

Viewpoint: Measuring the well-being of the poor with income or consumption: a Canadian perspective

Matthew Brzozowski *Department of Economics, York University*
Thomas F. Crossley *Faculty of Economics, University of Cambridge*

Abstract. There is a long tradition of using consumption measures derived from Statistics Canada's household expenditures surveys to study material well-being, inequality, and poverty. We offer an introduction to this research. Income and consumption measures give different pictures of the patterns of material well-being in Canada, but the differences are not as large as in the US. We also provide a comparison to Meyer and Sullivan's results on data quality. Canadian expenditure surveys are of high quality. Unique aspects of these surveys (variation in quality control measures over time and the possibility of comparing to income tax data) provide important insights into the quality of survey data on income and consumption. JEL classification: D31, C81

Mesurer le bien-être des pauvres à l'aide du revenu et de la consommation : une perspective canadienne. Il existe une longue tradition d'utilisation de mesures de consommation (dérivées de l'enquête des dépenses des ménages de Statistiques Canada) pour mesurer le bien-être matériel, l'inégalité et la pauvreté. Les auteurs présentent ces recherches. Les mesures de revenus et de consommation donnent des portraits différents des patterns de bien-être matériel au Canada, mais les différences ne sont pas aussi importantes qu'aux Etats-Unis. On présente une comparaison avec les résultats de Meyer et Sullivan sur la qualité des données. Les enquêtes sur les dépenses sont de haute qualité au Canada. Certains aspects uniques de ces enquêtes (variation dans les mesures de contrôle de qualité dans le temps et possibilité de comparer les données dérivées des rapports d'impôt sur le revenu) fournissent des aperçus importants sur la qualité des données d'enquêtes sur le revenu et la consommation.

Author order is alphabetical. Crossley is also affiliated with the Institute for Fiscal Studies, London. This work was supported in part by the ESRC-funded Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (grant number RES-544-28-5001.) We thank Sule Alan, Mike Brewer, Martin Dooley, Martin Lemire, Bruce Meyer, Kevin Milligan, Cormac O'Dea, John Sabelhaus, and James Sullivan for helpful discussions and comments. All errors are our own. Email: brzozows@yorku.ca; Thomas.crossley@econ.cam.ac.uk

1. Introduction

Meyer and Sullivan (2011; henceforth **MS**) extend a series of papers (2003, 2004, 2008) in which the authors make two important contributions. First, they demonstrate that findings regarding recent trends in material well-being in the US are quite sensitive to whether one measures well-being with income or with consumption. Second, **MS** argue that consumption should be preferred to income as a measure of material well-being or living standards. The conceptual reasons for preferring consumption to income as a measure of living standards are well known. Economists typically think that it is the consumption of goods and services (including service flows from durables) that generates utility, and that income is desired primarily because it affords consumption. Moreover, if households are forward looking, and have declining marginal utility, they will try to smooth consumption, by borrowing or dis-saving. This means that consumption gives an indication of households' assessment of their overall financial position and prospects, while income may reflect fluctuations that have little impact on living standards. What is more novel in **MS** is that the authors' preference for consumption-based measures of living standards is not based primarily on these conceptual arguments, but rather is based on an assessment of data quality. They contend that, especially at the bottom of the distribution, the available expenditure micro data (from which consumption measures are constructed) are of higher quality than the available income data. This proposition runs against the standard view that consumption is harder to measure, at least in developed countries (see, e.g., Deaton 2007), but they provide a range of supporting evidence for this position.

In the conclusion to their paper, **MS** call for similar research in other countries. As it turns out, there is a long tradition of using consumption measures derived from Statistics Canada's household expenditures surveys to study material well-being, inequality, and poverty. In this note we offer an introduction to this research. We also provide a comparison to **MS**'s results on data quality. In the next section we describe the main income surveys and expenditure surveys in Canada. We also summarize some Canadian research on poverty and living standards that exploits the expenditure data. In section 3, we take up the question of relative quality of the expenditure and income data in Canada, paralleling and extending the assessment of US data in **MS**. Section 4 concludes with some suggestions for future research.

2. Canadian income surveys and expenditure surveys

2.1. Income surveys

Until the middle of the 1990s, the main income survey in Canada was Statistics Canada's Survey of Consumer Finances (SCF). Unlike the US survey of the same name, this was not a wealth survey, but rather an annual supplement to the

(April) Labour Force Survey (LFS). This supplement collected detailed income from individuals and households (thus, it was much like the March Supplement to the Current Population Survey in the US). The SCF field work was timed to coincide with the tax-filing season in Canada (April), in the hope of maximizing the quality of responses to income questions. Household level data is available for every year between 1976 and 1997.¹

Statistics Canada introduced the Survey of Labour and Income Dynamics (SLID) in 1993. While it has a (rotating) panel structure,² it can also be used cross-sectionally and cross-sectional public use files are available. It became the primary source of cross-sectional information on individual and household incomes in Canada from 1998. Like the SCF, the SLID is based on the LFS sampling frame and is meant to be nationally representative. In the SLID, household characteristics are collected in the January following a reference year, and income data are collected in the May following a given reference year. SLID respondents are given the option to skip the May interview by giving Statistics Canada permission to access their income tax records. More than 70 percent of the records in the data have income information derived from tax records. Neither the SCF nor the SLID collect detailed expenditure information.

2.1. *Expenditure surveys*

Statistics Canada has a long history of household budget surveys, going back to 1938 (McWhinney and Champion 1974). The Family Expenditure Survey (FAMEX) was conducted and the resulting micro data made publically available at irregular intervals from 1969 through 1996. From 1997 a revised survey, re-named the Survey of Household Spending (SHS) has been conducted annually. Both the FAMEX and SHS are based (like the SCF and SLID) on the LFS frame. However, resource constraints sometimes dictated that the FAMEX be limited to an urban sample: nationally representative FAMEX surveys were collected in 1969, 1978, 1982, 1986, 1992, and 1996.³ The FAMEX and SHS collect detailed expenditure information at the household level as well as household demographics and household income information.

The FAMEX and SHS differ substantially from the US Consumer Expenditure Survey (CE) and other national budget surveys. As described by **MS**, the CE has both an interview sample and a separate sample, which keeps expenditure diaries. The interview sample is a short panel lasting five quarters. While

- 1 The SCF was also conducted every other year from 1965 to 1971, and some data from this period is publicly available. However, differences in the basis on which the data are organized and incompatibilities in survey weights mean that many researchers focus on the data from 1976 onward. See Brzozowski et al. (2010) for further discussion.
- 2 For each sampled household in the SLID, up to 12 interviews are conducted over a six-year period.
- 3 There are, of course, small changes over time in the surveys, so that even with the nationally representative FAMEX and SHS, some work is required to construct a consistent series. See Crossley and Pendakur (2006) or Brzozowski et al. (2010) for more details.

the first contact is face-to-face, an increasing number of the follow-up interviews (in which the expenditure data are collected) are conducted by telephone. MS focus exclusively on the data collected from the interview sample; many other researchers use a mix of the interview and diary data, as does the BLS in producing official tables (see Attanasio, Battistin, and Padula 2007 for more details). The use of diaries is common in national budget surveys.

In contrast, the FAMEX and SHS are based entirely on recall in face-to-face interviews. Interviews are conducted in the first quarter of a year to collect household expenditure and income information for the entire previous year. For example, the 2006 data were collected in January, February, and March of 2007 but refer to the 2006 calendar year. As is true for the SCF, the timing of data collection to coincide with tax-preparation season may improve the quality of the data if households are thinking about their finances at this time of year. Historically, data capture in the field was by paper and pencil; computer-assisted personal interviewing was introduced in 2006 (more on this below.)

The annual recall required by the SHS/FAMEX is unusually long (the CE interviews largely ask about expenditures over the previous quarter). This is likely to exacerbate recall problems. On the other hand, short recall questions and diary procedures (which typically collect expenditures on non-durable items over periods as short as one or two weeks) suffer from problems of purchase infrequency that are not a problem in the FAMEX and SHS.⁴

Statistics Canada expends considerable effort to enhance the quality of the data collected. During the in-home interview, households are asked to consult bills and receipts. A key quality control measure is the 'balance edit' by which annual income, annual expenditure and changes in assets are reconciled in the field. Households that are significantly out of balance are probed to identify possible errors questioning. The data are widely thought to be of high quality. For example, Bosworth, Burtless, and Sabelhaus (1991) compare the Canadian data favourably to the CE.

2.2. Research using the Canadian expenditure surveys

As is the case for other national budget surveys, the primary purpose of the FAMEX and the SHS is to provide inputs to the construction of goods baskets for price indices such as the Consumer Price Index. However, there is also a strong tradition of using the micro-data from the FAMEX and SHS for economic and social research and, in particular, to study material well-being, poverty, and inequality in Canada.⁵

4 Purchase infrequency refers to the fact that some goods and services are purchased by households periodically or irregularly; so in a diary covering one or two weeks, a household may (accurately) record zero *expenditure* on a good or service that they are, in fact, consuming.

5 These micro data have also been used to study household demands (e.g., Nicol 1989 or Blundell, Duncan, and Pendakur 1998); life-cycle patterns of consumption (e.g., Robb et al. 1989); heterogeneity and aggregation of consumption (e.g., Fortin 1995); saving behaviour (e.g.,

Pendakur (2001), Crossley and Curtis (2006), and Milligan (2008) all use the FAMEX/SHS data to study consumption poverty to Canada and to compare it with income poverty. Crossley and Curtis (2006) focus particularly on child poverty, while Milligan (2008) and Baker, Gruber, and Milligan (2009) compare consumption and income poverty among the elderly. Pendakur (1998), Pendakur and Crossley (2006), and Brzozowski et al. (2010) compare consumption inequality with income inequality.

Unsurprisingly, this literature finds that the level of consumption poverty is lower than income poverty and, similarly, consumption inequality is exceeded by income inequality. Incomes are simply more dispersed than consumption. A limitation of the Canadian data, relative to the CE, is that rental equivalents are not collected from households that own their homes. Thus, if housing services are to be included in consumption, rents must be imputed for owning households, usually by a regression model that relates rents to a small number of housing attributes among renting households. These imputations are typically not very variable, and hence they have the effect of reducing the dispersion of consumption. A number of authors (Crossley and Curtis 2006; Milligan 2008) have noted that the level of consumption poverty, and sometimes the magnitude of changes over time, are quite sensitive to the treatment of housing services.

With respect to the evolution of inequality, a typical finding has been that income inequality and consumption inequality move up and down together, but that movements in consumption inequality are more muted than corresponding movements in income inequality (Pendakur 1998; Brzozowski et al. 2010). Turning to measured poverty in Canada, Pendakur (2001) reports that consumption poverty fell by more than income poverty during the 1970s and 1980s, but rose by more in the 1990s. Poverty measures and inequality measures capture different aspects of the distribution (poverty measures being more sensitive to the bottom), so these results do not necessarily contradict the findings for inequality.

There are also interesting differences between income and consumption poverty over the life-cycle. For example, Milligan (2008) documents a sharp increase in income poverty in the immediate pre-retirement years which is not apparent in the consumption data.

Much of Meyer and Sullivan's previous work has focused on the living standards of lone parents and how these have been affected by welfare reforms in the US. For example, Meyer and Sullivan (2004) analyze the effects the US TANF reforms on the living standards of single mother households. They demonstrate that, relative to estimates based on expenditure data, the use of after-tax income leads to an overstatement of the severity of the policy effects.

Burbidge and Davies 1994; Carroll, Rhee, and Rhee 1994 or Engelhardt 1996); equivalence scales (e.g., Phipps 1998 or Pendakur 1999); the allocation of resources within households (e.g., Browning et al. 1994 or Phipps and Burton 1998); tax incidence (e.g., Curtis and Kingston-Riechers 2007); and more. It is difficult to overstate the research value of these surveys.

The FAMEX/SHS have been used for similar research on this particularly vulnerable group in Canada. Brzozowski (2007) analyzes the effects of welfare reforms in the province of Ontario in the second half of 1990s on the living standards of single-mother-headed households. In contrast to Meyer and Sullivan (2004), Brzozowski finds very similar results whether he uses income or consumption as his measure of material well-being. This is true even when he employs control groups that are very similar to those employed by Meyer and Sullivan.⁶

In summary, there is a strong tradition in Canada of using data from household budget surveys for economic research, including research on the questions considered by Meyer and Sullivan in the US. We now turn to assessment of the quality of the income and consumption data.

3. The quality of income and expenditure micro data in Canada

A novel aspect of **MS** is the authors' argument that consumption is to be preferred to income as a measure of material well-being on the basis of data quality considerations. They support their position with a range of evidence on the relative quality of the income and expenditure micro data, and in this section we provide a similar assessment of the Canadian data sources. Following **MS**, we consider sample sizes and the precision of estimates, survey, and item non-response and under-reporting. Because the FAMEX and SHS do not have a panel component, panel attrition, which **MS** investigate, is not relevant when the Canadian data are considered. We then go on to consider some unique aspects of the Canadian surveys (variation in quality control measures over time; the possibility of comparing with income tax data) that provide additional insights into the quality of survey data on income and consumption.

3.1. Sample sizes and precision

Table 1 provides information on sample sizes for Canadian income surveys and household budget surveys. The SCF sample size was typically around 40,000, while SLID cross sections are typically about 30,000 households. These are not as large as the comparable Current Population Survey CPS Annual Social and Economic Supplement (March Supplement) in the US, which currently includes approximately 100,000 households. Nevertheless, they are large surveys, particularly as Canada's population is about one tenth of the US population.

In the years that it was a nationally representative survey, the FAMEX has a sample size of about 10,000 households, about one-third larger than the CE.

6 Both Meyer and Sullivan (2004) and Brzozowski (2007) analyze the effects of welfare reform in a difference-in-difference framework. However, in Canada welfare policy falls under provincial jurisdiction, and there has been considerable provincial variation in welfare reforms. This allows Brzozowski to employ demographically identical but geographically separate control groups, in addition to the demographic control groups employed by Meyer and Sullivan.

TABLE 1
Sample sizes and response rates in selected surveys

Year	Consumption				Income			
	Family Expenditure Survey		Survey of Household Spending		Survey of Consumer Finances (Economic Family File)		Survey of Labour and Income Dynamics (EE FILE)	
	N	RR (%)	N	RR (%)	N	RR (%)	N	RR (%)
1978	9536	72						
1979					37,440	66		
1982	10938	81			37,765	78		
1984	4792	75			36,413	79		
1986	10333	77			32,822	79		
1990	4562	72			45,580	79		
1992	9492	74			40,007	81		
1996	10,417	77			35,689	82	31,793	87
1997			23,842	76			32,240	85
1998			20,236	76			32,721	84
1999			23,518	73			30,410	83
2000			20,989	70			30,212	79
2001			22,172	76			30,928	79
2002			20,861	71			29,436	79
2003			23,869	72			29,846	78
2004			20,446	69			28,936	75
2005			21,331	71			28,197	75
2006			20,436	72			28,524	75
2007			21,407	65			27,843	70

The CE, though, has multiple observations on each household. The transition from FAMEX to SHS involved doubling of the sample sizes to about 20,000 observations. These very large samples are motivated by the desire to provide provincial level estimates. Thus, the current Canadian budget survey has two-thirds the sample of the corresponding income survey and has almost three times as many households as the CE. Concerns about household budget survey sample sizes do not seem relevant in Canada.

MS report that measured consumption (CE) is much less variable than measured income (CPS) and that this means that even with significantly smaller sample sizes, estimates based on consumption may have smaller standard errors. While similar results can be obtained from Canadian data, we do not report them for several reasons. First, of course, the very large sample size of the SHS

makes this issue less important in Canada. Second, we believe that this kind of precision comparison, while interesting, is subject to several important caveats.

One concern is that some narrow consumption bundles (food, for example) are strongly necessities and so less variable than either income or total consumption. It's not clear this is a good thing – it may also make them less sensitive to important differences in broad living standards.

A second point is that broad consumption bundles often contain imputed service flows from housing and other durables. Imputed (or predicted) rents and service flows are typically not very variable (because they are based on a small number of measured characteristics of the stocks). Including them substantially reduces the variability of the consumption bundle. Again, this is not necessarily a good thing, as real variability is missed. While this might not matter for means or conditional means, it certainly does matter for inequality and poverty rates and other measures that depend on dispersion. This turns out to be a particular problem with the Canadian household budget surveys, because, unlike the CE, the Canadian budget surveys do not collect estimates of rental equivalents from home-owning households. Housing services, which are a significant fraction of consumption, must therefore be imputed. As noted above, both Crossley and Curtis (2006) and Milligan (2008) report sensitivity of poverty measurement to the treatment of household services when data from the Canadian budget surveys are used. The standard errors of income or consumption estimates can be reduced by imputing important components of the measure, but it is not clear when this would represent an improvement in data quality.⁷

3.2. Survey and item non-response

As noted by MS, survey (or unit) non-response and item non-response are important measures of data quality. Survey or unit non-response is the failure of units (here households) to participate in the survey; this can arise because the statistical agency fails to contact a household that has been selected for interview or because a contacted household refuses to participate. Item non-response occurs when the statistical agency fails to obtain a measurement on a particular survey item from a participating unit.

Survey response rates are included in table 1. The SCF response rate through the 1980s and 1990s was around 80%. The SLID was initially even better, at around 85%, but there has been a subsequent decline, to about 70%. The FAMEX had response rates of about 75%. This was initially matched by the SHS, but the SHS has exhibited a subsequent decline in response rates to about 65–70%. The SHS response rates are still quite good, particularly given the high respondent

⁷ A further concern is that the FAMEX, SHS, SCF, and SLID are multistage, stratified samples. Low-population regions (such as the Atlantic provinces) are over-sampled. Statistics Canada provides sample weights, but not cluster information, in the public use files. The omission of cluster information means that standard errors are likely to be underestimated. Although all of these surveys are based on the same sampling frame, it is not at all clear that the degree of underestimation of standard errors is the same across surveys.

load for this survey, and we do not believe that the difference in response rates represents a reason to favour the income data.

The FAMEX and SHS do involve some imputation in the case of item non-response. Unfortunately, there is no flagging of imputation in the public use files. Some sparse information on imputation rates is provided in the documentation that accompanies the public use files. With respect to the household budget surveys, total income is imputed in about 1% of cases and some imputation of expenditures is required in about 12% of cases (excepting some components of clothing expenditure). This is roughly comparable to what **MS** report for the CE.

3.3. *Under-reporting*

MS pay particular attention to the issue of under-reporting in the income and expenditure surveys. There are several ways to assess this. One approach is to compare estimates of aggregate income or expenditure, 'grossed up' from the surveys, to the National Accounts. There are several reasons why the aggregate estimates should be different, including differences in concepts and coverage. Nevertheless, comparison to National Accounts has become a standard way of assessing the quality of household budget surveys. Attanasio, Battistin, and Padula (2010) emphasize that the aggregate expenditures implied by the CE fall significantly short of personal sector consumption reported in the National Accounts, even when one makes the adjustments necessary to make household survey and National Accounts numbers as comparable as possible. Perhaps more worrying, the gap appears to be growing over time. **MS** acknowledge the growing disparity in between the CE and National Accounts in measured total expenditure, but they argue that, from the point of view of measuring the living standards of the poor, this may be less of a problem than it first appears. In particular, they point out that there is much less apparent under-reporting of expenditures in the CE, and less of a trend in apparent under-reporting if one focuses on the interview sample only and if one focuses on a subset of goods that make up a large portion of the expenditure basket of poorer households.

In figure 1 we compare aggregate income and expenditure, 'grossed up' from the Canadian surveys using appropriate survey weights, to the Canadian National Accounts. The data are drawn from Brzozowski et al., (2010).

Aggregate income estimates based on the SCF fall short of the National Accounts figure, and the deficit grows over time from about 10% to 15%. This is perhaps surprising, given the improvements in response rate over this period. The trend is reversed with SLID data. Here the survey coverage of the National Accounts benchmark improves over time and, paradoxically, this happens as response rates fall. Interestingly, the aggregate income figures derived from the expenditure surveys (FAMEX and SHS) tend to match national accounts better than the aggregate income figures derived from the income surveys (SCF and SLID).

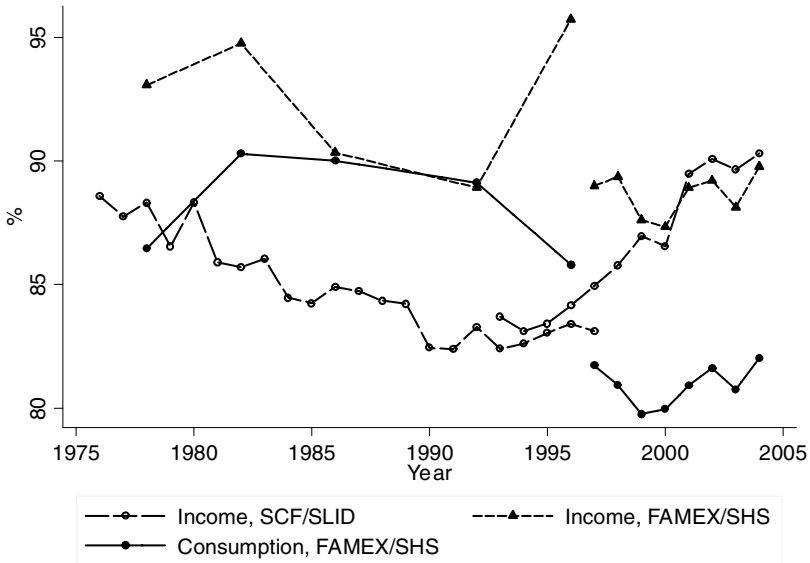


FIGURE 1 National Accounts coverage ratios
SOURCE: Brzozowski et al. (2010)

Concerning expenditures, figure 1 shows that the expenditure data in the FAMEX have about 90% coverage of the national accounts expenditure benchmark. This is *better* than the ratio of the income data in the SCF to national accounts income, but not quite as good, on average, as the income data in the FAMEX. The SHS does not do quite as well: survey coverage of the National Accounts expenditure benchmark is 80% to 85%.

In summary, the Canadian household budget surveys look quite good against this criterion. The income surveys do not appear to do significantly better against a National Accounts benchmark, and the expenditure coverage of the Canadian data appears to be significantly better than in the CE.

A second way to investigate potential under-reporting is to validate individual components of income or expenditure against external standards. MS are particularly concerned with potential under-reporting of transfer income among poorer households. This appears to be a problem with the income micro data in Canada as well. Aggregates of reported amounts of transfer income – particularly social assistance – are significantly below estimates derived from administrative sources (Dooley 1994; Kapsalis 2001; Warbarton and Warbarton 2004).

Finally, one can consider the cross-sectional distribution of income and expenditure in the household budget surveys and assess whether the apparent rates of saving (or dis-saving) in the data are plausible. Figure 2 is based on the 2005 SHS and plots the means of equivalized disposable income and equivalized total

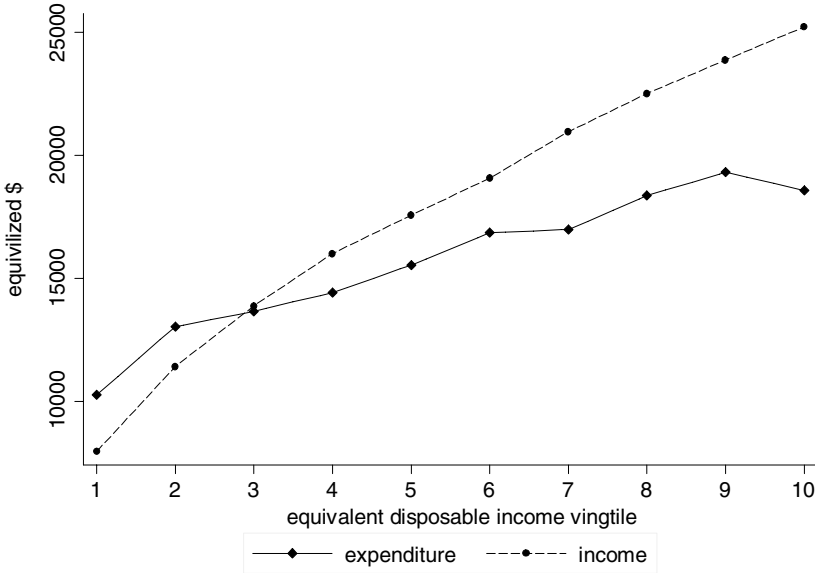


FIGURE 2 Equivalized disposable income and equivalent total expenditure by income vingtile, Survey of Household Spending, 2005

NOTES: Definitions of disposable income and total expenditure follow Meyer and Sullivan (2011) as closely as possible. Disposable income is defined as (total household income before taxes)-(taxes paid). Total expenditure is defined as total expenses incurred during the survey year for food, shelter, household operations, household furnishings and equipment, clothing, transportation, health care, personal care, recreation, reading materials, education, tobacco products and alcoholic beverages, games of chance, and a miscellaneous group of items, but excluding personal taxes, personal insurance payments, vehicle purchases, tuition, medical expenses and gifts and contributions. Both disposable income and total expenditure are equivalized using the equivalence employed by Meyer and Sullivan: $(Adults + 0.7 * Children)^{0.7}$.

expenditure by vingtiles of income.⁸ We find considerable apparent dis-saving among households with the lowest current disposable income. There are several possible explanations for this observation. First, it might be genuine dis-saving, as households with temporarily low incomes smooth consumption. It was one of the key insights of Friedman’s original work on the permanent income hypothesis that consumption might be relatively flat with respect to income in cross-section.

⁸ Definitions of disposable income and total expenditure follow MS. Disposable income is defined as (total household income before taxes) – (taxes paid). Total expenditure is defined as total expenses incurred during the survey year for food, shelter, household operations, household furnishings and equipment, clothing, transportation, health care, personal care, recreation, reading materials, education, tobacco products and alcoholic beverages, games of chance, and a miscellaneous group of items, but excluding personal taxes, personal insurance payments, vehicle purchases, tuition, medical expenses, and gifts and contributions. Both disposable income and total expenditure are equivalized or adjusted for differences in need resulting from differences in household size and composition of the household. This is standard in inequality and poverty measurement and is done here using the equivalence scale employed in MS: $(Adults + 0.7 * Children)^{0.7}$.

Alternatively, it may indicate mis-reporting: income may be under-reported or expenditure may be over-reported.

MS strongly favour interpreting the dis-saving at the bottom of income distribution as indicative of income under-reporting. They argue that over-reporting of expenditures is unlikely and is inconsistent with the fact that aggregate survey consumption falls short of the National Accounts. They further argue that the extensive dis-saving at the bottom of the income distribution suggested by the CE data is implausible, given the very low wealth levels of poorer households. This argument has been made more formally by Sabelhaus and Groen (2000). The degree of dis-saving at the bottom of the income distribution that **MS** report for the CE greatly exceeds the dis-saving we find at the bottom of the income distribution in the FAMEX and SHS. **MS** report that in the CE, the ratio of mean total expenditure to mean after-tax income in the bottom vingtile of the income distribution is about 7 (a saving rate of -600%). We find that the same ratio in the FAMEX and SHS, while greater than 1, is less than 2, and typically significantly so. We next consider a quasi-experiment in the conduct of the SHS that provides a clue to understanding the degree of apparent dis-saving at the bottom of the income distribution in the CE.

3.4. A quasi-experiment in survey design

As noted in section 2, the field methodology for the FAMEX and SHS has normally included a data quality control measure called the 'balance edit.' The edit identifies households in which reported expenditure is more than 20% different from the sum of disposable income and net cash money flows into assets. The interviewer is instructed to try to collect additional information from such households in order to balance expenditure with income and changes in assets and debts. Then, at the processing stage, household records that are still 'out of balance' (by more than 20%) are deemed unusable and deleted both from the public use data and from the calculations for standard tables that Statistics Canada releases. Because the balance edit is conducted in the field, it was not possible to examine the effects of the balance edit on an observation by observation basis. However, Statistics Canada has reported that most of the resulting field adjustments are to reported income and reported asset changes.

In 2006, Statistics Canada adopted computer-assisted personal interviewing (CAPI) for the SHS (prior to 2006, field data collection for the FAMEX, and then SHS, was by paper and pencil). In this first year of CAPI, the balance edit was not applied. Without the field balance edit, the number of 'out of balance' (by more than 20%) records increased from 546 in 2005 to 4,300 in 2006 (29% of completed questionnaires) Statistics Canada decided that it could not discard this many records, so unbalanced records are included in the data files for 2006. The balance edit was reintroduced (within CAPI) in 2007 (Statistics Canada 2008).

It is therefore possible to infer the effect of the balance edit by comparing 2006 data with data from 2005 and 2007. We do this in figures 3 through 5. Figure 3

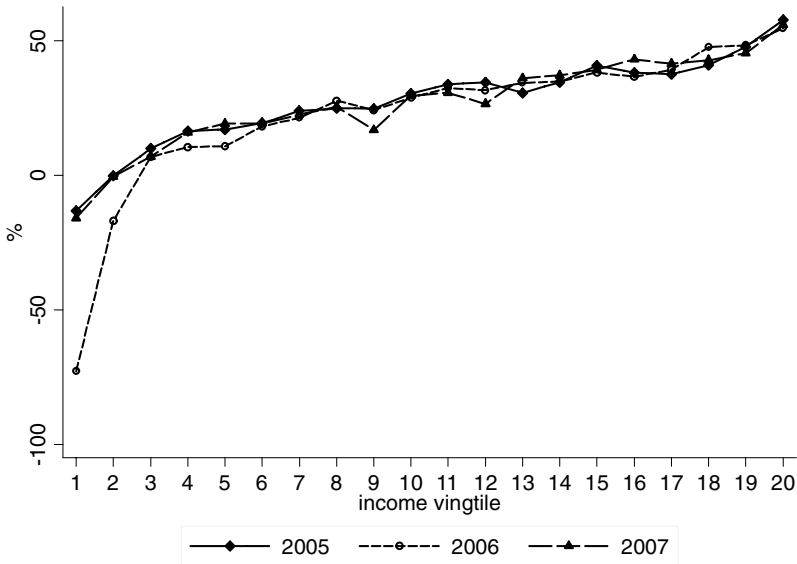


FIGURE 3 Mean savings rate by income vingtile, Surveys Of Household Spending, 2005, 2006, 2007

NOTES: Equivalent disposable income and total expenditure are defined as described in the notes to figure 2. The saving rate is defined as disposable income minus total expenditure, all divided by disposable income.

plots the average saving rate in the SHS (the difference between disposable income and total expenditure, divided by after-tax income) by vingtile of disposable income. The data for 2006 (the survey with no balance edit) is represented with open dots. The data from 2005 is represented with solid dots, and the data from 2007 plotted with triangles. There is much more apparent dis-saving in 2006 (when there was no balance edit).

In figure 4 we plot, for each year, and for each of the bottom five disposable income vingtiles, mean equivalized disposable income (left-hand panel) and mean equivalized total expenditure (right-hand panel). The left-hand panel shows under-reporting of disposable income in 2006. The right-hand panel shows higher total expenditure in 2006 than in 2005 and 2007 at the bottom of the *income* distribution. Note that this is consistent with (i) over-reporting of *expenditure* by low income households or (ii) under-reporting of *income* by households that are, in fact, moderately well-off (so that some households with moderate expenditures end up at the bottom of the income distribution in 2006). In figure 5 we plot mean equivalized disposable income (left-hand panel) and mean equivalized total expenditure (right-hand panel) by *expenditure* vingtile. There is no difference in the patterns of expenditure by expenditure vingtile between 2006 and 2005 or 2007. This is consistent with (ii) but not with (i). This evidence confirms that the key problem that the balance edit corrects is under-reporting of income (not

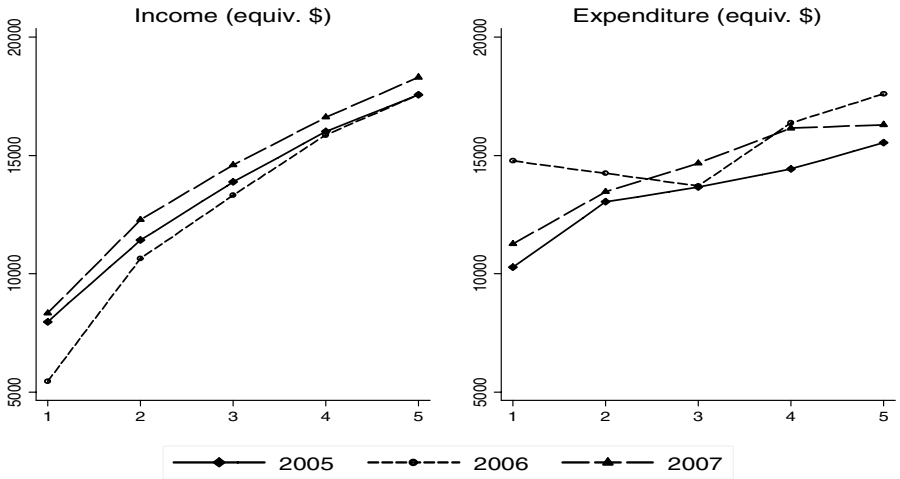


FIGURE 4 Equivalized disposable income and equivalized total expenditure, income quintiles 1 to 5, Surveys of Household Spending, 2005, 2006, 2007

NOTES: See notes to figures 2 and 3.

over-reporting of expenditure) and, in particular, it appears to eliminate some incorrectly reported very low incomes.

This in turn supports the view, expressed in *MS*, that the high rate of dis-saving among low-income households in the CE reflects a problem of income under-reporting. Note that even in 2006 SHS, with no field balance edit, the ratio of mean total expenditure to mean after-tax income in the bottom quintile of the income distribution is less than 2, so that there is much less apparent dis-saving than in the CE. One possible explanation is that the entire structure of FAMEX and SHS pushes the interviewer and respondent towards budget balance, and this in itself improves income reporting in the SHS and FAMEX. For example, in the SHS and FAMEX incomes and all expenditures are reported for identical reference periods. This is not a feature of the CE. This is a hypothesis that warrants future investigation.

Because the balance edit eliminates some income under-reporting among lower-income households in the FAMEX and SHS, the bottom percentiles of disposable income should be higher in the FAMEX and SHS than in the SCF. The SCF has no comparable check on reported incomes, as it does not collect information on expenditures. The first three columns of table 2 show that this proposition is correct. Table 2 compares percentiles of disposable income and expenditure distributions derived from various data sources for 1996. This is an interesting year to consider, because the SCF, SLID, and FAMEX were conducted in 1996. Column 1 reports mean disposable income by income quintiles in the SCF. Columns 2 and 3 report corresponding figures from the SLID and the FAMEX, respectively. Results for the 1996 SLID and 1996 SCF are taken

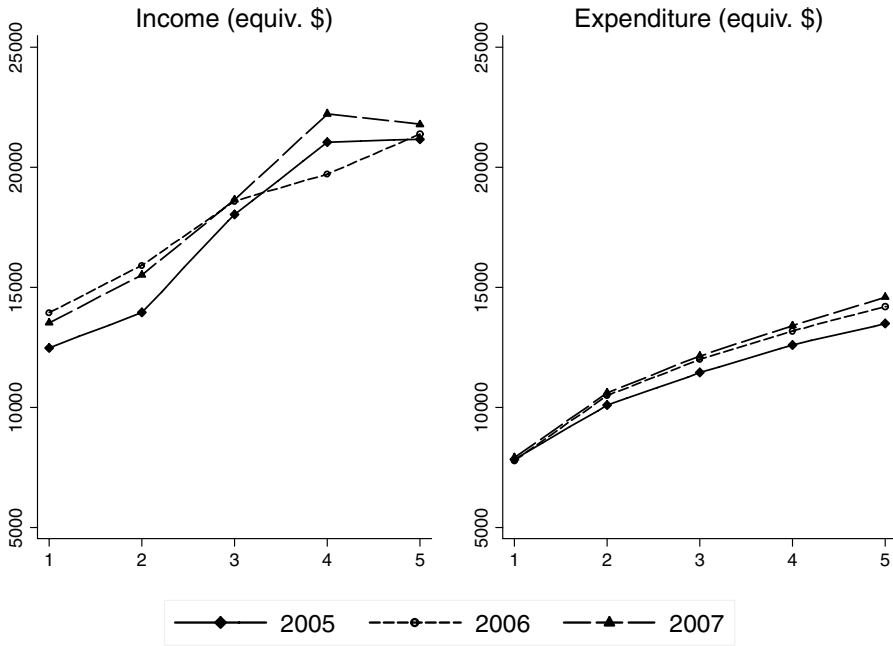


FIGURE 5 Equivalized disposable income and equivalized total expenditure, expenditure quintiles 1 to 5, Surveys of Household Spending, 2005, 2006, 2007

NOTES: See notes to figures 2 and 3

TABLE 2
Mean disposable equivalent income and expenditure by income quintile, 1996 (year 1996 Can\$)

	SCF (1)	SLID (2)	FAMEX		
			Disposable income (3)	Total expenditure (4)	Tax records (5)
<i>Vingtile</i>					
1	4,585	3,859	5,543	10,405	1,530
2	8,429	8,126	9,280	10,228	6,187
3	10,725	10,183	11,261	11,156	8,508
4	12,640	11,992	12,826	12,118	10,274
8	18,329	17,659	18,934	15,052	16,618
10	21,220	20,705	21,817	16,228	19,781
<i>Response rate</i>	82%	87%	77%		

NOTES: All amounts are expressed in 1996 Can\$. Survey weights used in all calculations use survey data. Following Frenette et al. (2004), the equivalence scale is the square root of household size. SOURCE: SCF, SLID, and Tax Record numbers are from Frenette et al. (2004, table 3b). FAMEX numbers are the authors' own calculations.

from Frenette, Green, and Picot (2004), while results for the 1996 FAMEX are our own calculations.

Differences between these estimates can arise in a number of ways. Income under-reporting, and survey (or unit) non-response are two important ones. As noted above, the balance edit should minimize income under-reporting in the FAMEX and lead to higher estimates of disposable income at the bottom of the distribution in the FAMEX than in the SCF, and this indeed seems to be the case. However, because a large fraction of the income information in the SLID is derived from tax records, we would also expect under-reporting to be less of a problem in the SLID than in the SCF, and estimates of mean disposable income in the bottom vingtiles of the income distribution to be *higher* in the SLID than in the SCF. Perhaps surprisingly, table 2 reveals that this is not the case: estimates of mean disposable income in the bottom vingtiles of the income distribution appear to be *lower* in the SLID than in the SCF (which is in turn lower than in the FAMEX).

Survey non-response is one possible explanation for this finding. Note that survey non-response was lower in the SLID than in either the SCF or the FAMEX (the response rate for the 1996 SLID was 87%, versus 82% in the SCF and 77% in the FAMEX). Survey non-response may be non-random, and it may be a particular problem among very low-income households. If the households that are 'missing' from the SCF (and the FAMEX), relative to the SLID, are drawn from the very bottom of the disposable income distribution, this could explain why the estimates of mean disposable income in the bottom vingtiles of the income distribution appear to be *lower* in the SLID than in the SCF (and the FAMEX). This brings us to a final issue that has received attention in the Canadian literature on poverty but not, to our knowledge, in the corresponding US literature: the population coverage of household surveys.

3.5. The population coverage of household surveys

While the response rates obtained by Statistics Canada's household surveys are very good, the survey non-response that does occur could seriously bias estimates of the tails of the distribution of material well-being if non-responding households are more likely to be drawn from those tails. Frenette, Green, and Picot (2004) and Frenette, Green, and Milligan (2007) have studied this problem in detail by comparing estimates of the income distribution derived from household surveys (the SCF and SLID) to estimates derived from data sources that do not suffer from survey non-response: tax records and the Canadian Census. Historically, participation in the Census has been compulsory, while refundable tax credits targeted at low-income households have made tax filing almost universal in Canada since the early 1990s.

Column 5 of table 2 reproduces, from Frenette, Green, and Picot (2004), mean disposable income by income vingtiles in the Canadian tax data. Estimates of mean disposable income in the bottom vingtiles of the income distribution are

significantly lower in the tax data than in either the SLID or the SCF. Frenette et al. (2004) conclude that survey non-response in the SLID and SCF is a particular problem in the lower tail of the income distribution (they also show that it is a problem in the upper tail, but that tail is not our focus here.)

To summarize: from the point of view of estimating well-being of the poor, the household income surveys seem to have two problems – under-reporting of income by responding low-income households and survey non-response by low-income households. The resulting biases appear to be offsetting: under-reporting leads to underestimates of the incomes of the poor, while non-random survey non-response leads to overestimates of the incomes of the poor.

Expenditure seems to be relatively well reported by the poor, and the balancing of income with expenditure and changes in assets that is done in the FAMEX and SHS seems to improve income reporting by the poor. However, table 2 suggests that, at least in 1996, the reduction of income reporting achieved in the FAMEX led to estimates of mean income at the bottom of income distribution that were *more* biased (relative to estimates based on tax records) than estimates from either the SCF or the SLID, perhaps because under-reporting offsets the effects of non-random survey non-response.

These issues clearly warrant further study, and it is hardly an attractive empirical strategy to choose data with more biases in the hope that they are offsetting. The results do suggest that when tax or other administrative data are available, they should be included in the set of sources that researchers consider.

4. Conclusions

Motivated by the work of **MS**, we have reviewed the evidence on consumption- and income-based studies of material well-being in Canada. Statistics Canada has a long-running program of household budget surveys, and data from those surveys have been the basis of a substantial body of research using consumption estimates to study poverty, inequality, and material well-being in Canada. Our reading of that literature is that income and consumption measures give different pictures of the patterns of material well-being in Canada, but that the differences are perhaps not as large as in the US.

We also follow **MS** in comparing the quality of the available data on income and consumption. Canadian expenditure surveys (SHS and FAMEX) are of high quality, with large samples, good response rates, and good coverage of benchmarks in the National Accounts. Variation in the use of budget balance as a quality control in the SHS suggests that income under-reporting at the bottom of the income distribution is likely an important problem in income surveys (which cannot check budget balance) in both Canada and the US. However, comparisons with tax data that are thought to be very accurate suggest that the Canadian surveys – both income and expenditure – suffer from non-random survey non-response, which leads to a failure to accurately capture the tails of

the income distribution. Presumably, estimates of the tails of the consumption distribution are similarly affected.

The availability of administrative data on incomes, but not on consumption, could be advanced as an argument for income over consumption as the preferred measure of the well-being of the poor. However, given the conceptual arguments for using consumption as a measure of well-being, which were rehearsed in the introduction, we would argue instead for a renewed effort to minimize the effects of survey non-response. Strategies for doing so might well make use of administrative data on incomes. Statistics Canada already calibrates the survey weights for the SHS to the tax data on incomes in a fairly crude way. The procedure for doing so could surely be improved, and a particular focus on the tails of the distribution would be welcome.

The issues considered here and in **MS** have also been taken up in other countries, notably the UK (Brewer, Goodman, and Leicester 2006; Brewer et al. 2009). Another promising avenue of research would be work based on international comparisons that fully exploited differences in survey design and administration across countries to draw lessons for the accurate measurement of well-being.

References

- Attanasio, O., E. Battistin, and M. Padula (2009) *Inequality in Living Standards since 1980* (Washington: American Enterprise Institute for Public Policy Research Press)
- Baker, M., J. Gruber, and K. Milligan (2009). 'Retirement income security and well-being in Canada,' NBER Working Paper No. 14667
- Blundell, R., A. Duncan, and K. Pendakur (1998) 'Semiparametric estimation of consumer demand,' *Journal of Applied Econometrics* 13, 435–61
- Bosworth, B., G. Burtless, and J. Sabelhaus (1991) 'The decline in saving: some microeconomic evidence,' *Brookings Papers on Economic Activity* 1, 183–256
- Brewer, M., A. Goodman, and A. Leicester (2006) 'Household spending in Britain: what can it teach us about poverty?' (Bristol: Policy Press)
- Brewer, M., C. O'Dea, G. Paull, and L. Sibieta (2009), 'The living standards of families with children reporting low incomes,' Department for Work and Pensions, Research Report No. 577
- Browning, M., F. Bourguignon, P. Chiappori, and V. Lechene (1994) 'Income and outcomes: a structural model of intrahousehold allocation,' *Journal of Political Economy* 102, 1067–96
- Brzozowski, M. (2007) 'Welfare reforms and consumption among single mother households: evidence from Canadian provinces,' *Canadian Public Policy* 33, 227–50
- Brzozowski, M., M. Gervais, P. Klein, and M. Suzuki (2010) 'Consumption, income, and wealth inequality in Canada,' *Review of Economic Dynamics* 13, 52–75
- Burbidge, J.B., and J.B. Davies (1994) 'Household data on saving behavior in Canada,' in *International Comparisons of Household Saving*, ed. J.M. Poterba (Chicago: University of Chicago Press for the NBER)
- Carroll, C.D., B. Rhee, and C. Rhee (1994) 'Are there cultural effects on saving? Some cross-sectional evidence,' *Quarterly Journal of Economics* 109, 685–99
- Crossley, T.F., and L. Curtis (2006) 'Child poverty in Canada,' *Review of Income and Wealth* 52, 237–59

- Crossley, T.F., and K. Pendakur. (2006). 'Consumption inequality,' in *Dimensions of Inequality in Canada*, ed. D. Green and J. Kesselman (Vancouver: UBC Press)
- Curtis, L., and J. Kingston-Riechers (2007) 'The GST and the family: evaluating the welfare effects of the GST'
- Deaton, A. (1997) *The Analysis of Household Surveys* (Baltimore and London: John Hopkins University Press for the World Bank)
- Dooley, M. (1994) 'The converging market work patterns of married mothers and lone mothers in Canada,' *Journal of Human Resources* 29, 600–20
- Engelhardt, G.V. (1996) 'Tax subsidies and household saving: evidence from Canada,' *Quarterly Journal of Economics* 111, 1237–68
- Fortin, N.M. (1995) 'Heterogeneity biases, distributional effects, and aggregate consumption,' *Journal of Applied Econometrics* 10, 287–311
- Frenette, M., D.A. Green, and K. Milligan (2007) 'The tale of the tails: Canadian income inequality in the 1980s and 1990s,' *Canadian Journal of Economics* 40, 734–64
- Frenette, M., D.A. Green, and G. Picot (2004) 'Rising income inequality in the 1990s: an exploration of three data sources,' Statistics Canada, Analytical Studies Research Paper Series, No. 219
- Kapsalis, C. (2001) 'An assessment of EI and SA reporting in SLID,' Statistics Canada, Analytical Studies Research Paper Series, No. 166
- McWhinney, I., and H.E. Champion (1974) 'The Canadian experience with recall and diary methods in consumer expenditure surveys,' *Annals of Economic and Social Measurement* 3, 411–35
- Meyer, B.D., and J.X. Sullivan (2003) 'Measuring the well-being of the poor using income and consumption,' *Journal of Human Resources* 38, S1180–220
- (2004) 'The effects of welfare and tax reform: the material well-being of single mothers in the 1980s and 1990s,' *Journal of Public Economics* 88, 1387–420
- (2008) 'Changes in the consumption, income, and well-being of single mother headed families,' *American Economic Review* 98, 2221–41
- (2011) 'Further results on measuring the well-being of the poor using income and consumption,' *Canadian Journal of Economics* 44, 52–87
- Milligan, K. (2008) 'The evolution of elderly poverty in Canada,' *Canadian Public Policy* 34, s79–s94
- Nicol, C.J. (1989) 'Estimating a third-order translog using Canadian cross-sectional micro-data,' *Canadian Journal of Economics* 22, 543–60
- Pendakur, K. (1998) 'Changes in Canadian family income and family consumption inequality between 1978 and 1992,' *Review of Income and Wealth* 44, 259–83
- (1999) 'Estimates and tests of base-independent equivalence scales,' *Journal of Econometrics* 88, 1–40
- (2001) 'Consumption poverty in Canada 1969 to 1998,' *Canadian Public Policy* 27, 125–49.
- (2002) 'Taking prices seriously in the measurement of inequality,' *Journal of Public Economics* 86, 47–69
- Phipps, S. (1998) 'What is the income cost of a child? Exact equivalence scales for Canadian two-parent families,' *Review of Economics and Statistics* 80, 157–64
- Phipps, S.A., and P.S. Burton (1998) 'What's mine is yours? The influence of male and female incomes on patterns of household expenditure,' *Economica* 65, 599–613
- Sabelhaus, J., and J.A. Groen (2000) 'Can permanent-income theory explain cross-sectional consumption patterns?' *Review of Economics and Statistics* 82, 431–8
- Statistics Canada (2008) *User Guide for the Survey of Household Spending, 2006*, Household expenditures research paper series, Ottawa.
- Warburton, R.N., and W.P. Warburton (2004) 'Canada needs better data for evidence-based policy: inconsistencies between administrative and survey data on welfare dependence and education,' *Canadian Public Policy* 30, 241–55