wealth*, Vol. 62, pp. 266-282.

Keane, Claire, Tim Callan, Michael Savage, John R. Walsh and Kevin Timoney, 2014. "Identifying Policy Impacts in the Crisis: Microsimulation Evidence on Tax and Welfare", *Journal of the Statistical and Social Inquiry Society of Ireland*, Vol. 42, p. 1.

Leahy, Eimear, Sean Lyons and Richard S. J. Tol, 2011. "The Distributional Effects of Value Added Tax in Ireland", *The Economic and Social Review*, Vol. 42, pp. 213-235.

Manning, Willard G., 1998. "The Logged Dependent Variable, Heteroscedasticity and the Retransformation Problem", *Journal of Health Economics*, Vol. 17, pp. 283-295.

Meyer, B. and J. Sullivan, 2011. "Further Results on Measuring the Well-Being of the Poor Using Income and Consumption", *Canadian Journal of Economics*, Vol. 44, No. 1, pp. 52-87.

Mirrlees, James, *et al*., 2011. *Tax by design: The Mirrlees Review*. Oxford University Press.

O'Connor, Brendan, Terence Hynes, David Haugh and Patrick Lenain, 2016. "Searching for the Inclusive Growth Tax Grail: The Distributional Impact of Growth Enhancing Tax Reform in Ireland", *The Economic and Social Review*, Vol. 47, pp. 155-184.

O'Donnell, Owen, Eddy van Doorslaer, Adam Wagstaff and Magnus Lindelow, 2008. *Analyzing Health Equity Using Household Survey Data : A Guide to Techniques and Their Implementation*. Washington DC: World Bank.

O'Donoghue, Cathal, Massimo Baldini and Daniela Mantovani, 2004. "Modelling the Redistributive Impact of Indirect Taxes in Europe: An Application of EUROMOD", EUROMOD Working Papers, EUROMOD at the Institute for Social and Economic Research.

Pestel, Nico and Eric Sommer, 2016. "Shifting Taxes from Labor to Consumption: More Employment and more Inequality?" *Review of Income and Wealth*, forthcoming, DOI: 10.1111/roiw.12232.

Picos-Sanchez, Fidel, 2011. "Consumption Taxation as an Additional Burden on Labour Income", OECD Taxation Working Papers, OECD Publishing.

Revenue Commissioners of Ireland, 2013. "Excise Net Receipts by Commodity", Statistical Report, Revenue Commissioners.

Revenue Commissioners of Ireland, 2013b. "Guide to Natural Gas Carbon Tax", Statistical Report, Revenue Commissioners.

Revenue Commissioners of Ireland, 2012. "Statistical Report for year ended 31 December, 2011: Excise", Statistical Report, Revenue Commissioners.

Roantree, Barra and Jonathan Shaw, 2014. "The Case for Taking a Life-Cycle Perspective: Inequality, Redistribution and Tax and Benefit Reforms", Tech. rep., Institute for Fiscal Studies.

Sanchis-Llopis, Juan A., 2001. "Infrequency of Purchase: A Model for Clothing Consumption with Panel Data", *Annales dEconomie et de Statistique* (GENES), pp. 49-71.

Savage, Michael, Tim Callan, Brian Nolan and Brian Colgan, forthcoming. "The Great Recession, Austerity and Inequality: Evidence from Ireland", *Review of Income and Wealth*.

Shonkwiler, J. S. and Steven T. Yen, 1999. "Two-Step Estimation of a Censored System of Equations", *American Journal of Agricultural Economics*, Vol. 81, pp. 972-982.

Sustainable Energy Authority of Ireland (SEAI), 2013. "Domestic Fuel Cost Archive", Tech. rep., Sustainable Energy Authority of Ireland.

Sutherland, Holly, Rebecca Taylor and Joanna Gomulka, 2002. "Combining Household Income and Expenditure Data in Policy Simulations", *Review of Income and Wealth*, Vol. 48, pp. 517-36.

Tax Strategy Group, 2015. "Distributional Aspects of the Tax System", Tech. rep., Interdepartmental Tax Strategy Group.

Webber, Dominic and Richard Tonkin, 2013. "Statistical Matching of EU-SILC and the Household Budget Survey to Compare Poverty Estimates Using Income, Expenditures and Material Deprivation", Tech. rep., Eurostat Working Papers and Methodologies Edition ISSN 1977-0375.

World Bank, 2003. "Value-added Taxation: Mechanism, Design and Policy Issues", *Course on Practical Issues of Tax Policy in Developing Countries*, by Le, Tuan Minh (World Bank).

APPENDIX A ADDITIONAL TABLES

Table A1: Sample Descriptives HBS data and SILC data

	HBS	*SILC*
Mean Disposable Income (€)	885.7	827.6
Mean N in Household	2.8	2.7
Proportion with Married Head of Household (%)	51.4	52.7
Mean Proportion Females in Household (%)	52.2	51.3
Highest Education Level in Household (%)		
Junior Cert or Still in Education	30.8	27.9
Leaving Cert, Higher Certificate or Ordinary Degree	45.6	46.4
Honours Degree or Higher	22.7	22.9
Other	0.9	2.8
Location (%):		
Dublin	28.3	26.7
Other Urban	35.1	34.0
Rural	36.7	39.2
Household Tenure (%)		
Own	33.6	44.6
Mortgage	34.3	27.8
Rent	20.1	14.1
Other	12.0	13.5
Household Reference Person LFS (%)		
Employed	53.6	45.6
Unemployed	11.5	8.8
Retired	14.5	14.5
In Education	3.8	2.0
Home Duties	10.4	22.3
Other	6.2	6.8
Predominant Labour Force Status of Adults in HH (%)		
Employed	40.3	34.2
Unemployed	7.1	4.9
Retired	9.8	7.6
In Education	4.9	4.8
Home Duties	4.0	7.1
Households with Equal N of Adults across LFSs	20.3	23.6
Other	13.6	17.7

Table A2: Estimation of Total Expenditure with HBS Data Across A Range of Samples

	(1)		(2)		(3)		(4)		(5)	
	β	*se*	*β*	*se*	*β*	*se*	*β*	*se*	*β*	*se*
Intercept	5.99	0.09	16.69	2.36	5.21	2.85	8.62	2.21	7.41	0.41
Log(Income)	–0.23	0.06	–5.40	1.08	–0.41	1.28	–2.01	1.01	–1.02	0.21
Log(Income Squared	0.02	0.02	0.84	0.16	0.12	0.19	0.36	0.15	0.16	0.04
Log(Income) Cubed	0.00	0.00	–0.04	0.01	–0.01	0.01	–0.02	0.01	0.00	0.00
N in Household	0.14	0.01	0.13	0.01	0.12	0.01	0.12	0.01	0.14	0.01
N under 5 in HH	–0.11	0.01	–0.09	0.01	–0.09	0.01	–0.08	0.01	–0.10	0.01
N 5-13 in HH	–0.09	0.01	–0.07	0.01	–0.07	0.01	–0.07	0.01	–0.08	0.01
N 14-20 in HH	–0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.01
LFS Unemployed	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
LFS Employed	0.09	0.02	0.05	0.02	0.02	0.02	0.05	0.02	0.07	0.02
LFS Retired	–0.21	0.03	–0.22	0.03	–0.24	0.03	–0.20	0.03	–0.21	0.03
LFS In Education	0.20	0.03	0.19	0.03	0.15	0.03	0.20	0.03	0.20	0.03
LFS Home Duties	–0.13	0.03	–0.14	0.03	–0.14	0.03	–0.11	0.03	–0.13	0.03
LFS Ill Disabled	–0.22	0.04	–0.22	0.04	–0.23	0.04	–0.18	0.04	–0.22	0.04
LFS Other	–0.23	0.08	–0.24	0.08	–0.23	0.08	–0.20	0.08	–0.24	0.08
LFS Equal -1 Emp	0.09	0.02	0.07	0.02	0.04	0.02	0.06	0.02	0.08	0.02
LFS Equal -No Emp	–0.02	0.03	–0.03	0.03	–0.07	0.03	–0.03	0.03	–0.02	0.03
Own Outright	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
Mortgage	0.21	0.01	0.21	0.01	0.20	0.01	0.20	0.01	0.21	0.01
Tenant Purchase	–0.12	0.10	–0.11	0.10	–0.12	0.09	–0.11	0.09	–0.11	0.10
Rent	0.05	0.02	0.05	0.02	0.04	0.02	0.05	0.02	0.05	0.02
Local Authority	–0.12	0.02	–0.11	0.02	–0.11	0.02	–0.10	0.02	–0.12	0.02
Rent Free	–0.13	0.05	–0.13	0.05	–0.08	0.05	–0.10	0.04	–0.13	0.05
Other Urban	–0.07	0.01	–0.07	0.01	–0.07	0.01	–0.06	0.01	–0.07	0.01
Rural	–0.08	0.01	–0.08	0.01	–0.07	0.01	–0.08	0.01	–0.08	0.01
Primary education	–0.25	0.02	–0.23	0.02	–0.21	0.02	–0.20	0.02	–0.24	0.02

Table A2: Estimation of Total Expenditure with HBS Data Across A Range of Samples (contd.)

	(1)		(2)		(3)		(4)		(5)	
	β	*se*	*β*	*se*	*β*	*se*	*β*	*se*	*β*	*se*
Lower secondary	–0.21	0.02	–0.19	0.02	–0.18	0.02	–0.17	0.02	–0.20	0.02
Higher secondary	–0.12	0.02	–0.10	0.02	–0.10	0.02	–0.10	0.02	–0.11	0.02
Post Leaving Certificate	–0.10	0.02	–0.09	0.02	–0.09	0.02	–0.07	0.02	–0.10	0.02
Higher Certificate	–0.05	0.03	–0.04	0.03	–0.06	0.03	–0.04	0.03	–0.05	0.03
Ordinary Degree	–0.03	0.02	–0.03	0.02	–0.03	0.02	–0.02	0.02	–0.03	0.02
Honours Degree	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
Postgraduate Degree	0.04	0.02	0.04	0.02	0.04	0.02	0.04	0.02	0.03	0.02
Other	–0.18	0.06	–0.16	0.06	–0.15	0.06	–0.14	0.06	–0.17	0.06
Married Head of Household	0.06	0.01	0.06	0.01	0.05	0.01	0.06	0.01	0.06	0.01
Prop Female	0.04	0.02	0.04	0.02	0.04	0.02	0.05	0.02	0.04	0.02
R-Squared	0.68		0.69		0.69		0.71		0.68	
N	5,891		5,838		5,715		5,796		5,867	

Notes:
(1) Full Sample
(2) Drop bottom 1% of Gross Income Observations
(3) Drop bottom 3% of Gross Income Observations
(4) Drop HHs with expenditure more than 4 times gross income (Pestel and Sommer 2013)
(5) Trimming as per Decoster *et al.* 2013 – if durable expenditure is greater than 2 times disposable income

Table A3: Comparison of actual, estimated and imputed total expenditure percentage in each expenditure category

Expenditure per Week (euro)	*Actual*	*Estimated*	*Imputed*
<200	6.6	8.2	9.5
<400	16.5	12.9	14.6
<600	19.5	16.7	17.0
<800	16.0	17.2	16.7
<1000	12.7	14.0	13.5
<1250	11.6	11.4	12.5
<1500	7.0	8.8	7.2
<1750	4.4	4.9	4.2
<2000	2.2	2.4	2.3
>2000	3.6	3.5	2.5
Mean (euro)	811	839	792
Median (euro)	683	738	700

Table A4: Frequency of Zero Budget Shares Actual, Estimated and Imputed

	HBS		*SILC*
	Actual (%)	*Estimated (%)*	*Imputed (%)*
Alcohol	29.0	30.7	32.8
Tobacco	66.0	68.0	69.8
Clothing	26.3	29.8	31.5
Health	38.0	44.1	44.4
Public Transport	60.9	63.7	65.4
Private Transport	14.3	13.8	13.3
Education	62.3	65.7	69.5
Durables	10.1	13.7	16.3

Table A5: Frequency of Zero-Expenditures with Standard OLS Estimation for All Goods

	HBS		*SILC*
	Actual (%)	*Estimated (%)*	*Imputed (%)*
Alcohol	29.0	0.7	1.9
Tobacco	66.0	5.2	4.5
Clothing	26.3	0.4	2.8
Health	38.0	2.9	5.8
Public Transport	60.9	4.4	3.0
Private Transport	14.3	0.3	2.7
Education	62.3	10.0	10.3
Durables	10.1	0.2	3.3

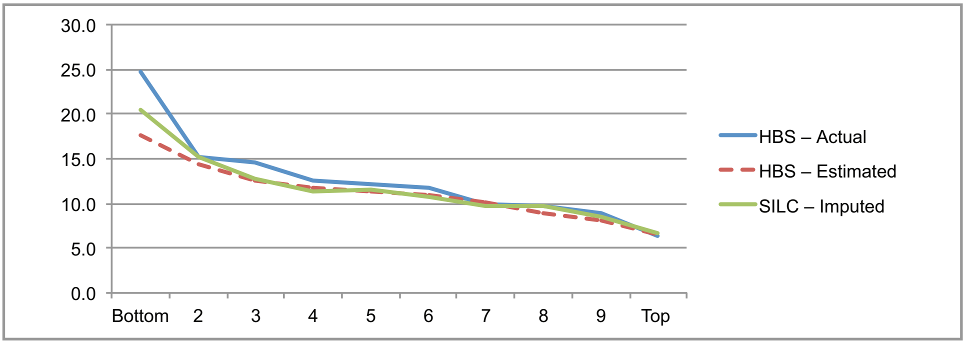
APPENDIX B USING EXPENDITURE TO ASSESS THE DISTRIBUTIONAL IMPACT OF INDIRECT TAXES

Table 2 in the main text compares indirect taxes as a proportion of income across income deciles using the actual, estimated and imputed expenditure data. The distributional impact of indirect taxes can also be measured relative to expenditure. Expenditure can be used as the measure of household resources in two ways here. First, households can be ranked according to total expenditure, so that deciles are calculated on the basis of total equivalised expenditure, rather than equivalised income. Second, indirect taxes can be expressed as a proportion of total expenditure, rather than as a proportion of income. Figure A1 shows that the distributional impact of indirect taxes varies quite substantially depending on whether income or expenditure is used as the measure of resource for the household. Figure A1a reproduces the numbers presented in Table 2, showing that indirect taxes decline as a proportion of income as income rises. Figure A1b presents indirect taxes as a proportion of expenditure in each income decile. The pattern is quite different: the curve becomes flatter, with indirect taxes as a proportion of expenditure in the bottom income decile comparable to deciles in the middle of the income distribution. The top income decile pays the lowest proportion of its expenditure in indirect taxes. Figures A1c and A1d rank households by equivalised expenditure rather than equivalised income. The Figures show that indirect taxes in the bottom expenditure decile represent a relatively low proportion of income, but a relatively high proportion of expenditure. The relationship between income and expenditure (and, indeed, consumption) is complex and it is unclear which is a more suitable measure of "resource" or "welfare" of a household (see Brewer and O'Dea (2012) and O'Donnell *et al*. (2008) for discussion). Resolving this issue is beyond the scope of this analysis – here we simply aim to show that the imputation approach produces a representative measure of expenditure, across both income and expenditure distributions.

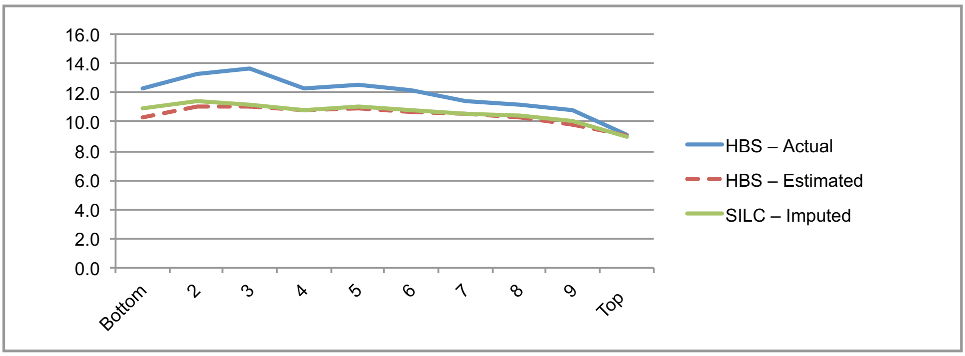
In each subfigure, the distribution of indirect taxes is shown based on the actual, estimated and imputed expenditure data, as in Table 2 in the main analysis. When indirect taxes are expressed as a proportion of income, the results are quite insensitive to the choice of data source. When indirect taxes are expressed a proportion of expenditure, indirect taxes are slightly higher (between 1 and 2 per cent) for low and middle deciles when based on actual expenditure data than on estimated or imputed expenditure data. This is due to the presence of a small number of households in the actual HBS data reporting relatively large budget shares for high-tax goods such as alcohol and tobacco (budget shares up to 60 per cent in both cases – particularly in the lower deciles), which are not captured in the modelling procedure.

Figure A1: Distribution of Indirect Taxes, 2010

(a) Indirect Taxes as % of Income, by Decile of Equivalised Disposable Income



(b) Indirect Taxes as % of Expenditure, by Decile of Equivalised Disposable Income



(c) Indirect Taxes as % of Income, by Decile of Equivalised Expenditure

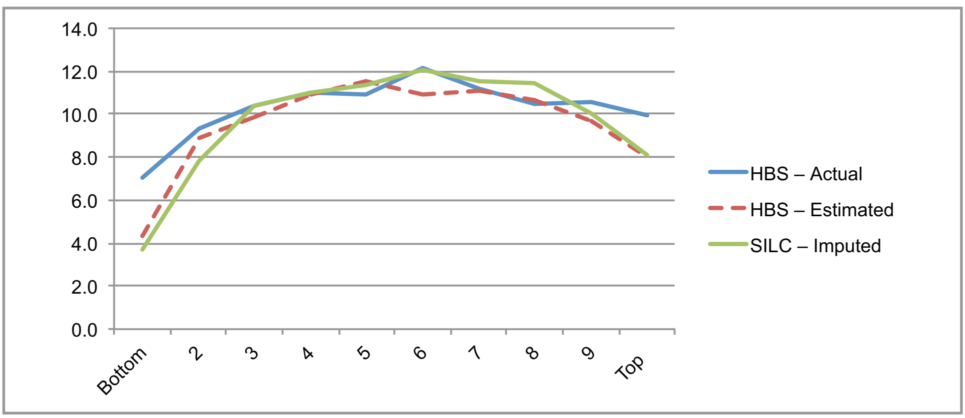


Figure A1: Distribution of Indirect Taxes, 2010 (contd.)

(d) Indirect Taxes as % of Expenditure, by Decile of Equivalised Expenditure