The Value of Registry Data for Consumption Analysis: An Application to Health Shocks

Jonas Kolsrud[*]   Camille Landais[†]   Johannes Spinnewijn[§]

Uppsala University   London School of Economics   London School of Economics

October 24, 2019

Abstract

This paper measures consumption expenditures using registry data on income and asset holdings in Sweden. We show how a registry-based measure complements traditional survey measures of consumption and can alleviate some critical limitations. We describe the construction of our measure, which builds on prior work and exploits the identity coming from the household budget constraint between consumption expenditures and income net of savings. We demonstrate the value of the registry-based measure to study consumption responses to shocks, also relative to surveyed consumption. In our application to health shocks, we find that Swedish household experience permanent earning drops, but generous social transfers provide substantial consumption smoothing. We document important heterogeneity in consumption responses and the limited role for private means.

Keywords: Consumption, Measurement, Registry Data, Health Shocks

1 Introduction

The access to administrative data from tax, social security and other government registries has been the driver of a real data revolution in economic research and in the field of public economics in particular. Researchers around the world have been doing enormous data collection efforts using these registers to document trends in inequality in income, wealth, health, social mobility, etc which are having an important impact on the public debate. This data revolution has lead to new empirical breakthroughs that should help evaluating policy: the analysis of income responses to tax incentives, unemployment responses to unemployment benefits, health expenditure changes in response to health coverage, etc. While the analysis of household consumption and how it responds

to income shocks has remained central in economic modelling, with a few recent exceptions the ever-growing literature has been mostly confined to the use of survey data on consumption.

Traditional measures of consumption based on surveys have some clear advantages and can provide information that may be impossible to attain otherwise, but they suffer from important limitations as well. Most evidently, the relative small samples of traditional budget surveys make aggregate or distributional analysis challenging and affect the precision with which consumption responses to adverse events or income shocks can be estimated. This paper builds on the important contribution of Koijen et al. [2014] to construct an alternative, registry-based measure of consumption using comprehensive, detailed information on income and wealth from Swedish tax registers for the *universe* of households in Sweden. The basic idea underlying the construction is that household consumption equals household income minus changes in the household's asset positions. Importantly, the universal coverage of the registry data clearly helps overcoming the small sample limitation of survey data.[1]

The registry-based measure we construct has additional advantages that can contribute positively to the quality of empirical research on consumption. First, beyond the universal coverage, administrative records enable the construction of long panels, which allow us to study consumption dynamics. Second, the measure allows to analyze precisely how different means are used to smooth consumption in response to shocks and reduce the variance in consumption. We can shed light on the use of debt and different types of assets, but also on the role played by other members of the household or by the government through taxes and transfers, etc. Third, the registers allow to identify income shocks or specific events underlying income shocks (e.g., unemployment, disability), by linking tax registers to social security registers. Finally, the registry-based measure benefits from the accuracy and reliability of the third-party reported administrative data. We will illustrate these advantages in an application and highlight also the complementary value the registry-based measure brings to standard survey measures.

Importantly, the quality of the registry-based measure critically depends on the comprehensiveness of the available data on income and changes in the asset portfolio. To identify consumption expenditures from a household's budget constraint, one requires data on total income and all changes in asset holdings. For example, the consumption out of bank accounts from one period to the next equals the interest earned minus the change in the account holdings, both of which are precisely recorded in the Swedish administrative registers. However, measuring the contribution of changes in other asset types to consumption can be challenging. In an ideal setting all asset transactions are observed and one can simply subtract purchases from sales to calculate the corresponding flow of consumption expenditures. In practice, the way asset variables are recorded in tax registers may not be directly relevant for determining the consumption flow, like when levying a tax on realized capital gains or on the stock of wealth. We exploit the tax registry - which supported the wealth tax Sweden has in place until 2007 - and actually recorded information on the end-of-year balances

---

[1]See Pistaferri [2015] for a broader discussion of the measurement issues with household consumption and the use of other data sources (e.g., scanner data, data from financial aggregators).

(i.e., quantities) of all listed financial assets. This allows us to calculate the re-balancing from one year to the other. Even when this data is missing, one can still try to exploit information on prices to impute the contribution from changes in balance values to the consumption flow.

In the first part of the paper, we present our approach on how to use registry information to construct a measure of consumption and explain in detail how these can be applied in Sweden. In addition to the challenges with the financial assets, we also discuss issues around durable goods and in particular the role of real estate and the imputation of rents for the measurement of consumption. As a first check, we confirm that aggregate consumption using the registry-based measure compares well to aggregate consumption measured in the national accounts. But we also exploit the availability of the annual Household Budget Survey (HUT, Hushållens Utgifter), which can be matched to the registry data and thus allows us to compare our registry-based measure with survey-based measures of consumption. The distribution of the registry-based measure is similar to the distribution of the survey measure and the correlation between the two is very strong. Still, the measures tend to differ more systematically in the tails. Importantly, the survey measure tends to underestimate consumption levels relative to the registry-based measure at the high end of the distribution. This reflects the well-known challenge for surveys to incentivize high-income households to accurately report their expenditures on the one hand and to capture the tail of Pareto distributions on the other hand.

In the second part, we demonstrate the value of the registry-based measure and its longitudinal dimension to analyze how consumption responds to health shocks. A large body of economic research has analysed how idiosyncratic income changes affect household consumption (e.g., Blundell et al. [2008], Heathcote et al. [2014], Blundell et al. [2016]). Stemming back to the Life Cycle-Permanent Income Hypothesis of Modigliani-Brumberg and Friedman, this literature relates movements in consumption, mostly relying on survey data, to anticipated and unanticipated income changes as well as persistent and non-persistent shocks to resources. Instead we focus on a well-specified event, a health shock, that has a substantial impact on individuals' income and expenditures. Our focus on this specific event circumvents the challenging identification of the specific nature of income shocks and is inspired by the direct relevance for policies that condition on these events (e.g., Baily [1978], Chetty [2006]). We use event-study designs which complement our related work in the context of unemployment, studying how household consumption responds to job loss (see Landais and Spinnewijn [2019]) and evolves during the unemployment spell (see Kolsrud et al. [2018]).

While a series of papers have investigated the effect of health shocks on labour supply, the literature on consumption responses to health shocks is still relatively small. This is mostly due to the difficulty of finding reliable panel data with information on both consumption and health status. Important exceptions are Stephens [2001], Meyer and Mok [2018] and Autor et al. [2019]. Our study contributes to this literature by using our registry-based measure of consumption to document how much households are insured against such shocks. In particular, we exploit the registry-based measure to provide a rich investigation of heterogeneity in these responses, according to the type

3

and severity of the shock, as in Meyer and Mok [2018], but also depending on various individual characteristics that may be relevant for targeting insurance and transfer policies. Our measure also allows for providing a detailed analysis of the underlying mechanisms of consumption smoothing, and the relative role of various forms of insurance or self-insurance such as social insurance and transfer programs, precautionary savings, spousal labor supply, etc.

We find that health shocks are associated with a substantial drop in earnings that persists over several years and is exacerbated by a drop in earnings for other household members. The drop in earnings results in a significant drop in consumption, but the latter drop is substantially smaller. We find that most cushion is provided by sickness and disability benefits, while financial assets play a very limited role in smoothing consumption, leading to (or consistent with) a permanent decline in consumption expenditures of around 5% relative to a 10% drop in household earnings. We also show how the average responses mask some important heterogeneity, for example by age and income. The consumption response is also larger when health shocks are larger, which lead to larger drops in earnings, both for the individual subject to the health shock and other household members. Overall, we are not only exploiting the panel nature and the possibility of demonstrating the different means of consumption smoothing in these applications, but we are also very much relying on the power of the registry-based measure. In comparison, the estimates obtained using consumption surveys are much more noisy and an order of magnitude less precise. The consumption surveys, however, provide complementary insights on the types of expenditures that respond to the health shock (e.g., health expenditures increase), which are important to translate the consumption responses into welfare conclusions.

All our programs and detailed documentation of the data are available online at the following address: `https://sticerd.lse.ac.uk/_new/research/pep/consumption/default.asp` and hopefully will help other researchers in their analysis of household consumption. The construction of the registry-based measure of consumption is ongoing work. While the Swedish context is quite ideal, we hope that our efforts can also spur more research to further improve our measurement and assess underlying measurement error (e.g., Eika et al. [2017], Baker et al. [2018]), but also extend its measurement beyond the period 1999-2007 in Sweden and to other countries with potentially less precise information on income or wealth. The registry-based measure of consumption will always be subject to measurement error, but the consequences of such measurement error fundamentally depend on the research question. It is less of a concern when consumption is used as a dependent variable, like in our application, and measurement error is arguably uncorrelated with the independent variable of interest.

The remainder of our paper is organized as follows. In the final part of this introduction, we explain how this paper relates to the existing literature dedicated to building consumption meaures out of registry data. In section 2, we lay out the methodology used to build residual measures of expenditures out of household budget constraints and information on income and wealth. In section 3, we explain the way we implemented this methodology in the Swedish context given the data available. In section 4, we show various descriptive statistics for our registry measure

of consumption expenditures. We show how it compares to other measures of consumption, from national accounts or survey data. In section 5, we further demonstrate the value of our measure to the analysis of consumption responses to health shocks. Section 6 concludes.

Related Literature   In our attempt to construct an alternative, registry-based measure of household consumption, our work relates directly to a literature initiated by Browning and Leth-Petersen [2003]. The original goal of this stream of papers was to investigate issues with existing consumption survey measures and how to improve upon them using registry data (cf. for instance Kreiner et al. [2014]). The comparison of consumption survey measures and registry-based measures of course is facilitated when the survey data and registers can be linked, like in Denmark and Sweden. In this context, Browning and Leth-Petersen [2003] were the first to measure consumption expenditures as the residual from what one earns minus what one saves, exploiting information on both income and wealth in Denmark. Being unable to properly impute unrealized capital gains, Browning and Leth-Petersen [2003] assumed away capital gains in the construction of their registry-based measure and that the benefits of this measure of consumption were limited.

To overcome this limitation, subsequent work has tried to obtain additional data in order to properly impute capital gains. In the context of Sweden, Koijen et al. [2014] have taken the first key step using disaggregated information on financial assets (stocks, bonds, mutual funds, etc.) to determine the proper contribution of changes in these financial assets to consumption. They also use some information on real estate transactions in order to disentangle unrealized capital gains from active rebalancing of households' real assets. Our paper directly builds on Koijen et al. [2014], using similar data in Sweden, extended for a longer time span and for the universe of the Swedish population.[2] We further improve upon their measure of consumption expenditures by properly imputing rents to homeowners and by refining the computation of real estate transaction flows. Since Koijen et al. [2014] do not consider imputed rents in households' capital income, they have to operate under sample restrictions to address the non-comparability of homeowners' and renters' expenditures. We also identify further opportunities with the available registers in Sweden, which should allow to reduce the measurement error even more. The value of these efforts will very much depend on the application the consumption measure is used for.[3]

Our paper also relates to two recent papers trying to build similar consumption measures from administrative data in Norway, the only Scandinavian country that still has a wealth tax in place. Fagereng and Halvorsen [2017] use account balances for individual stocks similar to Koijen et al. [2014].[4] Compared to Koijen et al. [2014], they include imputed rents from owner-occupied housing, but they also resort to sample restrictions to mitigate measurement error concerns regarding the

---

[2]Koijen et al. [2014] focus on a subsample of roughly 11,000 households that are both present in the LINDA panel ($\approx 300,000$ individuals) and in one cross-section of the consumption expenditure survey over the period 2003 to 2007.

[3]Similarly to Kreiner et al. [2014], Koijen et al. [2014] use registry information to inform the magnitude of potential biases in survey data. Using car registers, they investigate the importance of recall errors in survey data regarding purchases of large durables.

[4]The data on financial asset holdings is limited to stocks publicly traded in Norway, and is only available for years 2004-2011.

used market value of real estate wealth.[5] Interestingly, Fagereng and Halvorsen [2017] also use data on inheritances and gifts, which is available in Sweden as well. Most related to our paper is Eika et al. [2017], who further improve the comprehensiveness of the registry-based consumption measure. Compared to Fagereng and Halvorsen [2017], they use data on real estate transactions and also impute consumption flows for cars using registry data, which is also available in Sweden. Our paper is similar in spirit, both in the construction of the registry-based measure and the use of applications to demonstrate the value-added of registry-based measures of consumption expenditures. Unique to the Swedish context is that the Budget Surveys are linkeable to the registry data so that we can revisit our empirical applications using the survey measure and thus make the comparison with this traditional approach more explicit.

2   Residual Measure of Expenditure: Methodology

We start by revisiting the measurement of consumption using registry data (see also Browning and Leth-Petersen [2003], Koijen et al. [2014] and Eika et al. [2017]). Like the prior work, we start from the observation that any krona of income is either spent or invested. A household's budget constraint indeed leads to an accounting identity between expenditures on the one hand and income net of the change in assets on the other hand. Consumption expenditures C_{it} by household i in period t can thus be written as

$$C_{it} = Z_{it} - \sum_k p_{kt} \left[A_{ikt} - A_{ikt-1} \right], \tag{1}$$

where Z_{it} captures all sources of income and transfers, $\mathbf{A}_{it} = A_{i1t}, .., A_{iKt}$ denotes the portfolio of assets and $\mathbf{p}_t = p_{i1t}, .., p_{iKt}$ the corresponding vector of prices at which they are traded. The practical challenge in using this identity is twofold:

First, we require comprehensive information on all sources of income of a household and on all assets its members own. This challenge can be mostly overcome by the quality and scope of registry data. Indeed, data from administrative registries can provide detailed information on labor income, capital income, taxes paid and transfers received, including even private transfers like bequests. This information is mostly third-party reported by employers, financial institutions, government administrations, etc. The registries miss out on the income and wealth that is unreported, but the resulting measure error is small when tax enforcement is strong and the role of cash holdings is minor.[6]

Second, to construct how a household's asset portfolio contributes to the consumption by its members, one needs to know the asset transactions, $A_{ikt-1} - A_{ikt}$, and the corresponding transaction price p_{kt} for all assets k. Transaction data is often hard to come by. Exceptions exist when

---

[5]They use the so-called "capital market approach" and assume a single long-term rate of return to the value of real estate investment.

[6]The informal sector in Sweden is estimated to be one of the smallest in developed countries (see Medina and Schneider [2018]).

transactions are subject to a specific tax (e.g., housing, financial assets) or recorded for other reasons (e.g., cars). Depending on the tax treatment of capital, different types of data may be available at different levels of aggregation. This includes capital income $Y_{ikt} = r_{kt}A_{ikt}$ (e.g., interests, dividends), realized capital gains $G_{ikt} = \left[p_{kt} - p_{kt}^0\right]\left[A_{ikt-1} - A_{ikt}\right]$ (e.g., on sales of stocks), where p_{kt}^0 is the original purchase price of the asset sold at t, and/or on the balance value $W_{ikt} = p_{kt}A_{ikt}$ of an asset (or some aggregate of assets). We include capital income in our measure of income Z_{it}, but not the realized capital gains. The reason is that a krona of capital income translates one-to-one in a krona of expenditures, everything else equal, but a krona of realized capital gains does not, neither does a krona change in capital wealth. To see this, we can re-write the asset component of the identity in (1) as

$$p_{kt}\Delta A_{ikt} = \Delta W_{ikt} - \Delta p_{kt}A_{ikt-1},$$

where we use the difference notation $\Delta X_t = X_t - X_{t-1}$. Since only the active re-balancing of assets contributes to the flow of consumption, this suggests two approaches to measure the contribution of assets to the consumption flow: a price approach and a quantity approach.

Price Approach　The first approach we call the price approach as we use the change in asset prices to impute the consumption flow from the change in asset wealth. When the price of an asset changes, one cannot simply use changes in the balance value to construct the consumption flow out of assets. The reason is that any unrealized capital gain will lead to an identical increase in wealth W without contributing to the consumption flow. However, when we have information on prices p_{kt}, we can impute the increase in wealth due to the price change and subtract this from the overall increase in wealth,

$$p_{kt}\Delta A_{ikt} = \Delta W_{ikt} - \frac{\Delta p_{kt}}{p_{kt-1}}W_{ikt-1}.$$

Note that this is exactly how financial assets like bank accounts are treated in measuring the corresponding consumption flow: we subtract interest payments $r_{kt}W_{ikt-1}$ (considered as capital income) from the change in the asset balance ΔW_{ikt}. Hence, a krona of savings interests accrued on the savings account does not contribute to the consumption flow.

The accuracy of the price approach will depend on the level at which information is aggregated, both over time and across asset classes. Even with high-frequency data, observing only broad asset classes forces one to impute the unrealized gains using a price index for that asset class, not accounting for the exact holdings and price variation within that class. Heterogeneity in asset prices and portfolio holdings is substantial and missing out on this variation by only observing broader asset classes could be problematic, since

$$\frac{\Delta \bar{p}_{Kt}}{\bar{p}_{Kt-1}}W_{iKt-1} \neq \Sigma_{k \in K}[\frac{\Delta p_{kt}}{p_{kt-1}}W_{ikt-1}],$$

for asset types k in asset class K.

Quantity Approach   The second approach we call the quantity approach as we start from observed differences in quantities of assets held to measure the corresponding consumption flow. Even when for tax purposes only aggregate wealth is relevant, banks or other instances may still provide more detailed information on asset quantities. Linking a change in quantities with the asset price allows approximating the value of the underlying transaction $p_{kt}[A_{ikt} - A_{ikt-1}]$. The remaining shortcoming of this imputation is that in practice such information may only be available at low frequency and the change in observed quantities may correspond to several transactions made in between the times of observation. The accounting identity, however, presumes that an individual holds on to her assets throughout the period and trades at the end of the period at price p_{kt}. Using end-of-the-year balances thus introduces measurement error if transactions are timed in between or when multiple transactions happen within a year (potentially to exploit price variation at higher frequency). That is,

$$p_{kT}[A_{ikT} - A_{ikT-365}] \neq \Sigma_{t=T-365}^{T} p_{kt}[A_{ikt} - A_{ikt-1}].$$

Comparing the price and quantity approach, a key advantage of the latter is that active re-balancing is required for asset holdings to affect the imputed consumption flow. For the price approach, however, the imputed consumption flow could be fully driven by price variation.[7] Note that additional information can be used to identify active re-balancing. For example, realized capital gains in a given year do comprise the sum of all transactions. However, the orginal purchase value is subtracted from the sales. That is, the realized capital gains at time T over the past year equals

$$G_{ikT,365} = \Sigma_{t=T-365}^{T} [p_{kt} - p_{kt}^{0}][A_{ikt-1} - A_{ikt}].$$

Hence, while realized capital gains may seem useful, without information on the original purchasing price, we can not back out the resulting consumption flow.

Durable Consumption   An important issue in the measurement of consumption is the treatment of investments in durable goods. While such investments generate a consumption flow for several periods, our approach assigns the entire consumption value to the period of purchase. While consumption expenditures surveys often ask directly about durable investments, registry data may also provide information on particular durable assets owned (e.g., cars) and thus allow to impute the corresponding consumption flow. This seems to be particularly important for the case of real estate, where we account for the corresponding consumption flow by calculating the 'imputed rents'. This also improves the comparability of consumption expenditures for homeowners and renters.

The importance of making the distinction between durable and non-durable consumption will depend on the application one has in mind. For example, the ability to delay investments in durable goods in response to a temporary shock will improve the smoothing of the actual consumption flow

---

[7]See Baker et al. [2018] for an assessment of measurement error in the registry-based measure using high-frequency data of financial transactions.

and thus reduce the welfare impact of a shock. The distinction may be less relevant when for example studying consumption inequality.

3   Implementation in Sweden

We now describe the different components we use for the construction of a registry-based measure of consumption in Sweden. The main source of data is the longitudinal dataset LISA, which merges several administrative and tax registers for the universe of Swedish individuals aged 16 and above. In addition to socio-demographic information (such as age, family situation, education, county of residence), LISA contains exhaustive information on income and transfers. We link these data to other registers, including data on asset holdings. This data is collected by the tax administration as Sweden levied a tax on wealth (until 2007). Our data thus provides comprehensive information on income, transfers and assets, which are reported to the tax administration at the individual level, but can be aggregated up at the household level using household identifiers. The measure we construct based on this information captures annual expenditures between December of each year and is computable for the universe of households between 1999 and 2007. Appendix Table 2 provides an overview of the different income and wealth components and their corresponding magnitudes, following Eika et al. [2017].

3.1   Income and Transfers

The dataset LISA contains exhaustive information on income from labour and capital, transfers and taxes on an annual basis. For labour earnings, we use a measure of disposable income which contains all employment-related earnings including wages, vacation pay and severance payments, but also wage income for self-employed and business owners. Business losses for small business owners and independent contractors are included in LISA from 2004 onward.[8] For capital income, we include interests earned or paid and dividends received on all assets, including banks accounts, bonds, listed securities and closely-held businesses. Capital income also includes rental income, but we explicitly take out the realized capital gains (or losses) from selling financial or real assets. To the income from labour and capital, we also add all government transfers received, which include unemployment benefits, sickness and disability benefits, pension benefits, housing benefits, child benefits and other smaller programs. In our application we single out the receipt of sickness benefits to study the impact of a health shock. Any decomposition analysis is considered before taxes as we only have information on aggregate taxes paid. The aggregate taxes are, however, subtracted from the total income and transfers received.

---

[8]This results in slightly higher labour earnings before 2004. For instance, in 2004 labour earnings excluding business losses are 0.5% lower compared to the labour earnings measure which includes business losses. Throughout our sample period business losses can be carried forward for tax purposes. Hence, losses in one year affect tax payments and therefore disposable income in subsequent years.

3.2 Financial Assets

Data on wealth comes from the wealth register (Förmögenhetsregistret), which comprehensively covers the asset portfolios for the universe of Swedish individuals from 1999 to 2007. The register contains information on the balance value of all financial assets and listed securities without any top-coding of wealth. This includes information on aggregate holdings by asset class (bank accounts, bonds, stocks, mutual funds, etc.), but also on total outstanding debt, which includes mortgage debt, consumer credit, student debt, etc. All financial institutions were compelled to report this information directly to the tax administration for the purpose of the wealth tax.

The amount of bank account holdings reported in the wealth registry changes over our sample period. Before 2006 only bank holdings which carry returns exceeding 100 kronor are reported.[9] After 2006, all bank holdings exceeding 10,000 kronor, irrespective of the amount of interest they carry, are reported to the tax administration.[10] We may also miss some financial wealth held in foreign banks to the extent that these banks do not comply with the requirement to transmit information on the financial wealth of their Swedish customers to the Swedish tax authority.[11]

The balance values in the wealth registers allow us to compute the contribution to consumption from asset re-balancing for the asset classes with fixed prices. For example, for bank holdings, we calculate the change in bank holdings between year $t-1$ and t, which together with earned interests (included in our income measure) determines the consumption out of bank holdings. Similarly, for debt, we calculate the change in outstanding debt between year t and $t-1$, which net of paid interests (again included in our income measure) determines the consumption out of debt. The data is complemented with disaggregated data (KURU) on the balance positions for all financial securities by their International Securities Identification Number (ISIN). Swedish financial institutions are required to report the quantity and ISIN code of their clients' security holdings.[12] However, instead of using the reported prices, the wealth register held by Statistics Sweden only aggregates the listed securities for which it can find a trading price. The register thus excludes non-listed securities, like the ownership of small businesses or stocks in companies that are not publicly traded. Lundberg and Waldenström [2017] estimate that the unlisted securities amount to approximately 5% of all wealth held in securities. For the listed securities, we use closing prices from the last day of trade in December 2000 to 2007 based on the SIX data, which covers security prices from about 1,500 different securities exchanges and contributors. Just like Statistics Sweden,

---

[9]The exchange rate is 1 SEK≈0.11 USD

[10]Before 2006 about 70% of all bank holdings reported in national accounts are reported in the wealth statistics. The new censoring rule of 2006 results in 97% of all bank holdings found in national accounts to be found in the wealth register. Note also that the fraction of all households that lack bank holdings in the registry data goes from 45% on average for 1999-2005 to 27% on average for 2006 to 2007. How the change in censor rule is handled when estimating consumption expenditures is described in detail in Appendix B.

[11]Note that the fraction of foreign-held assets is low (about 3% of household assets according to the Savings Barometer of Statistics Sweden), and likely to be held by the wealthiest households (Alstadsæter et al. [2019]). Foreign assets held via Swedish financial institutions are required to be reported while assets held by foreign financial institutions do not meet the same requirement, since the third-party reporting obligation only applies to Swedish financial institutions.

[12]This information is pre-printed on every tax payer's income statement, but can be amended by the tax payers.

we currently ignore the financial securities from the KURU data for which we can not link the security to a price.[13] For options, warrants and the tax-favored capital insurance accounts, we do not observe the quantities owned, but only the balance values and use the price approach with an aggregate price-index to impute the consumption flows. Note that for the tax-deferred private pension accounts, we do not observe the balance value, but only the contributions and withdrawals. This is, however, exactly what we need to compute the corresponding consumption flows.

3.3 Real Assets

There are two main forms of real estate in Sweden that can be held by individuals: real estate that individuals own and real estate for which individuals own the occupation right. The first category includes all stand alone housing, semi-detached, bungalows, terrace housing and second homes, but also commercial real estate such as apartment buildings, industrial plants and farming property. The second category includes all apartments. In Sweden it is uncommon that individuals own their apartments. Rather individuals who live in apartments buy a share in a housing cooperation (co-op henceforth). Apartments can be bought and sold like any other real estate but from a legal point of view co-ops are different from regular housing as the occupier does not legally own the property, but only the right to occupy it.

The wealth tax register (Förmögenhetsregistret) contains detailed information on the stock of real estate wealth, estimated at market value as of December 31 of each year.[14] In principle, this allows for an imputation of the consumption flows following the price approach using housing price indices. For owned housing, however, we have data on all transactions from the housing price registry (Fastighetsprisstatistiken) from 1999 to 2007. Since we observe all real estate purchased and sold, we can directly calculate the contribution from re-balancing real estate holdings to the annual consumption flow for each household.[15]

A separate register records the sales of co-ops, but not the purchases. However, the purchasing price of the sold apartment is registered.[16] For individuals buying and selling coop apartments we

---

[13]The SIX data has wider coverage than the Swedish Stock Exchanges used by Statistics Sweden to assign prices to the individual securities. Hence, we manage to retrieve more financial wealth than in the wealth register. Note that for about 1% of stock holders in the wealth register, we are not able to match their records to the KURU data. For these individuals we use the balance values in the wealth registers and a price-index to impute the consumption flows. Since our coverage of financial assets is not complete, we are worried that recorded changes in asset positions may be offset by corresponding changes in unrecorded assets positions. For 0.03% of individuals, we have trimmed too large transactions without offsetting changes in calculating the consumption flow. In principle, we could could further extend our coverage by using the self-reported prices for the securities for which we cannot assign a price. Appendix B provides more detail on the KURU data, its relation with the wealth register and the trimming procedure.

[14]Statistics Sweden use data on the real estate tax value and regional sales statistics to compute market values for housing wealth. The real estate tax value, as used for determining the wealth tax, is a function of characteristics of the property and updated frequently. See Lundberg and Waldenström [2017], but also Appendix B for further details.

[15]From 2004 onward we observe all parties involved in a transaction. Before 2004 only one seller and one buyer is recorded, even if more individuals are involved. Before 2004 we calculate housing transaction flows on the household level under the assumption that only two households are involved in most transactions. After 2004 we calculate these flows on an individual basis. More detail is provided in Appendix B.

[16]The purchasing price of the sold apartment is recorded so that the tax administration can calculate the tax on sales profits.

combine the information from the wealth tax registers and the purchasing price of sold apartments to impute the corresponding consumption flows.[17] For individuals buying a coop without having owned one the previous year we use the coop wealth variable from the wealth registry.

Like for financial assets, investments in real estate generate capital gains (or losses), but can also generate other income. Rental income is part of the disposable income measure we use. For real estate that is not rented out, we impute rents to account for the consumption services a house provides. As discussed before, imputing such rents may be desirable for all durables which generate consumption flows beyond the moment of purchase, but clearly this value will be highest for housing services. Different approaches have been used for imputing rents, ranging from estimating the alternative income that would be generated when investing in a low-risk asset to estimating the potential rental income when subletting the house (see Poterba [1992]). The latter approach is followed by Statistics Sweden and used for the national accounts.[18] The imputed rent is calculated as the square-meter living space multiplied by a square-meter rental value which depends on the geographical region, construction period and dwelling category. While we do not have individual data on the imputed rents, we follow Eika et al. [2017] and assign the aggregate imputed rents to the households based on the value of their primary real estate in the wealth tax register.[19]

3.4   Household Definition

To construct a household identifier for registry data, Statistics Sweden uses the population register (Register över totalbefolkningen). In the register there is information on residence and relations. An individual's relations include spouse, registered partner (for gay couples), cohabitant with whom the individual has common children (either biological or adopted), biological parent(s), adoptive parent(s), legal guardian (for children under age 18), or foster parent. To form a household two (or more individuals) need to both live on the same address (place of residence) and have a registered relation.

An obvious caveat of the registry measure of households is that it misses cohabitants without children, who have no registered relation. Note however that cohabitants who have children together are defined as a household in the registry data. In the Swedish consumption expenditure survey (HUT), all individuals that live in the same dwelling and have a shared economy constitutes a household. This means that cohabitants without children are registered as cohabitants and will constitute one household instead of two separate households as would be the case using the registry data definition of a household. The HUT definition of households also includes friends sharing

---

[17]In particular, we observe individuals who buy a new coop apartment during the years for which we compute the residual consumption measure. Then we estimate the probability of buying a coop apartment based on variables such as household characteristics and changes in debt. We also use the same covariates to predict purchasing prices. For individuals where the probability of buying a coop apartment is sufficiently high we impute the purchasing value.

[18]We note that rents are controlled in Sweden according to a user cost principle. That is, rents are negotiated each year between the tenants association and the landlords and rents are only supposed to vary based on apartment characteristics. Thus, two apartments with equivalent standard within a geographical area cost the same regardless of the location in that area.

[19]This imputation assumes that the primary real estate is not (partially) subletted and ignores potential consumption flows from secondary real estates.

housing and economy, but it does not include lodgers. By merging the HUT data to the registry data we can investigate further the discrepancies between households defined in the registry data and in surveys. Using the matched HUT-registry data from 2003-2007 we find that the share of individuals that are found to be part of the same household in both HUT and LISA is large (82%). The fraction of individuals that are registered as being in the same household in the HUT but not in the registry data is 12%, and is mostly comprised of unregistered cohabitants (9%). Finally, 6% of individuals are found to be part of the same household in the registry data but not in the HUT data, and this group is almost entirely composed of children or separated partners.

3.5 Some Further Steps

All our data ranges from at least 1999 until 2007. In principle, the measure could be extended for a longer time period. However, the wealth tax was abolished in Sweden in 2007, after which the government collected only limited information on assets. Hence, any imputation of the consumption flows would need to rely on information from realized capital gains and income (see Saez and Zucman [2016]). The data could be extended more easily to earlier years. While the LISA data dates back to 1990, the wealth register (Förmögenhetsregistret) started only in 1999. The comprehensive information on income and wealth could, however, be retrieved from the Income and Taxation Register (Inkomst- och taxeringsregistret) from 1988 onwards.

The Swedish registers contain a wealth of additional information that could be integrated in the construction of the registry-based measure and further improve the quality of analysis. A first example is the vehicle register (Fordonsregistret), which contains data on all car transactions, allowing to connect cars to their owners and includes information on the car's general status (registration date, when the car became the current owner's, in use or not, etc.) and various specifics (make, model, odometer reading, fuel consumption per kilometer, etc.). The data does not contain any value assessment of the cars, but this could be approximated using Blue Book data. A second example is the inheritance register (Belinda), which contains data on inter vivos gifts and bequests. While gifts and bequests may be directly used to increase actual consumption expenditures, any resulting increase in the heirs' assets in fact contributes negatively to the registry-based measure. While the inheritance register allows to identifying these heirs, properly correcting the consumption measure is challenging since the inheritance data is aggregated across asset classes and lumped with the inter vivos gifts in the ten years before death. The data also ends in 2004 when the inheritance tax was abolished (see Appendix B for details). A final example is the migration records which is part of the population register (Registret över totalbefolkningen). It records all migration flows in and out of Sweden, at the individual level, with the exact date of migration (see again Appendix B for details). This allows to correct mismeasurement of consumption flows when individuals move in and out of Sweden. While we have not fully integrated this registry information in the construction of our consumption measure, we can already illustrate the use of this data to highlight some specific issues.

4 Registry-Based Measure: Descriptive Statistics

This section provides some descriptive statistics, comparing how the registry-based consumption measures compares to earnings and alternative measures of consumption.

4.1 Distribution & Age Profile

Figure 1 shows the distribution of household earnings and the distribution of our registry-based measure of household consumption for the same sample for the period 2000-2007. Note that each observation is a household-year. Overall, consumption exhibits less dispersion than earnings, except at the top of the distribution.[20] While the earnings distribution is bimodal, with a big mass at zero earnings, the density of consumption is much smoother. One can nevertheless see that there is a mass of households with zero consumption in our data. A significant fraction of these observations come from individuals who move in or out of the country in a given year. Panel A of Appendix Figure C.1 shows the same distribution when excluding these observations, and exhibits a much smaller excess mass at zero.[21] Another well-known issue with the registry-based measure (e.g., Eika et al. [2017]) is that for a substantial share of households the imputed consumption expenditures are negative. This can be due to misattributing wealth increases to savings out of income when they are due to price increases or due to inheritances or gifts. We use the inheritance register to investigate this issue in Sweden. While 6.2% of heirs have a strictly negative level of consumption compared to 3.8% of individuals that do not receive an inheritance, less than 1% of individuals receive a positive inheritance as recorded in the inheritance register. Hence, excluding observations of households in the years they have received bequest has a minimal impact on the consumption distribution, as shown in Panel B of Appendix Figure C.1.

We can further investigate the evolution of consumption over the lifecycle. Figure 2 shows the evolution of both earnings and consumption by age and birth cohort. To construct this graph, we split the sample into 6 birth cohorts (1930, 1940, 1950, 1960, 1970 and 1980) and follow each birth cohort as they age over the eight years of our consumption measure (2000-2007). This allows us to get a sense of a global age profile by "stitching" together the birth cohorts. Interestingly, the graph shows that earnings exhibit a much steeper hump shape profile than consumption. Consumption is smoothed over the lifecycle, by saving and accumulating assets over the age range 35 to 65.

---

[20]In Figure 1 we report the mean, standard deviation and interquartile range of the distribution of earnings and consumption, where we trim the top and bottom 1% of each distribution. The variance of the consumption distribution is nevertheless much larger than that of earnings at the very top, as wealth, which is an important determinant of consumption at the top end of the distribution, is much more unequally distributed than earnings. The standard deviation of the non-trimmed distribution of consumption is 561 versus 186 for earnings.

[21]The consumption measure equals zero for about 30.000 households per year. These are virtually all single-person households and most of them are immigrants (about 85%) who recently moved to Sweden.

4.2   Comparison to National Accounts and to HUT Surveys

We now compare the evolution and the distribution of our registry-based measure of consumption expenditures to consumption measured in the national accounts (NA) and in HUT surveys of consumption expenditures. There are significant conceptual differences between what is considered consumption expenditures in the national accounts, in the consumption surveys, and in our registry-based measure. While the NA measure of consumption relies on many different sources and possibly offers the most comprehensive and conceptually sound measure of aggregate consumption, it does not offer any guidance in terms of the distribution of consumption and its evolution. Economists therefore have to mostly rely on consumption survey measures to investigate consumption distribution and consumption dynamics. In Appendix C we describe in more details the sources and concepts used for measuring consumption in the NA and in the HUT.

The HUT surveys provide direct measures of bi-weekly consumption expenditures at the moment the household is surveyed. Information on consumption comes from shopping books where surveyed households are asked to report all their expenditures during the 2 weeks preceding their interview date. For a few durable goods, extra information is added from answers to interview questions, where individuals are asked to recollect some expenditures for up to 12 months prior to the interview date.

In the national accounts, consumption is computed from various sources. Information from the HUT surveys is used, but complemented with various information from the registry data and with VAT receipt data. One of the main differences between our registry measure of expenditures and the NA measure pertains to the treatment of some investments. Large renovations, home refurbishments, and replacement of home appliances is treated as investment (rather than consumption) in the national accounts, while they show up as expenditures in our registry measure. But compared with the survey data, both the NA and the registry measure of consumption include imputed rents. Both measures also include expenditures in the black market or on illegal goods and services.[22]

Figure 3 compares the evolution of total consumption expenditures per adult in Sweden in the HUT, the NA and our registry measure over the period 2000 to 2007.[23] For NA, we display two series: the official series, and one in which investments made by households (refurbishing of real estate, etc.) are added back to NA consumption. The Figure shows that the average level of the NA and the registry measures compares well over the period, especially when household investments are added to the NA series of consumption expenditures. The HUT measure, which does not include imputed rents, is systematically lower. We note that our consumption expenditure measure diverts from NA consumption in 2004 and 2005. This is driven mainly by large changes in asset holdings by wealthy individuals and changes in the deposit rates.[24]

---

[22]This type of consumption is recorded in the residual measure as long as the resources that are used to pay for it are accounted for.

[23]Adults are defined as individuals aged 16 or more

[24]The registry-based consumption measure increases by 50% between 2003 and 2004 for the top 10 percent wealthiest and even by 100% for the top 1 percent. Such large consumption increases are outside the range of what we perceive as plausible. It is worth noting that on January 1, 2004 the inheritance tax between spouses and cohabitants was abolished and on January 1, 2005 the inheritance tax was abolished for everyone, along with a tax on gifts. This

Figure 4 compares the density of the distribution of individual consumption expenditures in the HUT and with our registry-measure for all years 2003-2007 for all adults between 25 and 55. For both the HUT and the registry, individual consumption is computed by dividing total household consumption by the number of adults in the household. The vertical lines denote average consumption expenditures for each measure. As discussed before, in contrast with the survey measure, the registry-based measure takes negative values and has significant mass at zero. The density of the registry measure, which includes imputed rents, is also slightly skewed to the right compared to the HUT distribution. We also report the density of a registry measure excluding imputed rents, which lines up better with the HUT distribution. The averages of the two distributions are very close. Still, even when correcting for imputed rents, the registry-based measure still has a fatter right tail. One reason for this, as is well known from research on top income inequality, is that due to the Pareto structure of the income (and consumption) distribution at the top, survey sampling will underestimate the fat tail of the distribution, irrespective of the quality of the income or consumption measure. Another reason is that, in consumption surveys, mismeasurement of consumption seems to be more severe at the top end of the income distribution.[25] A critical consequence of this is that survey data may miss the dynamics of consumption inequality at the top, and also how it relates to the dynamics of income inequality.[26]

Importantly, Sweden offers the possibility to match HUT survey to the registry data for years 2003 to 2009. This gives the opportunity to directly compare consumption measures in the HUT and in the registry measure for the same households. In Figure 5, we plot log household consumption in HUT vs log expenditures from the registry measure. We restrict the sample to all household from the HUT surveys from 2003 to 2007 that are surveyed in the month of December to make sure that the annual expenditure measure from the HUT (which includes recall items such as durables from the last 12 months) corresponds to the annual measure in the registry data which spans all expenditures from December of year $n-1$ to December of year n. When the household composition in the HUT and registry data differs, we reconstruct households for the registry measure using the exact same individuals that are listed as members of the household in the HUT. To make the expenditure measure comparable, we exclude imputed rents from the registry measure. Results show that, within households, both measures of consumption expenditures compare quite well, with most observations scattered closely along the 45 degree line. However, as anticipated, this is different in the tails of the distribution. At the top of the distribution of household consumption,

may have led to large changes in asset transfers, also abroad. Transfers to oversees financial institutions show up as consumption in our consumption measure since we do not have access to Swedes' holdings abroad. The second factor is a sharp decline in deposit rates between 2003 and 2004, from 1.35% to 0.8%. This means that the minimum amount of bank account holdings to be recorded increase from 7,500 kronor to 12,500 kronor. All else being equal, this raises consumption since bank account holdings are being reduced in our data.

[25]As noted by Bee et al. [2013], in the U.S. Consumer Expenditure Survey, "there is strong evidence of under-representation at the top of the income distribution and under-reporting of income and expenditures at the top." Among the many reasons for such mismeasurement at the top, it is often argued that consumption surveys have more difficulty getting top income individuals to participate in the survey, not drop out, fill properly their diaries and track their expenditures, especially irregular or infrequent ones.

[26]On the many issues of measuring consumption inequality in surveys see for instance Attanasio and Pistaferri [2016]

most observations are below the 45 degree line and consumption expenditures are systematically underestimated in the survey compared to the registry data. The opposite is true at the bottom of the distribution.

5   Consumption Responses to Health Shocks

This section uses our registry-based measure of consumption to document consumption responses to health shocks and identify how much households are insured against such shocks. This application also illustrates the strengths and complementarities of the registry-based measure relative to using surveyed consumption.

5.1   Related Literature

While a series of papers have investigated the effect of health shocks on labour supply, the literature on consumption responses to health shocks is still relatively limited. This is mostly due to the difficulty of finding reliable panel data with information on both consumption and health status.

A couple of papers have documented the impact of disability shocks on consumption, using survey data in the US. Stephens [2001] was the first paper to analyse the dynamics of consumption around the onset of self-reported disability using PSID data. Meyer and Mok [2018] use a similar approach with a larger number of waves of the PSID and investigate the heterogeneity of consumption responses by disability types. Both papers rely on measures of food consumption (or food plus housing expenditures) and find that household consumption of food drops after the onset of a disability but that the drop is much less severe than the drop in earnings and post-tax income. Some related structural papers have also analysed the extent to which individuals are insured against disability shocks, and the role played by Disability Insurance (DI) system using survey data. In Ball and Low [2014], the structure of the model allows for self insurance through spousal labor supply and savings / intertemporal smoothing. Identifying the model on data from the British Household Panel Survey (BHPS), they find that work-limiting disabilities result in a 9% drop in food expenditures, absent any self-insurance response. Self insurance offers a 2-4 percentage point mitigation of the consumption drop. Low and Pistaferri [2015] use data from the PSID to identify the welfare effects of changes to the design of the DI system in the US, in a life-cycle model of consumption. Finally, a recent literature uses a registry-based measure of consumption in Norway similar to ours, and document the impact of DI receipt on consumption for individuals at the margin of DI eligibility (Autor et al. [2019], Kostøl and Mogstad [2015]).

Compared to Autor et al. [2019], our focus is on understanding how well individuals are insured against a broad set of health shocks, and on understanding the mechanisms (e.g., social transfers vs. self-insurance) they use to insure themselves against these shocks. Conceptually, while our study focuses on consumption drops in response to the bad health shock, which is a measure of the level of insurance that individuals have against these shocks, Autor et al. [2019] identify the marginal propensity to consume out of disability insurance benefits when individuals are observed

in the bad health state. The two are different objects of interest, but can both be related to the value of insurance against health shocks (see Landais and Spinnewijn [2019]).

As we show, compared to survey data, our measure of consumption provides precise estimates of the consumption responses to an health shock. It also allows for a rich investigation of heterogeneity in these responses, according to the type and severity of the shock, as in Meyer and Mok [2018], but also depending on various individual characteristics that may be relevant for targeting insurance and transfer policies. Our measure also allows for providing a detailed analysis of the underlying mechanisms of consumption smoothing, and the relative role of various forms of insurance or self-insurance such as social insurance and transfer programs, precautionary savings, spousal labor supply, etc. We also note that compared to self-reported disability indicators, our measure of health shocks, based on the first occurrence of a work incapacitation longer than 14 days, is likely to be more unanticipated and lends itself more to a credible event study design. This type of health shock is also directly policy relevant as a work incapacitation of 14 days is the minimum threshold above which individuals can become eligible to specific sickness and disability insurance programs.

5.2 Methodology

We are interested in identifying the dynamics of some outcome Y_{it} around the time of a health shock. Households are indexed by i and $t = 1, .., T$ denote the calendar year of observation. For each household, the health shock event happens at some time E_i. We restrict our attention to the following class of models:

$$Y_{it} = \alpha_i + \nu_t + \sum_{j=-N_0}^{N_1} \beta_j \cdot \mathbb{1}[J_{it} = j] + \varepsilon_{it} \tag{2}$$

where $[-N_0; N_1]$ is the window of dynamics effects around the event, α_i is a household fixed-effect and ν_t is a time effect. $J_{it} = t - E_{it}$ denotes event time, that is the time in year relative to the occurrence of the event. We are interested in retrieving the full path of dynamic effects $\{\beta_j\}_{j=-N_0}^{N_1}$, which means we are interested in pre-event as much as in post-event effects.[27]

If we restrict our attention to a sample of households who all receive treatment at some time E_i, an OLS specification with two-way fixed effects (household and time fixed effects) can only identify the true dynamic effects $\{\beta_j\}_{j=-N_0}^{N_1}$ up to a linear trend (see Borusyak and Jaravel [2016] for a recent discussion). More fundamentally, when introducing household fixed effects, and focusing only on treated households, it is impossible to jointly and non-parametrically identify the full sequence of dynamic effects $\{\beta_j\}_{j=-N_0}^{N_1}$ and of any other time-related variable, such as calendar time, or age. The reason is that, within household, the passage of event time and of any other time related variable is collinear.[28]

---

[27] The model in equation (2) abstracts from heterogeneity in the dynamic treatment effects but such effects are certainly important, and the framework can be extended to accommodate such heterogeneity.

[28] In equation (2) we have focused on calendar time fixed effects, but in some instances, one might be more interested in controlling for other time related variables such as age, tenure, or experience fixed effects. This is easily done by denoting age (or tenure or experience) by $t = 1, .., T$, and age (or tenure or experience) fixed effects by ν_t. The same

We deal with this identification issue by constructing a control group of individuals who never experience the event. The introduction of such control group allows to identify the time effects ν_t independently of the dynamic treatment effect of the event $\{\beta_j\}_{j=-N_0}^{N_1}$, because control individuals never experience treatment. Identification relies on the assumption that the control individuals offer a good counterfactual of the time trends of the treated individuals absent dynamic treatment effects. In settings where the timing of the event is random and anticipatory effects can reasonably assumed to be zero (i.e. all coefficients $\{\beta_j\}_{j=-N_0}^{0}$ are zero), this identification assumption can partially be tested by looking at parallel pre-trends between the control and treatment groups prior to the event. But it should be noted that in fully flexible settings where dynamic treatment effects can be non-zero prior to the event, the identification assumption of parallel trends is fundamentally untestable.

We create our control group using nearest-neighbor matching based on pre-event characteristics. One of the main challenges is to select observable characteristics that can reasonably be assumed to be unaffected by the event prior to its occurrence. We adopt the following matching strategy: for each calendar year t, we take all individuals who receive the event in that particular year ($E_{it} = t$), and find a nearest neighbor from the sample of all individuals who never receive treatment. Individuals are matched exactly on age, gender, region of residence in $t-1$ (21 cells), level of education in $t-1$ (10 cells) and family structure in $t-1$ (12 cells), and by propensity score on their number of dependent children in $t-1$, 12 industry dummies in $t-1$ and their earnings in $t-1$, $t-2$ and $t-3$. Note that by matching exactly on age, we make sure that age effects, which could not be identified in specification (2), are not confounding our estimated dynamic effects.

Household composition For household level outcomes, we fix the composition of the household as of the last year prior to event, meaning that household growth and re-composition is not accounted for, neither before nor after the event year. Note also that we do not adjust household consumption for household equivalence scales. Because household structure is fixed, household size differences are subsumed in the fixed effects in specification (2).

Event studies using surveys Some outcomes of interest come from survey data that does not have a panel structure. This is the case for some of our health outcome measures (which come from the HEK survey), or for consumption measures coming from the HUT.

Yet, these surveys can be matched to registry data based on personal identifiers, so that we know for each individual surveyed the timing of the event E_i (if it occurs) and can therefore compute their event time J_{it} at the moment they are surveyed. This enables us to also provide pseudo-event studies estimates using these survey measures. We follow again a control group approach. Using nearest-neighbor matching based on pre-event characteristics, we create a control group of individuals who are not subject to health shocks in the entire period of our data. For individuals in the control group, we allocate a placebo event time E_i corresponding to the event time experienced

---

collinearity problem obviously arises in such settings.

by their nearest neighbor in the treated group. We then estimate the following specifications to retrieve the dynamic effects:

$$Y_{it} = \sum_{j=-N_0}^{N_1} \beta_j \cdot \mathbb{1}[J_{it} = j] \cdot \mathbb{1}[T_i = 1] + \gamma \cdot \mathbb{1}[T_i = 1] + \sum_{j=-N_0}^{N_1} \kappa_j \cdot \mathbb{1}[J_{it} = j] + \nu_t + X'_{it}\alpha + \varepsilon_{it} \quad (3)$$

where T is an indicator for being in the group of treated individuals experiencing the event.

5.3 Identifying Health Shocks

In Sweden, employees who become sick are eligible to a paid sick leave. To this effect, they need to obtain a certificate from a physician providing evidence of their incapacitation. For the first 14 days of incapacitation, individuals get a "sickness wage" of 80% of their actual wage, paid directly by the employer. When incapacitation lasts longer than 14 days, individuals start receiving benefits from the Social Insurance Agency (SIA). These sickness benefits amount to 80% of actual wage, up to a cap.[29] Sickness benefit receipts therefore identify the occurrence of a health shock large enough to involve a work incapacitation larger than 14 days. We define a health shock event as the first occurrence of sickness benefits receipt by an individual over the period 1997-2011.[30] We restrict the sample to individuals aged 25 to 55 at the time of the event. Appendix Table 3 displays descriptive statistics for our sample of analysis, breaking it down between treated individuals and the control group, using nearest-neighbor matching, of individuals who are never treated. Individuals experiencing a health shock are around 38 years old on average, and have little net wealth. The median individual has only 77k SEK in total net wealth ($\approx$ 25% of total annual household earnings), and zero liquid assets (bank holdings) prior to the onset of the health shock. A long tail of individuals have already negative net wealth when hit by the shock. This obviously limits consumption smoothing opportunities.[31]

Our approach to identifying the effects of health shocks relies on two assumptions: (i) that the occurrence of sickness benefit receipt strongly correlates with health outcomes and (ii) that this event is exogenous, i.e., that the shock is not anticipated. We can probe into the validity of our approach by using HEK survey data, which provides direct information about health outcomes, and that can be matched with our registry data.[32] In Figure 6, we report estimates β_j from specification

---

[29]Individuals with partial incapacitation after 14 days can work part time and will receive reduced sickness benefits.

[30]Before 2008, there was no clear rule regarding the time individuals could receive sickness benefits, or on the issue of knowing when they should be transferred to the disability insurance system instead. In 2008 the sickness insurance (SI) system and its transition of individuals from SI to DI was made more rigorous and strict with the chain of rehabitalization. After 90 days of SI benefit take-up the individual must be deemed unfit to carry out any work at the current employer in order to keep SI benefits. After 180 days the individual must be deemed unfit to carry out any work on the labor market as a whole to keep SI benefits. Individuals who are assessed to be able to work are transferred to the public employment service. Individuals with a permanent inability to work are given DI. The assessment of individuals' ability to work is done by a specialized officer of the SIA based on a physician's statement.

[31]Treatment and control groups are well balanced in terms of demographics although treated individuals have larger debt levels and lower level of net wealth on average.

[32]The HEK-survey, which is compiled by Statistics Sweden (SCB), is a detailed annual survey of the household

(3) for various health outcomes.[33] Results confirm that our event is associated with a sharp, large and significant increase in health expenditures (panel A) and the probability to report suffering from pain, reduced work capacity, or long term illness (panels B, C and D).[34] It is also associated with a sudden surge in the probability to visit a physician or a physiotherapist (panels E and F).

Figure 6 overall suggests that our event (the receipt of sickness benefits corresponding to a work incapacitation of at least 14 days) is associated on average with significant and somewhat permanent degradation in health. There is however interesting heterogeneity in the severity and persistence of the shock, that we explore and exploit below.

Furthermore, the various panels suggest that there is little anticipation on average: for all outcomes except the number of visits to a physiotherapist, there is no significant pre-event dynamic effects. This lack of anticipation brings support to our identification assumption.

5.4 Earnings and Consumption Response

Our central object of interest is how health shocks affect earnings and whether this shock is smoothed in the consumption expenditures. Figure 7 reports event study estimates of the earnings and consumption responses to a health shock. We use specification (2) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. All estimates $\hat{\beta}_j$ are expressed as a fraction of total average household expenditures of the treated group as of event year -1.

Panel A shows that health shocks are on average associated with a sharp, significant and permanent decline in earnings equivalent to about 10% of household consumption levels in year -1. Panel B shows that the permanent drop in earnings following a health shock is also associated with a significant and permanent decline of household consumption expenditures, of around 5%. Anticipatory effects on earnings or on consumption prior to the health shock seem limited, which confirms the evidence of Figure 6 that these shocks are largely unanticipated.

The evidence from Figure 7 confirms that health shocks negatively and permanently affect earnings capacity of individuals (e.g., Stephens [2001]), and translate into a sharp and significant, but smaller decrease in consumption of about 5%. In other words, our results suggest that only 50% of the total shock to earnings is insured. This evidence is in line with survey evidence from the US documenting drops in household food consumption of about 5% on average following the

---

sector with data on income, debt and wealth. Furthermore, approximately half of the participating households are selected for interviews, where various questions on health are asked to each individual in the household. The survey can be matched with registry data based on personal identifiers.

[33]For each panel, we also report the long term (6 years) dynamic treatment effect $\bar{\beta}_{0-5}$ corresponding to the following modification to specification (3):

$$Y_{it} = \sum_{j=-N_0}^{N_1} \bar{\beta}_{0-5} \cdot \mathbb{1}[5 \geq J_{it} \geq 0] \cdot \mathbb{1}[T_i = 1] + \gamma \cdot \mathbb{1}[T_i = 1] + \sum_{j=-N_0}^{N_1} \kappa_j \cdot \mathbb{1}[J_{it} = j] + \nu_t + X_{it}'\alpha + \varepsilon_{it}$$

[34]Panel A focuses on the probability to have a medical card, which is a measure of total drug expenditures. In Sweden, when annual spending on drugs exceeds a certain threshold (e.g. 5646 kr in 2018), individuals qualify for a medical card and received 100% coinsurance on their health expenditures.

onset of a disability (Stephens [2001], Meyer and Mok [2018]).

5.5 Smoothing Mechanisms

We now turn to the means households use to smooth their consumption following a health shock. Clearly, our findings will depend on the Swedish context, and in particular the use of private means will depend on the public system in place.

Social Insurance & Transfers Figure 8 explores the role played by social insurance and transfers in mitigating the earnings drop. We use again specification (2), using various measures of benefits and transfers as our outcomes of interest. Panel A shows that sickness benefits cover initially a large fraction ($\approx 80\%$) of the drop in earnings in the immediate aftermath of the health shock, but slowly decline over time after the initial shock. Sickness benefits are not the only source of social insurance benefits available to individuals subject to health shock. A small fraction of individuals ends up transitioning into disability insurance. This is visible in panel B, where we see a small increase in the amount of disability benefits received two to five years after the initial health shock.

Insurance Within the Family Figure 9 investigates the role played by the family in insuring individuals against health shocks.

First, we explore the sensitivity of our results to the assumption regarding income pooling within the household. By focusing so far on total household consumption, we maintained the implicit assumption that there is perfect pooling of resources across household members. Because all income and assets are observed at the individual level, the polar case of a total absence of pooling of resources within the household can be investigated by focusing instead on a measure of individual consumption expenditure. For this purpose, we split household consumption between consumption stemming from the individual subject to the health shock and consumption of all other members of the household.

Figure 9 panel A shows estimates of specification (2) using individual consumption of the affected household member as an outcome. It shows that individual consumption drops significantly after the health shock, and that the magnitude of the percentage drop is relatively similar to the drop of household consumption.

Beyond the pooling of resources across household members, the family can provide additional sources of consumption smoothing. One source that has been found to provide significant consumption smoothing opportunities in other contexts is spousal labor supply (e.g., Cullen and Gruber [2000] in the context of job loss). By increasing labor supply at the intensive or extensive margin, other family members can partly alleviate the negative effects of the health shock on household consumption. Panel B studies the role played by this added-worker effect. The graph shows the evolution, around the occurrence of a health shock, of the total earnings of all members of the household, excluding the earnings of the individual subject to the shock. The panel shows that earnings

from other members of the households actually decrease slightly and permanently after the health shock and therefore do not provide much opportunities for additional consumption smoothing. This decline in earnings could be explained by correlated shocks to labor supply: other members of the household may need more time to take care of the individual exposed to the health shock. The limited added worker effects could also be due to a context where health shocks are covered by relatively high insurance from sickness benefits and transfers. These results suggest that the family offers little additional insurance beyond the opportunity of pooling consumption across household members.

Intertemporal Smoothing A potentially important source of self-insurance against health shocks is intertemporal consumption smoothing, which consists in reallocating resources over time. This can be done in particular through the use of debt and savings, but also using durable consumption.

Figure 10 displays the evolution of consumption out of debt and consumption out of all non-real estate assets, using our baseline event study specification (2). Consumption out of debt is defined as the change in the total stock of debt of the household between $t-1$ and t: $D_t - D_{t-1}$. We find that consumption out of debt significantly decreases after the onset of a health shock. The fact that debt is not used to smooth consumption could be indicative of the presence of credit or liquidity constraints. In panel B, we report the estimates for the evolution of consumption out of assets defined as: $\sum_k p_{kt} [A_{ikt} - A_{ikt-1}]$, where we exclude consumption out of real estate assets. If anything, consumption out of assets seems to increase slightly just after the health shock, and then reverts back to the reference level at $t = -1$, just before the pre-health shock. Note also that in some pre-health years the deviation relative to this reference level is as large. Overall, the contribution of existing financial assets to consumption smoothing appears very limited.

We also investigate changes in durable consumption by studying the transactions of cars using the car registry. Figure A.1 displays the evolution of the changes in ownership, i.e., whether one buys minus sells a car in a given year, using our baseline event study specification (2). The Figure suggests that the consumption out of cars decreases after the onset of a health shock and the effect is persistent. Even five years after the shock, individuals are more likely to sell than to buy a car compared to before the shock.[35]

Decomposition To get a sense of the relative importance of the different consumption insurance mechanisms available to households, in Figure 11 panel A, we put the potential sources of consumption smoothing together and provide a full decomposition of the consumption response to

---

[35]Since we do not control for car values, we do not know how much individuals subject to the shock draw down their car holdings to increase their expenditures by downgrading their cars. While Figure A.1 displays the difference, we note that actually both the probability of selling and buying a car increases after the onset of the health shock, but this could also be driven by the change in work and ability. We could in principle correct our registry-based measure of consumption accounts for the contribution of real asset holdings (cars, boats, etc.) to the consumption flow, assuming that the annual depreciation in the value of these real assets captures the annual consumption flow. While households are required to report real assets for the wealth tax (see Appendix Table 4), but as there is no third-party reporting, under-reporting seems a major issue.

health shocks into six main components: own earnings, earnings of other members of the household, sickness and disability benefits, other transfers and taxes, consumption out of debt and consumption out of assets. The graph confirms that most of the consumption smoothing is taken care of by the social safety net, through an increase in benefits and transfers and the decline in taxes. Interestingly, variation in taxes and transfers offer a level of insurance comparable in magnitude to that of sickness benefits. The graph also confirms that assets, debt, and household labor supply contribute little to consumption smoothing on average.

5.6 Heterogeneity

So far we have focused on the average consumption response to a health shock. Our data enables us to provide a rich analysis of the heterogeneity underlying these consumption responses.

We start by investigating the response of the conditional distribution of consumption by running quantile event study regressions modelled on specification 2. In Appendix Figure B.1 we plot estimates for the $10 - th$ percentile, the median and the $90 - th$ percentile of the conditional distribution of consumption. The figure suggests that the bottom of the conditional distribution of consumption experiences little decline following a health shock, while the decrease in average consumption seems to be driven by the upper end of the conditional distribution of consumption. In other words, individuals at low level of consumption prior to the health shock appear to be more insured against the shock than individuals who have high level of consumption. This may reflect the lower value of insurance at high level of consumption, if utility is strongly concave in consumption, but is also due to the progressivity of the tax and transfer system, which offers more insurance against adverse earnings shocks at the bottom of the income distribution than at the top of the income distribution. This is confirmed if we compare the consumption responses for different income quantiles. In Table 1 we report for different subsamples, the average long term (6 years) dynamic treatment effect $\bar{\beta}_{0-5}$ from our baseline specification (2). The consumption drop is smaller than 1% for individuals in the lowest income quartile ($\bar{\beta}_{0-5} = -.006(.006)$).

In Table 1, we investigate how the drop in consumption correlates with other observable characteristics as well. We divide our sample between young individuals (who are between 25 and 35 years old at the onset of the shock) and older individuals (who are between 35 and 55 years old). Younger individuals experience a significantly larger drop in consumption ($\bar{\beta}_{0-5} = -.052(.006)$) than older individuals ($\bar{\beta}_{0-5} = -.027(.005)$), reflecting a lower level of insurance against health shocks, consistent with the fact that younger individuals have lower levels of assets/savings and that a permanent earnings shock represents a larger decrease in potential lifetime earnings when experienced earlier in life. The Table also indicates the importance of indebtedness. Individuals with high level of debt experience a larger drop in consumption, again indicating the potential importance of credit constraints, although the difference is only marginally significant (p-value = .091). We do not find a similar pattern when splitting the sample by net wealth.

We also investigate how consumption responses vary with the severity of the underlying health shock. We use two proxies to measure the severity of shock. We first split the sample between

individuals who report a long vs a short sickness spell, measured as the number of days spent receiving sickness benefits (i.e. number of days spent in work incapacitation). Results show that more persistent health shocks are indeed correlated with larger and more permanent drops in household consumption ($-.043(.007)$ relative to $-.021(.006)$). We also split the sample between individuals who end up disabled and those who do not, where disability is measured by the fact of applying to disability benefits. The drop is $1ppt$ higher for those who end up disabled, but the difference is not significant. The small difference may of course be driven by the fact that the disability is actually insured. Figure 11 panel B offers further insights into this by running a decomposition of the household response into various consumption insurance mechanisms. The panel shows that a large health shock leading to disability is associated with larger drop in earnings, both for the affected individuals, and for his or her spouse. This suggests that large health shocks require significant time for spousal care and increase the opportunity cost of spousal labor supply, therefore increasing the welfare cost of the shock. The panel nevertheless suggests that sickness and disability benefits indeed offer insurance against this larger earnings drop.

5.7   Comparison with Consumption Surveys

We compare in Figure 12 the estimates of consumption responses to health shocks obtained from our registry-based measure of consumption with the estimates from the consumption surveys. To make the consumption measures comparable, we take out imputed rents from our registry-based measure of consumption. The Figure highlights the superiority of the registry measure over the survey measure for this application. Using the latter measure, we find a drop of around 6 to 7% in consumption following the event, but estimates are very imprecise and noisy. The registry measure estimates deliver a slightly lower drop, but with very tight confidence intervals and a much less noisy dynamic pattern.[36]

One important interest of the survey measure though, is its ability to break down expenditure responses into various categories and gain further insights into smoothing mechanisms.[37] In Figure 12 panel B, we show the estimated drop in consumption for various expenditure categories. The graph shows that the health shock is associated with a clear increase in health expenditures. It also shows that there is significant heterogeneity in the way various expenditures react to the shock. While food consumption declines relatively little, most of the adjustment in consumption seems to be coming from more recreational expenditures (clothing, travels, etc.).

5.8   Value of Insurance

We started by showing that health shocks in Sweden negatively and persistently affect the earnings capacity of individuals and this shock to earnings is only partially insured. Permanent shocks to

---

[36]On the graph, we report a test of equality of the estimated average drop in consumption over the four years following the health shock, estimated with the registry measure (β^R) and with the HUT survey (β^S). The t-stat rejects that the estimates are equal across both sources.

[37]We discussed before how by construction the registry-based measure allows to study different means to smooth consumption, but surveys can and have been used in prior work to study this as well.

one's earnings capacity are by definition hard to insure as they cannot be fully insured through intertemporal smoothing (i.e., depleting assets or deferring expenses on durable goods). Instead, they require a permanent decrease in consumption. Consistent with this, we found that households get limited consumption insurance out of their financial assets and seem to decrease their consumption of cars permanently.[38] In addition, we uncovered the inability of other household members to smooth the earnings shock, potentially due to the need for extra care when sick or disabled.

When self-insurance is not effective, the role of social insurance is arguably even more important. Social insurance can help providing protection against permanent health shocks, and it does so through generous social transfers in Sweden. However, because of informational constraints (e.g., moral hazard, crowd-out of self-insurance), it rarely provides full insurance. An important trade-off arises between the value and cost of providing extra insurance. To determine the value of providing extra insurance, an influential literature (e.g., Baily [1978], Chetty [2006]) has proposed to study the drop in consumption in response to adverse events as it captures the remaining exposure, accounting for all the available means of insurance. A difficult task, however, is to determine how exactly consumption drops translate into the marginal value of insurance (e.g., Landais and Spinnewijn [2019]). This task is particularly challenging when preferences and expenditures are state-dependent. In addition to the drop in consumption, we have documented important changes in pain and work capacity, and in labor supply of both the individual itself and other members in the household in response to health shocks. All these changes will affect the marginal value of social transfers depending on the complementarities in preferences between consumption, health and leisure time (e.g., Finkelstein et al. [2013]). Even more directly, a health shock triggers the need for additional health expenditures, so our registry-based measure underestimates the drop in consumption expenditures net of health expenditures, while the latter are arguably more relevant for determining the value of insurance. This bias is, however, small in Sweden. We have documented a large relative increase in health expenditures using the HUT surveys, but the average out-of-pocket health expenditures only take 2.2% percent of overall consumption due to the generous public health insurance.

In general, our findings are specific to the Swedish context. From a policy perspective, the observed consumption consumption smoothing, conditional on the system in place, are what is needed to evaluate that system. However, it remains an open question to what extent our findings and conclusions would change in different contexts, especially when the health and social insurance system is less generous.

6 Conclusion

The Scandinavian countries received a lot of attention in recent economic research, not just because of their interesting history, culture and policies, but also because of the exceptionally detailed and

---

[38]The lack of consumption insurance out of financial assets can also be driven by the issue that credit constraints are more likely to bind after a health shock reducing one's earnings capacity, or simply by an overall lack of liquidity.

comprehensive administrative data available for research. Individuals can be linked to other members of the household and across all registries through the same personal identification number. In addition, Scandinavian countries levy or levied a tax on wealth, which makes that excellent data on asset holdings tends to be available as well. So this offers the best context to construct a residual measure of expenditures from information on income and wealth. The data requirements for constructing valuable registry-based measure are still quite formidable. So, even within Scandinavia, important differences exist across countries in terms of data coverage that have important implications for the computation of residual measure of consumption.

This paper has focused on the Swedish context, providing excellent data on financial and real assets to credibly measure their contribution to consumption flows, but only over a small time period, for example compared to Norway. Having laid out the potential challenges and remedies, we have attempted to further improve the registry-based measurement of consumption and illustrated how powerful this consumption measure can be to study consumption smoothing. We hope that future research will build on our efforts to extend the measurement to other time periods and countries, when the challenges may be even more binding. The ultimate goal is not just to revisit questions previously approached using consumption surveys, but also to exploit complementarities between the two alternative measure and hopefully also answer questions that could not be addressed before due to the limitation of consumption surveys.

The role of expenditures and how they respond to changes in income and wealth is at the heart of policy design, ranging from the design of fiscal stimuli to the design of social insurance programs. Still, how expenditures translate into welfare remains an underexplored question for which the rich information from registry data can provide new complementary insights to standard surveys. An important next step for the evaluation of welfare is also to take the role of public goods seriously and account for the consumption of public goods in the measurement of consumption. Registry data on education, public transportation, health expenditures, etc. can again be particularly useful here and further improve our understanding of consumption smoothing and inequality.

References

Alstadsæter, Annette, Niels Johannesen, and Gabriel Zucman, "Tax Evasion and Inequality," *American Economic Review*, June 2019, *109* (6), 2073–2103.

Attanasio, Orazio P. and Luigi Pistaferri, "Consumption inequality," *The Journal of Economic Perspectives*, 2016, *30* (2), 3–28.

Autor, David, Andreas Kostøl, Magne Mogstad, and Bradley Setzler, "Disability benefits, consumption insurance, and household labor supply," *American Economic Review*, 2019, *109* (7), 2613–54.

Baily, Martin N., "Some Aspects of Optimal Unemployment Insurance," *Journal of Public Economics*, 1978, *10* (3), 379–402.

Baker, Scott R, Lorenz Kueng, Steffen Meyer, and Michaela Pagel, "Measurement Error in Imputed Consumption," Working Paper 25078, National Bureau of Economic Research September 2018.

Ball, Steffan and Hamish Low, "Do self-insurance and disability insurance prevent consumption loss on disability?," *Economica*, 2014, *81* (323), 468–490.

Bee, Adam, Bruce D. Meyer, and James X. Sullivan, "The Validity of Consumption Data: Are the Consumer Expenditure Interview and Diary Surveys Informative?," in "Improving the Measurement of Consumer Expenditures" NBER Book Series Studies in Income and Wealth, University of Chicago Press, October 2013, pp. 204–240.

Blundell, Richard, Luigi Pistaferri, and Ian Preston, "Consumption Inequality and Partial Insurance," *The American Economic Review*, 2008, *98* (5), 1887–1921.

_ , _ , and Itay Saporta-Eksten, "Consumption Inequality and Family Labor Supply," *American Economic Review*, February 2016, *106* (2), 387–435.

Borusyak, Kirill and Xavier Jaravel, "Revisiting Event Study Designs," 2016. Available at SSRN: https://ssrn.com/abstract=2826228.

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

Chetty, Raj, "A General Formula for the Optimal Level of Social Insurance," *Journal of Public Economics*, 2006, *90* (10-11), 1879–1901.

Cullen, Julie Berry and Jonathan Gruber, "Does unemployment insurance crowd out spousal labor supply?," *Journal of Labor Economics*, 2000, *18* (3), 546–572.

Eika, Lasse, Magne Mogstad, and Ola Vestad, "What Can We Learn About Household Consumption From Information on Income and Wealth," 2017. Working Paper.

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

Finkelstein, Amy, Erzo F. P. Luttmer, and Matthew J. Notowidigdo, "What Good is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption," *Journal of the European Economic Association*, January 2013, *11*, 221–258.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," *American Economic Review*, July 2014, *104* (7), 2075–2126.

Koijen, Ralph, Stijn Van Nieuwerburgh, and Roine Vestman, "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," University of Chicago Press, July 2014, pp. 308–346.

Kolsrud, Jonas, Camille Landais, Peter Nilsson, and Johannes Spinnewijn, "The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden," *American Economic Review*, 2018, *108* (4-5), 985–1033.

Kostøl, Andreas Ravndal and Magne Mogstad, "Earnings, disposable income, and consumption of allowed and rejected disability insurance applicants," *American Economic Review*, 2015, *105* (5), 137–141.

Kreiner, Claus Thustrup, David Dreyer Lassen, and Søren Leth-Petersen, "Measuring the accuracy of survey responses using administrative register data: evidence from Denmark," in "Improving the measurement of consumer expenditures," University of Chicago Press, 2014, pp. 289–307.

Landais, Camille and Johannes Spinnewijn, "The Value of Unemployment Insurance," Technical Report, CEPR working paper 2019.

Low, Hamish and Luigi Pistaferri, "Disability insurance and the dynamics of the incentive insurance trade-off," *American Economic Review*, 2015, *105* (10), 2986–3029.

Lundberg, Jacob and Daniel Waldenström, "Wealth Inequality in Sweden: What can we Learn from Capitalized Income Tax Data?," *Review of Income and Wealth*, 2017, *64*, 517–541.

Medina, Leandro and Friedrich Schneider, "Shadow economies around the world: what did we learn over the last 20 years?," *IMF Working Paper*, 2018.

Meyer, Bruce D. and Wallace K.C. Mok, "Disability, earnings, income and consumption," 2018.

Pistaferri, Luigi, "Household Consumption: Research Questions, Measurement Issues, and Data Collection Strategies," Mimeo, Stanford University 2015.

Poterba, James M., "Taxation and Housing: Old Questions, New Answers," *The American Economic Review*, 1992, *82* (2), 237–242.

Saez, Emmanuel and Gabriel Zucman, "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *The Quarterly Journal of Economics*, 2016, *131* (2), 519–578.

Stephens, Melvin Jr., "The long-run consumption effects of earnings shocks," *Review of Economics and Statistics*, 2001, *83* (1), 28–36.

Figure 1: DISTRIBUTION OF HOUSEHOLD EARNINGS & HOUSEHOLD REGISTRY-BASED CONSUMPTION EXPENDITURES (2000-2007)
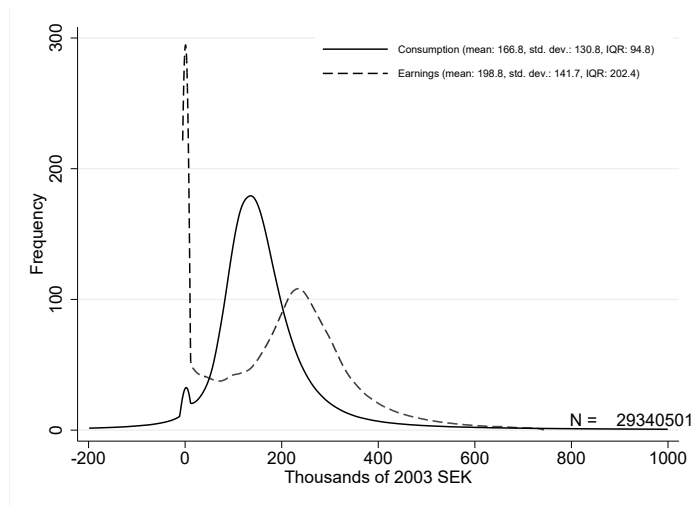

Notes: The Figure shows the distribution of household earnings and the distribution of our registry-based measure of household consumption for the same sample for the period 2000-2007. Each observation is a household-year. All variables are expressed in thousands of 2003 SEK. The IQR is the interquartile range.

Figure 2: INDIVIDUAL EARNINGS & INDIVIDUAL CONSUMPTION EXPENDITURES PROFILES BY AGE & BIRTH COHORT
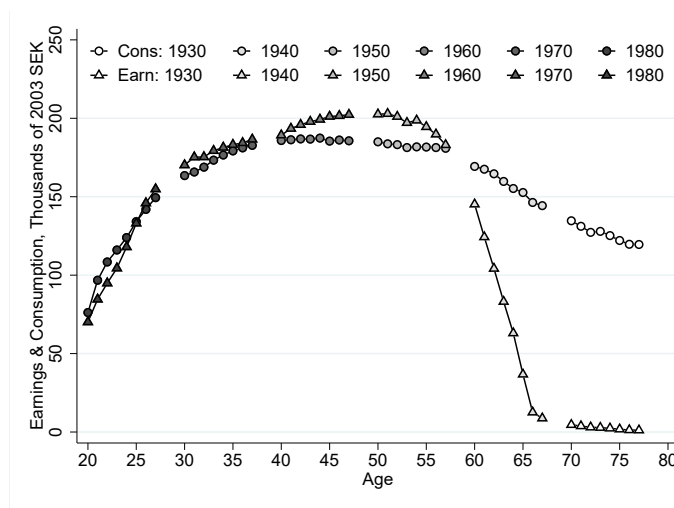
Notes: The Figure shows the evolution of both earnings and consumption by age and birth cohort. To construct this graph, we split the sample into 6 birth cohorts (1930, 1940, 1950, 1960, 1970 and 1980) and follow each birth cohort as they age over the eight years of our consumption measure (2000-2007).

Figure 3: NA, HUT and Registry-Based Consumption Expenditures
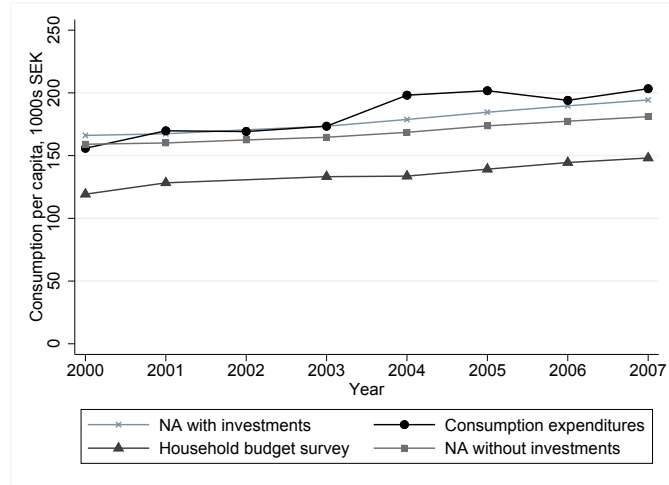
Notes: The Figure reports the evolution of consumption expenditures per adult (defined as being 16 or over) in the national accounts (NA), in the HUT consumption surveys, and in our registry-based measure. For NA, we display two series: the official series, and one in which investments made by households (refurbishing of real estate, etc.) are added back to NA consumption. The HUT survey measure of consumption does not include imputed rents while both the registry measure and the NA measure do. In Appendix C we describe in more detail the sources and concepts used for measuring consumption in the NA and in the HUT.

Figure 4: Distribution of HUT and Registry-Based Consumption Expenditures: 2003-2007
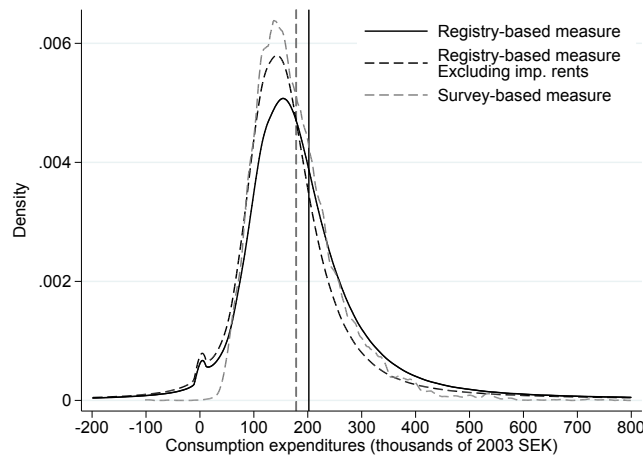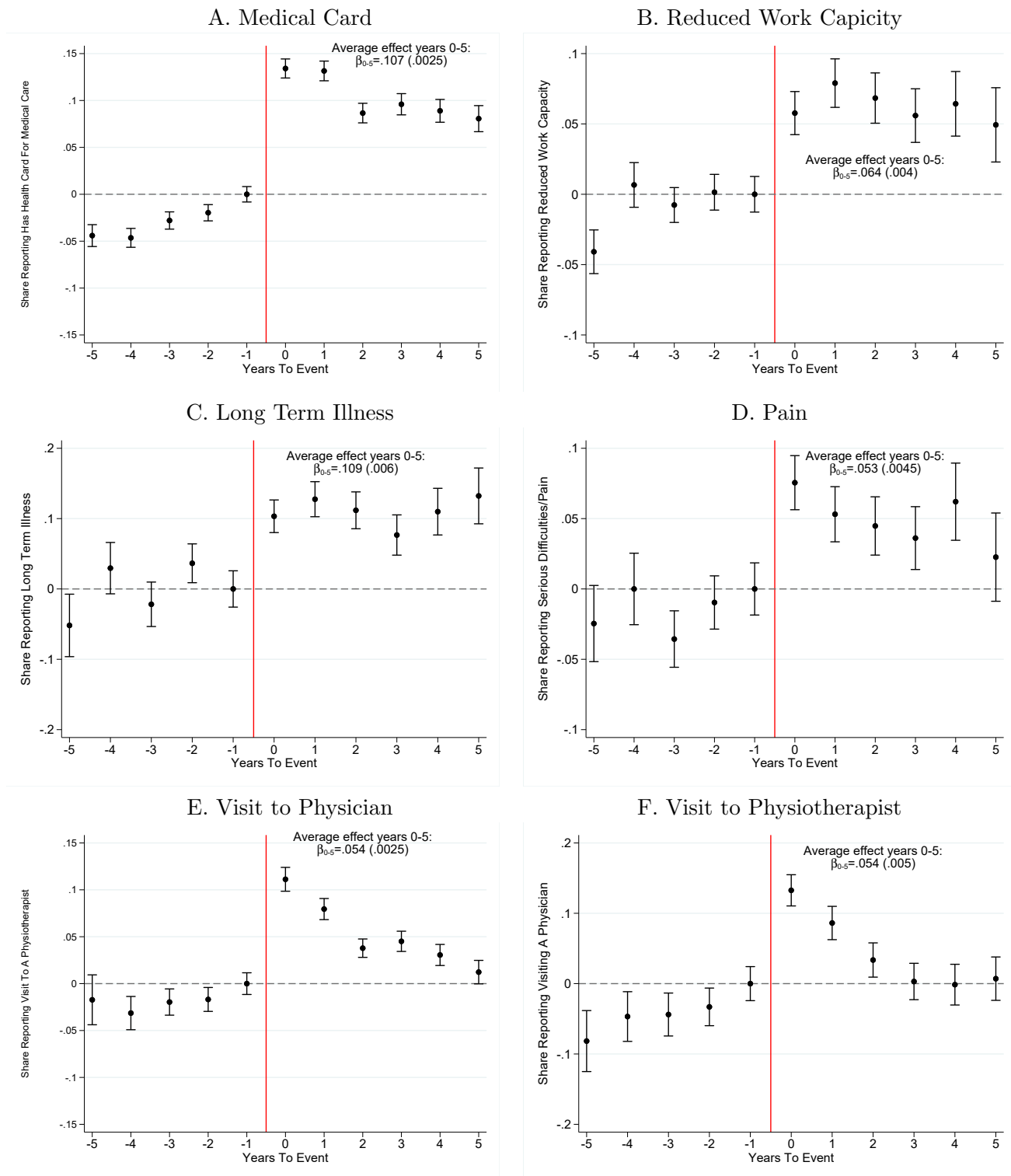

Notes: The Figure compares the distribution of consumption expenditures per adult aged 25 to 55 in the survey and in the registry data over the period 2003 to 2007. For all adults aged 25 to 55, consumption is computed as total household consumption divided by the number of adults in the household. To make the expenditure measure comparable in both samples, we also show the distribution of the registry based measure when we exclude imputed rents from household expenditures. All figures are in thousands of constant 2003 SEK. The vertical lines denote the average for each distribution. The HUT and registry-based measure excluding imputed rents have similar averages: 178.5 kSEK for the HUT and 178.8 for the registry-based measure excluding imputing rents

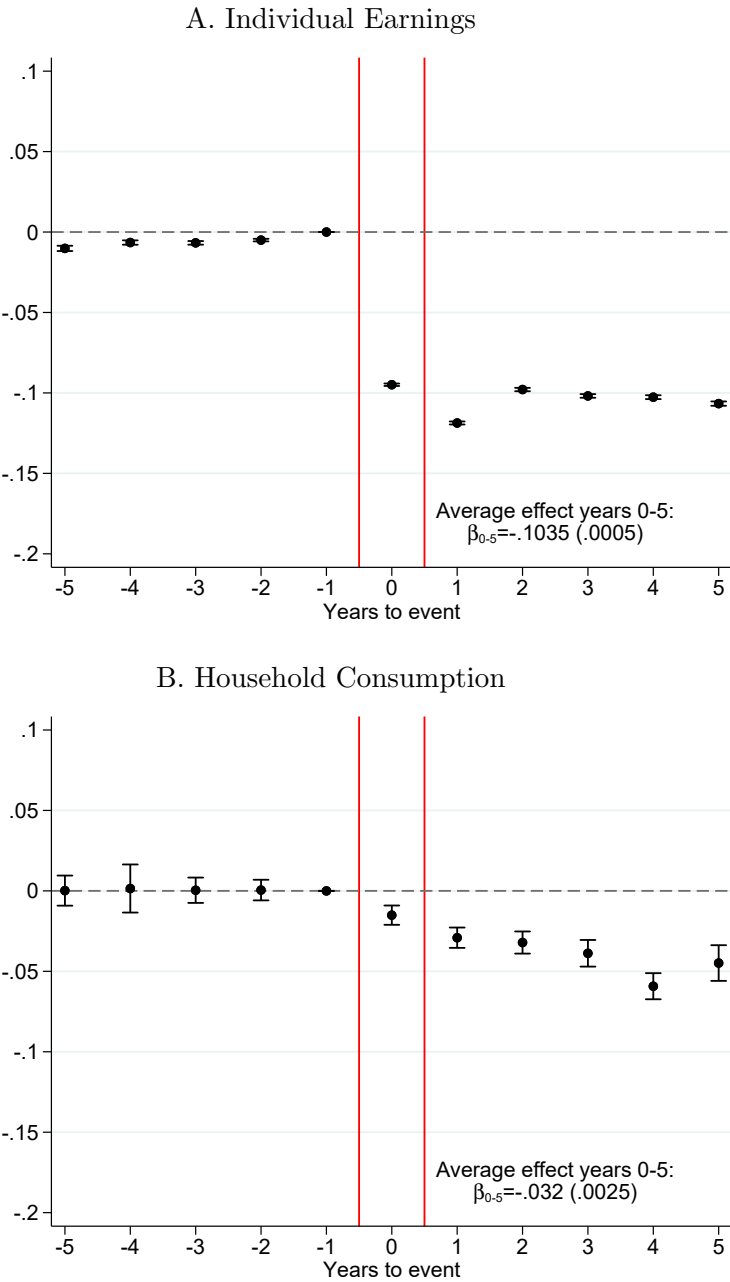Figure 5: HUT and Registry-Based Consumption Expenditures
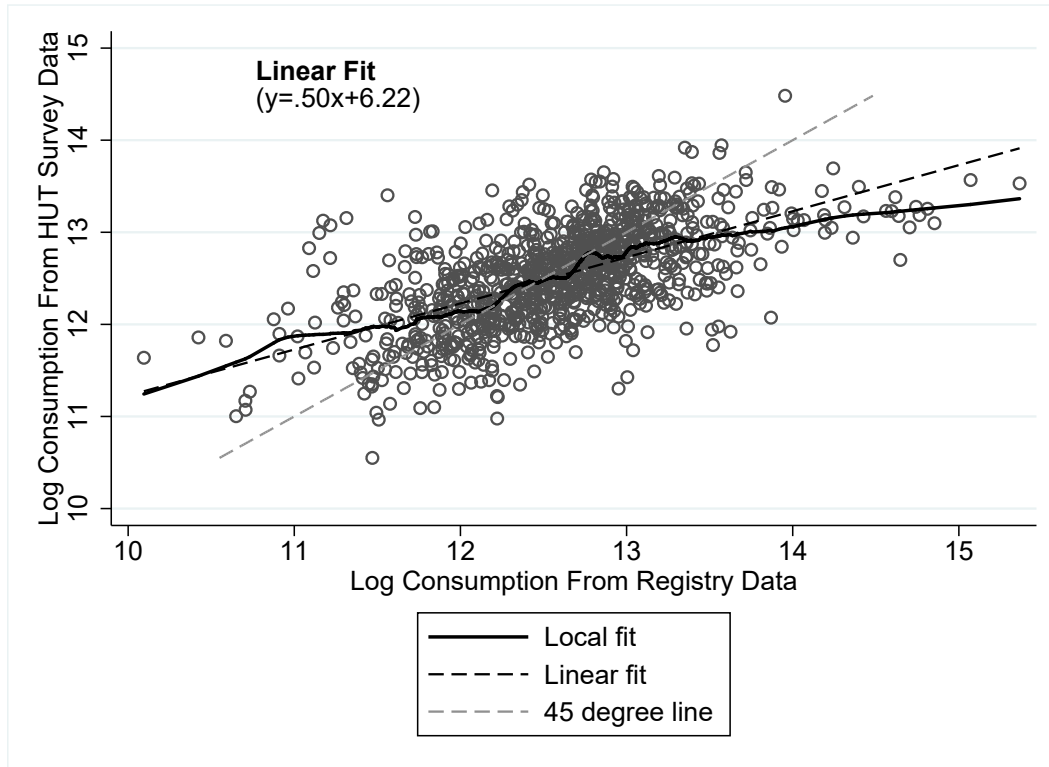
Notes: For years 2003 to 2009, the identifiers in the HUT survey can be used to match surveyed individuals to the registry data. The Figure directly compares consumption measures in the HUT and in the registry measure for the same households. It plots the log of total household consumption in HUT vs log of household expenditures in the registry. We restrict the sample to all household from the HUT surveys from 2003 to 2007 that are surveyed in the month of December to make sure that the annual expenditure measure from the HUT (which includes recall items such as durables from the last 12 months) corresponds to the annual measure in the registry data which spans all expenditures from December of year $n-1$ to December of year n. When the household composition in the HUT and registry data differs, we reconstruct households for the registry measure using the exact same individuals that are listed as members of the household in the HUT. To make expenditure measure comparable in both samples, we exclude imputed rents from the registry measure. Sample size for this graph is 1013 household observations.

Figure 6: Health Shocks: Health Outcomes Around First Occurrence of Sickness Benefit Receipt
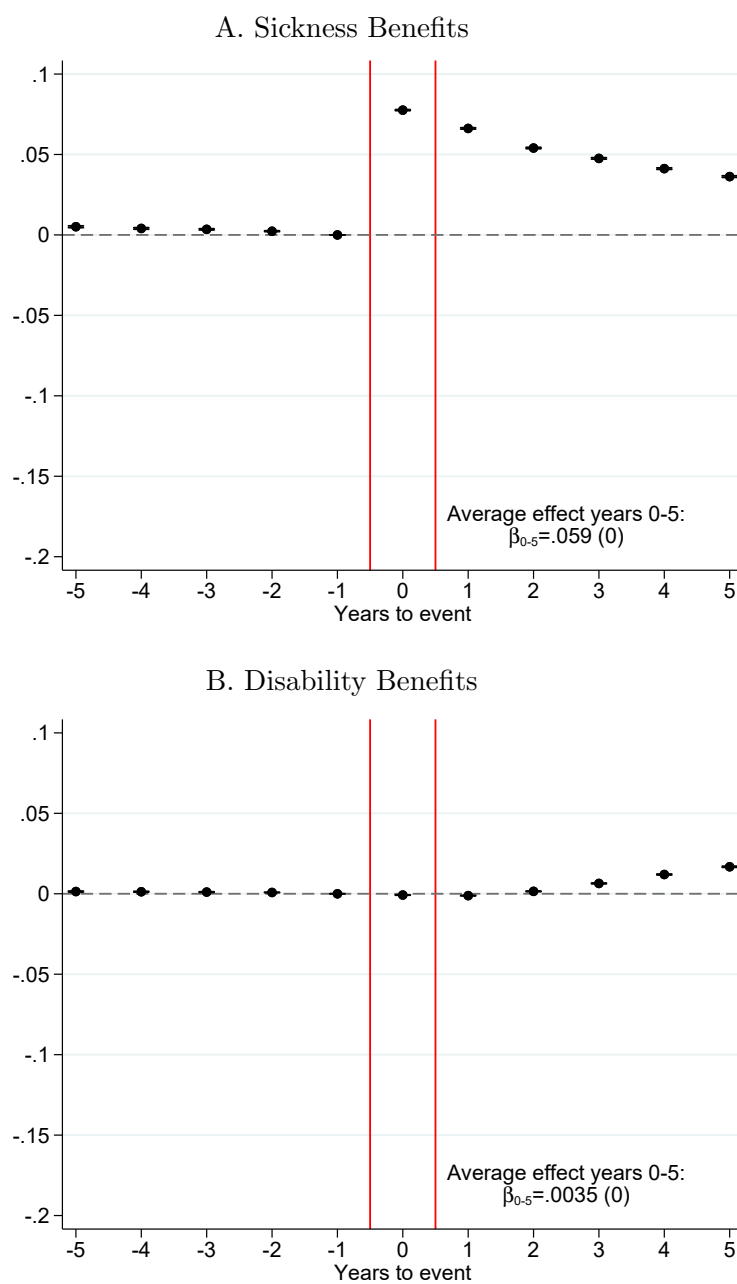
A. Medical Card

B. Reduced Work Capicity

C. Long Term Illness

D. Pain

E. Visit to Physician

F. Visit to Physiotherapist

Notes: The Figure reports event study estimates of various health outcomes from the HEK survey, around the occurrence of a health shock, using specification (3) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. All estimates are expressed relative to event year -1. Health shocks are defined as the first occurrence of sickness benefits receipt by an individual (i.e. a work incapacitation larger than 14 days). We restrict the sample to individuals aged 25 to 55 at the time of the event. The medical card in Panel A is used for drugs expenditures. Individuals qualify for this card when annual drugs spending exceeds a threshold. See text for details.

Figure 7: EARNINGS AND CONSUMPTION RESPONSES TO HEALTH SHOCKS

A. Individual Earnings

Average effect years 0-5:
β_{0-5}=-.1035 (.0005)

B. Household Consumption

Average effect years 0-5:
β_{0-5}=-.032 (.0025)

Notes: The Figure reports event study estimates of the earnings, and consumption responses to a health shock, using specification (2) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. All estimates are expressed as a fraction of total household expenditures as of event year -1. Health shocks are defined as the first occurrence of sickness benefits receipt by an individual (i.e. a work incapacitation larger than 14 days). We restrict the sample to individuals aged 25 to 55 at the time of the event. Individuals may transition to disability benefits after a year of sickness benefit receipt. In Panel A, earnings are the gross earnings of the individual experiencing the health shock. In Panel B, we report total household consumption where the composition of the household is fixed as of the last year prior to event. For each panel, we also report the average long term dynamic treatment effect for event years 0 to 5: $\bar{\beta}_{0-5}$. The number of household×year observations, including control households is N=16,364,347. See text for details.

Figure 8: Social Insurance & Transfer Responses to Health Shocks

A. Sickness Benefits

Average effect years 0-5:
$\bar{\beta}_{0\text{-}5}$=.059 (0)

Years to event

B. Disability Benefits

Average effect years 0-5:
$\bar{\beta}_{0\text{-}5}$=.0035 (0)

Years to event

Notes: The Figure reports event study estimates of the benefits responses to a health shock, using specification (2) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. All estimates are expressed as a fraction of total household expenditures as of event year -1. Health shocks are defined as the first occurrence of sickness benefits receipt by an individual (i.e. a work incapacitation larger than 14 days). We restrict the sample to individuals aged 25 to 55 at the time of the event. Individuals may transition to disability benefits after a year of sickness benefit receipt. For each panel, we also report the average long term dynamic treatment effect for event years 0 to 5: $\bar{\beta}_{0-5}$. See text for details.

Figure 9: CONSUMPTION RESPONSES TO HEALTH SHOCKS: INSURANCE WITHIN THE HOUSEHOLD

A. Individual Residual Consumption (Affected Individual)
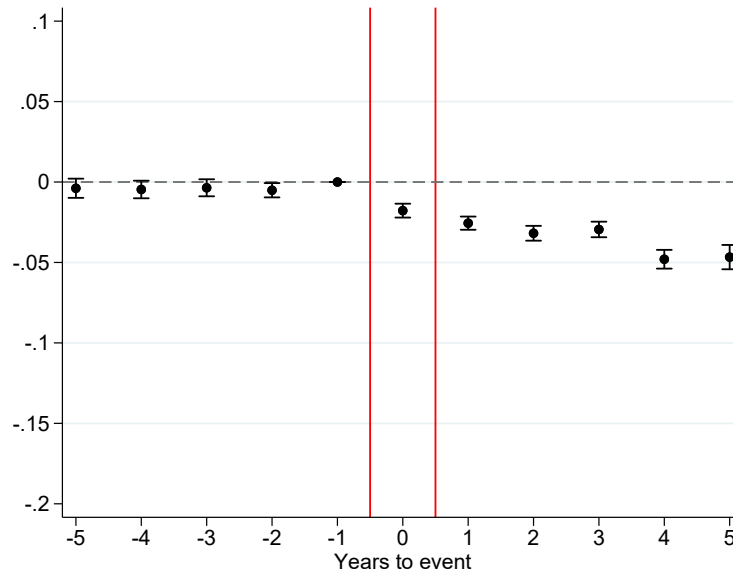

B. Added Worker Effect - Earnings of all Other HH Members

Average effect years 0-5:
β_{0-5}=-.0105 (.0005)
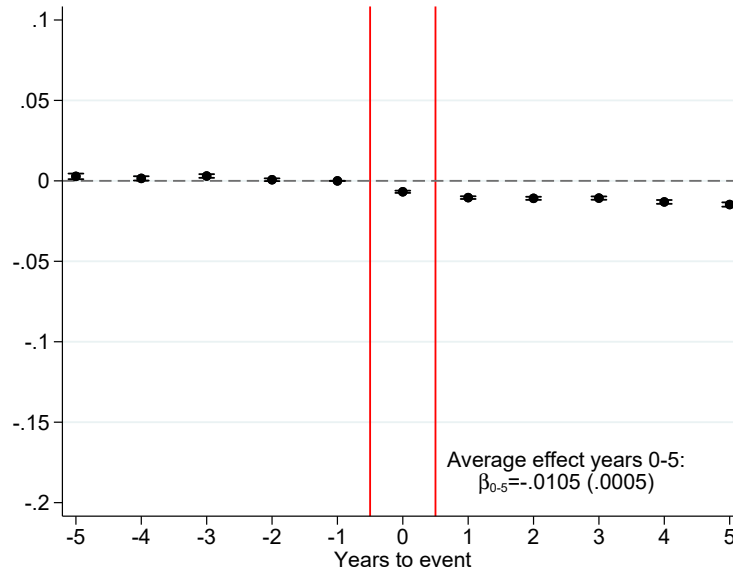
Notes: The Figure reports event study estimates of the consumption responses to a health shock, using specification (2) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. All estimates are expressed as a fraction of total household expenditures as of event year -1. Health shocks are defined as the first occurrence of sickness benefits receipt by an individual (i.e. a work incapacitation larger than 14 days). We restrict the sample to individuals aged 25 to 55 at the time of the event. Individuals may transition to disability benefits after a year of sickness benefit receipt. In Panel A, we focus on the individual consumption of the individual experiencing the health shock. In Panel B, we report total household consumption of all other members of the household, where the composition of the household is fixed as of the last year prior to event. See text for details.

Figure 10: INTERTEMPORAL CONSUMPTION SMOOTHING IN RESPONSE TO HEALTH SHOCKS

A. Consumption Out of Debt
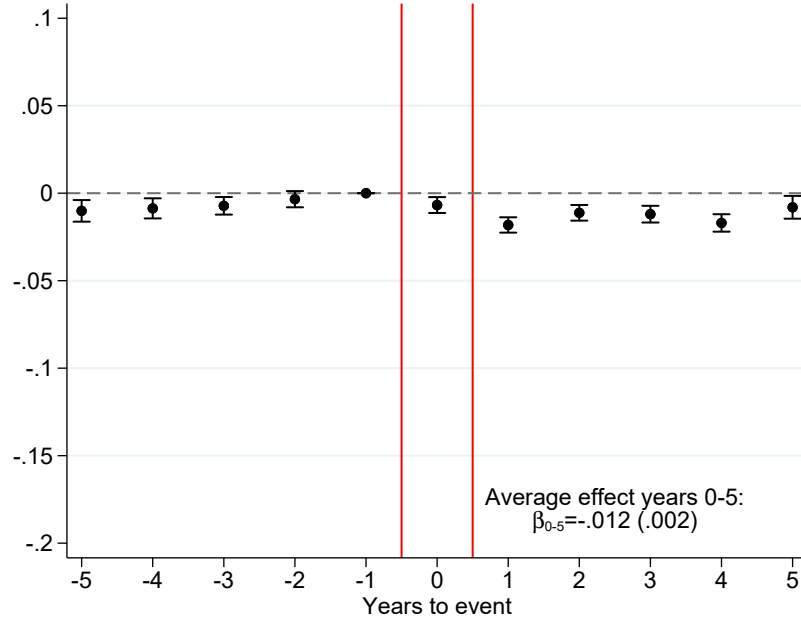
Average effect years 0-5:
β_{0-5}=-.012 (.002)

B. Consumption Out of Assets

Average effect years 0-5:
β_{0-5}=.0015 (.0025)
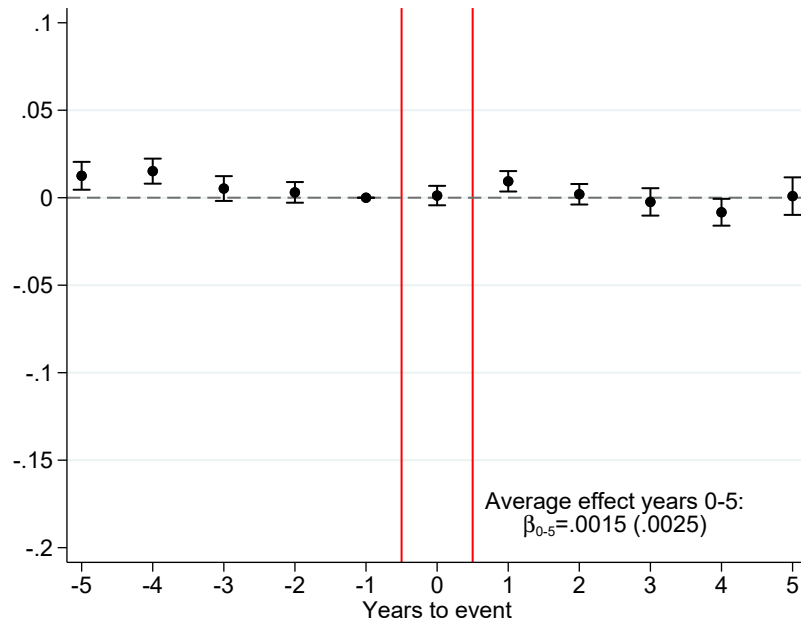
Notes: The Figure reports event study estimates of responses of consumption out of debt and consumption out of assets, around the occurrence of a health shock, using specification (2) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. Consumption out of assets excludes consumption out of real estate. All estimates are expressed as a fraction of total household expenditures as of event year -1. Health shocks are defined as the first occurrence of sickness benefits receipt by an individual (i.e. a work incapacitation larger than 14 days). We restrict the sample to individuals aged 25 to 55 at the time of the event. Individuals may transition to disability benefits after a year of sickness benefit receipt. For each panel, we also report the average long term dynamic treatment effect for event years 0 to 5: $\bar{\bar{\beta}}_{0-5}$. See text for details.

Figure 11: Decomposition of Consumption Expenditure Responses to Events
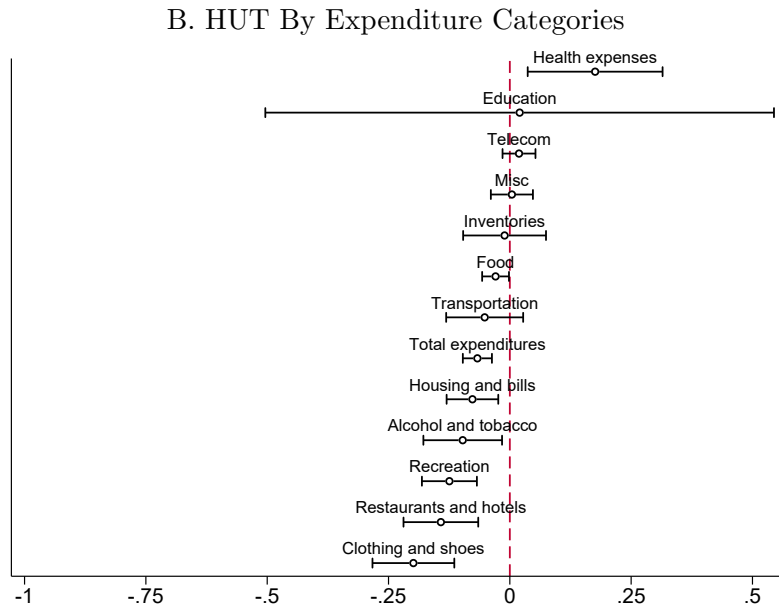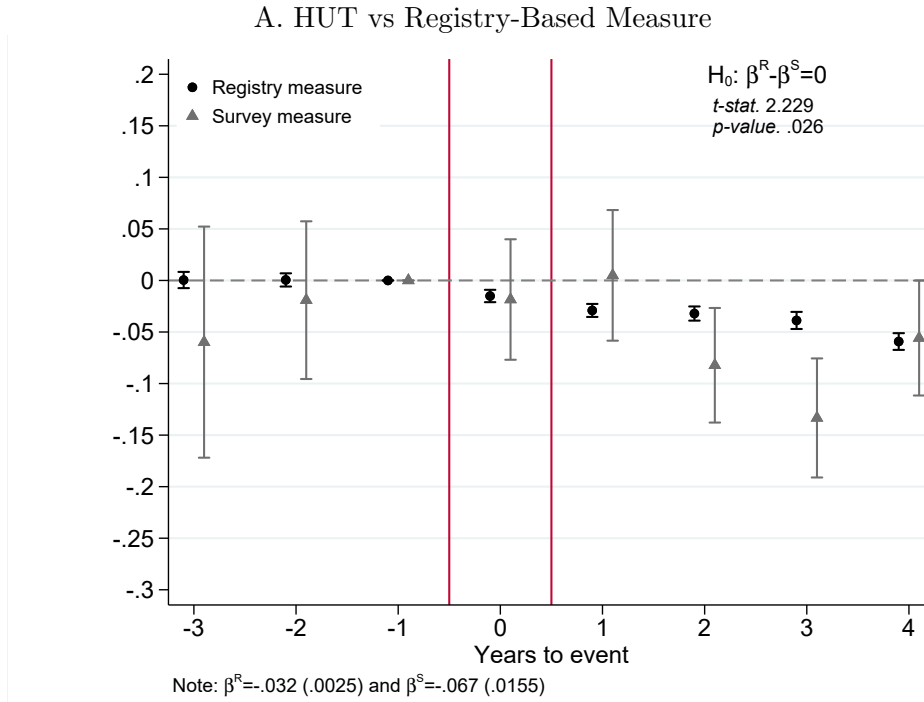
A. All Health Shocks

B. Large Health Shocks
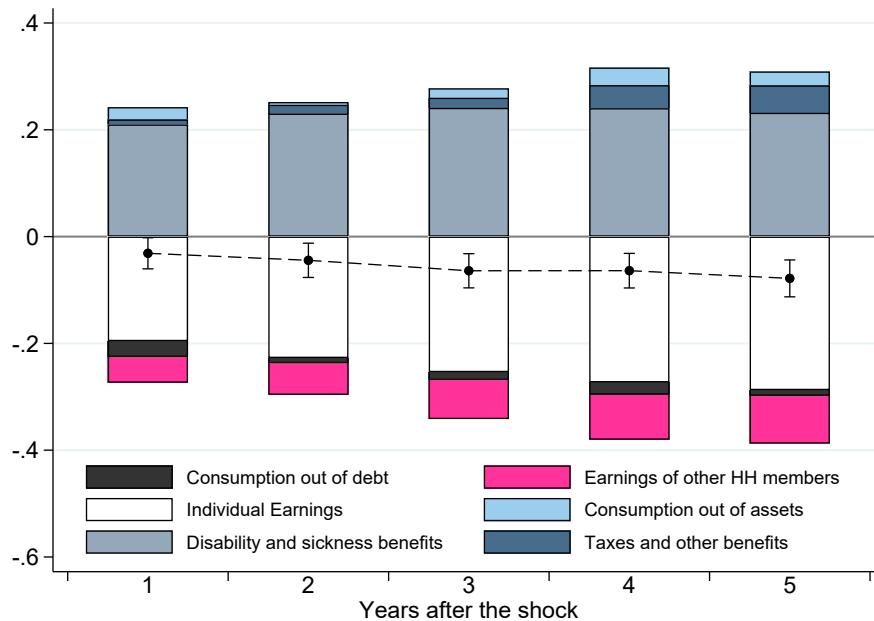
Notes: The Figure decomposes the household consumption expenditure response to health shocks for event time years $J_{it} = 1$ to $J_{it} = 5$. We decompose consumption into six main components: own earnings, earnings of other members of the household, sickness and disability benefits, consumption out of debt, consumption out of assets, and other transfer and taxes. The black dot represents the total change in consumption expenditures relative to event year -1, with its 95% confidence band. All estimates are expressed as a fraction of total household expenditures as of event year -1. Each component of the bar represents the contribution of each component to the total change in consumption expenditures. These components are estimated using specification (2) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. Panel A decomposes the response for the entire sample of individuals subject to a health shock. Panel B focuses on individuals subject to larger health shocks who end up becoming disabled, measured by their transitioning to disability benefits.

Figure 12: Consumption Dynamics Around Event: Survey Measure Estimates

A. HUT vs Registry-Based Measure

Note: β^R=-.032 (.0025) and β^S=-.067 (.0155)

B. HUT By Expenditure Categories

Notes: The Figure reports event study estimates of consumption dynamics around health shocks. Panel A compares estimates from the HUT survey with estimates from the registry measure of consumption where we remove imputed rents to make consumption measure consistent with the HUT survey. For the HUT estimates, we use specification (3). For the registry measure the event study estimates are obtained using specification (2). The number of household×year observations, including control households is: N=8980 for HUT, N=16,364,347 for registry. We report the $t-$stat and p-value of a test of equality of the survey estimate (β^s) and the registry estimate (β^r) of the average change in consumption in the 5 years post event. Panel B decomposes the average 5 years consumption response post-event from the HUT data into various expenditure categories. See text for details.

Table 1: RELATIVE CHANGE IN CONSUMPTION DUE TO SICKNESS, BY CATEGORY

	Estimate	Standard Error	p-Value of Equality Test
Panel A - Demographics			
Age At Event			
25-35	-.052	.006	
45-55	-.027	.005	.001
Married/Cohabitation Status			
With Partner	-.035	.004	
Single	-.033	.004	.724
Panel B - Income & Wealth			
Relative Income level			
Upper Quartile	-.048	.008	
Lower Quartile	-.006	.006	<.001
Relative Net Wealth			
Upper Quartile	-.032	.010	
Lower Quartile	-.021	.004	.286
Relative Debt level			
Upper Quartile	-.049	.008	
Lower Quartile	-.034	.005	.091
Panel C - Severity of Health Shocks			
Relative Duration of Sickness Spell			
Upper Quartile	-.043	.007	
Lower Quartile	-.021	.006	.014
Disability Status			
Benefits Claimed	-.040	.013	
Benefits Not Claimed	-.031	.003	.499

Notes: The Table shows the average long term (6 years) dynamic treatment effect $\bar{\beta}_{0-5}$ from the baseline specification (2) for different subsamples. The final column reports the p-value testing the equality of the estimated effect for the two corresponding subsamples.

Appendix A   Additional Graphs And Tables

Figure A.1: CHANGES IN CAR OWNERSHIP IN RESPONSE TO HEALTH SHOCKS

Average effect years 0-5:
β_{0-5}=-.085 (.0115)
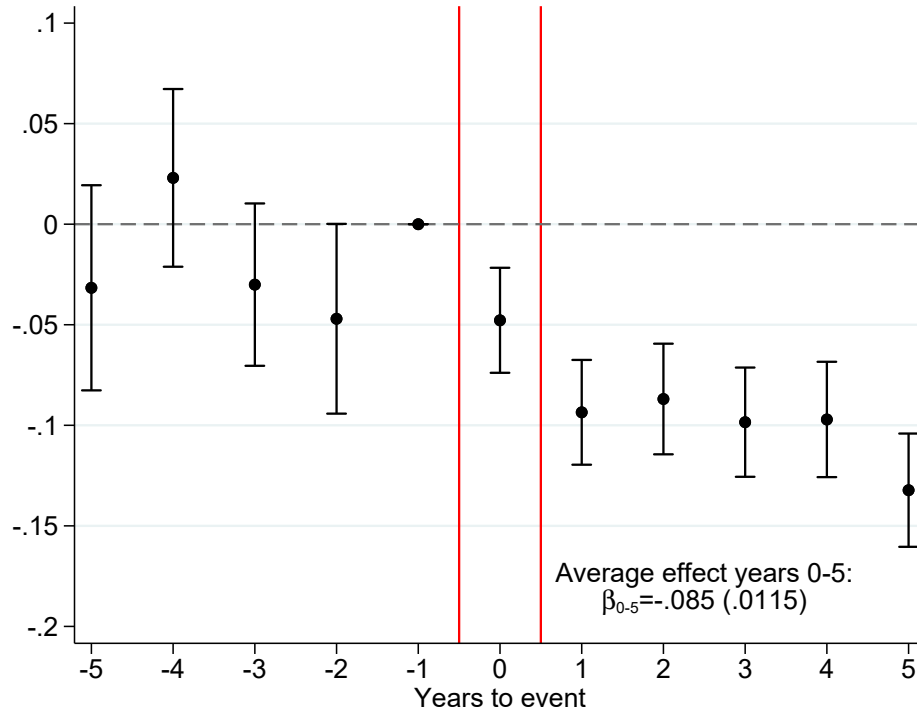
Notes: The Figure reports event study estimates of responses of net car purchases, around the occurrence of a health shock, using specification (2) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. Net car purchases is defined as total cars bought minus total cars sold by an individual in a given year. All estimates are expressed as a fraction of total household expenditures as of event year -1. Health shocks are defined as the first occurrence of sickness benefits receipt by an individual (i.e. a work incapacitation larger than 14 days). We restrict the sample to individuals aged 25 to 55 at the time of the event. Individuals may transition to disability benefits after a year of sickness benefit receipt. For each panel, we also report the average long term dynamic treatment effect for event years 0 to 5: $\bar{\beta}_{0-5}$. See text for details.

Figure B.1: DISTRIBUTIONAL EFFECTS: QUANTILE REGRESSIONS ESTIMATES
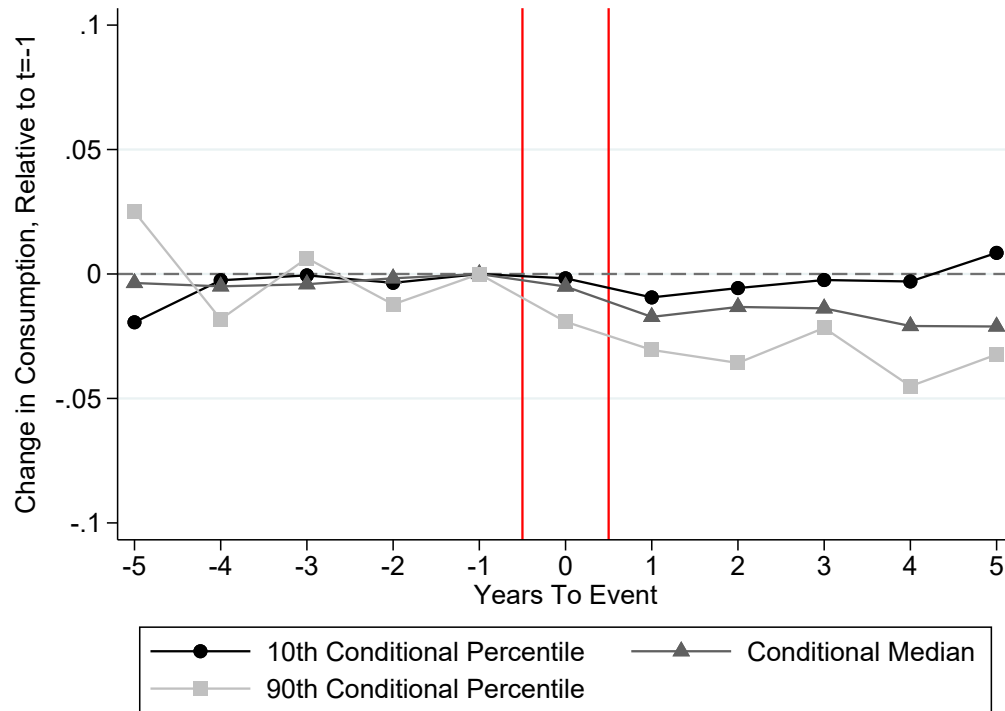
Notes: The Figure reports quantile regression event study estimates of responses of various conditional quantiles of the consumption expenditure distribution, around the occurrence of a health shock, using specification (2) on the sample of treated individuals and a control group of individuals obtained from nearest-neighbor matching on pre-event characteristics. All estimates are expressed as a fraction of total household expenditures as of event year -1. Health shocks are defined as the first occurrence of sickness benefits receipt by an individual (i.e. a work incapacitation larger than 14 days). We restrict the sample to individuals aged 25 to 55 at the time of the event. See text for details.

Table 2: Overview of income and wealth components

Disposable income (per capita)	160,930 SEK

Income from labor and transfers	Percentage of gross income
Labor income	55.66
Benefits replacing labor income	9.27
Pensions and means tested benefits	15.95
Tax deductions and other transfers	6.67
Taxes	-27.91

Income from holding capital	
Imputed rents	9.00
Dividends from securities	2.18
Other capital income	0.83
Interest on deposits	0.44
Interest on liabilities	-3.77

Net capital gains (per capita)	56,263 SEK

Capital gains components	Percentage of gross capital gains
Positive gains from real assets	66.09
Negative gains from real assets	-1.39
Positive gains from financial assets	33.91
Negative gains from financial assets	-10.27

Net wealth (per capita)	523,119 SEK

Wealth components	Percentage of gross wealth
Real wealth	69.32
Securities	19.84
Bank account holdings	9.07
Other wealth	1.77
Debt	-29.24
Observations	58,396,928

Notes: Income and wealth are measured for years 2000-2007 in 2003 year's Swedish kronor (1 SEK2003 ≈ 0.11 USD2003). The sample is all individuals residing in Sweden aged 16 years and above. "Gross income" is disposable income from labor and capital plus paid taxes and imputed rents. "Labor income" is income from employment and self-employment. "Benefits replacing labor income" are benefits suchs as unemployment and sickness insurance and parental leave benefits. In "Pensions and mean tested benefits" social aid and housing aid are included. "Tax deductions and other transfers" includes benefits such as child support. "Other capital income" includes income from subletting homes, gains from selling real estate and tax on lottery wins (primarily wins from outside of the EEA). "Other wealth" consists of cars, boats, art etc. that individuals have reported to the tax authority.

Table 3: SUMMARY STATISTICS: HEALTH SHOCK SAMPLE

	Mean	P10	P50	P90
Panel A: Treated individuals				
I. Demographics				
Age	40.0	29	40	52
Frac. Tertiary	0.43	-	-	-
Frac. Men	0.43	-	-	-
Frac. Married	0.36	-	-	-
II. Income and Wealth SEK 2003(K)				
Individual Gross Earnings	178.6	0	194	328
Household Gross Earnings	309.6	58	269	565
Household Consumption	296.2	93	237	546
Household Net Wealth	385.1	-229	61	1,264
Household Bank Holdings	45.0	0	0	117
Household Debt	434.5	0	249	1,052
Household Sickness Benefits	5.2	0	0	2
Panel B: Control individuals				
I. Demographics				
Age	40.0	29	40	52
Frac. Tertiary	0.42	-	-	-
Frac. Men	0.44	-	-	-
Frac. Married	0.37	-	-	-
II. Income and Wealth SEK 2003(K)				
Individual Gross Earnings	188.7	0	199	352
Household Gross Earnings	327.3	65	283	594
Household Consumption	311.4	93	248	574
Household Net Wealth	490.5	-211	115	1,448
Household Bank Holdings	53.8	0	0	135
Household Debt	464.2	0	269	1,108
Household Sickness Benefits	7.8	0	0	13

Notes: All values, except sickness benefits, are reported as averages in the year before the shock, expressed in thousands of kronor, in 2003 values. Consumption expenditures, net wealth, bank holdings and debt are expressed at the household level, gross earnings and sickness benefits are those of the treated individual (or its matched counterpart). For sickness benefits, we report the sum of the benefits received during the five years prior to the shock. Higher education refers to college/university less than 2 years, college/university of 2 years and more and PhD.

Appendix B   Further Details on the Construction of the Registry-Based Measure

Disposable Income   Our consumption measure builds on two main blocks. One is net-of-tax income from labor and transfers and the other is the flow from asset holdings. Net-of-tax income from labor and transfers are stored in the variable disposable income (DispInk) in LISA. The disposable income measure includes all labor income, both from employment ans self-employment, all public transfers and transfers from collective agreements such as occupational pensions and workers' compensations and annuities. The disposable income measure also includes capital income and student loans. From the sum of gross-of-tax labor income and transfers, Statistics Sweden subtracts the sum of taxes paid.

To avoid double counting we subtract changes in student loans and most of the capital income. Specifically we subtract interest payments from bank account holdings, dividends, gains from selling real estate and negative capital income from paid interest rates. However, the tax paid on these gains are kept in the disposable income measure. The only income capital from capital to still be included are incomes from subletting and lottery wins, primarily such from outside the EEA.

Bank Account Holdings   Before 2006 bank account holdings are only reported if their interest payments exceed 100 kr. In 2006 and 2007 all bank account holdings surpassing 10,000 kr are reported, irrespective of the interest payments. Smaller amounts are also reported if their interest payments surpass 100 kr. The regime change in 2006 led to a large increase in bank account holdings: the variable increases on average by about 35,000 kr. In 2005 about 70% of the bank account holdings that can be found in national accounts data can be found in the wealth register. In 2006 the bank account holdings in the wealth data is 97% of the holdings reported in national accounts. To avoid a large drop in the residual consumption measure in 2006 bank account holdings are censored in the same way as in 2005. This leads to a smaller increase in bank account holdings, about 22,000 kr, and to consumption increasing by about 10,000 kr on average. This adjustment improves the correspondence between consumption in national accounts and the residual consumption measure significantly. Using the 2006 censoring rule on the 2005 data does not address the problem that a much larger share of bank account holdings are reported in the data. As interest rates in 2005 was below 1% means that almost all reported holdings surpass 10,000 kr. The 2006 censoring rule is applied for 2007s consumption measure. Using the 2005 bank account censoring rule also for 2006 improves the match between survey consumption in HUT and the residual consumption measure.

KURU data   To assign a price on the individual securities found in the KURU data we use closing prices from the last day of trade in December 2000 to 2007. We use data provided by SIX which covers listed securities from about 1,500 different securities exchanges and contributors.[39]

---

[39]Securities which are not traded on Swedish stock exchanges, but are traded abroad, are reported in their respective currencies. To convert these prices into Swedish kronor we use exchange rates for the last day of trade each year.

This ensures that not just listed securities found on Swedish stock exchanges such as the Stockholm Stock Exchange (SSE) and First North (an exchange for smaller firms and start-ups) are used. In the data we find 40% of all the stocks and mutual funds from the KURU data and 43% of all bonds.[40] When Statistics Sweden constructed the wealth statistics they used price information only on securities listed on the SSE and first North. This means that listed securities which were not traded at either SSE or First North were not included in the wealth statistics.

Table 4 reports some descriptive statistics from the KURU data. The first column shows the share of securities held by the individuals in our sample that we find in the asset price data. As mentioned previously we find about 13 percentage points or almost 50% more securities compared to Statistics Sweden. The main explanation for this is that we use data on securities that are not exclusively traded on Swedish stock exchanges. The second column shows the share of individual-security cells we can match to our price data. For stocks and funds these are about the same as the numbers in column 1. For bonds, however, we are able to assign a price to almost 90% of all bonds held by the individuals in our sample. In column 3, we compare the total worth of price assigned securities to the value of securities found by Statistics Sweden. Notably we can attach a price to a significantly larger amount of stocks and bonds which translates into a much larger reported stock and bond wealth compared to the original wealth statistics. The sum of stock wealth is almost 35 times higher compared to what is found by Statistics Sweden. However, for funds we only find 8% of the wealth that Statistics Sweden found. Finally, column 4 shows the share of individuals with imputed security flows for each type of security. 9.4% of the individuals have imputed consumption expenditure flows from stocks while the corresponding figures for bonds and funds are 2.7 and 47.4% respectively. Note that the imputations are only made for cases where the sum of held security wealth in the wealth statistics is larger compared to what we find.

The significant surge in stock and bond wealth is primarily due to large amounts of foreign traded stocks held by a small number of wealthy individuals. But when comparing our calculations with statistics Sweden's we find about twice as many stock owners and twice as many bond owners (for which the securities can be assigned with a market price). All else being equal, this would however only explain a small share of the large difference in stock and bond wealth found by us and Statistics Sweden.

Outliers Wealthy individuals sometimes make very large changes in their asset positions without them being offset by corresponding changes in other types of assets. We are worried that these are assets we miss out on because of the incomplete coverage of the KURU data. As a result, these non-offset position changes affects aggregate consumption. For instance, in 2006 a single individual sells stocks to a worth of 3 billion kr which is not offset by a comparable change in real estate or bonds. Just one such transaction adds about 500 kr to average consumption. To manage

---

The currency data is provided by Sveriges Riksbank (the Swedish central bank) and are available for download on its site: `www.riksbank.se/en`.

[40]We have also used data from ThomsonReuters' service DataStream. In this data we found 11% of all stocks and mutual funds and 8% of all bonds. The main explanation seems to be that ISIN codes which are not currently used are also not retained in the data.

Table 4: Coverage ratios and wealth assessments

	Share of ISINs found	Share of individual-security cells found	Security wealth found relative to SCB	Share with imputed security flows
Stocks	0.400	0.433	34.54	0.094
Bonds	0.429	0.886	7.73	0.027
Funds	0.400	0.350	0.08	0.474

Notes: Stocks and funds cannot be separated in the KURU data. Therefore the shares found in column 1 are the same. Data spans 2000-2007.

very large changes in wealth the following rule for outliers is applied: 1. Outliers are defined as individuals with changes in stock or bond wealth surpassing three standard deviations; i.e, extreme outliers. However, some years there are only a few very large position changes taking place which results in a low standard deviation of the change in stocks and bonds. Therefore an absolute rule is also used: the standard deviation must be larger than 500,000 kr for the observation to be considered as an outlier. The number is chosen as this is what the standard error is for those years which have a larger number of large trades that seem to affect the average residual consumption. 2. To make sure that there is no offsetting change in another asset class lagged consumption is used to see if a change in, for instance, stock wealth leads to a similar change in consumption. For very large amounts (above 1.5 million kr) it is likely that the individual has invested in some asset which is not picked up in the data. Therefore an outlier must also have a change in consumption which is at least 75% of the change in either stocks or bonds. When computing consumption, the outliers will have a transaction amount of zero instead. This means that we retain the observation but leave out the large non-offset change in asset holdings. For the whole period 2000-2007 the share of outliers is on average 0.03% or about 2,200 individuals annually. The years 2006 stands out as mentioned. This year 0.14% or 10,734 individuals made large transactions which has an impact on overall consumption. When the asset changes of all individuals in 2006 are considered the average change in bonds is -10,441 kr. When outliers are left out, the change is instead 2,345 kr. All else being equal, excluding these outliers removes more than 12,500 kr from average consumption, or a decrease by about 6.5%. Removing outliers leads to an improved correspondence between national accounts and the residual consumption measure.

Housing Values   Housing in Sweden can be divided into two broad categories: owned housing and coop apartments. Below we describe how these two categories of housing are valuated in the wealth registry. We also describe in detail how housing wealth flows are calculated for owned housing and coop apartments.

Owned housing consists of several types of property, mainly primary residences (stand-alone houses, terrace buildings, bungalows etc.), vacation homes, apartment buildings, commercial property and farms. Their value is assessed for taxation purpose every sixth year. When tax values are assessed the tax administration collects characteristics of the property such as living space, size of the lot, construction year, when larger refurbishments were made etc. The taxation value is

determined by hedonic regressions run on actual sales prices using the collected characteristics as explanatory variables. Tax values also consider the market value of properties in the neighborhood. For primary housing the tax administration divide Sweden into about 9000 different "value areas". A simplified tax value assessment is also made three years after the assessment described above. This is also made every sixth year which means that the tax vale of a house is assessed every third year. The simplified tax value assessment only consider how prices in the value area has evolved the last three years, it does not collect any property specific characteristics.

The tax value is supposed to reflect 75% of the market value two years back. To determine market values statistics Sweden calculates municipal-level sales price ratios by dividing sales prices by taxation values for sold houses. For other categories of owned housing such as apartment buildings and commercial property this is done at the county level as there are often not a sufficient amount of transactions taking place in a single municipality during a year.

The other broad category of housing is coop apartments. Coop apartments are bought and sold just like owned housing but the buyer does not buy the apartment per se, he/she only buys the right to occupy the apartment. The apartment, as the whole apartment building (a coop can consist of several apartment buildings as well), is instead owned by a coop housing association. The coop housing association is run by a board, mostly consisting of individuals living in the building. Data on coop wealth is available 2004-2007.

Coop apartment wealth is determined on the apartment building level in the same way as the market value of primary housing determined. Statistics Sweden then divides the worth of the coop apartment building and divides it by the number of households residing in the apartment building. This overstates coop wealth for households living in small apartments and understates coop wealth for households living in large apartments (relative to the average apartment size of the building). In addition, individuals who rent apartments in coops wrongly get coop wealth assigned to them. These different measurement errors are caused by the lack of an apartment registry which assigns each individual to a specific apartment. Such a register is available in Sweden from 2013.

Consumption flows from housing   Whenever possible we use data on actual housing transactions to determine the flow from housing wealth to consumption. For owned housing we use data on actual transactions from the housing price registry (Fastighetsprisregistret) to calculate these flows. From 2004 we observe all parties involved in a transaction and can calculate ownership shares depending on the number of buying or selling parties involved in each transaction to assign the appropriate transaction value to each individual. The ownership shares assume that each involved party owns or buys an equal share. According to the wealth register, which records ownership shares for each type of real estate (primary housing, second homes, farming properties and apartment buildings), 96.1% of all jointly owned primary housing is owned at 50%. This suggests that most ownership is split equally among the owners.

Before 2004, only one buyer and seller per transaction is recorded. Who is recorded as buyer and seller is random. For individuals of the same households this is not an issue. But for transactions

involving two or more households the censoring of buying and selling parties results in housing transactions on the individual level being unrecorded. For instance, consider two siblings selling a jointly owned house. Only sibling A is recorded as a seller while sibling B is not, which means that sibling B will have no recorded revenue from selling in the house price registry. Sibling A's share of the revenue is calculated as the total revenue times the ownership share, from the wealth registry, he/she has for the particular type of real estate. According to 2004's statistics, 23.5% of all sales involved two selling households or more, and the average number of households involved in the sales which involves more than one selling household are 2.31. This means that $23.5 \times (1-1/2.31) = 13.3\%$ of all transactions on average go undetected. 128,480 individuals or 1.7% of all individuals sold housing in 2004. Given that the 2004 figures are similar to those in previous years we can expect about $.0.017 \times 13.3 = 0.2\%$ of our sample to have missing income from selling housing during 2000 to 2003.

For coops there is no data such as the housing price registry. This means that the amount of coop housing bought and sold cannot be detected in the same way as with owned housing. For individuals who sell their property, the purchase transaction is fully observed (date, price, etc.). But for individuals who do not, we need to impute these transactions. Note, however, that because we observe the stock of coop wealth in each period, we can restrict the sample of individuals for which these transactions need to be imputed. The purchase transactions in time t are only needed to be imputed for individuals who are observed with some positive stock of co-op wealth in both time t and $t-1$. Individuals whose stock of coop wealth is always equal to zero cannot have been buying at any time in our period of observation, so for them there will be no transaction flows. For individuals whose stock of coop wealth goes from zero to a positive amount at time t, we use the figure reported in the wealth registry. Finally, for individuals whose coop wealth goes from positive to zero at time t, the change in coop wealth is just equal to the amount of the sale which we observe.

To impute purchase transactions for those where the transaction is not observed, we use a statistical model, estimated on the sample for which we observe a purchase. Concretely, we create a sample based on all the purchases observed (i.e. for individuals who did resell). Then we create a balanced panel for years 2005-2007 of all the individuals for which we have a purchase record and who had a positive stock of coop wealth in both time t and $t-1$ (note that individuals will be included in this sample even if their purchase record falls outside the 2005-2007 window). For this sample we estimate a statistical model on the probability of buying a coop:

$$Pr[\text{BUY}_{it}] = X'_{it}\beta + \varepsilon_{it}. \tag{4}$$

The variables used are age-bins, income, household size and change in household size, county or municipality, indicator for moving address in year t (and/or t+1), change in coop wealth (in levels), indicator for selling a coop, indicator for selling an owned house, interaction between selling a coop and change in coop wealth and change in debt and year fixed effects.

Then we estimate a model of the transaction amount for the observed coop purchases (Y_{it}):

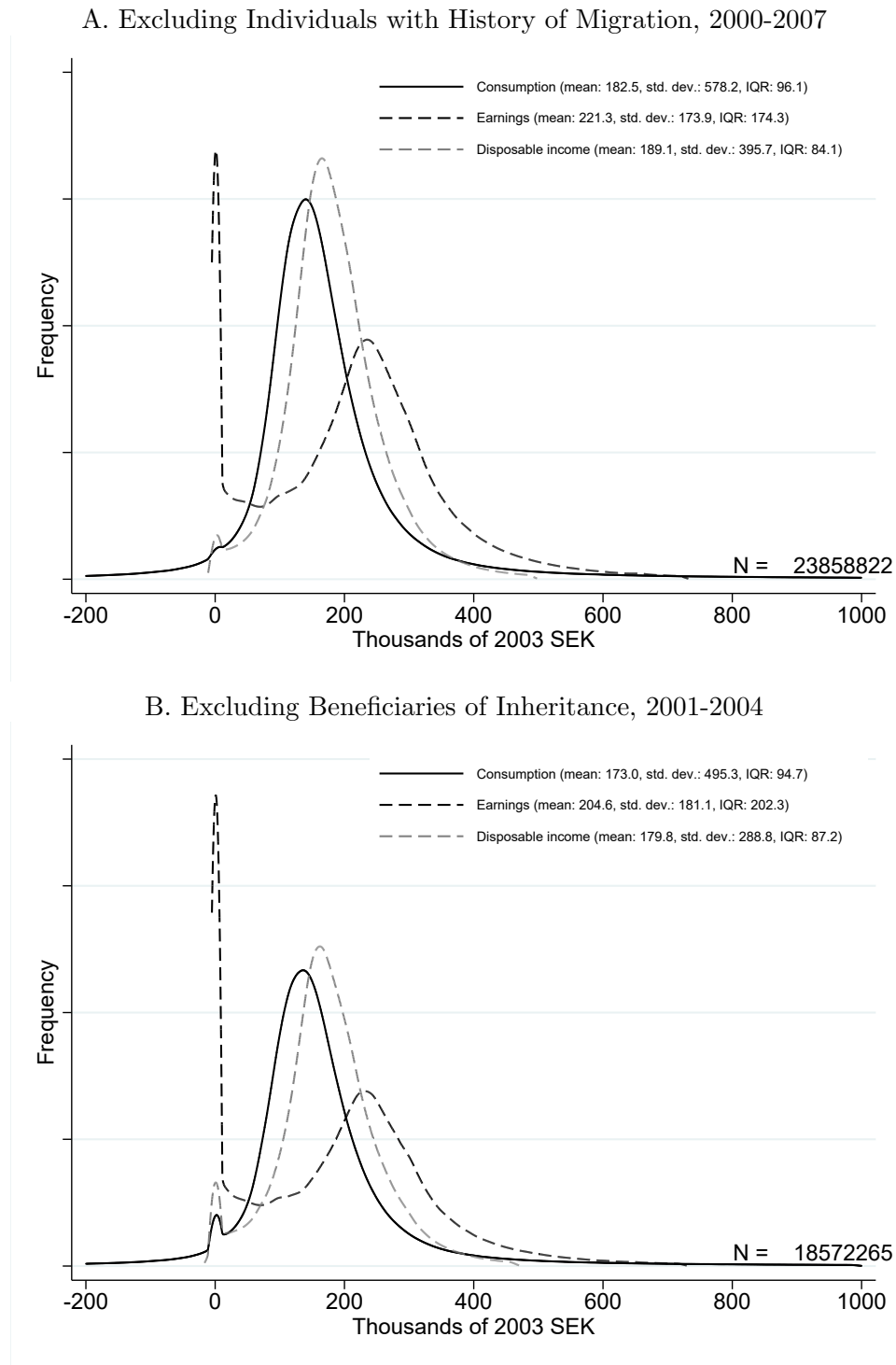$$Y_{it} = X'_{it}\beta + \varepsilon_{it}. \tag{5}$$

The same right-hand side variables as above are used.

Based on these two models, we predict the probability of buying a coop in the sample of individuals observed with positive stock of coop wealth in both time t and $t-1$ in the period 2005-2007. When this probability is large enough we impute the amount of the purchase based on model 5.

Migration Individuals who migrate in or out of Sweden in a given year will also have mismeasured consumption flows in the registry data. To handle this issue, we have added to our data the exhaustive migration register kept by the Swedish authorities. This register records all migration flows in and out of Sweden, at the individual level, with the exact date of migration. All individuals need to register to their local council to validate their personal identification number, which is necessary for almost all aspects of daily life (signing a labor contract, renting or buying an apartment, renting or buying a car, paying taxes, receiving benefits, etc.) There are 798,166 migration events in our period of interest from 1999 to 2007. In Figure C.1 Panel B, we display the distribution of earnings and consumption in 2003 when removing all individuals who migrate in or out of the country. Compared to Figure 1, we see that the inclusion or exclusion of migrants has relatively little effect on the overall distribution of consumption. But it is worth noting that removing individuals who migrate in any given year decreases significantly the fraction of individuals with negative or zero consumption, as can be clearly seen from the graphs. Overall, the average consumption therefore increases from 176k to 182k with this correction.

Inheritance A further challenge to the measurement of consumption flows using registry data comes from inheritances. Individuals receiving an inheritance in t register an one-off increase in assets. This results in a downward bias of the consumption flow measure. To alleviate this concern, we use the inheritance register for the years 2001-2004. It contains all beneficiaries of inheritances and inter vivos gifts. The inter vivos gifts are only reported at the time of death and aggregated together with the inheritance over the 10 prior years, as this formed the base for the inheritance tax. On average .8% of individuals receive a positive inheritance (including inter vivos gifts over the 10 years before death) per year. 6.2% of heirs have a strictly negative level of consumption compared to 3.8% of individuals that do not receive an inheritance. In Figure C.1 Panel B below, we display the distribution of earnings and consumption in 2001-2004 when removing all individuals who received an inheritance in a given year. There is a minor change in the share of individuals with strictly negative consumption from 3.89% to 3.87%.

Figure C.1: DISTRIBUTION OF EARNINGS & OF REGISTRY-BASED CONSUMPTION EXPENDITURES: SAMPLE RESTRICTIONS

A. Excluding Individuals with History of Migration, 2000-2007

B. Excluding Beneficiaries of Inheritance, 2001-2004

Notes: The Figure shows the distribution of household earnings and the distribution of our registry-based measure of household consumption with the following exclusions: Panel A excludes individuals who migrate in or out of the country in a given year; Panel B excludes individuals that did not receive an inheritance in a given year. All variables are expressed in thousands of 2003 SEK. Disposable Income is earnings plus taxes and transfers plus all realized capital income. The IQR is the interquartile range.

Appendix C Comparing Consumption in National Accounts, Survey, and Registry Data

This Appendix reviews how consumption is computed in the Swedish national accounts for the years 2000-2007. It also points out significant differences between what is viewed as consumption in national accounts, in household budget surveys and in a residual consumption measure derived from individual data on income and wealth.

C.1 Computing consumption in national accounts

Consumption in national accounts comprise all consumption by the residents of a country. This means that domestic consumption is only the consumption made by the residents of the country while consumption made by foreigners is left out. Foreigners include both tourist but also asylum seekers who have not yet been granted residence permit. Consumption made by these groups are booked as consumption by foreigners and is a part of the country's exports. Consequently consumption made by the country's residents abroad (as tourists or diplomatic personnel, for instance) is therefore also included in the national account's consumption measure. Consumption in national accounts are meant to include both white market and black market consumption. Black market consumption includes consumption that is not taxed for, when it should, but also activities such as consumption of illegal drugs and money spent on prostitution services (which are illegal to buy in Sweden).

Statistics Sweden use various sources to compute consumption for the national accounts (NA).[41] The household budget survey HUT is one of the main inputs, and the HUT of 1995/1996 was used as a benchmark for NA consumption up until 2004 when the HUT wave of 2003/2004 was meant to replace it. However, HUT, like many other household budget surveys, has experienced higher and higher attrition bias over time and it has been found by Statistics Sweden to severely underestimate various consumption categories. Consequently Statistics Sweden has begun to use more information from business sector to adjust for the decreasing quality of the HUT survey. The perhaps most important data is sales taxes (VAT) and turnover statistics from the retail sector. From 2004 onward, HUT is, together with surveys to the retail sector, mostly used to spread total consumption expenditure between the different COICOP (classification of individual consumption by purpose) categories. Statistics Sweden also use the vehicle registry to compute consumption of vehicles and special sales statistics from the three major operators on the Swedish grocery market to compute food consumption.

---

[41]The details in this review is based on two documents provided by Statistics Sweden: "Sweden GNI Inventory", (2016) and "Hushållssektorn i nationalräkenskaperna", (2014). `http://www.scb.se/contentassets/c89bb85e14184e92a4d5e4eec5ce4b98/sweden-gni-inventory-2016_public_rev_oct2016.pdf` and `http://www.scb.se/Statistik/_Publikationer/NR0103_2014A01_BR_X100BR1401.pdf`

C.2 Differences between measures of consumption

Consumption is not always the same thing in NA, household budget surveys and in residual consumption expenditures derived from registry data on income and wealth. Certain types of consumption that are treated differently in the three measures. The first is imputed rents, the shadow price of occupying one's own home, which is a part of consumption in NA, but is not a part of consumption in a household budget survey. This can however, be made into a consumption object in a residual consumption measure applying the same techniques used when computing the variable in NA.

The second special category is "larger" renovations, where larger is referred to such refurbishing not normally undertaken by tenants. In NA this is viewed as investment and household budget surveys do not normally ask for expenses on larger renovations. In a residual consumption expenditure measure it is not possible to separate expenses on refurbishing from other expenses. Therefore renovation expenses are treated as consumption which, all things being equal, makes registry data consumption larger compared to consumption in NA and surveyed consumption.

The third category is replacement of larger home appliances, such as stoves, dishwashers etc., in individuals' dwellings. This is treated as renovations, i.e. investments, in NA. Consumption of such goods are only incorporated into NA's consumption statistics if a household buys an additional stove or dishwasher.

The fourth category is insurance expenses and expenses on lotteries and gambling. In NA these are measured as the net of what is paid for insurance, lotteries and gambling and what is received by the households in terms of insurance payments or wins. In a residual consumption expenditures measure insurance premiums and lottery tickets are accounted for as they are not considered as saving but insurance payments and wins are in most cases exempt from tax which means that they are not recorded as income. In a sense, insurance payments and wins can in registry data be viewed as unaccounted gifts. Such insurance payments may reduce debts and increase savings but in such as case residual consumption expenditure decreases.

The fifth category is interest payments on liabilities. In NA these are not measured as consumption, rather households consume a service which is the banks' provision of liquidity. The price of this provision is the difference between the borrowing and lending interest rate. Therefore this consumption item is larger in the residual consumption expenditure measure compared to NA. In household budget surveys expenses on interest is a separate category from expenses on housing.

Finally, black market and illegal consumption is estimated by the NA division at Statistics Sweden. This type of consumption is recored in the residual measure as long as the income that is used to pay for it is accounted for. Black market expenses on legal goods is accounted for in household budget surveys but consumption of goods such as illegal drugs and prostitution services are not found in household budget surveys.