On Measurement Error in Imputed Consumption Data

Scott Baker* Lorenz Kueng† Steffen Meyer‡ Michaela Pagel§

very preliminary and incomplete

November 29, 2017

Abstract

Survey-based measures of consumption are the basis of a great deal of empirical work despite suffering from considerable measurement error. For that reason, Koijen et al. (2014) and a number of follow-up papers develop and use an alternative measure commonly referred to as "imputed consumption." The basic idea is to measure consumption as a residual from the household's budget constraint: consumption is the part of total income that was not invested. As another alternative, a number of recent papers use high-frequency transaction-level spending as a measure of consumption (Kuchler, 2015; Gelman et al., 2014; Baker, 2014; Olafsson and Pagel, 2016; Kueng 2017). In this paper, we use transaction-level data on balances, asset trades, and asset holdings to document the short-comings in using even the most detailed annual registry snapshots of wealth in imputing consumption. We show that substantial discrepancies between imputed and actual spending arise whenever investors buy and sell assets at different points within the year and whenever they incur trading costs. The reason is that asset prices fluctuate considerably over any given year and we know since Barber and Odean (2000) that many retail investors 1) have very large turnover, 2) have significantly different levels of returns, and 3) incur substantial trading costs that eat up to the entire historical risk premium.

*Northwestern University, Kellogg
†Northwestern University, Kellogg and NBER
‡Leibniz Universität Hannover
§Columbia Business School, NBER and CPER

1 Introduction

Economists need accurate measures of spending to research household decision making, on consumption and saving behavior, on inequality, poverty, and standards of living, as well as aggregate fluctuations in consumption-based asset pricing models for instance. Moreover, economists need to understand how accurate the measurement of spending really is. Spending data are difficult to come by especially when seeking measurement error free and detailed spending data over a long enough period of time.

Consumption surveys use paper or phone interviews to ask stylized questions on spending in a few broad consumption good categories over a particular recall period or households are asked to keep track of recurrent expenditures, such as groceries, for a short period of time in a diary. An existing literature has found basic problems with survey-based measures of consumption (Koijen et al., 2014a; Ahmed et al., 2006). Ahmed et al. (2006) compare two measurements for the same set of households and find that recall food consumption data, which is the basis of a great deal of empirical work, suffers from considerable measurement error, while diaries records are found to be somewhat more accurate. For obvious reasons, survey respondents may have little incentive to answer the questions accurately, may not understand the wording of the questions, or may behave differently in practice and forget their reported behavior. Moreover, such measurement error or noise in the data generated by surveys that simply ask about past purchases can increase with the length of the recall period (de Nicola and Giné, 2014). Additionally, surveys can produce biased (rather than merely noisy) data if respondents have justification bias, concerns about surveyors sharing the information, or stigma about their consumption habits (Karlan and Zinman, 2008). Other work has compared consumption measures across different surveys or across different waves of the same survey (Pudney, 2008; Bound et al., 2001). The measurement error in household-level consumption data, and the difficulty of estimating nonlinear models in the presence of such error, have led some to call for abandoning Euler equation estimation altogether (Carroll, 2001).

Koijen et al. (2014a) and a number of follow-up works (Browning and Leth-Petersen, 2003; Fagereng and Halvorsen, 2017; Bach et al., 2015; Fagereng et al., 2016; Sodini et al., 2016; Fagereng et al., 2016, 2017) develop and use an alternative measure of consumption, which avoids many of the problems with standard survey-based data, and is commonly referred to as "imputed consumption." The basic idea is to measure consumption as a residual from the household's budget constraint: consumption is the part of total income that was not invested. This approach imposes heavy data requirements on the measurement exercise because one needs comprehensive measures of income as well as comprehensive asset holdings and asset price data. Sweden, Norway, and Denmark collects that information as part of its tax registry. For the Swedish case, the tax registry is most comprehensive, containing data on every stock, bond, mutual fund, and bank account each household owns at the end of the year. Moreover, homeownership and household permanent address can be tracked via the housing registry and the data also contains information on labor, transfer, and financial income. In comparison to the Swedish data, the Norwegian and

Danish data lack information on actual assets as well as transfers.

The registry-based or imputed consumption measure covers all spending on non-durables or durables at annual frequency. Sweden also runs a standard Household Budget Survey that Koijen et al. (2014a) can match up with the households in the survey to the registry data. This setup allows them to compare registry-imputed and survey-based measures of consumption. The authors uncover discrepancies between registry-and survey-based consumption measures that increase with income and wealth. While the mean and median of the consumption distribution are similar, the survey understates the consumption of wealthy and high-income households, while slightly overstating consumption of the poorest quintile of households. Moreover, Koijen et al. (2014a) show that registry-based consumption is sensitive to an accurate imputation of returns that households are earning on their assets. The authors show that incorrectly applying a broad total return measure to a household's financial asset holdings leads to substantial deviations from the properly imputed registry measure and that these discrepancies are increasing in wealth. This critique applies to a number of published papers such as (Browning and Leth-Petersen, 2003; Fagereng et al., 2017, 2016).

Instead of using survey-based measures of consumption or imputed consumption, a number of recent papers use high-frequency transaction-level spending as a measure of consumption (Kuchler, 2015; Gelman et al., 2014; Baker, 2014; Olafsson and Pagel, 2016; **?**). In this paper, we use transaction-level data on balances, asset trades, and asset holdings to document the short-comings in using even the most detailed annual registry snapshots of wealth in imputing consumption. More specifically, discrepancies between imputed and actual consumption arise whenever investors buy and sell assets at different points within the year and whenever they incur trading costs. Since Barber and Odean (2000), it is well-known that some retail investors have very large turnover and significantly different levels of returns and may incur substantial trading costs that eat up to the entire historical risk premium. More recently, the same has been shown to be the case for Scandinavian and German investors (Bach et al., 2015; Koestner et al., 2017).

Our study accomplishes the following goals. First, we show that intra-year trading is a significant driver of variation across households in overall capital gains, realized and unrealized. We can compare our findings to what would be computed using only year-end snapshots of household portfolios to show the importance of more detailed data. Second, we show that these errors are not normally distributed across households. That is, we show that intra-year trading and variation in capital gains is correlated with both income and wealth. Third, we can employ the richness of the asset holdings, trades, and spending data to understand the effects of realized and unrealized capital gains on household spending. We can look at the marginal propensity to consume out of realized and unrealized capital gains as well as dividends and capital gain distributions when the respective assets are held for a short or long period of time. We also examine how trading behavior, portfolio composition, and returns respond to changes in household income within individuals not only cross-sectionally. We are interested in whether receiving a

negative income shock, for instance, getting laid off, leads to more or less trade and increases or decreases in risk-taking.

Finally, we can demonstrate a simple measurement error model that allows for both error in imputed and transaction-level measures of consumption and we compare the relative magnitudes of the error, following Koijen et al. (2014a).

2 Literature review

The Swedish data (Koijen et al., 2014b; Sodini et al., 2016; Bach et al., 2015, among others) includes information on the change in bank accounts (checking, savings, certificates of deposit, etc) in 2006 and 2007 if the balance is greater than 10,000 SEK (˜1,224 USD) and in prior years if the earned interest exceeds 100 SEK (˜12.24 USD). Furthermore, funds, stocks, and bonds are observed including their actual sale prices.

The Norwegian data (Fagereng and Halvorsen, 2017; Fagereng et al., 2016, 2017, among others) includes broad asset classes of bank deposits, bonds including money market funds, mutual funds, and stocks. Moreover, there is a self-reported category cash holdings that is commonly ignored though. In turn, the authors take the historical annual return of the Oslo Stock Exchange (30%) and MSCI World Index (70%) to impute the returns of mutual funds and the Treasury bill to impute the returns on bonds. Moreover, for the period 2005 to 2011, the portfolio details are observed, allowing for the same measurement as in Swedish data.

3 Data

To investigate the extent of measurement error in imputed consumption data, we use a unique panel dataset on the daily trading of 103,000 private investors in Germany spanning the years 2003 to 2015. We thus precisely measure each trader's daily activity from his log in to every single trade. We are able to identify quasi-automatic trades, such as savings plan transactions, and trading decisions in our sample are not moderated by any influence from third parties, such as financial advisors. Our sample includes only self-directed online investors. Moreover, we obtain data on customer demographics such as gender, age, and occupation as well as detailed information on traded securities such as asset class, risk class, issuer or issue date of a security. This data set has been used in other published papers such as Schmittmann et al. (2014). Beyond all portfolio holdings and trades, we have the time series of checking and savings account balances for all investors. We measure wealth as the sum of all assets an investor holds plus his checking and savings account balances. We measure consumption as the sum of all outflows that are not related to trades or transfers and compare this measure when we either use all trades or use only the year-end holdings.

In online brokerage data, silent attribution is a common phenomenon, as usually there is no charge for having an account. Thus, it is important to note that the online bank, from

which we obtained the data, is not only one of the largest retail brokerages in Germany but also a multi-services retail financial institution offering checking and savings account services as well as overdraft facilities, credit cards, mortgage and auto credits, insurance, and retirement savings vehicles. Moreover, because cash payments are still prevalent in Germany, the bank offers a dense network of ATM machines. Additionally, in Germany, holding multiple checking accounts is discouraged as this hurts individual credit scores. Furthermore, the bank offers online transfers of automatic checking account transactions as well as a financial account aggregation facility. Nevertheless, our measure of consumption may be biased if investors transfer money into other accounts and save rather than consume it. Therefore, we undertake multiple steps to ensure that we restrict the sample to individuals who use their accounts not only for trading but also for all everyday transactions.

Table 1 displays summary statistics for our sample including mean and median checking and savings account balances. Moreover, the table shows monthly turnover, as defined by Barber and Odean (2000), and the sum of monthly financial fees paid, including all load fees. As one can easily see, turnover and fees paid can be substantial and are likely to introduce measurement error in imputed consumption.

Table 1: Summary statistics for checking and savings account balances, portfolio values, monthly fees, and monthly turnover

	mean	standard deviation	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
checking acct balances	4,167	40,535	-276	43	552	3,501	12,886
savings acct balances	30,197	86,554	.54	101	6,339	28,710	71,344
portfolio value	669,599	6851768	3,110	12,450	39,505	103,191	435,098
monthly fees	118	5,830	0	0	0	26	194
monthly turnover	3.8	81	.00043	.00092	.0026	.025	.18
observations	1.1e+07						

Financial fees paid would be counted as consumption and thus consistute a direct source of measurement error. With respect to capital gains and losses, we can compute the measurement error by simply computing the costs of each trade (excluding fees) and comparing the actual costs to the imputed costs based on the price of the asset on the 1st of July in each year. Table

2 displays summary statistics for the monthly discrepancies arising between actual costs and imputed costs. Moreover, we aggregate the discrepancies by subtracting the discrepancy in buys from that in sells for the monthly and also to the yearly level.

Table 2: Summary statistics for the sum of discrepancies between actual trade costs and imputed trade costs of all monthly buys and all monthly sells

	mean	standard deviation	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
monthly positive discrepancy	362	35,389	0	0	0	0	0
monthly negative discrepancy	-280	14,174	0	0	0	0	0
monthly absolute discrepancy	641	40,091	0	0	0	0	0
monthly total discrepancy	83	36,028	0	0	0	0	0
yearly total discrepancy	1,052	224,109	0	0	0	0	0
observations	5406357						

Table 3 displays summary statistics for our sample income, savings, and spending. Because we do not observe all transactions, we proceed as follows: we first eliminate all the changes in balances because of trades. Then, we define income as all positive changes in checking account balances minus all negative changes in savings account balances and spending as all negative changes in checking account balances minus all positive changes in savings account balances. Thus, we decompose all non-trade-related changes in balances into income, spending, and (de-)saving.

Table 3: Summary statistics for monthly income (sum of inflows into checking account minus outflows from savings account) and monthly spending (sum of outflows from checking account minus inflows to savings account)

	mean	standard deviation	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
monthly income	15,637	288,950	-91	0	43	2,222	15,311
monthly spending	9,032	147,372	-375	0	0	1,437	11,948
monthly (de-)saving	-562	164,288	-5,562	0	0	91	4,675
observations	7708344						

References

Ahmed, N., M. Brzozowski, and T. F. Crossley (2006). Measurement errors in recall food consumption data. Technical report, IFS Working Papers, Institute for Fiscal Studies (IFS).

Bach, L., L. E. Calvet, and P. Sodini (2015). Rich pickings? risk, return, and skill in the portfolios of the wealthy.

Baker, S. (2014). Debt and the consumption response to household income shocks. Working paper, Stanford University.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance 55*(2), 773–806.

Bound, J., C. Brown, and N. Mathiowetz (2001). Measurement error in survey data. *Handbook of econometrics 5*, 3705–3843.

Browning, M. and S. Leth-Petersen (2003). Imputing consumption from income and wealth information. *The Economic Journal 113*(488), pp. F282–F301.

Carroll, C. D. (2001). A theory of the consumption function, with and without liquidity constraints. *The Journal of Economic Perspectives 15*(3), 23–45.

de Nicola, F. and X. Giné (2014). How accurate are recall data? evidence from coastal india. *Journal of Development Economics 106*(0), 52 – 65.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2016). Heterogeneity in returns to wealth and the measurement of wealth inequality. *The American Economic Review 106*(5), 651–655.

Fagereng, A., L. Guiso, and L. Pistaferri (2017). Portfolio choices, firm shocks, and uninsurable wage risk. *The Review of Economic Studies*, rdx023.

Fagereng, A. and E. Halvorsen (2017). Imputing consumption from norwegian income and wealth registry data. *Journal of Economic and Social Measurement 42*(1), 67–100.

Fagereng, A., M. B. Holm, and G. J. Natvik (2016). Mpc heterogeneity and household balance sheets.

Gelman, M., S. Kariv, M. D. Shapiro, D. Silverman, and S. Tadelis (2014). Harnessing naturally occurring data to measure the response of spending to income. *Science 345*(6193), 212–215.

Karlan, D. and J. Zinman (2008). Lying about borrowing. *Journal of the European Economic Association 6*(2/3), pp. 510–521.

Koestner, M., B. Loos, S. Meyer, and A. Hackethal (2017). Do individual investors learn from their mistakes? *Journal of Business Economics 87*(5), 669–703.

Koijen, R., S. Van Nieuwerburgh, and R. Vestman (2014a). Judging the quality of survey data by comparison with" truth" as measured by administrative records: Evidence from sweden. In *Improving the Measurement of Consumer Expenditures*, pp. 308–346. University of Chicago Press.

Koijen, R., S. Van Nieuwerburgh, and R. Vestman (2014b). Judging the quality of survey data by comparison with "truth" as measured by administrative records: Evidence from sweden. In *Improving the Measurement of Consumer Expenditures*. National Bureau of Economic Research, Inc.

Kuchler, T. (2015). Sticking To Your Plan: Empirical Evidence on the Role of Present Bias for Credit Card Debt Paydown. *Working Paper*.

Olafsson, A. and M. Pagel (2016). The liquid hand-to-mouth: Evidence from personal finance management software. Technical report, Working Paper, Columbia University.

Pudney, S. (2008). Heaping and leaping: Survey response behaviour and the dynamics of self-reported consumption expenditure. Technical report, ISER Working Paper Series.

Schmittmann, J. M., J. Pirschel, S. Meyer, and A. Hackethal (2014). The impact of weather on german retail investors. *Review of Finance 19*(3), 1143–1183.

Sodini, P., S. Van Nieuwerburgh, R. Vestman, and U. von Lilienfeld-Toal (2016). Identifying the benefits from home ownership: A swedish experiment. Technical report, National Bureau of Economic Research.