

Consistency between Household-level Consumption Data from Registers and Surveys

Kim Abildgren
Danmarks Nationalbank
kpa@nationalbanken.dk

Andreas Kuchler
Danmarks Nationalbank

America Solange L. Rasmussen
Statistics Denmark

Henrik Sejerbo Sørensen
Statistics Denmark

High-quality household-level data on savings and consumption with detailed household characteristics play an important role in a wide range of empirical research. Traditionally, such micro data were only available for rather limited samples sizes from on interview-based surveys. However, surveys are expensive and it has been subject to discussion whether the data quality has declined (Meyer et al., 2015). One way forward might be more focus on exploiting data from administrative registers. In some countries, the administrative registers now allow for imputation of household-level panel data with detailed household characteristics covering the entire population at relatively low costs.

Browning and Leth-Petersen (2003) and Kreiner et al. (2015) imputed household-level consumption figures 1994-1996 via information from the Danish administrative registers on income and changes in wealth. They found that the imputed consumption data were rather noisy but nevertheless performed reasonably well in terms of matching the median and mean level of consumption reported for the same households in the interview-based Danish Expenditure Surveys 1994-1996. The worst match was found for households with significant holdings of bonds or equities due to lack of information that allowed for control for capital gains and losses. Broadly similar studies have been made for Norway over the period 1994-2014 (Fagereng and Halvorsen, 2017; Eika et al., 2017) and Sweden 2003-2007 (Kojen et al., 2015).

All the studies mention above have focussed on the consistency between register-imputed and survey-based consumption figures in levels. There have not previously been any

studies on the consistency of consumption figures at a household level in first differences. Consumption figures in first differences are frequently used in empirical analyses of consumer behaviour. Furthermore, there have not been any studies on the consistency between register-imputed and survey-based consumption figures in Denmark based on data from the last twenty years although this period has seen several advances in the coverage of the administrative register on households. A new register on mortgage loans now allows the user to control for capital gains and losses in relation to housing loans and another new register contains population wide household-level information on car purchases.

The paper at hand has two main aims. The first is to explore the consistency at a household-level between register-imputed and survey-based consumption in levels as well as in first differences. As basis for the analyses, we will impute several different measures on household-level consumption for the entire population of households based on data from Danish administrative registers over the period 1994-2015. We will then compare the figures with detailed consumption data for the same households in the interview-based Danish Expenditure Surveys over the same period and analyse the characteristics (family composition, life cycle position, income, wealth, homeownership, self-employed versus wage and salary earners, consumption pattern, asset and liability composition, region of residence, etc.) of those households where the largest differences occur. Furthermore, we will explore whether the differences are related to macroeconomic factors such as the state of the business cycle or household-level events such as income shocks due to unemployment, illness or inheritance. The analysis might be used to identify possible suggestions for future improvements of the information content in the administrative registers or the content and design of the interview-based expenditure surveys.

The second aim of the paper is to explore the consistency between consumption functions in levels and first differences estimated on the basis of register data and similar consumption functions estimated on the basis of survey data. The sizes of the parameters of the consumption function are of crucial importance for an assessment of the marginal propensity to consume out of income or wealth or the impact of leverage on consumption and thereby for analyses of macroeconomic stability.

References

Browning, Martin and Søren Leth-Petersen. 2003. Imputing consumption from income and wealth information, *Economic Journal*, Vol. 113(488), pp. F282-F301.

Eika, Lasse, Magne Mogstad and Ola L. Vestad. 2017. What can we learn about

household consumption expenditure from data on income and assets?, unpublished working paper, March.

Fagereng, Andreas and Elin Halvorsen. 2017. Imputing consumption from Norwegian income and wealth registry data, *Journal of Economic and Social Measurement*, Vol. 42(1), pp. 67-100.

Koijen, Ralph, Stijn Van Nieuwerburgh and Roine Vestmanz. 2015. Judging the quality of survey data by comparison with 'truth' as measured by administrative records: evidence from Sweden. In: Christopher D. Carroll, Thomas F. Crossley and John Sabelhaus (eds.). *Improving the Measurement of Consumer Expenditures*, Chicago: University of Chicago Press, pp. 308-346.

Kreiner, Claus Thustrup, David Dreyer Lassen and Søren Leth-Petersen. 2015. Measuring the Accuracy of Survey Responses Using Administrative Register Data: Evidence from Denmark. In: Christopher D. Carroll, Thomas F. Crossley and John Sabelhaus (eds.). *Improving the Measurement of Consumer Expenditures*, Chicago: University of Chicago Press, pp. 289-307.

Meyer, Bruce D., Wallace K. C. Mok, and James X. Sullivan (2015), Household Surveys in Crisis, *Journal of Economic Perspectives*, Vol. 29(4), pp. 199-226