



Banc Ceannais na hÉireann
Central Bank of Ireland

Eurosystem

Research Technical Paper

Macro and Micro Estimates of Household Wealth

Mary Cussen, Reamonn Lydon & Cormac O'Sullivan

Vol. 2018, No. 11

Macro and Micro Estimates of Irish Household Wealth

Mary Cussen*

Reamonn Lydon†

Cormac O'Sullivan‡

Abstract

This paper compares estimates of the financial wealth of Irish households from the Irish Household Finance and Consumption Survey (HFCS) with estimates from statistical and administrative data sources. This cross-checking exercise yields three key results. First, like most wealth surveys in other countries, the Irish HFCS significantly under-records the total value of deposits held by Irish households: we estimate the HFCS captures around one-third of the aggregate figure. Second, this level of under-recording is broadly similar across the distribution of deposits. Thus, the HFCS appears to provide an accurate picture of the *distribution* of household deposits. Third, the degree of under-recording does not appear to be heavily concentrated in any one population group, i.e. by region or age group. With regard to the reason for the under-recording, we find no one single factor behind the under-recording. Issues identified in the survey design literature around sample design (sampling the wealthy), item and unit non-response all appear to play some role.

JEL classification: D12, E21 . *Keywords:* Assets, deposits, balance sheet, households, wealth.

*Corresponding author, Central Bank of Ireland, Statistics Division; mary.cussen@centralbank.ie

†Central Bank of Ireland, Economics and Statistics Directorate; reamonn.lydon@centralbank.ie

‡Central Bank of Ireland, Statistics Division; cormac.osullivan@centralbank.ie. We are grateful to Joe McNeill, Tara McIndoe-Calder, Juha Honkkila, seminar participants at the Central Bank of Ireland and participants at both the June 2017 HFCN meeting in Cyprus, the January 2018 Expert Group on Linking Micro and Macro Statistics for the Household at the ECB and participants at the 2018 Irish Economics Association Annual Conference. Sector for comments on an earlier version. We would like to thank Paul Crowley, Gerry Reilly and Edel Flannery (CSO) for comments. The CSO does not take any responsibility for the views expressed or outputs generated from this research. The views in this paper are those of the authors only and not the Central Bank of Ireland or the ESCB.

1 Introduction

The last fifteen years have seen a resurgence of interest in household wealth: how it is measured and distributed, why it changes over time, and why it matters for understanding how households respond to economic shocks. For many countries, including Ireland, the financial crisis, and the depth of the ensuing recession, can, in large part, be attributed to the real estate asset and debt bubble that had built-up in earlier years. As well as aggregate information on household balance sheets, such as the [Quarterly Financial Accounts](#), household-level wealth data facilitates research at the micro-level on topics such as household asset accumulation, portfolio choice, wealth distribution and financial stability. In 2015, the results of the first Irish Household Finance and Consumption Survey (HFCS) were published (see CSO 2015 and Lawless et al. 2015). This survey contains data on wealth, debt, income and consumption for a representative sample of almost 5,500 households. A potential issue with the HFCS however, is that, like all surveys, it may contain measurement errors which could materially distort the interpretation of its results. Measurement errors in wealth surveys largely stem from the refusal by households to participate in the survey, by households providing inaccurate replies to interviews and failure to adequately sample the entire population, in particular the wealthy.

There are a number of existing studies that compare estimates of household wealth across different data sources. Alvaredo et al. (2016) discuss the strengths and shortcomings of different data sources in the UK context, including Financial Accounts, revenue data, household wealth surveys and data from a 'Rich List' of the wealthiest families. They conclude that due to the issues of non-response and under-reporting at the top mean, that the wealth survey data cannot, on its own, provide a fully satisfactory representation of the upper tail of the UK wealth distribution. Bricker et al. (2016) explore issues that arise when comparing estimates of household wealth from aggregate, survey, and tax sources using US data. They examine variations in concepts, definitions of population, coverage, and units of analysis. Looking at European estimates, Honkkila and Kavonius (2013) discuss errors in estimation and measurement that may drive differences between micro and macro data. This article adds to the existing literature on measurement issues in household wealth surveys, both by presenting evidence for Ireland, and by using an administrative dataset of deposits to identify potential pockets

of miss-measurement and validate the distribution of deposit wealth.

This article cross-checks the assets and liabilities results from the HFCS with aggregate and granular-level data sources to identify whether measurement errors exist. The key analytical contribution is an in-depth analysis of deposit assets – by which we mean current and savings accounts – which are by far the largest category of financial asset held by most households. The results for deposit assets from the HFCS are compared to a micro-level administrative dataset which contains information collected from Irish banks on the holdings of deposits by households. This allows us to assess whether measurement issues for household deposits are larger for particular household cohorts. Our comparison of the HFCS deposits with administrative data leads us to tentatively conclude that, despite under-recording of deposits *in aggregate*, the *distribution* of these deposits across households, as recorded in the HFCS survey, is broadly representative of the estimated population distribution from the administrative data. Furthermore, we use survey meta-data to analyse the factors which could cause measurement errors, for example distinguishing between survey and item non-response. In addition to providing support for the survey as a reliable source for understanding the distribution of wealth, the results of our analysis can contribute to potential improvements in survey methodology for future HFCS waves to reduce potential measurement errors.

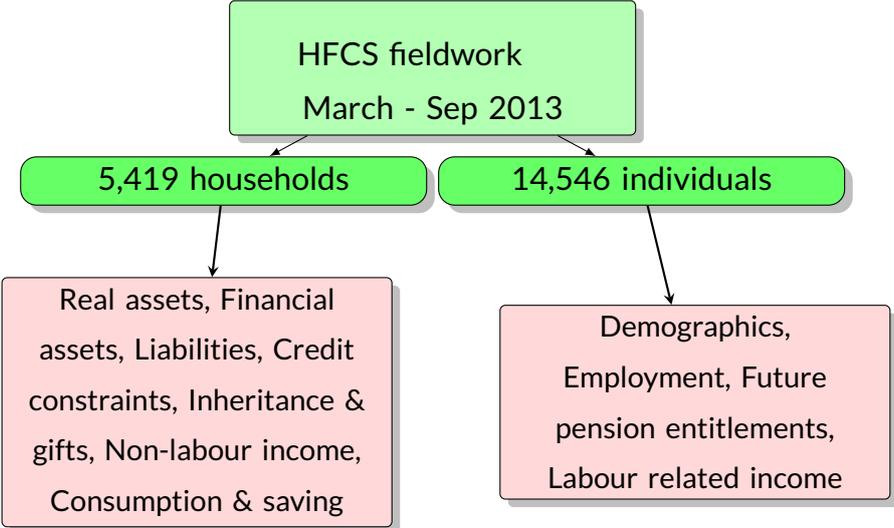
The remainder of this paper proceeds as follows. Section 2 gives some background on the Irish HFCS and summarises recent research based on the findings of the survey. Section 3 compares aggregate balance sheet data in the HFCS with Central Bank *Money and Banking Statistics* and the *Quarterly Financial Accounts*. Section 4 compares the distribution of deposit assets with administrative micro data collected by the Central Bank of Ireland. Section 5 looks at technical issues regarding survey design, such as weighting and respondents' ability to recall information with ease or accuracy. Section 6 concludes.

2 Background on HFCS

The Irish HFCS was carried out by the Central Statistics Office (CSO) between March and September 2013. Figure 1 provides an overview of the survey structure and size: 5,419 households, comprising 14,546 individuals – just under 11,000 of whom are aged 16 or older. The *household* module in the survey contains information on assets,

debts (linked to assets), inheritances, non-labour income and consumption; the *personal* module covers demographics, employment characteristics, work history, pensions and labour-related income. Compared to existing CSO household surveys which cover income (*Survey on Income and Living Conditions*) and employment (*Labour Force Survey*), the major innovation in the HFCS is the collection of data on gross wealth and debt. CSO (2015) and Lawless et al. (2015) provide detailed overviews of Irish household debt and wealth across different cohorts of the population. They find considerable distributional differences between groups of households, with the top 20 per cent of households holding 70 per cent of net assets. In addition, more than half of households hold some form of debt, with mortgage indebtedness proving to be a particularly heavy burden for younger cohorts. The survey has also been used for researching Irish households resilience to shocks. Fasianos et al. (2017) used the survey to examine debt and to model the potential impact of interest rates changes across the life-cycle. Lydon and McIndoe-Calder (2017) simulate the impact of the financial crisis across different household types. Le Blanc and Lydon (2016) combine the HFCS with repeated waves of the *Household Budget Survey* to analyse how households with different levels of indebtedness respond to income shocks. Le Blanc (2016) uses the HFCS to compare the savings levels and motives of Irish households with similar households in the Euro area.

FIGURE 1. HFCS structure



Internationally, similar types of household surveys have been used to examine issues impacting economic performance and financial stability. For example, Carroll et al.

(2014) find that different rates of marginal propensity to consume across countries may be explained by differing wealth distributions between countries. Disney et al. (2016) examine labour supply responses to house prices changes for the UK. Ampudia and Ehrmann (2014) analyse the extent to which euro area households' experiences influence their willingness to take financial risks.

Clearly, the HFCS is an extremely valuable resource for policy-makers and economists. However, accurate measurement of the components of the survey is vital for the overall credibility of the results. As such, we attempt to validate both the total amount of wealth and debt and the imputed distributions, comparing the results with aggregate data on household wealth and Central Bank administrative data.

3 Comparing the HFCS and Quarterly Financial Accounts

Quarterly Financial Accounts (QFA) present a complete and consistent set of quarterly balance sheet and financial transactions data for all sectors of the Irish economy, including households. The QFA is part of the integrated system of national accounts. The HFCS and QFA are largely comparable for most major household assets and liabilities. This section presents a comparison between the two datasets for Ireland. The paper then contrasts this comparison with other euro area countries for which these data are available.

In comparing the QFA and HFCS, it is important to note differences between the definition of households in each set of datasets. Households surveyed for the HFCS were only comprised of private households. QFA results also include non-profit institutions serving households (NPISH), non-incorporated enterprises, sole-traders and partnerships which are not large enough to be considered quasi-corporations and institutionalized households. Using the Central Bank's private households series, QFA results can be adjusted to only include private households and institutionalized households. This series will be referred to in the paper as 'QFA adjusted'. Institutionalised households largely consist of nursing homes, prisons and long-term medical facilities. While the inclusion of these institutionalised households inflates the QFA relative to the HFCS, the impact is not likely to be substantial.

Table 1 presents a comparison of all significant household assets and liabilities. The figures for QFA adjusted are from Q2 2013. Unquoted equity, insurance technical reserves and other assets/liabilities are excluded from the table as the level of comparability of these instruments between the HFCS and the QFA is low. Honkkila and Kavonius (2013) outline the definitional differences between items reported in the HFCS and financial accounts.

Some substantial differences are observed between results reported by the HFCS and by the QFA adjusted (Table 1). By far the largest reported difference between the two datasets is for household deposits. These assets are nearly 3.5 times larger in QFA adjusted compared to what is reported in the HFCS. Loans reported in QFA are roughly 1.3 times larger than those reported by HFCS. Quoted shares are also somewhat smaller in the HFCS, though this may in part reflect difficulties faced by households in accurately valuing these assets at market value.

TABLE 1. Comparison between HFCS results and QFA results

HFCS	€bn	QFA adjusted* - Q2 2013	€bn	Difference €bn
<u>Assets:</u>				
Sight and savings deposits	33.3	Deposits (excl. currency)	112.0	78.7
Bonds	1.6	Bonds**	0.5	-1.1
Quoted shares	6.3	Quoted shares	10.7	4.4
Mutual funds	3.7	Mutual funds**	0.5	-3.2
<u>Liabilities:</u>				
Loans	119.8	Loans	161.0	41.2

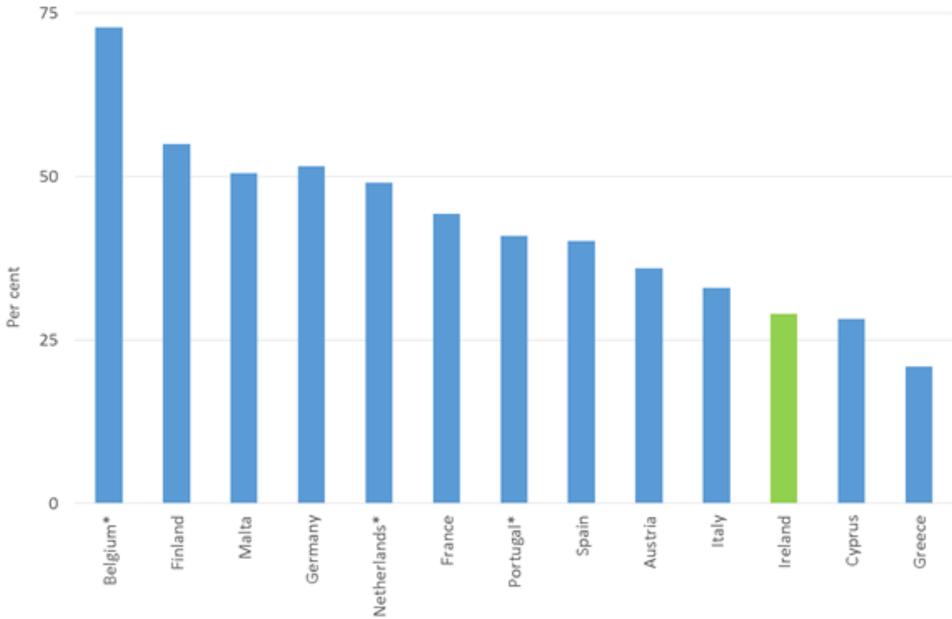
Sources: HFCS and internal estimates.

*adjusted to exclude NPISH. **based on Q4 2013 due to the introduction in that quarter of enhanced data sources for securities holdings.

Under-reporting of deposits in the HFCS is an issue for many countries. Figure 2 compares the ratio of household deposits reported in the HFCS to the estimates from the QFA across those countries that participated in wave 1 of the HFCS. Wave 1 was carried out by several euro area countries throughout 2010. The Irish results from wave 2 are included for comparison. The figure shows that there were significant differences between the two data sets in nearly all of the countries examined. In all cases the HFCS

results were lower.¹ Ireland had the third lowest HFCS as a proportion of financial accounts ratio, at 30 per cent. While under-reporting of deposits in the HFCS is clearly an issue in all countries that have undertaken the survey, the comparison with macro data reveals little regarding the nature of this under-reporting. To do so, Section 4 compares the distribution of deposits in the HFCS with an administrative dataset collected by the Central Bank of Ireland.

FIGURE 2. Deposits reported in the HFCS as a proportion of household deposits results from financial accounts –Cross country comparison



Source: Honkkila and Kavonius (2013), Andreasch and Linder (2014), ECB and Eurostat
 Note: For Austria, life insurance is subtracted from deposits in the HFCS to ensure comparability with financial accounts (see Andreasch and Lindner (2014)). For Belgium and Portugal, the comparison is between the HFCS and the sum of households and NPISH from financial accounts, as data for just households was not available.

4 Comparing HFCS and the Deposit Level dataset

In order to further understand the under-reporting of deposits, this section compares the HFCS results for households’ holdings of deposits with an administrative deposit-level dataset (DLD). As the DLD is reported on a deposit-by-deposit level basis, we can

¹Wave 2 of the HFCS includes Hungary, Poland and all euro area countries with the exception of Lithuania. Most countries carried out the wave 2 survey during 2014. Comparisons between wave 2 results and financial accounts are broadly similar to those revealed in Figure 2.

go beyond the comparison of aggregate totals to a comparison of the distribution of households' holdings of deposits. Information on size of deposit holdings by age and by region of residence of deposit holders also enables us to identify potential areas of under-reporting by population cohorts in the survey data.

The DLD contains records of 7 million deposit accounts held with the largest banks in the Irish market. These banks comprise 80 per cent of households' deposits with resident banks and 50 per cent of total deposits held by Irish households². As such, the DLD is a sub-sample of the entire population of household deposits. Total household deposits in the DLD amounted to €58.9 billion from which the DLD was collected, compared to a figure of €59.4 billion reported in the *Money and Banking Statistics* for the same group of institutions, a difference of less than 1 per cent. *Money and Banking Statistics* are the source of household deposits held with Irish monetary financial institutions in the QFA. The DLD was collected in June 2014, compared to the HFCS May-September 2013 survey period.³ It should be noted that the DLD cannot be assumed to be free from errors, so discrepancies between the DLD and the HFCS may not be entirely ascribed to problems with the survey estimates. Furthermore, there is no unique identifying information that would allow us to directly compare the survey results for a household with actual deposit data from the DLD, so that one-for-one matching is not possible.

An important issue to note is the difference in the basic unit of analysis between these data sources. While, the basic unit of the HFCS dataset is the household, the DLD is based on individual accounts. Using internal bank customer identifiers, we can aggregate multiple accounts held by the same customer with any single bank. We are, however, unable to link an individual's accounts across multiple banks, nor combine the accounts of multiple members of the same household. An implication of using customer accounts rather than households is that average deposits would be expected to be lower, even if the datasets were otherwise identical. We do not know how many Irish people have multiple accounts across credit institutions. Andreasch and Lindner (2014) found in the case of Austria, 91 per cent of Austrian households use only a single bank and only 2 per cent of households have accounts with more than two different banks.

²Irish households also hold deposits in other credit institutions, credit unions, government savings accounts and non-resident institutions. There is unfortunately no distributional data related to these holdings, and it is unknown how representative of total household deposits the distribution of the DLD is. With a view to filling this data-gap, for the 2018 wave of the the Irish HFCS additional questions have been added on the institutional make-up of household savings and deposits.

³We also compared the HFCS with a mid-2013 drop of the DLD, the results are very similar to those presented here. The main reason we use the mid-2014 DLD drop is because some of the key fields, including age, region and borrower ID are better populated in the later drop.

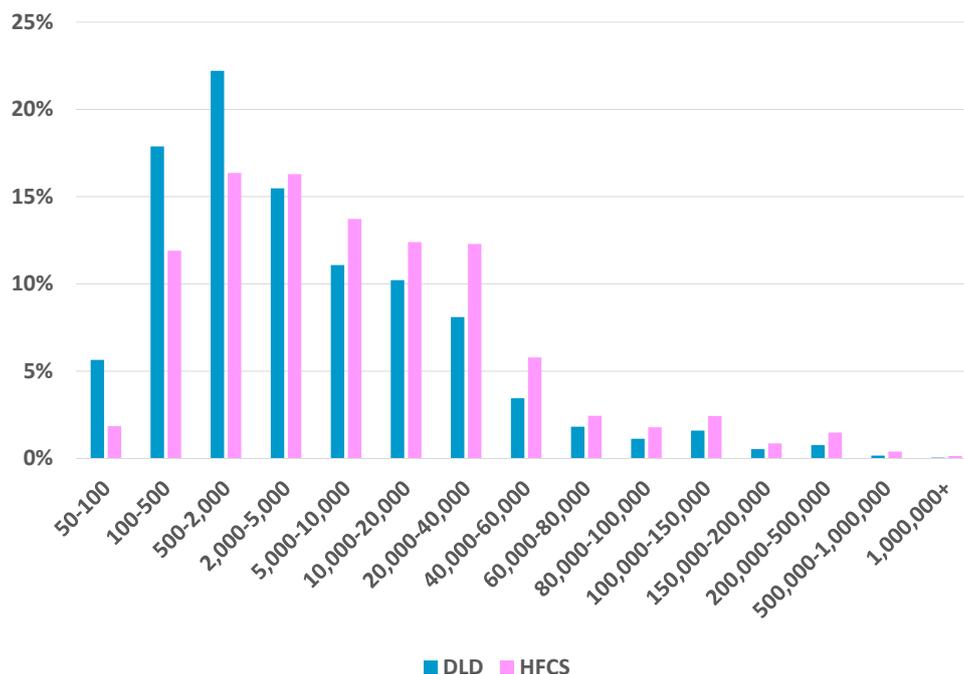
The deposit protection cap may mean that individuals with large deposits are more likely to hold more than one account. Given that most Irish institutions charge not insubstantial quarterly fees on a *per account* and *per card* basis however, the incentive to hold multiple deposit accounts across different institutions is reduced.⁴ A further issue with comparing the HFCS and DLD is that the DLD contains many accounts with very small amounts in them. These may be accounts that are dormant.

Comparison of the distribution of deposits

As an initial step, we divide households in both datasets into segments based on the size of their account holdings. For the purposes of this comparison, we exclude households from the HFCS that report that they do not own a deposit account. We also exclude those who hold deposits of less than €50. These deposit holders comprise 25.9 per cent all deposit holders in the DLD, compared to just 8.8 per cent of the HFCS. This is likely due to the nature of the unit of account; it would be expected that many of these smaller customer deposit holdings in the DLD would be aggregated together with other deposits to form household deposit holdings. As such, customer accounts in this deposit range are unlikely to be directly comparable to households reporting very low deposits in the HFCS, so we exclude from both datasets those who hold deposits of less than €50 in order to better compare the remaining distributions. The proportion of households in each segment is shown in Figure 3 for both data sets.

⁴According to bonkers.ie current account fees range from €3 to €6 *per month* in Irish institutions (as at 14 May 2018).

FIGURE 3. Distribution of Deposits by Size of Holdings (€)



Even with the exclusion of deposit holdings of less than €50, the proportion of those with smaller holdings is somewhat larger in the DLD than in the HFCS. This is an anticipated finding, as the deposit holdings of a household may include multiple customer accounts. Correspondingly, a higher proportion of households holding larger amounts of deposits is recorded in the HFCS than in the DLD, particularly in the segments between €2,000 and €60,000, albeit with less pronounced differences for any given segment. Some 55 percent of surveyed households lie in this deposit range, compared to 36 percent of deposit holders in the DLD.

It is noteworthy that differences between the two datasets are less pronounced among the segments of very wealthy deposit holders; 4 percent of deposit holders in the DLD have deposits holdings of €60,000 or more, compared to 9 percent of households in the HFCS. This is important as difficulty in sampling the rich is often cited as a potential weaknesses of wealth distributions based on household survey data. Based on this comparison with an administrative dataset, however, the Irish HFCS appears to have been successful in capturing some households with very large deposit holdings.

The distribution of deposits for both sets of data are depicted using Lorenz curves in Figure 4, showing cumulative deposit shares in ascending order of deposit holdings. The greater the degree of inequality in the distribution of the underlying data, the farther

the Lorenz curve will lie from the 45° ray. The Lorenz curve derived from the DLD lies below the curve derived from the HFCS, indicating that the distribution of deposits in the former is more unequal. The Gini Coefficient⁵ for HFCS deposits above €50 is 77.7, while the DLD Gini Coefficient for deposits is 80.5.

The Gini Coefficient measures the degree of inequality in a distribution, but does not reveal what part of the distribution is driving inequality. To do this, we compare the average deposit holdings of those at the top of the distribution with those in the middle and the bottom using an inter-quintile share analysis. Excluding those holding deposits of €50 or less, the average deposit holdings of the top 20 per cent of households in the HFCS is 318 times the size of the average deposit holdings of the bottom 20 per cent, while in the DLD a factor of 391 is observed. These figures can be further broken down into ratios between the top 20 per cent and the middle 20 per cent, and between the middle and the bottom 20 per cent. While the ratio of the top to the middle is greater in the DLD than in the HFCS (a factor of 25 as compared to a factor of 18), the situation is reversed when looking at the ratio of the middle to the bottom, with the difference being slightly larger for the HFCS than for the DLD (a factor of 17 in the HFCS compared to 15 in the DLD).

⁵The Gini Coefficient is a numerical measure of distributional inequality, calculated as the ratio of the area between a Lorenz curve and the 45° ray to the entire area beneath the 45° ray. The Gini Coefficient takes a value between zero and 100, where zero indicates perfect equality of deposit holdings across the distribution and 100 indicates that one household or customer holds all deposits.

FIGURE 4. Distribution of Deposits in the HFCS and DLD, Including and Excluding Deposits Less Than €50

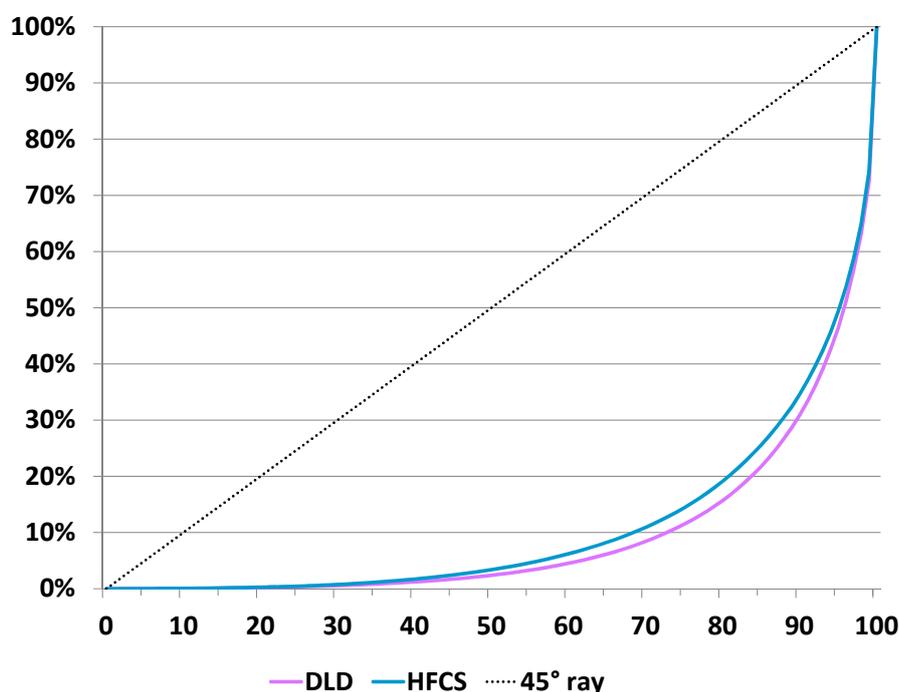


TABLE 2. Quintile Share Ratios

	top/bottom	top/middle	middle/bottom
HFCS	318	18	17
DLD	391	25	15

Although the two distributions are not identical, there is no evidence that the distribution of deposits in the HFCS is materially distorted by under-reporting. Given the similarity between the two distributions, we conclude that there is no definitive prima facie evidence from the DLD that the distribution of deposits as described by the HFCS is not representative of the true distribution of household deposits.

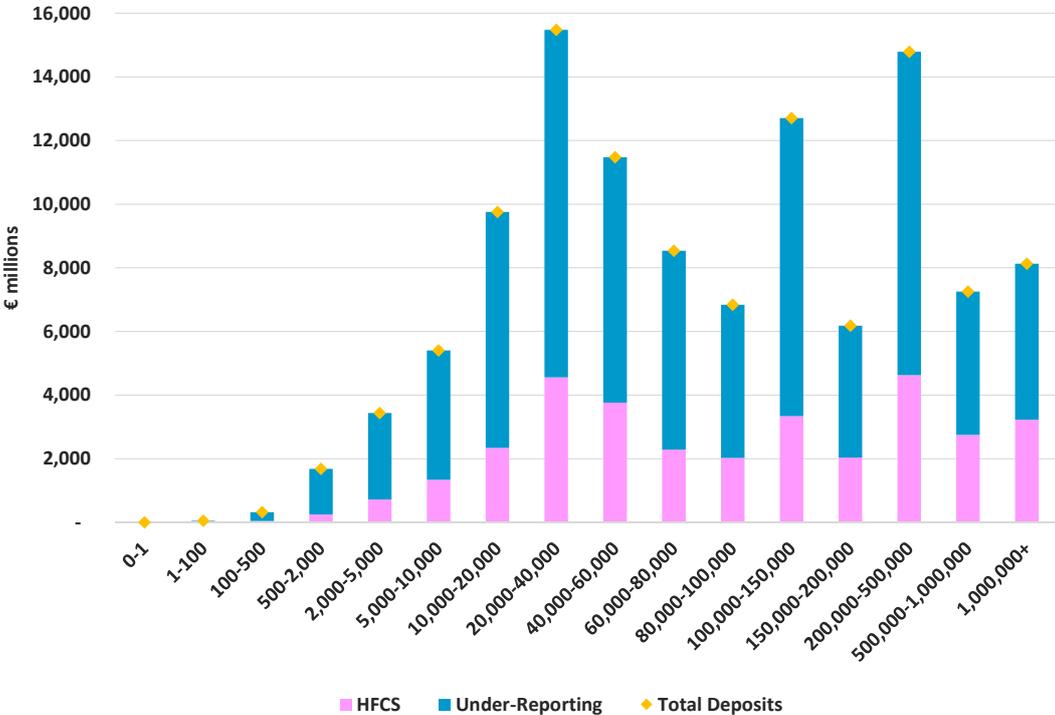
Under-reporting by cohort of population

In addition to comparing the distribution of deposits, we can use the DLD to identify patterns of under-reporting in the HFCS. As outlined in Section 3, aggregated HFCS deposits were 30 percent the size of QFA deposits. Using the DLD, we go beyond this

aggregate finding by comparing deposits held by different household groups.

We look first at patterns of under-reporting by size of deposit assets. As shown in Figure 3, there is a higher proportion of customers in the DLD in lower deposit categories as compared to the HFCS. However, when aggregated together, cumulative deposits from these deposit categories are relatively very small and so cannot be the explanation for the under-reporting of deposit wealth in the HFCS relative to the Financial Accounts. Even though households with deposit holdings of less than €5,000 make up 53 per cent of all households with deposits in the HFCS, it can be seen in Figure 5 that the amount of under-reporting in deposit categories of less than €5,000 is relatively small.

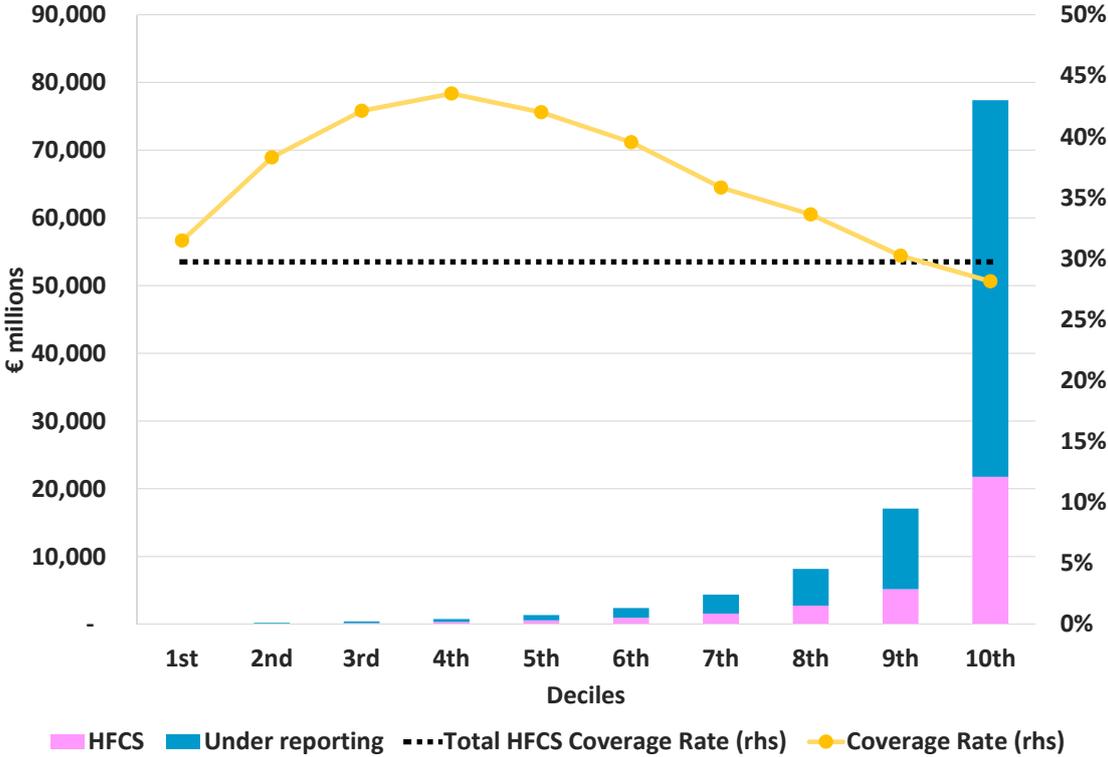
FIGURE 5. Comparison of HFCS and DLD by Deposit Size



A much more significant proportion of under-reporting is accounted for by households at the top of the distribution. This can be seen in Figure 6 in which we divide households in both datasets into deciles based on the size of their deposit holdings and then sum all deposits within each decile. As previously mentioned, the DLD includes a large number of small deposits, so in order to achieve greater comparability between the two datasets we again exclude deposit holdings of less than €50. As Figure 6 shows, the majority of under-reporting is attributable to the tenth decile, which accounts for 71 per cent of total under-reporting as compared to just 2 per cent for the entire bottom half of the

distribution. Given the high degree of inequality in the distribution of deposits discussed in Section 2, this result is perhaps unsurprising. Yet it should also be noted that the level of deposits reported in the HFCS relative to the DLD – measured by a per-decile coverage rate – is lowest for the tenth decile. The coverage rate for the tenth decile is 29 per cent, compared to a high of 44 per cent in the fourth decile, and 30 per cent for the HFCS as a whole.

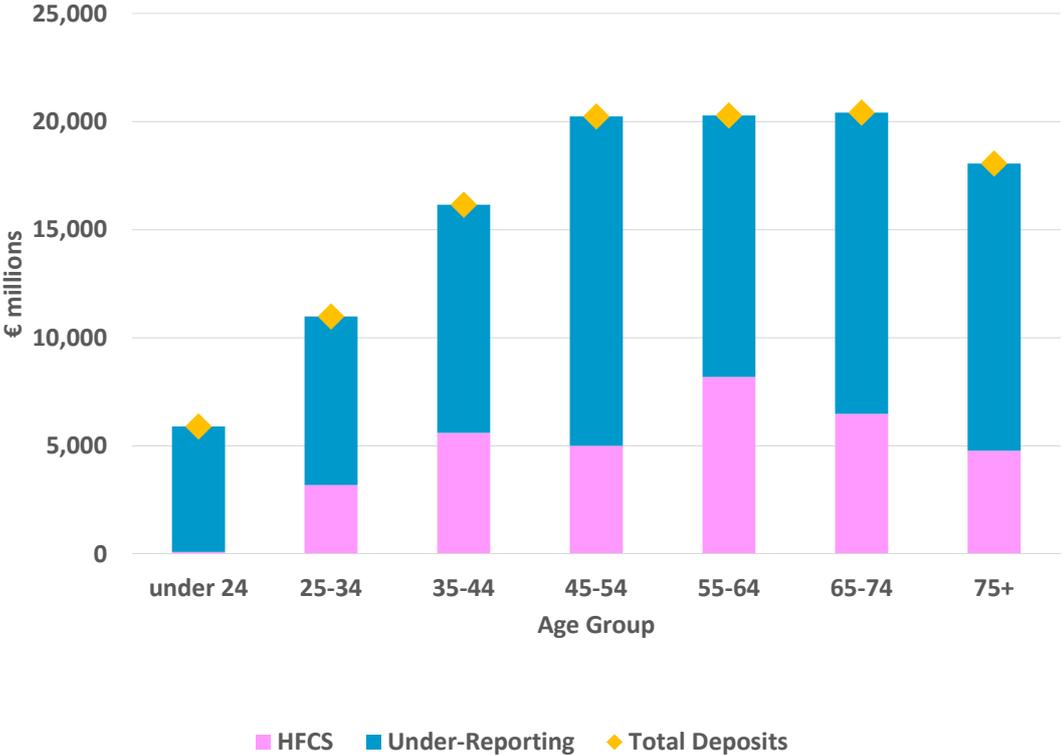
FIGURE 6. Comparison of HFCS and DLD by Deposit Size



The HFCS and the DLD also contain age and geographic information, which allows for analysis of under-reporting by age group and by region. As can be seen in Figure 7, deposits are not as unevenly distributed between age groups as they are between deposit wealth deciles. The age groups 45-54, 55-64, and 65-74 each held deposits of €20 billion according to the DLD. The youngest age group, those aged under 24 years, held the least amount of deposits with holdings of just under €6 billion. There is a noticeable lack of reporting in the HFCS of deposits held by those aged 24 or younger, but this is likely to be largely because many in this age group live in the family home and so their deposits would be amalgamated with those of the rest of the household during the survey interview. For the other age groups, the pattern of reporting in the HFCS is not dissimilar to what is found in the DLD, although reporting by those aged between 45

and 54 stands out as being particularly low. There is some limited support for the findings in the literature expressed in Section 5 of greater under-reporting by the elderly as the 75+ age group had a relatively low coverage rate of 26 per cent.

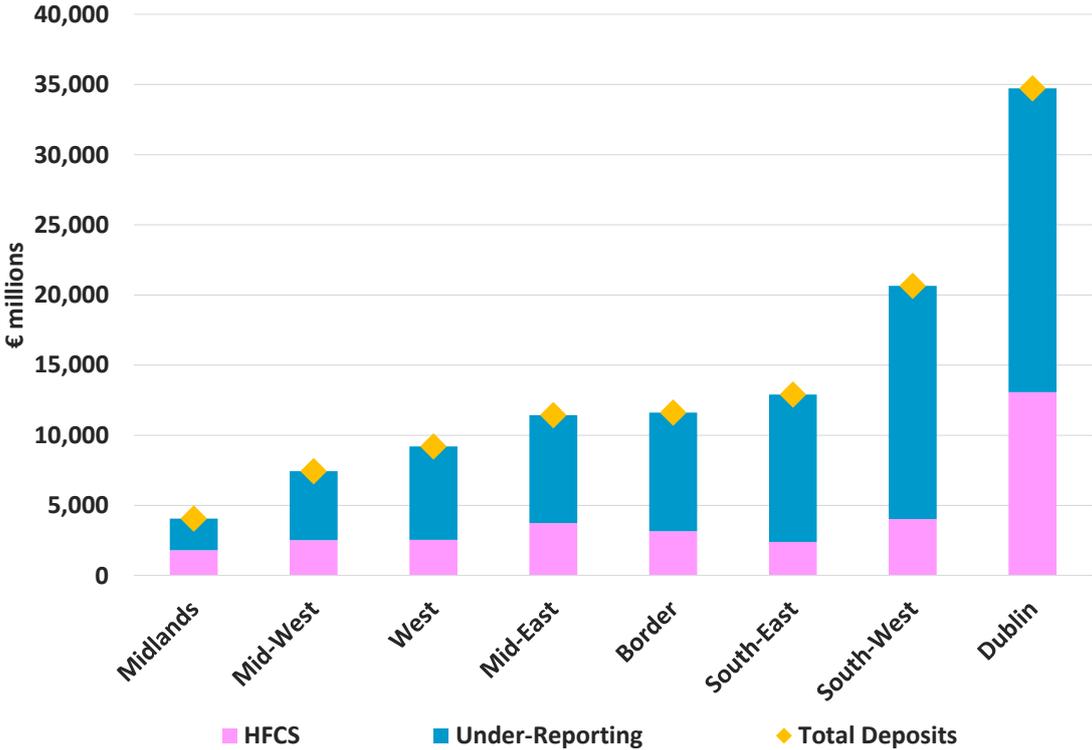
FIGURE 7. Comparison of HFCS and DLD by Age Group



Looking at deposits by region, the distribution is more unequal than by age, as can be seen in Figure 8. The Dublin region contains more deposits than any other region in both datasets, and is also the region with the most under-reporting of deposits. However, the fraction of deposits in the DLD that are reported in the HFCS is higher in Dublin than for the country as a whole, with a coverage rate of 38 per cent. According to the DLD, the South-West (comprising the counties of Cork and Kerry) and the South-East (Carlow, Kilkenny, Tipperary, Waterford and Wexford) have the second and third most deposits respectively. Additionally, these regions have the second and third highest amount of under-reporting in the HFCS. In contrast with the Dublin region, the coverage rates of the South-East and South-West are significantly lower than for the country as a whole, with reported deposits in the HFCS amounting to just 19 per cent of the totals available from the DLD for both regions. The Midlands region (Laois, Longford, Offaly and Westmeath) contained the least amount of deposits in both the HFCS and the DLD, but had the highest coverage rate, with deposits reported in the HFCS amounting to 45

per cent of deposits in the DLD. It is not possible to determine precisely why there are such wide differences in the level of reporting across regions, but a potential explanation is varying levels of trust. This is analysed further in Section 5, in which we consider paradata collected by HFCS interviewers relating to the levels of suspicion observed in interviewees.

FIGURE 8. Comparison of HFCS and DLD by Region



5 Why differences may arise between the HFCS and aggregate data sources

Though measurement issues can arise in both the QFA and the HFCS, generally the QFA results are considered to be more reliable. This is because a number of issues may lead to inaccurate results in surveys, particularly surveys that touch on potentially sensitive topics such as wealth. Firstly, it can often be difficult to contact and to then persuade some households to participate in the survey, particularly those households which hold the most wealth (unit non-response). Secondly, households may be reluctant to discuss the value of their assets and liabilities with an interviewer and may therefore

not respond to some questions (item non-response) or provide inaccurate answers. In addition, some households may find it difficult to fully comprehend the questions they are being asked. Finally, households may find it difficult to accurately value or recall the value of their assets and liabilities. Though it is difficult to ascertain the extent to which these issues may have resulted in under-reporting in the HFCS, aspects of the survey results may provide some indications.

Survey bias due to unit non-response

Overall, the response rate of those surveyed in the Irish HFCS was quite high, at 51.5 per cent. Portugal and Finland were the only countries with higher response rates and, in the case of Portugal, completion of the HFCS by sampled households was compulsory (ECB 2013). However, unit non-response rates are most likely to be the highest for very wealthy households. D'Alessio and Faiella (2002) find that in a survey of clients of a private bank, the response on wealth was 26 per cent for those with net financial wealth up to 20 million lira. However, the response rate fell to 10 per cent for those with wealth of over 1 billion lira. Similarly, Kenneckill (2007b) finds that when a sample of wealthy households was selected from income tax returns for the Survey of Consumer Finances (SCF), the response rate for this group was only 10 per cent. Moreover, wealthy households probably hold a considerable proportion of total household wealth. For example in the US, the wealthiest 1 per cent of households hold 30 per cent of total US household wealth (Kenneckill, 2007b).

In order to reduce the impact of non-response by wealthy households on results in wealth surveys, these households are often over-sampled. This is usually done by identifying them through income tax or wealth tax returns or by assuming they are most likely to live in a particular area. Vermeulen (2017) finds that in countries where the HFCS oversamples wealthy households, there are proportionally more households with net wealth of over €2 million in the collected results. The paper finds that this is particularly the case where oversampling was based on individual wealth tax returns.

In the case of Ireland, an index of affluence/deprivation by small area was used by the CSO to identify potentially wealthy households. The index was compiled based on information from the 2011 Census on 15 attributes relating to demographic profile, social class composition and labour market situation. Trutz, Pratschke and Gleeson (2015) provide further details on the methodology. Proportionally more households classified as affluent in the index were selected to be interviewed for the HFCS. Of those selected to be interviewed, 42 per cent were classified between the 5th and 9th decile of affluence/deprivation, while 20 per cent of selected households were in the 10th

decile. Moreover, affluent households were well represented in the final results of the survey. Of those that responded, 19 per cent were in the top income decile and 40 per cent were between the 5th and 9th income deciles. It is interesting to note, however, that success of the affluence/deprivation index in identifying wealthy households was somewhat mixed. Table 3 compares the decile ranking of households by their total assets from the data collected in the HFCS with the affluence/deprivation index ranking. It shows that 11 per cent of households ranked as being in the top decile according to their affluence/deprivation index score actually were amongst the least wealthy 10 per cent of households according to the survey. Moreover, almost two-thirds of them were in the bottom half of the wealth distribution.

TABLE 3. The Affluence/Deprivation ranking of households in the HFCS survey compared to their decile ranking based on reported results

Affluence/ Deprivation Decile	Wealth decile									
	1	2	3	4	5	6	7	8	9	10
1	27	12	8	9	10	10	12	3	4	3
2	15	9	11	13	9	9	9	9	8	6
3	6	8	9	15	15	13	13	6	8	7
4	8	8	7	6	8	10	11	13	15	13
5	8	7	9	9	10	12	12	10	11	11
6	9	6	8	13	11	10	9	9	9	16
7	8	8	9	9	10	12	12	10	11	11
8	6	8	8	5	9	12	10	12	17	15
9	7	13	10	7	4	7	11	14	16	12
10	11	15	16	11	10	8	7	9	6	7

Source: Irish HFCS.

An additional method of identifying the successfulness of wealth surveys in obtaining results from rich households is to measure how many extremely wealthy households responded to the survey. Table 4 contrasts the results of the Irish HFCS with the wave 1 HFCS results from other euro area countries and with the results of the USA’s SCF. The table shows that fewer households with a very high net worth were surveyed in Ireland compared to the other countries where wealthy households were oversampled based on individual’s tax information. This may reflect an issue with non-response from very wealthy households. However, the substantial decline in the value of household assets since the housing bubble burst in 2007 may also be a contributing factor in comparatively

low number of very wealthy households surveyed. By Q2 2013 Irish household net wealth had declined by 35.5 per cent from its peak in Q2 2007.

TABLE 4. High net worth households surveyed as a percentage of total survey and oversampling method used, cross country comparison

	% of sample with net worth over €2m	% of sample with net worth over €1m	% of sample with net worth over €0.5m	Oversampling of wealthy households method used
USA	14.9	19.4	26.1	Based on individual's income tax information
Spain	8.7	18.2	33.7	Based on individual's taxable wealth information
France	4.3	11.4	23.5	Based on individual's taxable wealth information
Belgium	3.1	8.9	25.7	Based on average regional incomes
Germany	2.4	6.9	18.3	Based on taxable income of the region
Austria	2.0	4.7	11.4	Vienna oversampled
Italy	1.0	3.8	13.5	No oversampling
Ireland	0.7	3.6	10.5	Affluence/deprivation index
Finland	0.5	2.7	11.2	Individual income information from register
Portugal	0.5	2.0	5.7	Lisbon and Porto oversampled
Netherlands	0.2	2.4	13.2	No oversampling

Source: Vermeulen 2017.

It would be interesting to identify some additional characteristics of the households which refused to participate in the survey in order to draw further conclusions about whether unit non-response by extremely wealthy households may have impacted the survey results. An approach adopted by D'Alessio and Faiella (2002) was to examine the results of those households who initially refused to provide results or were uncontactable but who later agreed to be interviewed. In the case of the Irish HFCS, there are no recorded instances of households who initially refused to participate but later agreed after they were re-contacted. In addition, interviewers' observations on the 'dwelling rating' where households were not contactable or refused to reply were not recorded. These may have allowed users to make inferences about whether these households may have had extremely high net worth.

TABLE 5. Imputations for sight and savings deposit accounts in the HFCS, cross country comparison

	Has sight deposits and reported amount	Has sight deposits and amount imputed	Has savings account and reported amount	Has savings account and amount imputed
France	98.1	0.0	87.3	0.1
Italy	92.7	0.0	74.1	0.2
Luxembourg	96.8	0.0	74.1	0.2
Portugal	92.7	0.0	42.6	0.2
Slovakia	92.4	0.0	26.4	0.2
Germany	99.2	0.0	81.2	0.3
Spain	96.6	0.3	30.9	0.1
Austria	98.6	0.3	85.7	1.6
Greece	72.3	0.6	3.6	0.1
Malta	71.2	1.8	78.4	5.3
Netherlands	70.7	2.0	85.3	1.6
Belgium	93.7	3.7	78.1	0.6
Slovenia	86.3	7.0	29.7	0.3
Ireland	89.4	8.0	56.3	12.6
Cyprus	69.8	8.6	36.5	0.2
Finland	0.0	100.0	0.0	100.0

Sources: Irish HFCS and HFCS country surveys metadata information.

Survey bias due to item non-response

Where item non-response occurs in the HFCS, a value is assigned for the item. In the case of Ireland, the approach taken is to assign a value for the item which matches values for similar households. In the Irish HFCS, imputations for sight deposit and savings deposits were 8 per cent and 12.6 per cent, respectively (Table 5). Compared to other countries, imputations for sight deposits were the third highest. Irish imputations to savings accounts were the second highest.

In the case of the Irish HFCS, there was a correlation between the proportion of imputations made and the age category of the interviewee (Figure 9). This may be partly because of increased suspicion about the motivation behind the interview or increased recall issues amongst older interviewees. Interviewers reported that both interviewee suspicion and inaccuracy of answers increased as the age category of the reference person increased. As the value of deposits generally increases with age (Figure 10), item non-reporting by these household may be problematic. This may be why when savings were compared in the DLD and HFCS in Section 4, we found significant under-reporting for older households. Future surveys could consider if there may be additional ways to

further reassure older interviewees about the motivation of the interview and to further encourage them to consult bank statements.

FIGURE 9. Imputations made decomposed by age category

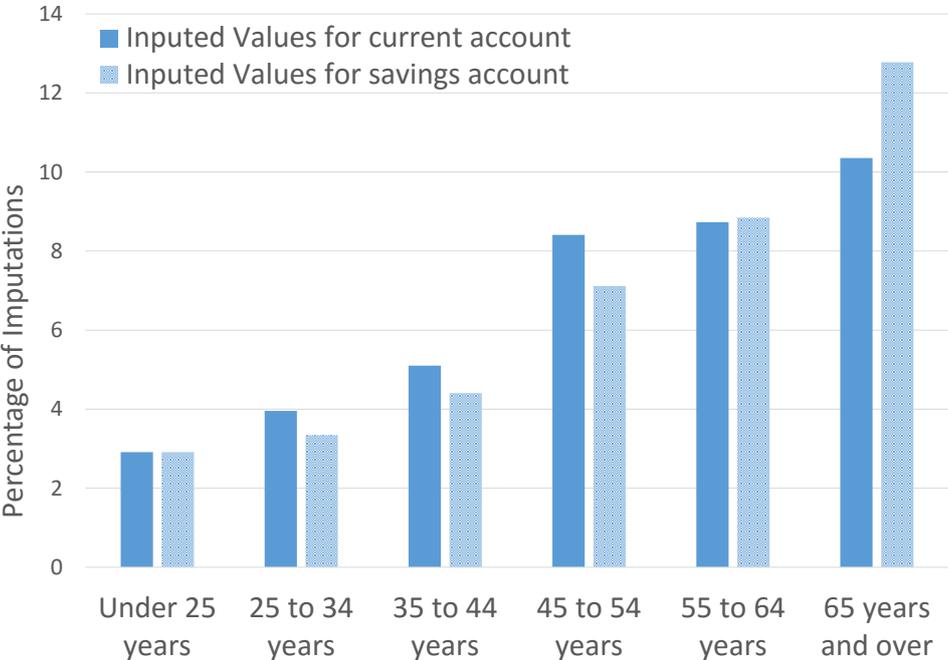
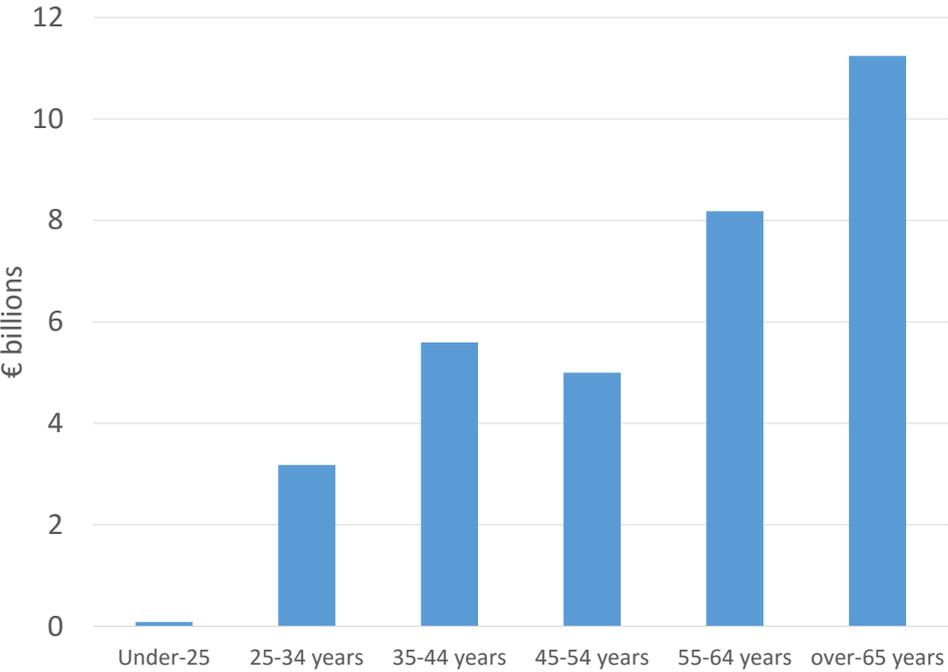


FIGURE 10. Total deposits by age category



Source: Irish HFCS

Survey bias due to mis-reporting by participants and comprehension difficulties

One possible way of improving the survey results could be by excluding the responses of households who may have provided inaccurate responses. In the HFCS, interviewees were asked how accurately they keep track of their bank account balances and cash. In addition, interviewers also recorded information on whether households consulted bank account records during the interview. By eliminating from the sample households who did not consult bank accounts records and who also said that they did not keep track of how their cash bank account balances, this paper explores if the HFCS results became closer to the QFA results. The results were reweighted once these households were excluded from the sample.

Table 6 shows that when only households who knew exactly how much was in their bank accounts/cash balances after withdrawals or who consulted bank statements were included in the sample, the re-weighted amount recorded for total household deposit accounts in the HFCS fell from €33.3bn to €28.5bn. The decline in the amount is because households who did not accurately keep track of the balances had relatively higher bank account balances than those who did. When households who said they roughly kept track of how much they had in bank accounts/ cash were also included in the sample, the total for household deposits accounts increased to €33.5bn. The reweighted totals for household debt liabilities and quoted shares assets were not significantly impacted by either change in sample. Therefore even after excluding households who may have not provided accurate responses, the significant differences between the QFA and HFCS presented in Table 1 remain.

TABLE 6. Weighted household responses once households who may have reported inaccurate results are excluded

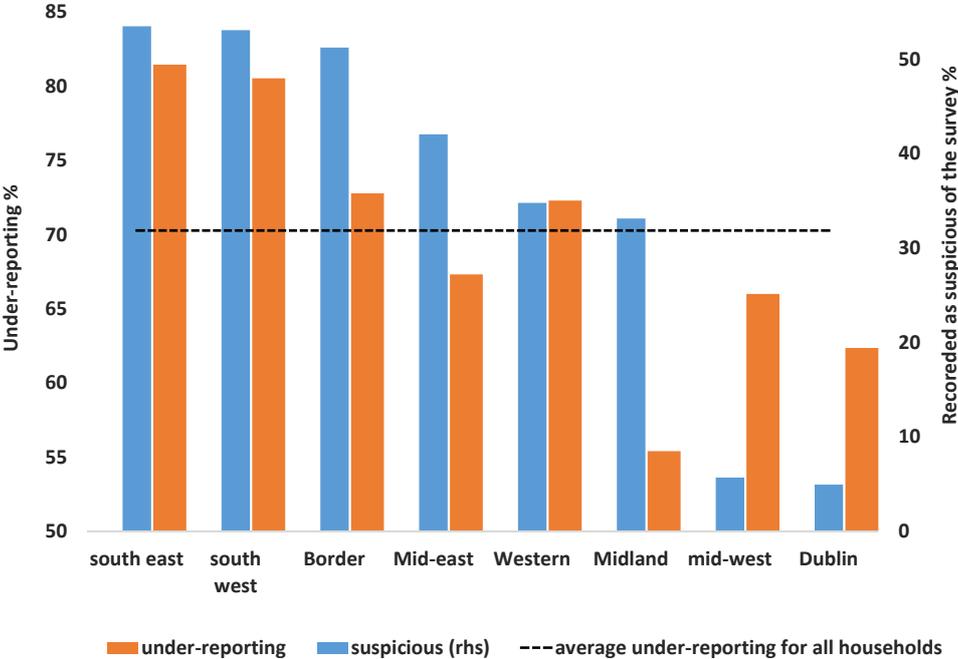
	Total sample of households	Household knows exactly how much is in accounts/cash after withdrawing or spending money or has consulted statements	Household knows exactly or roughly how much is in cash/accounts after withdrawing or spending money or has consulted statements
Number of households	5419	3193	5220
Reported sight and savings account balances	€33.33m	€28.54m	€33.49m
Reported quoted shares	€6.3m	€5.9m	€6.3m
Reported debt	€119.8m	€122.0m	€120.4m

Source: Irish HFCS.

Information recorded on interviewers' observations during the interview could help to assess the extent to which mistrust of the survey and comprehension issues may have been a problem. Kennickill (2006, 2007a) emphasises the extent to which paradata collected from interviews can be useful. It is important to note however that paradata reflect the observations of interviewers and can therefore be subjective. Figure 11 compares the under-reporting by region with the paradata collected by region. The paradata shows the proportion of interviewees in each region where the interviewer recorded that they were 'to some extent' or 'absolutely' suspicious of the survey. The figures find a correlation between regions when under-reporting was the largest and where interviewers reported the highest level of suspicion. This suggests that there may be a link between under-reporting and interviewee suspicion. Figures 12 - 14 contrast the paradata collected by the interviewers conducting the Irish HFCS with those in other countries. Figures 12 and 13 show that the percentage of interviewees' who were perceived to be suspicious of the survey and their apparent comprehension of the survey questions were largely in line with those recorded by interviewers in other countries. Interviewers' observations of the reliability of responses were the second highest for Ireland, after Italy (Figure 14). However at just 3.2 per cent it is not high enough to materially impact the quality of the results or to explain the level of under-reporting in deposits.

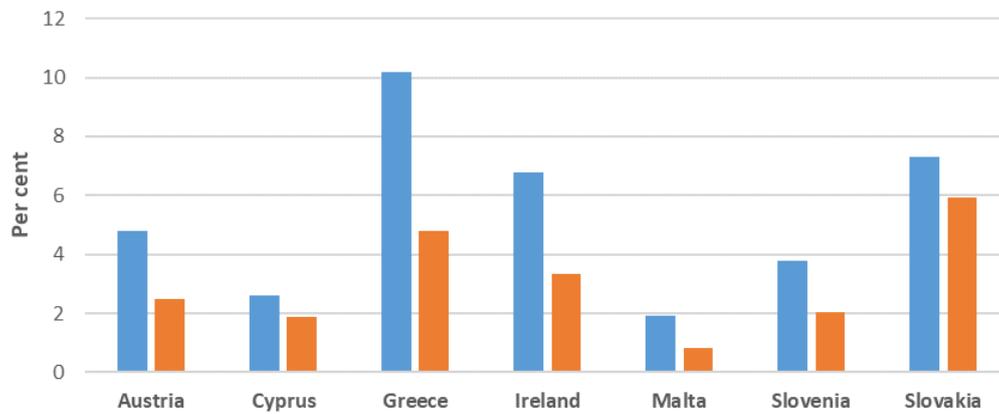
An important factor in reducing comprehension issues is identifying possible changes to questions which could reduce ambiguity. Country specific questions could be important in this regard. For example in the case of Ireland, households held €18 billion of deposits in government accounts at Q2 2013. This would include accounts such as 'prize bonds', which should be classified as deposits but the name of the account may lead households to include it under securities or other assets. Asking households specifically about deposits held with a Post Office (including prize bond accounts) could assist households with recalling in these types of deposit accounts and ensure they are recorded consistently as deposits across households.

FIGURE 11. Comparison between interviewees reported as suspicious and estimated under-reporting by region



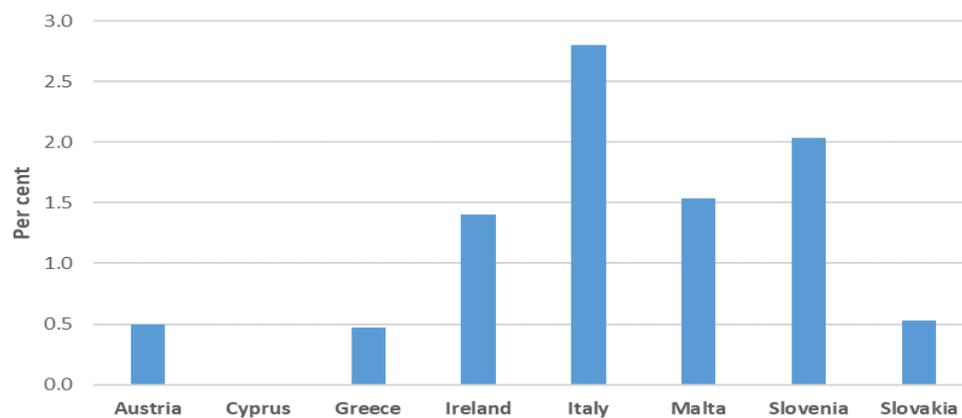
Source: Irish HFCS.

FIGURE 12. Proportion of interviewees reported by interviewers as suspicious before and after the survey



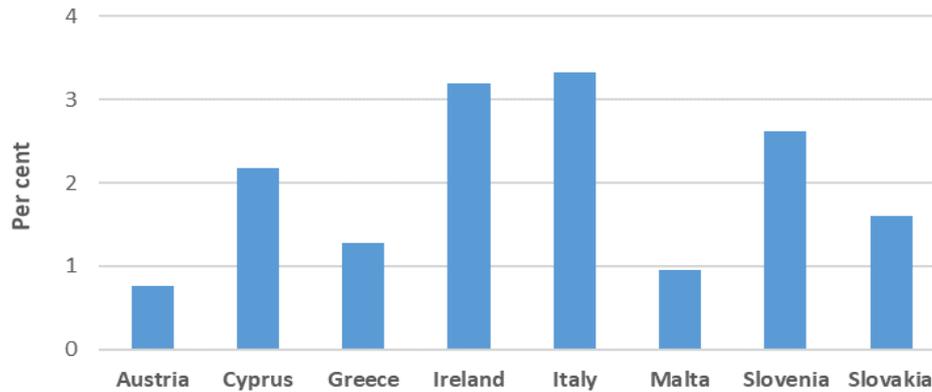
Source: Irish HFCS.

FIGURE 13. Proportion of interviewees reported by interviewers as having a 'poor' understanding of the questions



Source: Irish HFCS.

FIGURE 14. Proportion of interviewees reported by interviewers as providing 'unreliable' information



Source: Irish HFCS.

6 Conclusion

This paper provides a detailed assessment of the quality of household balance sheet data in the Irish HFCS, with a focus on household deposits. This is an important contribution for researchers who use this data for policy design and analysis. We find that the Irish HFCS significantly under-records the total value of deposits held by Irish households, capturing around one-third of the aggregate figure. However, drawing on administrative data on the population of household deposits in credit institutions, we show that the under-recording is similar across the distribution. Thus, the HFCS appears to provide an accurate picture of the distribution of household deposits. We do not find the degree of under-recording to be heavily concentrated in certain cohorts of the population. Although, in-line with results from other countries, the degree of under-recording appears to be slightly higher for older households.

With regard to the reasons for under-recording, we find no one single factor can explain all of the under-recording. Our ex-post analysis of the sample design based on deprivation indices shows a high proportion of 'low-deprivation' households in the bottom half of the wealth distribution, contrary to expectations. This might lead to under-recording of very wealthy households (unit non-response). However, this alone cannot explain the overall level of under-recording of household deposits. Another candidate is the relatively high degree of item non-response in the Irish data (albeit still a

small proportion of households). However, whilst the share of households with imputed values for deposits – between 8 and 13%, depending on the type of account – is amongst the highest in the sample of countries carrying out the HFCS, the average deposit value for these imputed households would need to be very high (over €350,000, compared to a mean of €70,000) to explain a large share of the observed under-recording. We think such a difference is implausibly high when one considers that the imputation is done by matching households on observable characteristics, including income.

The analysis presented in this paper offers some important learning points for the next wave of the Irish HFCS, currently scheduled for mid-2018. For the analysis of non-response, further information on the characteristics of the households who did not respond would assist further with understanding characteristics of those households who did not take part. For example, paradata on dwelling appearance. From a sample design perspective, investigating additional methods of identifying ‘wealthy’ households, for example, from linked-administrative data on other real wealth of the household. Another option which could be explored is using income tax records to identify wealthy households. However, as it would be important to identify high-earning households who are now retired, tax records would need to be examined over an extended period, say, 20 years. Given that administrative records in Ireland only go back a decade-or-so, this may be something to consider in surveys beyond the next wave. To the extent that it is not already part of the field work, respondents could also be reassured about the motivations behind survey or encouraged to consult statements. There could also be a specific question on deposits in post office accounts, as some households may consider these as debt securities, other assets or may not remember to include them at all when replying. Post Office savings certificates and deposits accounted for around €18 billion of household deposits in Ireland in 2013.

References

Alvaredo, F., A. B. Atkinson, and S. Morelli (2016). “The challenge of measuring UK wealth inequality in the 2000s”. *Fiscal Studies*, 37.1, 13-33.

Ampudia, M., & Ehrmann, M. (2017). “Macroeconomic experiences and risk taking of euro area households”. *European Economic Review*, 91, 146-156.

Andreasch M. and P. Lindner (2014), “Micro and macro data: A comparison of the Household Finance and Consumption Survey with Financial Accounts in Austria”, Working Paper Series no. 1673, May.

Bricker, J., A. Henriques, J. Krimmel and J. Sabelhaus (2016). “Measuring income and wealth at the top using administrative and survey data”. *Brookings papers on economic activity*, 261-331.

Carroll, C.D., Slacalek, J. and K. Tokuoka (2014), “The Distribution of Wealth and the MPC: Implications of New European Data”, *The American Economic Review* 104(5):107–111, May.

Central Statistics Office (CSO) (2015), “Household Finance and Consumption survey 2013”, www.cso.ie, January.

D’Alessio G. and I. Faiella (2002), “Non-response behaviour in the Bank of Italy’s Survey of Household Income and Wealth”, *Temi di discussione del Servizio Studi*, Number 462, December.

Disney, R. and J. Gathergood (2013), “House Prices, Wealth Effects and Labour Supply”, Discussion Papers 13/02, University of Nottingham, School of Economics, July.

European Central Bank (2013), “The Eurosystem Household Finance and Consumption Survey, Methodological report for the first wave”, European Central Bank, Statistics Paper Series, April.

Fasianos, A., R. Lydon and T. McIndoe-Calder (2017), “The Balancing Act: Household Indebtedness Over the Lifecycle”, Central Bank of Ireland, Quarterly Bulletin No. 1, January.

Honkkila J. and I. K. Kavonius (2013), “Micro and Macro Analysis on Household Income, Wealth and Saving in the Euro Area”, Working Paper Series no. 1619 / November.

Kennickell, A. B. (2007a), “Look and Listen, but Don’t Stop: Interviewers and Data Quality in the 2007 SCF”, In Proceedings of the Survey Research Methods Section, Annual Meetings of the American Statistical Association. American Statistical Association, pp. 3646-3656, May.

Kennickell, A.B. (2007b), “The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances”, Federal Reserve Board, Survey of Consumer Finances Working Papers, July.

Lawless M. and R. Lydon and T. McIndoe-Calder (2015), “The Financial Position of Irish Households”, Central Bank of Ireland, Quarterly Bulletin No. 1., January.

Le Blanc, J. (2016), “Household Saving Behaviour in Ireland”, *Economic Letter Series*, Vol. 2016, No. 5.

Le Blanc, J. and R. Lydon (2017), “Indebtedness and Spending: What Happens when the Music Stops”, *Mimeo*, November 2016.

Lydon, R. and T. McIndoe-Calder (2017), “The Great Irish (De)Leveraging”, ECB Working Paper No. 2062, May 2017

Tourangeau, R., L. J. Rips, and K. Rasinski (2000), “The Psychology of Survey Response”, Cambridge University Press: Cambridge, March.

Trutz, H. Pratschke, J. and J. Gleeson (2014), “Consistent Deprivation Measures for the Island of Ireland, based on the 2011 Census of Deprivation” [Report by Trutz-Haase](#), November 2014.

Vermeulen P. (2017), "How fat is the top tail of the wealth distribution?", *Review of Income and Wealth*, Feb(1).