

A comparison of micro and macro expenditure measures across countries using differing survey methods

March, 2012

Garry Barrett
School of Economics, University of Sydney
garry.barrett@sydney.edu.au

Peter Levell
Institute for Fiscal Studies
peter_1@ifs.org.uk

Kevin Milligan
Department of Economics, University of British Columbia
kevin.milligan@ubc.ca

Paper prepared for “Conference on Improving the Measurement of Consumer Expenditures”, sponsored by the Conference on Research in Income and Wealth and the National Bureau of Economic Research, December 2011. This paper builds on joint ongoing work with Thomas F. Crossley. Cormac O’Dea assisted in the collection of the UK data. Bruce Meyer and Jim Sullivan provided some assistance with our US data. We thank the aforementioned for their contributions to this work. We also thank the editor John Sabelhaus and conference participants for their suggestions for the paper.

1.0 Introduction

Household-level consumption lies at the center of research into many important economic questions. The measurement of microeconomic phenomena such as household poverty requires the observation of consumption choices made by households to provide useful information on economic hardship. At the macroeconomic level, the understanding of responses to booms and busts is enhanced by observing household consumption responses. Reliable consumption data are necessary to engage in meaningful empirical research in these areas.

However, there are ongoing concerns about the reliability of expenditure surveys in many countries. These concerns have led to efforts to renew expenditure survey methodology. In the United States, this activity centers on the ‘Gemini Project’ of the Bureau of Labor Statistics, tasked with improving the Consumer Expenditure Survey.¹ In Canada, the Survey of Household Spending was revised in 2010 with similar goals in mind.²

In this paper, we aim to contribute to these discussions by providing an international comparison of the performance of household expenditure survey data across four ‘Anglosphere’ countries: Australia, Canada, the United Kingdom, and the United States. Our international comparison is a useful way to gather some evidence on the potential sources of problems with expenditure surveys, as differences in experience and methodology provide sources of variation that may give insights into the importance of factors influencing the performance of expenditure surveys.

¹ Edgar and Safir (2011) provide an overview of the Gemini Project.

² Tremblay, Lynch, and Dubreuil (2010) report results from a pilot project from 2007 evaluating several changes. Many of these changes have been implemented for the 2010 Survey of Household Spending.

Our strategy is to compare household expenditure survey data to expenditure measured in the national accounts of each country. While this ‘coverage’ approach is frequently adopted in country-specific studies of expenditure behaviour, the novelty of our contribution is to produce comparable results across four countries.³ Attanasio, Battistin, and Leicester (2006), in assessing the expenditure behaviour of poor households in the US and UK, provide an assessment of micro survey evidence benchmarked against the national income. In comparison to their paper, we provide more recent years of data, two more countries with differing methodology, and a more detailed accounting of the survey differences. Deaton (2005) provides a comparison of a vast array of countries, with analysis of the same kind of ‘survey vs. national accounts’ comparison we perform here.

The paper proceeds first by reviewing the survey methods employed in the four target countries. We then discuss in more detail the construction and interpretation of household survey vs. national account comparisons, and examine the trends in aggregate ratios of survey to national account data across countries. Next, we consider how survey response rates have varied across countries and relate them to our aggregate coverage measures. We then compare the coverage measures to high income concentration through time and across our countries. Finally, we look at selected subcategories of expenditure to observe how trends vary across countries.

³ Some well-known examples of this measurement approach are Slesnick (1992), Garner et al. (2006), Garner, McClelland and Passero (2009) for the United States. Adler and Wolfson (1988) perform a similar exercise for Canada. Passero et al. in this volume also provide an updated approach to comparing the CE survey with PCE.

2.0 Expenditure survey methodology

In this section, we provide some background on the methodology employed for the household expenditure surveys in each of the four countries in our focus. We describe the target population, survey design, and other special features for each country. We begin with Australia, and proceed through Canada, the United Kingdom and the United States. At the end of these descriptions, we provide a summary table of the key elements of the survey methodologies.

2.1 Australia: Household Expenditure Survey

The Australian Household Expenditure Survey (HES) has been carried out seven times: 1975-76, 1984, 1988-89 and 1993-94, 1998-99, 2003-2004 and 2009-10. The HES is conducted over a 12 month period, typically coinciding with the financial (July-June) rather than calendar year, with households enumerated evenly over the survey period. The primary purpose of the HES is to collect comprehensive information on household expenditures, along with household income and, since 2003/ 04, wealth. The original objective of the HES program was to provide information for the construction of commodity weights in the consumer price index (CPI) - for more details on the HES background and methodology see Australian Bureau of Statistics (2011).

Expenditures are recorded in HES on an acquisition basis, with details on most regular expenditures collected using diary methods. Regular expenditure items for each household member aged 15 years or older are recorded in a personal diary covering a two week reporting

period.⁴ The fineness of the expenditure categories used in the survey has increased over time, with 660 items separately recorded in the 2003-04 and 2009-10 surveys.

Expenditures on infrequent, irregular or expensive items are recorded by personal interview with each household member aged 15 years or older. The recall period for irregular purchases varies, ranging from three months for major household furniture and appliances, 12 months for motor vehicle registrations, and three years for house purchases. Items such as insurance, rent payments and utilities bills are recorded in the interview with respondents asked the value of the last payment and the period of time that payment covered. Given the recall periods for items recorded in the household interview questionnaire, some of these expenditures will refer to time periods prior to the reference year. The public release HES reports average weekly expenditures for all items, with expenditures on some items converted to average weekly amounts. Additional information on household demographics and income are also collected during the household interview on a recall basis.

The scope of the HES includes “usual residents of private dwellings in urban and rural areas of Australia.” Excluded from the survey are residents of “non-private dwellings” such as hotels, boarding schools, boarding houses and institutions. Further exclusions are residents of very remote districts (or Indigenous Communities).⁵ In addition, “non-usual” residents of a private dwelling (e.g. visitors) are not included in the survey. Approximately 97-98 percent of the Australian population are within the scope of HES.

⁴ The HES records regular expenditures using one-week diaries for two consecutive weeks. In the 1975-76 and 1984 HES, the reporting period for rural respondents was four weeks.

⁵ Non-Australian defence forces stationed in Australia and the diplomatic personnel of overseas governments located in Australia are also excluded.

Sampling is based on a stratified multistage cluster design. The strata are based on census collector districts. Individual household records are weighted according to the probability of initial selection into the survey adjusted according to population benchmarks based on the demographic characteristics of household size and age composition, geographic location and labour force status.⁶ The sample size of the individual HES collections is typically 7,000 households, though has ranged from 4,492 in 1984 to 9,774 in 2009-10. For the most recent survey, the response rate in the HES was 71.9 percent.

2.2 Canada: The Survey of Household Spending

The Survey of Household Spending (SHS) has been the primary household expenditure survey in Canada since it replaced the Family Expenditure Survey (FAMEX) starting in 1997.⁷ The methodology is described in detail in Statistics Canada (2001). When relevant, we also referred to the methodological description in the User Guide from the most recent SHS from 2009 (Statistics Canada 2011). A detailed comparison of the SHS with the American Consumer Expenditure Survey is provided in Brzozowski and Crossley (2011). These sources provide the foundation for the description of the SHS below. We also use the FAMEX surveys for some of our analysis, but the primary focus is on the more recent SHS.⁸

⁶ The two initial HES did not use population benchmarking.

⁷ The differences between the SHS and FAMEX are outlined in Statistics Canada (2000). The sample size increased, the survey became annual, population coverage broadened, and some minor changes to survey content were implemented. We include some data from the FAMEX in our work here, but primarily focus on the SHS.

⁸ FAMEX surveys were conducted in 1969, 1974, 1978, 1982, 1984, 1986, 1990, 1992, and 1996. The 1984 and 1990 surveys are less comparable because in those years only residents of certain large cities were surveyed.

The SHS targets individuals living in Canadian private households, as well as residents of Indian reserves and Crown Lands. This definition excludes those who are official representatives of foreign countries living in Canada, as well as those who are representing Canada abroad. It also leaves out residents of institutions, hotels and rooming houses, religious orders and members of the Canadian Forces living in camps. For the lower 10 provinces, the coverage is around 98 per cent of the population. For the sparsely populated northern territories, coverage is over 80 per cent.

Sampling is based on the Labour Force Survey sample design, which uses stratified clusters. The strata are based on geography within each province. Special strata of households in areas with geographical concentrations of high and low income residents are also used. Clusters are chosen, and then a sample of households is chosen from each cluster. Extensive follow-up is engaged for households who refuse to comply, including further phone calls, visits, and letters. Sample size started at 18,031 in 1997. From then until 2007, the number of observations slid down to 13, 939. For 2008 and 2009, budget cutbacks meant a jump down to samples of 9,787 and 10,811.

The SHS attempts to gather information on the 12 month period from January 1st to December 31st. The information is gathered via a face-to-face recall survey of one household member in the January, February, or March following the end of the target calendar year. The survey respondents are encouraged to gather source documents such as credit card statements, mortgage statements, and their income tax records. The average survey takes one hour and forty minutes to complete. A ‘balance edit’ is applied when the difference between expenditure, income, and

savings, exceeded a 20 percent tolerance level.⁹ Item non-response is countered by imputing data based on ‘nearest neighbor’ imputations.

For 2009, weights are provided to account for non-response according to cells defined by province, age, household size, and family income as measured by administrative tax data. This weighting strategy has changed several times. Importantly, starting in 1999 tax-filing data from the Canada Revenue Agency were used to match on wage and salary income.¹⁰ This is helpful if there is a concern that lower response rates are particular to certain parts of the income distribution, as the weights can account for such systematic patterns.¹¹

Major changes to the SHS were implemented in 2010, although the data have not yet been released. Dubreuil et al. (2011) report that the 2010 SHS removes the calendar year focus and now has an interview-diary format. Because income and expenditure periods no longer will match, the balance edit will no longer be used. For 2009, both the old and new SHS methodologies were employed, which will allow researchers to study the impact of the change in methodology.

2.3 United Kingdom: Living Costs and Food Survey

The information in this section is drawn from Office for National Statistics (2010). The Office for National Statistics (ONS) in the UK has carried out some form of annual survey of household

⁹ Brzozowski and Crossley (2011) look into the impact of this balance edit in detail, by examining the data from 2006 when no balance edit was imposed.

¹⁰ The income weights account for incomes in the following percentile ranges: 0-25th percentile, 25th-50th, 50th-65th, 65th-75th, 75th-95th, 95th-100th).

¹¹ Sabelhaus, et al (this volume) show that response rates in the U.S. Consumer Expenditure Survey are in fact much lower at the top of the income distribution.

expenditures since 1957. From 2008 this survey has been known as the Living Costs and Food Survey (LCFS). Prior to this it was known as the Expenditure and Food Survey, which brought together what were two separate surveys for food and expenditure – the Family Expenditure Survey and the National Food Survey – in 2001. The survey is conducted continuously throughout the year.

Participation in the survey is voluntary. In 2009, the survey selected over 12,000 addresses, but only 5,825 of these were included in the survey. The remaining addresses were either ineligible to be included (because, for instance, the addresses were for businesses), refused to participate or were not possible to contact. Households in Northern Ireland are sampled separately and oversampled relative to the rest of the UK in order to achieve the sample size required for separate analyses. The response rate among eligible addresses was 56% in Northern Ireland and 50.4% in the rest of the UK.

Households who are surveyed are first asked a series of questions on income, demographic characteristics, large purchases over the last year or so (on white goods, vehicles, holidays etc.) and committed expenditures such as magazine subscription costs. Each household member over 16 is then given a spending diary in which they record all purchases made over the next two weeks. Simplified diaries for children aged 7-15 have also been included since 1998. At the end of the two weeks, each adult who kept a diary is paid £10 (\$16) for completing the survey (children who kept a diary are paid £5 (\$8)). Spending is grossed-up using weights from the most recent population Census (which have in the past been carried out once every 10 years—although 2011 may be the last.)

Data collected in the LCFS are used for a number of official purposes. As well as being used for the construction of the National Accounts, the LCFS is used to calculate expenditure weights for headline inflation measures.

2.4 United States: Consumer Expenditure Survey

The Consumer Expenditure Survey (CE) has been collecting information about American expenditure patterns on an ongoing basis since 1980.¹² The Bureau of Labor Statistics publishes the Handbook of Methods, of which Chapter 16 is devoted to the Consumer Expenditure Survey (Bureau of Labor Statistics 2011a). A short summary is also provided in Bureau of Labor Statistics (2011b). A review of changes to methodology through time is provided by Goldenberg and Ryan (2009). We draw on these sources in forming our description of the CE survey in this section.

The CE survey combines two one-week diaries of around 7,000 households with a series of five quarterly recall surveys of another 7,000 households. The target is the total US civilian non-institutional population, which excludes military personnel living on base, nursing home residents, and people in prisons. Sampling takes place by choosing households from a list within each of 91 clusters. The list of addresses comes from the most recent Census file, augmented by new construction permits. For the 2010 survey, the response rate was 73.4 percent.

¹² There were antecedents to the ‘modern’ CE in 1960-61 and 1972-73, as well as earlier years.

The diary component starts with an interview for demographic information on the first day. The diary of expenditure is to be completed every day during the week. The diary is collected at the end of the first week and a second diary is delivered. When the second diary is picked up, more questions are used to collect information on work and income from the previous year. The data are put through edits and adjustments when being processed. Some imputations are performed as well.

The quarterly recall survey component aims to gather information on less-frequently purchased items, with a three-month recall window. The raw data from the surveys is put through various checks, with imputed values being imposed for missing data. Other adjustments, such as the splitting of mortgage payments into principal and interest components, are made. With the switch from pencil and paper to Computer Assisted Personal Interview in 2003, the time to complete the interview survey fell from about 90 minutes to around 65 minutes.

The survey is available annually from 1980 to 2010. For several quarters in the early 1980s, rural households were not surveyed. In our analysis below, we retain these years but they do stand out on several of the graphs for this reason.

Weights in the CE survey are calibrated to 24 population counts, including age, race, household tenure, region, and rural/urban. The target population counts are updated quarterly, and the demographics of the sample are assigned weights to match the population on these 24 factors. Of note, there is no adjustment for income.

2.5 Comparison

In Table 1, we summarize the main features of the survey data in each country. The data from Canada are different in a number of ways, including the annual focus, having no diary, weighting based on administrative income data, and featuring a balance edit. In Australia, there is some weighting by income—but just the source of income is used. The recall window for the surveys varies across countries. In Australia, it goes back up to three years for some items. In the UK, one interview goes back for a period of a year. For the United States, the survey is a sequence of five quarterly-focused questionnaires.

3.0 Aggregate Coverage Rates

The first step in our assessment of the performance of the household expenditure surveys is the examination of coverage rates of aggregate expenditure for each of the four countries. That is, we take the ratio of expenditures observed in the household survey, grossed up to the aggregate level, to the total expenditures taken from the national accounts. We compare this ratio across time and across countries.

There are several well-known reasons to expect this ratio not to be 100 percent. (See, for example, Meyer and Sullivan 2009.) The population covered by each source may differ. For example, foreigners living in the host country and nationals living out of the country receive different attention in the national accounts and the expenditure surveys, as do military personnel and native peoples. In addition, the categories of expenditure available in the national accounts may not match those available in the expenditure surveys. For example, imputed housing rent is

included in the national accounts, but not in the expenditure surveys. Finally, expenditure in the household sector of the national accounts includes spending by non-profit institutions serving households (such as charities), which does not appear in the expenditure surveys.

To make the best possible comparison, we adjust both the national accounts data and the expenditure survey data to remove items where they do not overlap.¹³ For example, non-cash items such as imputed rent and food grown and consumed at home are taken out of the national accounts measure of household expenditure. Similarly, some items from the expenditure surveys, such as insurance purchases, are removed. With these adjustments made, we calculate the ratio of the grossed up expenditures from the household expenditure survey to the aggregate from the national accounts. This ratio is referred to as the ‘coverage rate.’ This coverage calculation is performed for expenditures in aggregate (as we do here in this section) as well as category by category comparisons (some of which are presented in a later section).

The coverage rates are graphed in Figure 1 for each of the four countries. In order to emphasize the nature of the decline, we have adjusted the y-axis to start at 0.5. Both the levels and trends differ sharply across countries. The Australian coverage rate stays in the 60 to 75 percent range, with no discernible trend. For Canada, the coverage rate is close to 1.0 for both the FAMEX (1969-1996) and the SHS (1997-2009) periods. There is no sign of an aggregate drop in coverage. The coverage rate for the UK drops steadily over the years, from 90 to 67 percent.

¹³ For a detailed description of the methodology used for our UK sample, please see Crossley and O’Dea (2010).

Finally, the United States shows coverage rates lower than Canada, but follows a very similar trend to the United Kingdom.¹⁴

In the two sections of the paper that follow, we investigate two aspects of this decline in our four countries. First, we look at the impact of declining response rates and increasing income inequality for the expenditure surveys on coverage. Following that, we compare different categories of expenditures, looking across diary and survey categories, as well as frequently and less frequently purchased items.

4.0 Candidate Explanations for Declining Coverage

Many possible explanations for declining coverage rates have been offered. Here, we use our four countries to explore two possibilities. First, we look at declining survey participation rates. A decline in survey participation rates has been observed around the developed world, a trend that began accelerating around 1990.¹⁵ This trend coincides with the decline in the coverage rate in the CE survey in the United States, making non-response a candidate explanation for the decline in coverage. Response rates are relevant for the representativeness of samples, and reliability of micro-level evidence on expenditures. In particular, if the decline in response rates is not random across the population (and cannot be corrected adequately by sampling weights) then the results of the survey will no longer be representative of the population. For example, if

¹⁴ The extra dip down in 1982-1983 is likely related to the discontinuation of rural data collection from the 3rd quarter of 1981 to the first quarter of 1985. We have checked our calculations against those in Meyer and Sullivan (2009) and found our coverage rates to be very similar.

¹⁵ See de Leeuw and de Heer (2002) for international evidence. Tourangeau (2004) provides a discussion of the trends.

high expenditure households have become increasingly less likely to respond, and if weighting did not account for this change, then coverage rates would be expected to decline.¹⁶

The second possibility we examine is the impact of income inequality on survey accuracy.¹⁷ The large trends in the concentration of income are documented across countries in Atkinson, Piketty, and Saez (2011). This concentration has been especially acute in the ‘Anglosphere’ countries on which we focus. None of the four countries we study oversamples high income households for the expenditure surveys.¹⁸ If increasing concentration of income is leading to an increasing concentration of expenditures, an increasing share of expenditure may be missed if the upper tail of expenditure is not adequately included in the survey sample. We also investigate this possibility. In addition, it is possible that the income inequality effect interacts with survey non-response. If the change in non-response is occurring more at the top of the income distribution, then the two effects (declining response rates and increasing income distribution) would reinforce each other.

With either survey response rates or income inequality, we will be comparing time series trends that happen to coincide with the change in coverage rates. It is prudent to be cautious in the interpretation of these results as causal. That said, we do get some mileage out of our cross-

¹⁶ Tourangeau (2004) reviews the evidence on the causes of declining survey participation, but does not discuss how non-participation is correlated with characteristics such as income. D’Alessio and Faiella (2002) find that non-response in the Bank of Italy’s Survey of Household Income and Wealth is more frequent among wealthier households. Finally, Sabelhaus et al. in this volume show that response rates in the CE are much lower at the top of the income distribution.

¹⁷ We thank Angus Deaton for suggesting this possibility to us.

¹⁸ Canada uses the Labour Force Survey sampling frame, which does target certain high income areas when choosing strata from which to sample. However, there is not explicit oversampling of high income households within the survey.

country comparison by including in our regression specification common time trends, allowing us to exploit the cross-country variation in the coverage, response rate, and inequality trends.

4.1 Response rates

Figure 2 shows the basic response rates for the expenditure surveys for the four countries, with the y-axis starting at 0.5. Each country exhibits declining responses rates, with the steepest occurring in the United Kingdom—where the drop is from 72 percent to 53 percent. The decline begins in the early 1990s in Australia, the United Kingdom, and the United States, but is not observable in Canada until the 2000s. While Canada was later starting downward, the decline exceeded 10 percentage points over the last decade.

To compare coverage and response rates, we graph the data from Figure 1 and Figure 2 together for each country as a scatter plot in Figure 3. The axes are different for each country in order to highlight the nature of the relationship in each country. For Canada, the United Kingdom, and the United States, there does appear to be a positive relationship between the response rate and the coverage rate. For Canada, the positive relationship in the figure is perhaps deceptive—the variation in the coverage rate is quite small—it ranges only from just under 1.0 to just under 1.07. The UK shows a fairly tight positive relationship across the 35 years available. In the US, the data are clustered in two groups that together suggest a similar positive relationship between coverage rates and response rates. For Australia, in contrast, there is no apparent relationship between response rates and coverage rates, although the limited number of surveys makes any conclusion difficult.

Figure 4 stacks together the data for all four countries in one plot with common axes. Looking across countries, the data display little clear relationship. However, within-country the United States and United Kingdom reveal positive relationships. Later in the paper, we can confirm these relationships in regressions.

4.2 Trends in high income concentration

The other trend we compare to declining coverage rates is the increase in income inequality. We draw on data from the high incomes database maintained at the Paris School of Economics (Alvaredo et al. 2012). We use the proportion of income earned by those in the top one percent of the income distribution for our analysis here, although other high income measures showed similar results.

Figure 5 shows how the top one percent income shares have evolved in our four countries. Through the mid-1980s, there is little to be seen—although the top income share does start to rise in the UK following 1980. From around the beginning of the 1990s, there is a strong upward trend in each of the countries. The weakest of these upward trends is in Australia and the strongest is in the US. This timing does correspond to the decline in coverage rates which accelerated in the 1990s.

We compare the trends in top income shares to the trends in coverage rates across all four countries in Figure 6, with separate scales for each country's axes. All four countries show signs of a negative relationship. Canada, again, has little variation in the coverage rate across years, so

looks a bit different from the others. In the United States and the United Kingdom, there is a clear negative relationship between income inequality and the coverage rate.

Some parallels may be drawn here between our findings and those of Deaton (2005). In that paper, he finds that the coverage rate across countries is declining in the log of GDP, so higher income countries are experiencing worse coverage.¹⁹ One of Deaton's explanations is that higher income countries tend to have higher income concentration, which may be captured less well in surveys. This is consistent with our findings here.

4.3 Regression analysis

The relationships from these figures can be summarized with some basic regressions. The coverage rate is regressed on the response rate, with country and time controls using Ordinary Least Squares. The equation takes the following form.

$$Coverage_{it} = \beta_0 + ResponseRate_{it}\beta_1 + Top1\%share\beta_2 + Country_i\beta_3 + Year_t\beta_4 + e_{it}$$

We report these results in Table 2. The dependent variable in all cases is the country-year coverage rate, and each column reports the results from a different specification. We report the regression coefficient, with the standard error beneath in parentheses. In column (1), we include no controls other than the constant term and the response rate variable. This effectively estimates a best-fit line through the data points as seen in Figure 4. The small and insignificant estimated coefficient confirms the lack of relationship across countries. The second column of the table

¹⁹ When comparing to our results, though, it must be remembered that much of the impact Deaton finds is concentrated among those countries with very low incomes.

includes country fixed effects. Here, the within-country relationships are used in the estimation, essentially taking an average of the relationships shown in the country-specific scatter plots in Figure 3. The coefficient swings strongly positive, at 0.779. This suggests that for every percentage point increase in the response rate, there is a 0.779 percentage point increase in the coverage rate. Taking the US as an example, the response rate dropped by 11.86 points from 1990 to 2008, so this coefficient explains a (0.779×11.86) 9.24 percentage point drop in coverage, which is larger than the 6.85 percent drop that occurred. On this basis, we interpret this coefficient as large.

In column (3) of Table 2 we include the top one percent income share variable. The coefficient on the response rate drops, but remains statistically significant and positive. The coefficient on the top income share is -1.006, which suggests that a one percentage point increase in the top income share is associated with a 1.006 percentage point decrease in the coverage rate, all else equal. To interpret these magnitudes differently, consider that the top one percent share in the US increased by 4.69 percentage points from 1990 to 2008. Over that same period, coverage dropped by 6.85 percentage points. The -1.006 coefficient means that the decline in top income share predicts a (1.006×4.69) 4.72 point drop in coverage, which is 68.9 percent of the 6.85 point drop.

Column (4) includes an interaction of the response rate and the top income share. This change leads to negative (but insignificant) coefficients on the top one percent share term and on the interaction term. The large standard errors on both estimated effects indicate that the interaction term is not well identified from the linear effect of the top one percent share on coverage rates.

Indeed, a joint test for significance of these two variables shows they are highly jointly significant.²⁰ Further, the magnitude of the partial effect of an increase in the top one percent share evaluated at the mean US survey response rate of 0.81 based on the estimates in column (4) is numerically very similar to the linear partial effect of -1.006 in column (3). Together, the insignificance of the interaction terms and the comparable estimated partial of income inequality on coverage rates with the two specifications indicate no evidence that the effects seen in column (3) were driven by an interaction of the two factors.

In the last two columns of Table 2 we try alternative controls for time. Column (5) has a linear time trend. This time trend accounts for any global trend that is common to the four countries in our study. The coefficient on the response rate changes slightly to 0.345, while for the top income share the coefficient jumps up to a larger (in absolute value) magnitude. Finally, the last column includes dummies for each year of the sample, which controls flexibly for any common calendar time effects across countries. This is a fairly demanding specification given the number of observations we have and the nature of the variation we are using. Since there are only four observations per year, it may be difficult to detect any effect in this specification. The resulting coefficients remain fairly stable---but both lose statistical significance in this final specification.

This graphical and regression analysis shows that the trend downward in response rates is common across all four countries, and that the decline in expenditure coverage in surveys compared to national accounts has a positive relationship with changes in survey response rates. Top Income shares are also shown to be negatively related to coverage rates. Taken together, our

²⁰ The calculated *F*-statistic for the interaction of the top one percent variable and the response rate and the one percent variable itself is 12.19. For both main effects and the interaction, the calculated *F*-statistic is 50.83. For the response rate and the interaction the joint test yields an *F*-statistic of 4.69. All of these are highly significant.

results suggest that declining survey response rates and increasing income inequality may prove to be important determinants of the decline in expenditure coverage rates.

5.0 Coverage rates within expenditure categories

The next step in our analysis is to compare different categories of expenditure across countries, looking for evidence that aligns with differences in survey methodology. Canada here is the most noticeable outlier in survey methodology, as the SHS uses an annual recall survey for both frequently purchased and infrequently purchased items—with no diary component. There is also a balance edit, and substantial income weighting. The four categories we consider are food at home, alcohol purchased in stores, new and used motor vehicles, and furniture appliance and household equipment.

The first category we examine is food consumed in the home. These data are collected through a diary in Australia, the UK, and the US, but with recall in Canada. Food for consumption at home is a basic non-durable commodity that has been used as a summary measure of household welfare, and has been the focus of many studies testing predictions of consumption smoothing at the household and aggregate level. We graph the coverage rates in Figure 7.²¹ The UK shows a decline of 10 percentage points since the early 1990s. However, there is little evidence of a similar trend in the other three countries.

²¹ For Canada, we now show only the SHS results, as the category-by-category analysis tends to exhibit seams between the FAMEX and SHS survey years.

The second expenditure category considered is alcohol purchased in stores. This category is collected using the same methods as for food consumed at home. This category is of interest because alcohol consumption is generally viewed as socially undesirable which may lead individuals to underreport these expenditures in household survey. As Figure 8 shows, it is the case that survey coverage of this item is very low—being around 50% for Australia²² and Canada, and substantially less for the US. However, conditional on the lower level of coverage, the coverage rates for this item are remarkably stable in each of these three countries. For the UK, the coverage rate is higher and has declined through time.

The final two graphs show more infrequently purchased items. In all countries, these data are collected with recall surveys. Figure 9 shows new and used vehicles, while Figure 10 has household equipment, furniture and appliances. There are no easily discernible patterns for vehicles. For Australia, coverage rates for these categories are neither consistently rising nor falling. In Canada, there is an upward trend for new and used vehicles, and perhaps small downward movement for the other two. For the UK, coverage of vehicles appears quite cyclical, but do not show a long term decline. The series for furniture, household equipment, and appliances shows a fairly slow and steady decline, although Australia does rebound at the end.

This examination of category-by-category patterns has revealed little clear evidence about differences across countries. In all countries, the frequently consumed product (food) seems to have high coverage. In contrast, the less frequent bigger purchases appear to be much more volatile year to year, and have a more pronounced downward trend on average. This is consistent with the evidence shown previously in Meyer and Sullivan (2009) and elsewhere for the US. The

²² Apart from atypically high coverage in the Australian HES in 1975-76.

income elasticity of demand for the goods likely plays a role as well. As income concentration increases, coverage rates for goods consumed more by higher income households may decline.

6.0 Conclusions

In this paper we provide a comparative assessment of the performance of the household expenditure survey programs in Australia, Canada, the UK and US. The survey methodologies employed in each country share a number of common features while containing distinct elements. There are also differences in survey response rates and income concentration across the countries. We use this variation across countries to assess the implications for the performance of the household surveys.

After first outlining the key features of the household expenditure surveys for each country, we assess the coverage of aggregate expenditure relative to the national account benchmark. Both the survey expenditure aggregate and the national accounts data are adjusted to ensure the expenditure concepts are comparable. Coverage rates are highest in Canada and the UK; for Canada, and Australia coverage remained fairly stable over the past three decades. In contrast, in the UK and the US coverage rates have sharply declined over the past three decades.

Next, survey response rates and top income shares were considered in tandem with coverage rates. This analysis is motivated by the widely observed decline in participation rates for household surveys over time, and the strong concentration of income that has occurred in many countries. From a series of graphical comparisons and regression models it is found that the fall

in response rates and top income shares over time is quite predictive of changes in coverage rates within countries.

The last component of the analysis examined coverage rates for specific components of expenditure. Individual expenditure items considered were food at home, alcohol purchased in stores, new and used motor vehicles, and furniture appliance and household equipment. This list included categories which were collected using the divergent methodologies (e.g. food by diary in Australia, UK and US; by annual recall in Canada) or by comparable methods (e.g. motor vehicles, furniture and recreational equipment collected by recall in interviews in all countries). From this, there was no clear pattern across country. Rather, most evident is the high and stable coverage of regularly purchased items, along with the more volatile coverage of irregular and larger expenditure items.

Overall, our comparative assessment of the household expenditure surveys across the four Anglosphere countries studied has shown the sharpest differences between Canada and Australia vs. the United Kingdom and the United States. However, the many unique aspects of the Canadian survey methodology make it difficult to identify single features of the methodology that are pivotal to its performance. Given the Canadian methodology changes that were put in place for 2010, some further information may soon be available about the reasons for the relative success of the Canadian data.

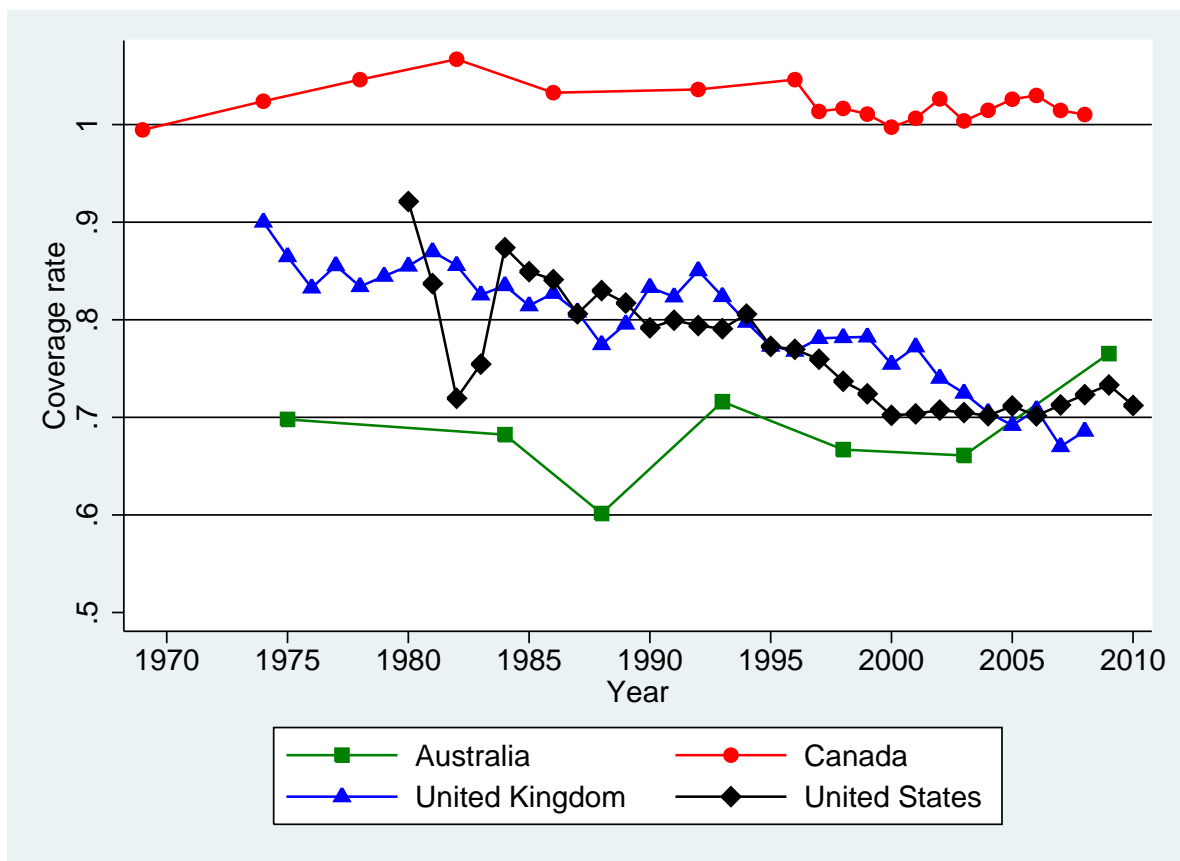
.

References

- Adler, Hans J. and Michael Wolfson (1988), "A prototype micro-macro link for the Canadian household sector," *Review of Income and Wealth*, Vol. 34, No. 4, pp. 371-392.
- Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty and Emmanuel Saez (2012), "The World Top Incomes Database," <http://g-mond.parisschoolofeconomics.eu/topincomes>, accessed February 5, 2012.
- Atkinson, Anthony B., Thomas Piketty and Emmanuel Saez (2011). "Top Incomes in the Long Run of History," *Journal of Economic Literature*, Vol. 49, No. 1, pp. 3-71.
- Attanasio, Orazio P., Erich Battistin, and Andrew Leicester (2006), "From Micro to Macro, from Poor to Rich: Consumption and Income in the UK and US." Draft, National Poverty Center, Gerald R. Ford School of Public Policy, University of Michigan.
- Australian Bureau of Statistics (2011), *Household Expenditure Survey, Summary of Results Australia*, Australian Bureau of Statistics, Canberra, Catalogue No. 6530.0.
- Brzozowski, Matthew and Thomas F. Crossley (2011), "Measuring the well-being of the poor with income or consumption: A Canadian perspective," *Canadian Journal of Economics*, Vol. 44, No. 1, pp. 88-106.
- Bureau of Labor Statistics (2011a), *Bureau of Labor Statistics Handbook of Methods*. Internet version, accessed November 2011.
- Bureau of Labor Statistics (2011b), "Appendix A: Description of the Consumer Expenditure Survey," Consumer Expenditure Survey Anthology, 2011. US Bureau of Labor Statistics Report 1030.
- Crossley, Thomas F. and Cormac O'Dea, July 2010, "The wealth and saving of UK families on the eve of the crisis," IFS Reports , R71
- D'Alessio, Giovanni and Ivan 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.
- de Leeuw, Edith and Wim de Heer (2002), "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison," in Robert M. Groves, Don A. Dillman, John L. Eltinge, Roderick J. A. Little (eds.) *Survey Nonresponse*. New York: Wiley; 2002. p. 41-54.
- Deaton, Angus (2005), "Measuring poverty in a growing world (or measuring growth in a poor world)," *The Review of Economics and Statistics*, Vol. 87, No. 1, pp. 1-19.

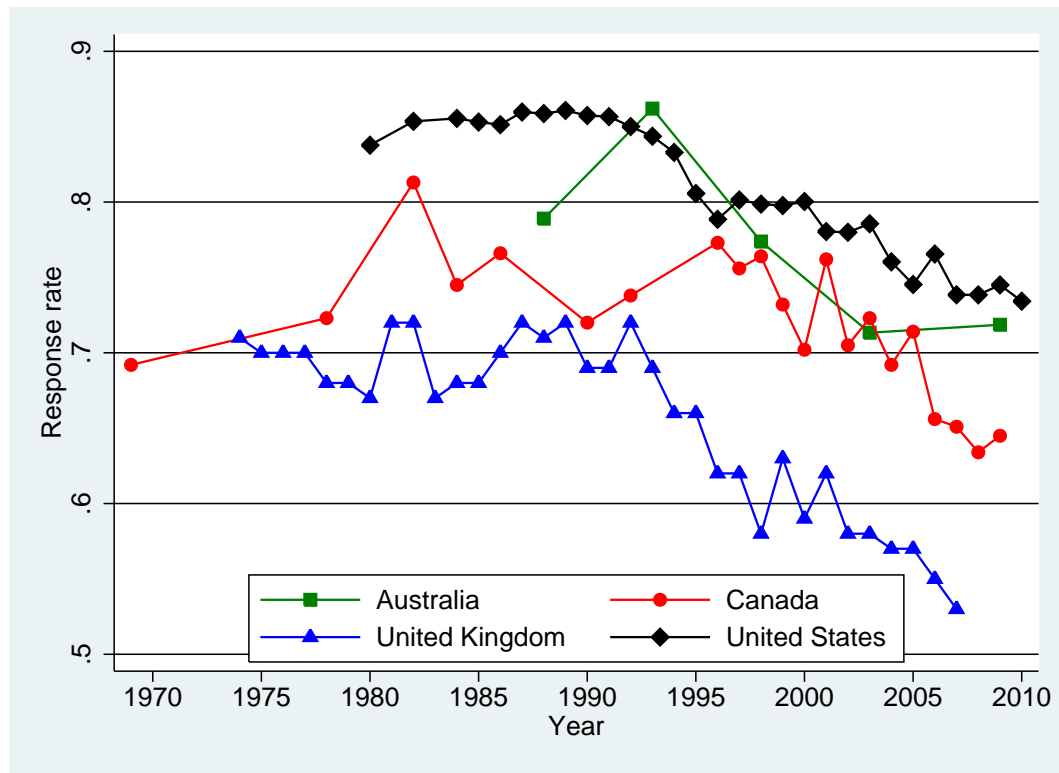
- Dubreuil, Guylaine, Johanne Tremblay, Jenny Lynch, and Martin Lemire (2011), “Redesign of the Canadian Survey of Household Spending,” Paper presented at Household Survey Producers Workshop, Washington DC, June 1-2, 2011.
- Edgar, Jennifer and Adam Safir (2011), “Gemini Project Overview,” Consumer Expenditure Survey Anthology, 2011. US Bureau of Labor Statistics Report 1030.
- Garner, Thesia I., George Janini, William Passero, Laura Paszkiewicz, and Mark Vendemia (2006), “The CE and the PCE: A comparison,” *Monthly Labor Review*, Vol. 129, No. 9, pp. 20-46.
- Garner, Thesia I., Robert McClelland, and William Passero (2009), “Strengths and weaknesses of the Consumer Expenditure Survey from a BLS perspective.” Draft presented at National Bureau of Economic Research Summer Institute, July 2009.
- Meyer, Bruce D. and James X. Sullivan, “Five decades of consumption and income poverty,” National Bureau of Economic Research Working Paper No. 14827.
- Office for National Statistics (2010), “Living Costs and Food Survey 2009 User Guide”.
- Passero, Bill, Caitlin Blair, Rob Cage, Thesia I. Garner, and Clint McCully (2012), “The Devil’s in the Details: A Concordance for the CE, CPI, and PCE,” This volume.
- Sabelhaus, John, David Johnson, Stephen Ash, David Swanson, Thesia Garner, John Greenlees, and Steve Henderson (2012), “Is the Consumer Expenditure Survey representative by income?” This volume.
- Slesnick, Daniel T. (1992), “Aggregate consumption and saving in the Postwar United States,” *Review of Economics and Statistics*, Vol. 74, No. 4, pp. 585-597.
- Statistics Canada (2000), “Note to former users of data from the Family Expenditure Survey,” Income Statistics Division, Statistics Canada. Catalogue No. 62F0026MIE-00002.
- Statistics Canada (2001), “Methodology of the Survey of Household Spending,” Income Statistics Division, Statistics Canada. Catalogue No. 62F0026MIE-01003.
- Statistics Canada (2011), “User guide for the Public-use Microdata file, Survey of Household Spending, 2009,” Income Statistics Division, Statistics Canada. Catalogue No. 62M0004XCB.
- Tremblay, J., Lynch, J., and Dubreuil G. (2010), “Pilot Survey Results from the Canadian Survey of Household Spending Redesign.” Paper presented at the 2010 Joint Statistical Meetings of the American Statistical Association, Vancouver, British Columbia.
- Tourangeau, Roger (2004), “Survey Research and Societal Change,” *Annual Review of Psychology*, Vol. 55, pp. 775-801.

Figure 1: Coverage Rates



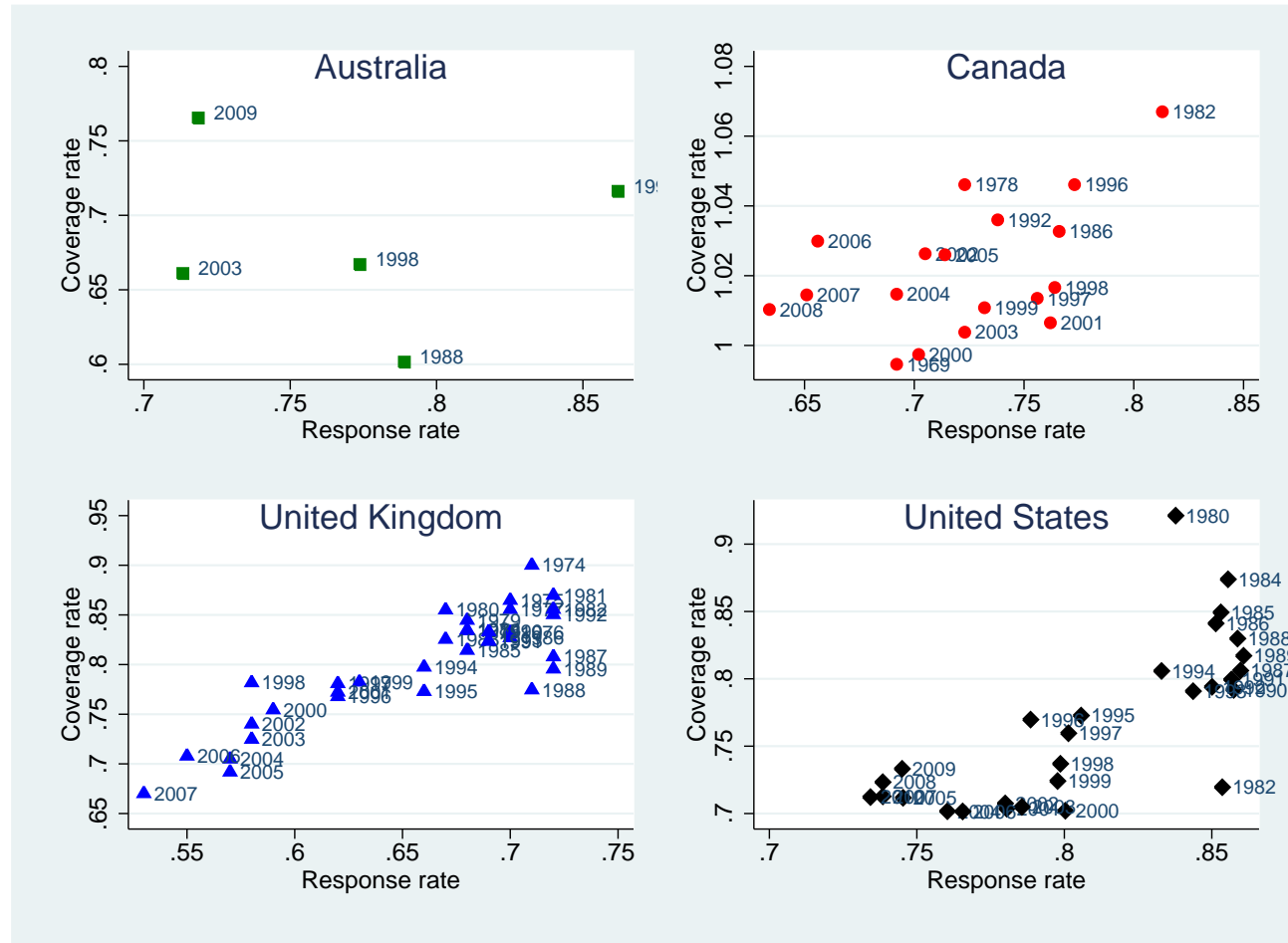
Notes: Coverage rate is the proportion of consumer expenditure in the national accounts that is accounted for in the household surveys. Calculations by authors.

Figure 2: Response Rates



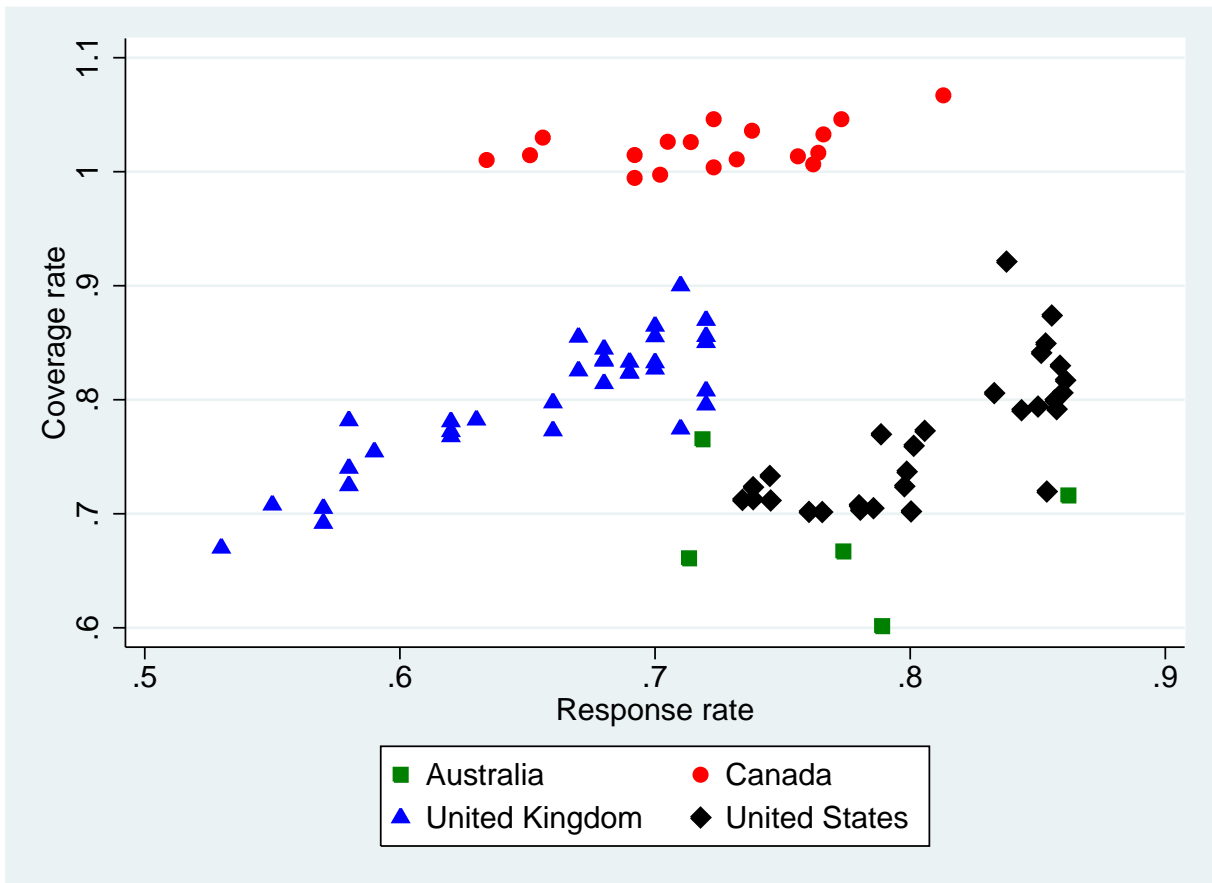
Notes: Response rate is the proportion of contacted households with completed surveys. Source is the documentation for the surveys in each of the four countries, as referenced in the text.

Figure 3: Response Rates vs. Coverage Rates



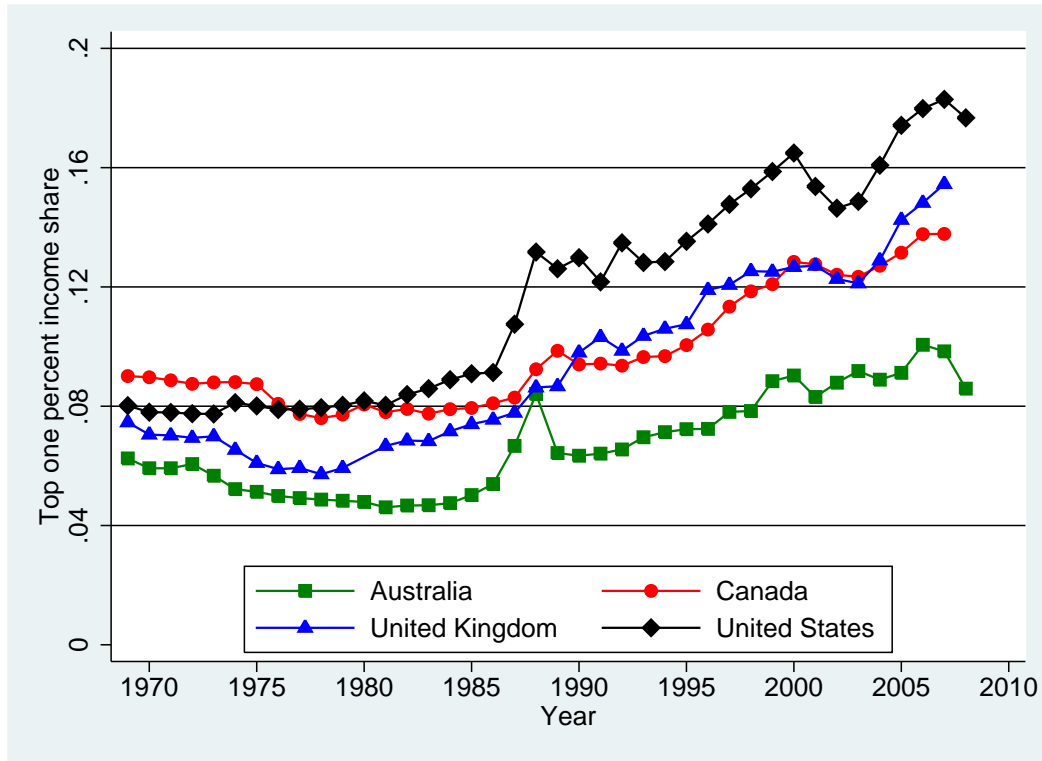
Notes: Coverage rate is the proportion of consumer expenditure in the national accounts that is accounted for in the household surveys. Response rate is the proportion of contacted households with completed surveys. Calculations by authors.

Figure 4: Response Rates vs. Coverage Rates, All Countries



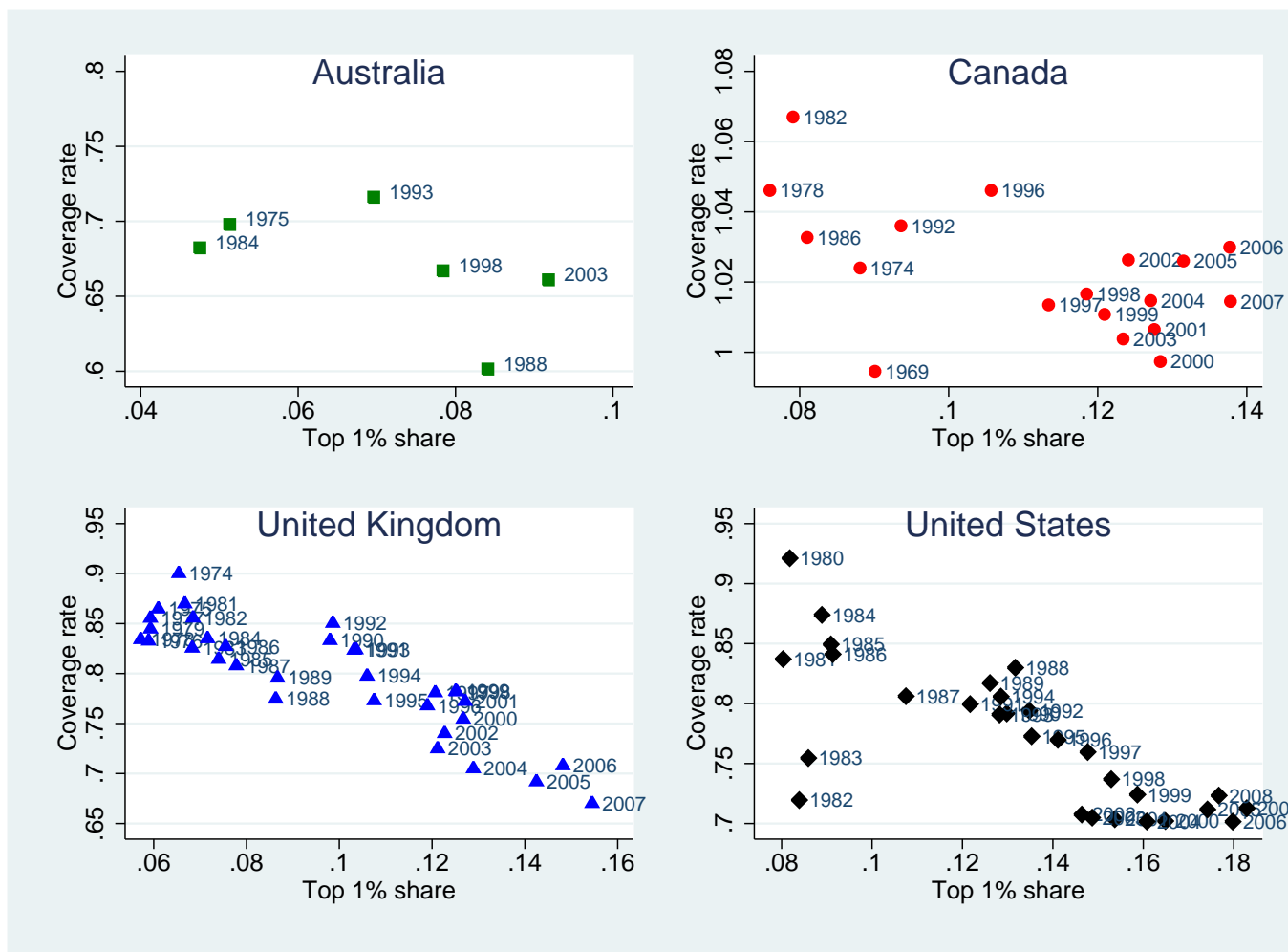
Notes: Coverage rate is the proportion of consumer expenditure in the national accounts that is accounted for in the household surveys. Response rate is the proportion of contacted households with completed surveys. Calculations by authors.

Figure 5: Top 1 Percent Income Shares, All Countries



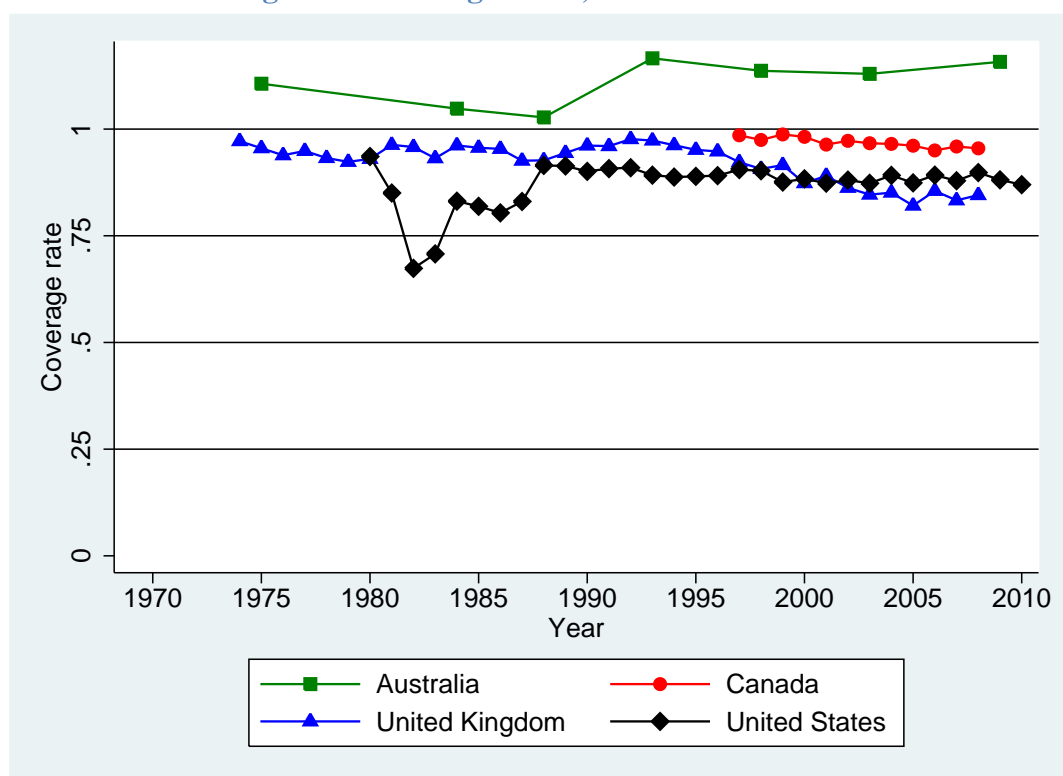
Notes: Top one percent income share is the share of total income received by those in the top one percent. Source is Paris School of Economics World Top Incomes Database, Alvaredo et al. (2012).

Figure 6: Coverage Rates vs. Top Income Shares



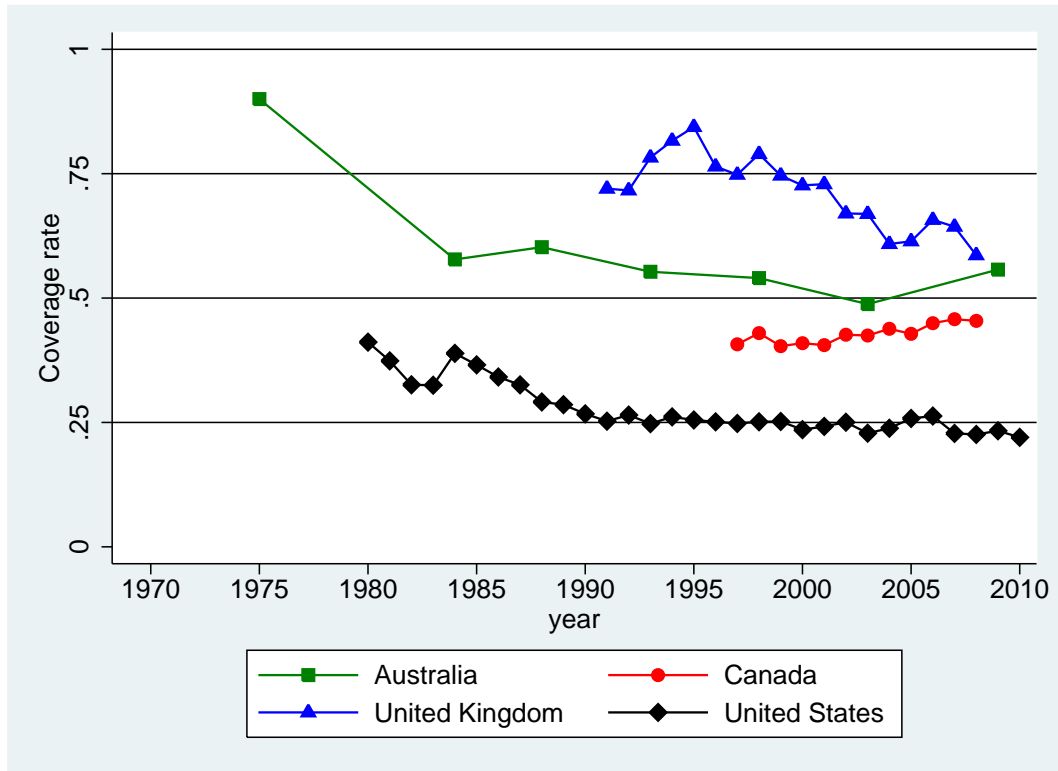
Notes: Top one percent income share is the share of total income received by those in the top one percent. Source is Paris School of Economics World Top Incomes Database, Alvaredo et al. (2012). Coverage rate is the proportion of consumer expenditure in the national accounts that is accounted for in the household surveys. Source is calculations by the authors.

Figure 7: Coverage Rates, Food in the Home



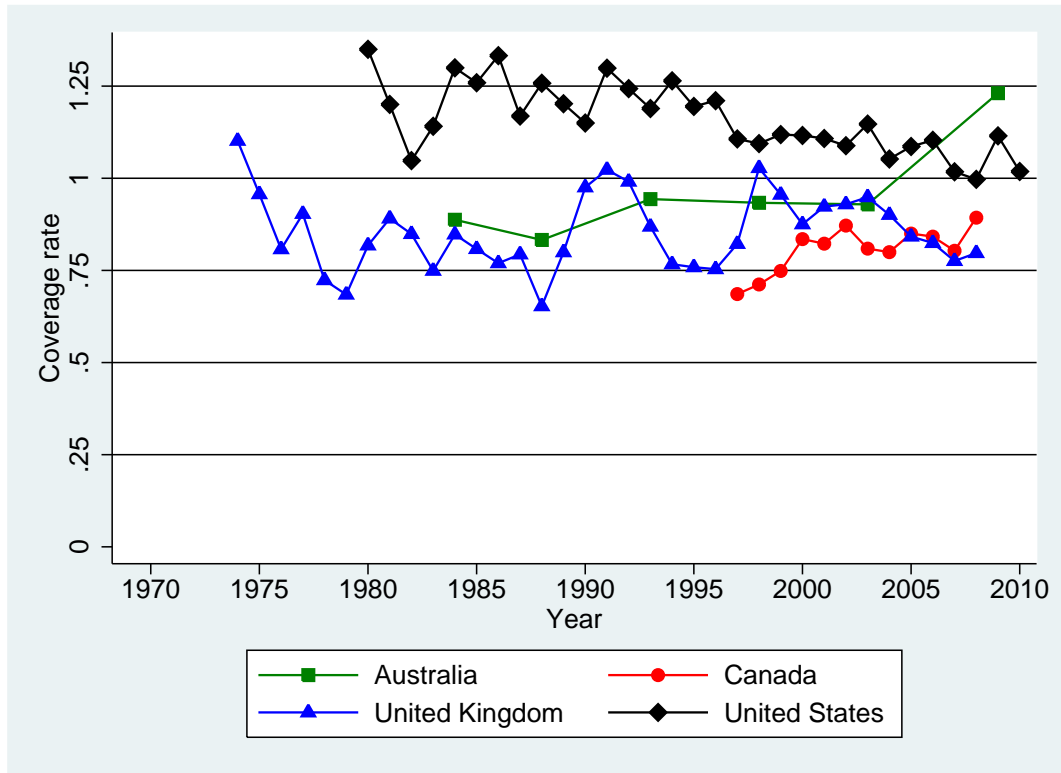
Notes: Coverage rate is the proportion of consumer expenditure in the national accounts that is accounted for in the household surveys in this category. Source is calculations by the authors.

Figure 8: Coverage Rates, Alcohol Purchased in Stores



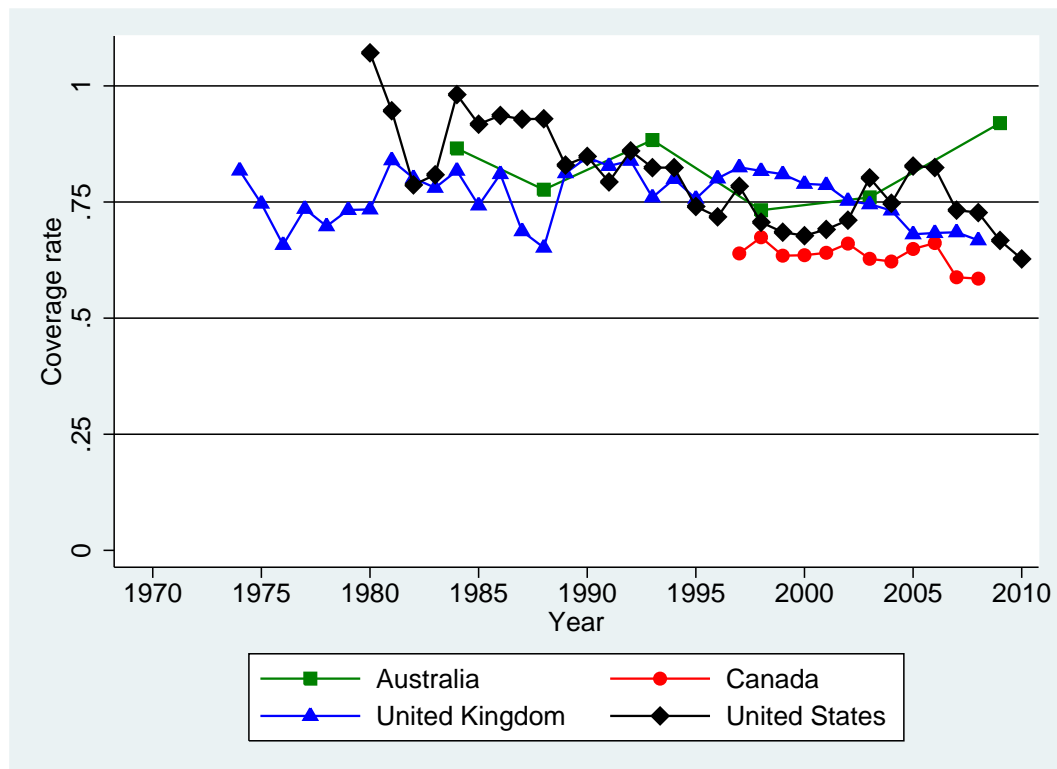
Notes: Coverage rate is the proportion of consumer expenditure in the national accounts that is accounted for in the household surveys in this category. Source is calculations by the authors.

Figure 9: Coverage Rates, New and Used Vehicles



Notes: Coverage rate is the proportion of consumer expenditure in the national accounts that is accounted for in the household surveys in this category. Source is calculations by the authors.

Figure 10: Coverage Rates, Furniture, Household Equipment, and Appliances



Notes: Coverage rate is the proportion of consumer expenditure in the national accounts that is accounted for in the household surveys in this category. Source is calculations by the authors.

Table 1: Features of the Data Sets

	Australia	Canada (SHS)	United Kingdom	United States
Recall vs. Diary	Diary (regular) Recall (irregular or large items)	Recall	Diary (regular) Recall (irregular or large items)	Recall/ (and diary)
Interview recall period	Varies; up to 3 years	Annual	About a year	Five quarterly surveys
Balance Edit?	No	Yes	No	No
Weighting benchmarks	Age, household size, state, labour force status, income source	Age, province, earnings, household size	Age, region and sex	Age, race, region, urban/rural status
Typical Sample Size	7,000	10,000 to 18,000	6,000	7,000/ (7,000)

Notes: Source is the documentation for the surveys in each of the four countries, as referenced in the text.

Table 2: Coverage and Response Rates

Dependent Variable: country-year coverage rate						
	(1)	(2)	(3)	(4)	(5)	(6)
	no controls	add country fixed effects	add top 1%	add interaction	with linear trend	add year fixed effects
Response Rate	0.084 (0.154)	0.779 (0.078)	0.342 (0.111)	0.407 (0.255)	0.345 (0.112)	0.337 (0.207)
Top 1% income share			-1.006 (0.203)	-0.642 (1.307)	-1.232 (0.362)	-1.026 (0.688)
Response X top share interaction				-0.487 (1.731)		
Canada		0.406 (0.020)	0.413 (0.017)	0.414 (0.018)	0.420 (0.020)	0.416 (0.028)
United Kingdom		0.237 (0.021)	0.197 (0.020)	0.199 (0.021)	0.204 (0.022)	0.191 (0.036)
United States		0.086 (0.019)	0.153 (0.021)	0.154 (0.022)	0.166 (0.027)	0.147 (0.047)
Linear trend					0.0007 (0.0010)	
Year fixed effects	no	no	no	no	no	yes
Adjusted R-Squared	-0.009	0.912	0.933	0.932	0.933	0.928
Number of Observations	81	81	81	81	81	81

Notes: Unit of observation is a country-year. Excluded country dummy is Australia. Standard errors below each parameter estimate in parentheses. Each column represents a separate regression using the coverage rate as dependent variable.