RUHR
ECONOMIC PAPERS

Johannes Ludwig

Income Shocks or Insurance – What Determines Consumption Inequality?

Ruhr Graduate School in Economics
RUHR UNIVERSITÄT BOCHUM RUB

#550

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #550

Responsible Editor: Thomas K. Bauer

ISSN 1864-4872 (online) – ISBN 978-3-86788-628-4

Ruhr Economic Papers #550

Johannes Ludwig

Income Shocks or Insurance – What Determines Consumption Inequality?

Ruhr Graduate School in Economics

Ruhr Universität Bochum RUB

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788628
ISSN 1864-4872 (online)
ISBN 978-3-86788-628-4

Johannes Ludwig[1]

Income Shocks or Insurance – What Determines Consumption Inequality?

Abstract

Contrary to the implications of economic theory, consumption inequality in the US did not react to the increases in income inequality during the last three decades. This paper investigates if a change in the type of income inequality - from permanent to transitory - or a change in the ability to insure income shocks is responsible for this. A measure of household consumption is imputed into the Panel Study of Income Dynamics to create panel data on income and consumption for the period 1980-2010. The minimum distance investigation of covariance relationships shows that both explanations work together: the share of transitory shocks increases over time, but the capability to insure permanent and transitory shocks to income also improves. Together, these phenomena can explain the lack of an increase in consumption inequality.

JEL Classification: D12, D31, E21

Keywords: Consumption inequality; income inequality; consumption insurance

April 2015

1 RGS Econ and RUB. – The author is grateful for helpful comments and suggestions by Mathias Klein, Michael Roos as well as by the participants of the 8th Ruhr Graduate School Doctoral Conference in Economics and the IZA Workshop on Inequality 2015. *– All correspondence to: Johannes Ludwig, Chair of Macroeconomics, Building GC 2/149, Ruhr-Universität Bochum, 44780 Bochum, Germany, e-mail: johannes.ludwig@rgs-econ.de*

1 Introduction

Much research has been devoted to the problem of rising income inequality in the US since the beginning of the 1980s. Far less authors have analyzed the development of consumption inequality over the same time frame. In case it is examined, consumption inequality has mostly been found to increase surprisingly little (see e.g. Heathcote et al. 2010, for a different view see Aguiar & Bils 2013, Attanasio et al. 2013). From a theoretical perspective this lack of a response to increasing income inequality is stunning: When the variance of income increases in a permanent income-setting, households should adjust their consumption expenditure and consumption inequality should increase proportionally.

Two explanations for this phenomenon have been suggested: Krueger & Perri (2006) assume that households have increased their ability to insure consumption expenditures against shocks to income with the help of credit. Blundell et al. (2008) argue that it is the type of the income shocks that has changed - from less insurable permanent shocks to better insurable transitory shocks. The aim of this paper is to test these two explanations against each other and investigate by how much changing insurance possibilities and changing income shocks can explain the transition of income inequality to consumption inequality during the last three decades.

The development of consumption inequality over time is of its own interest, since consumption is generally considered to be a better indicator for current household welfare compared to current income (Meyer & Sullivan 2013b). Most researchers use only income as an indicator, but to be able to determine the welfare consequences of increasing income inequality it is necessary to know how changes in income inequality translate into consumption. Moreover, estimating consumption insurance parameters is also relevant for computational macroeconomists, since Kaplan & Violante (2010) have argued that the parameters estimated by Blundell et al. should become a benchmark for macroeconomic incomplete-markets models.

Krueger & Perri (2006) were the first to highlight the different dynamics of income and consumption inequality in the US since the 1980s. To explain this phenomenon they develop a model where financial markets react endogenously to increasing income inequality by extending credit supply to households that are hit by large shocks. Thus, when income inequality increases over time, credit supply reacts and households do not need to decrease spending on consumption. In this way households are able to insure consumption against larger and larger income shocks. Krueger & Perri use repeated cross-section data on income and consumption from the Consumer Expenditure (CE) Survey to estimate inequality trends in the period 1980-2003 and then try to match the trend in consumption inequality via simulation. They simulate several models with different specifications for the financial sector and show that the development of consumption inequality is determined by the amount of insurance that the financial sector offers.

Blundell et al. (2008), on the contrary, emphasize that the type of income shock is crucial for the reaction of consumption and criticize Krueger & Perri for not differentiating between transitory and permanent shocks to income. Blundell et al. first generate a new panel dataset on income and consumption by imputing information on consumption from CE data into the Panel Study of Income Dynamics (PSID) and then estimate empirically how consumption reacts to the two kinds of income shocks. Their results show that household consumption does not react at all to transitory income shocks, but that it is only partially insured against permanent shocks. Contrary to the assumptions of Krueger & Perri they do not find the degree of insurance to change over time. Blundell et al. show that during their period under review it is a change in the composition of aggregate income inequality - from permanent to transitory - that drives the response of consumption inequality.

It has to be noted that Blundell et al. (2008) only examine the period 1980-1992. It is intuitive that changes in the degree of consumption insurance cannot be detected if the observation period is relatively short. The first and foremost contribution of this paper is to demonstrate that the results of Blundell et al. break down when the observation period is prolonged. By examining PSID data up to 2010 it becomes possible to show that both the insurance against permanent and against transitory shocks improve significantly at the aggregate level. Second, I use a new imputed measure for consumption proposed by Attanasio & Pistaferri (2014) which allows to check if the results of Blundell et al. are robust to other measures of household consumption.

In general, evidence on consumption inequality is scarce mainly due to data restrictions. Up to now, the literature heavily relied on consumption data from the CE. However, the quality of CE data has recently been under discussion. This paper circumvents CE-related problems by exploiting information on consumption from the PSID collected between 1999 and 2011. This information is used to impute a measure of consumption for all sample years of the PSID following the example of Attanasio & Pistaferri (2014). The resulting dataset is a household panel that includes information on income and consumption which is needed to study the effects of transitory and permanent income shocks on consumption.

The empirical framework of Blundell et al. (2008) is used to identify the size of transitory and permanent income shocks and the consumption insurance parameters. This framework is parsimonious and flexible - it allows that the variances of the shocks as well as the insurance parameters change over time. I use their techniques to check if the size of transitory and permanent income inequality has changed and whether the degree of insurance against permanent and transitory shocks has remained constant over the last thirty years.

The imputation procedure of Attanasio & Pistaferri (2014) yields a consumption variable that shows only a slight increase in consumption inequality compared to the drastic increase in income inequality during the same time. The minimum distance estimation of empirical covariance relationships shows that the variance of transitory shocks to household income increases significantly whereas the variance of permanent shocks has been rather flat over the last three decades. Moreover, the parameters indicating insurance against income shocks also change substantially over time. Both, the insurance against permanent and transitory shocks improve. Permanent shocks, though, are found to be less well insured throughout the sample period, but its insurance parameter improves to a larger extent. Thus, the type of shock that is better insured becomes more important over time and any shock to household income translates less and less into consumption over time. Combining these two phenomena can explain why consumption inequality does not react much to income inequality.

The paper is organized as follows: Section 2 gives an overview over the sources of data and the imputation procedure while Section 3 introduces the methodological framework. The results are presented in Section 4 whereas Section 5 discusses them in the light of the related literature. Finally, Section 6 concludes.

2 The Data

One major problem encountered when trying to study empirically the response of consumption to permanent and transitory income changes is the lack of panel data on income *and* consumption. Although containing some panel elements, the Consumer Expenditure Survey that contains information on household income and consumption is primarily a repeated cross-section data set. The Panel Study of Income Dynamics - a panel data set renowned for high quality information on household income - includes only information on food expenditure, but not on other consumption items for most of the time. The majority of studies concerned with consumption inequality therefore uses information from CE data (e.g. Fisher et al. 2013, Meyer & Sullivan 2013a). However, the quality of CE data has recently been under discussion. It has been shown that CE data and tables from the National Income and Product Accounts (NIPA) on aggregate consumption diverge over time. The CE Survey seems to suffer from an increasingly worsened coverage of nondurable consumption. Therefore, Aguiar & Bils (2013) as well as Attanasio et al. (2013) argue that the use of raw data from the CE yields biased results on consumption inequality. Meyer & Sullivan (2013b), in contrast, state that the largest expenditure categories in the CE are measured well and that most of the differences to NIPA data stem form definitional differences.

Attanasio & Pistaferri (2014) have recently shown how to circumvent CE-related problems by exploiting information from the PSID on consumption collected between 1999 and 2011. They use this information to impute consumption “backwards” for all sample years. In the remainder of this paper, I will follow their example and generate a panel dataset on income and consumption for the years 1978-2010.

2.1 The Panel Study of Income Dynamics

The Michigan Panel Study of Income Dynamics (PSID) is the world's longest running household panel dataset. Since its start in 1968 it has collected very detailed information on household income as well as on socioeconomic characteristics, household composition and health status. Interviews have been conducted yearly until 1996 and biennially since 1997. For most of the time information on consumption expenditure has been confined to few variables on food expenditure (food at home, restaurant meals and value of food stamps). Since 1999, though, the PSID has started to collect more detailed information on nondurable consumption, namely expenditure on household utilities, gasoline, car maintenance, transportation, health, education and child care. These variables sum up to around 70% of NIPA nondurable consumption. In 2005 the PSID has added further categories like expenditure on clothing, vacation and entertainment. Blundell et al. (2012) have shown that the new PSID consumption variables align well with aggregate NIPA data and do not seem to suffer from shrinking coverage.

My sample includes all years between 1978 and 2011. Households with a head who is not in working age (i.e. younger than 25 or older than 65 years) were dropped from the dataset. The central income variable used in the analysis is household net income since it is the type of income that determines household consumption and savings. To create a variable for household net income information on federal and state tax payments are necessary which are not available in the PSID in many cases. They therefore have to be simulated with the help of NBER TaxSim (Appendix A1 contains further information on sample selection, tax simulation and the construction of the central variables). All income variables are first deflated using Consumer Price Index variables from the U.S. Bureau of Labor Statistics and then equivalence-weighted according to the modified OECD scale. The final sample comprises 12,029 households and 130,692 observations.

2.2 The Imputation Procedure

Attanasio & Pistaferri (2014) have recently shown how to use the new PSID variables on nondurable consumption to create a consistent panel data set that includes income and consumption. They first pool all new consumption expenditure variables that are continuously available for the period 1999-2011 to create a measure of "net consumption" $N_{i,t}$ for household i in year t. Attanasio & Pistaferri then regress this indicator on a set of variables that are available for all years in the PSID according to:

$$\ln N_{i,t} = Z'_{i,t}\beta + p'_t\gamma + g\left(F_{i,t};\theta\right) + u_{i,t}. \tag{1}$$

$Z_{i,t}$ denotes observable socio-economic characteristics, p_t is a set of different prices and $F_{i,t}$ stands for consumption variables that are available for all years in the broad sample (i.e. food consumption). Note that all of the variables used here need to be available for all sample years in order to be able to impute consumption backwards[1]. The imputed value for total household consumption $\widehat{C_{i,t}}$ is then equal to the sum of the consumption variables that are continuously available ($F_{i,t}$) and the predictions for net consumption ($\widehat{N_{i,t}}$):

$$\widehat{C_{i,t}} = F_{i,t} + \exp\left\{Z'_{i,t}\widehat{\beta} + p'_t\widehat{\gamma} + g\left(F_{i,t};\widehat{\theta}\right)\right\}. \tag{2}$$

For further details on the regression results and the variables included see Appendix A2.

Attanasio & Pistaferri (2014) argue that this procedure has three important advantages compared to other approaches. First of all, there is no need to rely on information from the CE which is valuable in the light of the debates on data quality. Second, due to the availability of some consumption variables over the whole time span it is possible to control for preference heterogeneity to some extent. Finally, compared to other methods imputation quality can be assessed conveniently for the last decade by comparing the forecasts with actual PSID values.

3 Methodological Framework

3.1 The Income Process

The empirical framework of this paper is built around the seminal work of Blundell et al. (2008). The income process of the model is specified in a well-known permanent-transitory fashion - the log of real household net income $Y_{i,t}$ can be separated into a part that is explained by observable household characteristics $X_{i,t}$ and two unexplained components:

$$\log Y_{i,t} = X'_{i,t}\varphi_t + P_{i,t} + \nu_{i,t}. \tag{3}$$

The first of these components $P_{i,t}$ collects all shocks to income that exert a permanent influence on household income whereas the latter $\nu_{i,t}$ includes all shocks that only affect $Y_{i,t}$ temporarily. Therefore, $P_{i,t}$ is modeled as a random walk with serially uncorrelated

[1]The PSID did not collect information on food expenditure in 1988 and 1989. Thus, total consumption expenditure cannot be imputed in these years.

innovations $\zeta_{i,t}$:

$$P_{i,t} = P_{i,t-1} + \zeta_{i,t}. \tag{4}$$

This assures that permanent innovations to household income $\zeta_{i,t}$ do not vanish over time. The transitory component, on the contrary, needs to be modeled in a way that the effect of the innovations on household income disappears after some time. Hence, $\nu_{i,t}$ is represented by an MA(q)-process where the innovations $\epsilon_{i,t}$ are also serially uncorrelated:

$$\nu_{i,t} = \epsilon_{i,t} + \sum_{j=1}^{q} \theta_j \epsilon_{i,t-j}. \tag{5}$$

The order q of the MA-process is a priori unknown and will be estimated empirically.

As the two residual components are of main interest, the remainder of this article is concerned with residual income $y_{i,t}$ and the growth in residual income $\Delta y_{i,t}$:

$$y_{i,t} = \log Y_{i,t} - X'_{i,t}\varphi_t = P_{i,t} + \nu_{i,t}, \tag{6}$$

$$\Delta y_{i,t} = \zeta_{i,t} + \Delta \nu_{i,t}. \tag{7}$$

The variable $\Delta y_{i,t}$ is also needed to identify the order q of the MA-process. By assumption, the innovations $\zeta_{i,t}$ and $\epsilon_{i,t}$ are mutually uncorrelated. Then, all covariances of $\Delta y_{i,t}$ and $\Delta y_{i,t+s}$ with $s > q+1$ are equal to zero, whereas for $s \leq q+1$ all covariances need to be significantly different from zero. The order q can, thus, be identified by estimating $Cov\left(\Delta y_{i,t}, \Delta y_{i,t+s}\right)$ for $s = 1,...,T-t$ and checking at which value for s the covariances cease to be significantly different from zero. Table 1 depicts the autocovariances for $s = 0, 1, 2, 3$ for the years 1979 - 1996. After 1996 the frequency of PSID data is biennial so that first differences of income cease to be available.

The values in Table 1 show that - as expected - the variances of the first difference of residual income are positive and significantly different from zero. The first autocovariance of income change is always negative and significantly different from zero. The second and the third autocovariance are very close to zero, sometimes positive and sometimes negative. Statistically, they cannot be distinguished from being zero in nearly all cases. This pattern is in line with $q = 0$, i.e. with a transitory income component that is a simple idiosyncratic shock and does not show any moving average-behavior. Thus, PSID data in my sample can be described best by the following income process:

$$\log Y_{i,t} = X'_{i,t}\varphi_t + P_{i,t} + \epsilon_{i,t}. \tag{8}$$

This income process has often been used for PSID data (e.g. Heathcote et al. 2010). It is

Table 1: Variance and autocovariance of income growth

Year	$Var\,(\Delta y_{i,t})$	$Cov\,(\Delta y_{i,t}, \Delta y_{i,t+1})$	$Cov\,(\Delta y_{i,t}, \Delta y_{i,t+2})$	$Cov\,(\Delta y_{i,t}, \Delta y_{i,t+3})$
1978	-	-	-	-
1979	0 1120***	-0.0377***	-0.0007	-0.0032
1980	0.1079***	-0.0460***	0.0032	-0.0010
1981	0.1182***	-0.0475***	-0.0034	0.0013
1982	0.1203***	-0.0483***	-0.0002	0.0013
1983	0.1236***	-0.0457***	-0.0021	-0.0002
1984	0.1228***	-0.0512***	-0.0018	-0.0013
1985	0.1461***	-0.0488***	-0.0081**	0.0004
1986	0.1356***	-0.0507***	-0.0030	-0.0037
1987	0.1351***	-0.0517***	0.0060*	-0.0032
1988	0.1358***	-0.0567***	-0.0029	-0.0018
1989	0.1324***	-0.0450***	0.0007	0.0014
1990	0.1302***	-0.0494***	-0.0055	-0.0041
1991	0.1389***	-0.0541***	-0.0012	0.0044
1992	0.1734***	-0.0764***	-0.0029	-0.0015
1993	0.2017***	-0.0825***	0.0016	0.0002
1994	0.1930***	-0.0680***	-0.0058	-
1995	0.1828***	-0.0664***	-	-
1996	0.1880***	-	-	-

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

a special case of the more general income process of Blundell et al. (2008) since there is no difference between the transitory shock $\nu_{i,t}$ and the transitory innovation $\epsilon_{i,t}$.

3.2 The Response of Consumption to Income Shocks

The second central equation in the model of Blundell et al. (2008) is the equation that describes how transitory and permanent shocks are translated into (residual) household consumption. Blundell et al. (2008) do not want to presuppose any kind of theoretical model, but prefer to allow the amount of insurance with respect to both kinds of income shocks to be determined freely. Hence, two parameters $\phi_{i,t}$ and $\psi_{i,t}$ are introduced that govern how much of the permanent and transitory shocks is transmitted into consumption. Therefore, the change in residual consumption is determined by

$$\Delta c_{i,t} = \phi_{i,t}\zeta_{i,t} + \psi_{i,t}\epsilon_{i,t} + \xi_{i,t} \tag{9}$$

where $c_{i,t}$ is residual real household consumption, $\phi_{i,t}$ is the parameter for insurance against a permanent income shock and $\psi_{i,t}$ is the counterpart with respect to a transitory shock. Both parameters theoretically lie between zero and one: If a parameter is zero, the shock does not affect consumption at all. If a parameter is one, the shock fully translates into consumption. If the parameter lies between zero and one, Blundell et al. speak of "partial insurance". Finally, the term $\xi_{i,t}$ allows for unobserved heterogeneity and captures all innovations to residual consumption growth that are not related to income changes.

Economic theory predicts that the reaction of household consumption to income shocks is driven by the state of the financial markets of the economy. In the case of perfect insurance markets it is assumed that households can buy insurance against any kind of shock and, thus, their consumption does not react at all to transitory or permanent income shocks ($\phi_{i,t} = \psi_{i,t} = 0$). In a situation where the financial markets allow only self-insurance via a non-contingent bond, households accumulate savings that enable them to buffer temporary shocks to income, but permanent shocks will affect the level of consumption ($\psi_{i,t} = 0, 0 < \phi_{i,t} \leq 1$). Finally, when credit markets are imperfect and households face tight borrowing constraints, transitory and permanent shocks will both be fully transmitted into consumption ($\phi_{i,t} = \psi_{i,t} = 1$).

Since the consumption variable used is imputed, it is very likely that it is measured with error. Blundell et al., therefore, expand equation (9) so that it allows for measurement error in $c_{i,t}$. Thus, for estimation the following variant of equation (9) will be used:

$$\Delta c_{i,t} = \phi_{i,t}\zeta_{i,t} + \psi_{i,t}\epsilon_{i,t} + \xi_{i,t} + u^c_{i,t} - u^c_{i,t-1} \tag{10}$$

where $u^c_{i,t}$ stands for the measurement error in consumption in period t.

The advantage of this framework is its flexibility - both the variances of transitory and permanent shocks as well as the insurance parameters are allowed to change over time. Due to this fact the framework can also be used to analyze whether the prediction of Krueger & Perri (2006) on the development of consumption inequality can be confirmed by the data. Their model assumed that financial markets react endogenously to increasing income inequality by extending credit supply. Thus, they assume that over time the level of insurance to income shocks increases. As they do not differentiate between transitory and permanent shocks to income, it is not possible to tell which one of the two population mean values ϕ_t and ψ_t needs to decrease, but it is clear that if the mechanism that Krueger & Perri (2006) describe is present, a decrease in either ϕ_t or ψ_t needs to be found.

On the other hand, Blundell et al. (2008) stress the fact that it is the kind of shock that determines the evolution of consumption inequality. They explain the lack of a response of consumption inequality to the rise in income inequality by an ever larger fraction of well-insured transitory shocks. If their result for the 1980s is a general finding than we should find (i) $\phi_t > \psi_t$, i.e. transitory shocks are better insured than permanent shocks, and (ii) ϕ_t and ψ_t do not vary over time.

It has to be noted that these are clearly extreme characterizations of the transmission of income inequality to consumption inequality. Of course, both explanations can be present at the same time and work together.

3.3 Minimum Distance Estimation of the Model Parameters

The model parameters are estimated by collecting the variances, covariances and autocovariances of $\Delta y_{i,t}$ and $\Delta c_{i,t}$ for all t and then minimizing the distance between the empirical covariances and the parameters predicted by the parametric income-consumption model. Using the first differences of variables in estimation has a long tradition especially in labor economics that traces back to Abowd & Card (1989). Thus, it first needs to be analyzed what the two central equations (8) and (10) imply for the empirical variances and covariances. Specifying the income process with $q = 0$ has the consequence that only the first autocovariance of income growth is different from zero:

$$\begin{aligned}
Var\left(\Delta y_{i,t}\right) &= Var\left(\zeta_{i,t}\right) + Var\left(\epsilon_{i,t}\right) + Var\left(\epsilon_{i,t-1}\right), \\
Cov\left(\Delta y_{i,t}, \Delta y_{i,t+1}\right) &= -Var\left(\epsilon_{i,t}\right), \\
Cov\left(\Delta y_{i,t}, \Delta y_{i,t+s}\right) &= 0 \quad \forall s > 1.
\end{aligned}$$

The first autocovariance of income change identifies the variance of the transitory shocks and together with $Var\left(\Delta y_{i,t}\right)$ also the variance of permanent income shocks can be identified. Hence, variances and autocovariances of the change in residual income suffice in general to identify $Var\left(\epsilon_{i,t}\right)$ and $Var\left(\zeta_{i,t}\right)$. However, adding also the consumption moments helps to improve the estimation.

The autocovariances of the change in residual consumption have a similar structure. Due to the consumption-martingale property changes in consumption today are related to changes in the past or the future only via measurement error. The autocovariances of (10) can be summarized by:

$$\begin{aligned}
Var\left(\Delta c_{i,t}\right) &= \phi_t^2 Var\left(\zeta_{i,t}\right) + \psi_t^2 Var\left(\epsilon_{i,t}\right) + Var\left(\xi_{i,t}\right) + Var\left(u_{i,t}^c\right) + Var\left(u_{i,t-1}^c\right), \\
Cov\left(\Delta c_{i,t}, \Delta c_{i,t+1}\right) &= -Var\left(u_{i,t}^c\right), \\
Cov\left(\Delta c_{i,t}, \Delta c_{i,t+s}\right) &= 0 \quad \forall s > 1.
\end{aligned}$$

The autocovariance with respect to $t + 1$ identifies the variance of measurement error in consumption $Var\left(u^c_{i,t}\right)$, but the variance of unobserved heterogeneity in consumption growth $Var\left(\xi_{i,t}\right)$ can only be identified if the insurance parameters ϕ_t and ψ_t are known.

To identify ϕ_t and ψ_t the covariances between income and consumption changes are of crucial importance. The covariances between $\Delta c_{i,t}$ and $\Delta y_{i,s}$ are as follows:

$$
\begin{aligned}
Cov\left(\Delta c_{i,t}, \Delta y_{i,t}\right) &= \phi_t Var\left(\zeta_{i,t}\right) + \psi_t Var\left(\epsilon_{i,t}\right), \\
Cov\left(\Delta c_{i,t}, \Delta y_{i,t+1}\right) &= -\psi_t Var\left(\epsilon_{i,t}\right), \\
Cov\left(\Delta c_{i,t}, \Delta y_{i,t+s}\right) &= 0 \quad \forall s > 1 \text{ and } s < 0.
\end{aligned}
$$

It can be seen that the three sets of second moments $Cov\left(\Delta y_{i,t}, \Delta y_{i,t+s}\right), Cov\left(\Delta c_{i,t}, \Delta c_{i,t+s}\right)$ and $Cov\left(\Delta c_{i,t}, \Delta y_{i,t+s}\right)$ for all t and s suffice to identify the key model parameters ϕ_t, ψ_t, $Var\left(\zeta_{i,t}\right)$, $Var\left(\epsilon_{i,t}\right)$, $Var\left(u^c_{i,t}\right)$ and $Var\left(\xi_{i,t}\right)$. As in principle all parameters of the model are allowed to vary over time[2], this method enables us to analyze how much consumption inequality is driven by changes in the insurance parameters and how much by changes in the income shock variances.

Note that the covariance relationships are set up for annual data. As PSID data has a biennial frequency after 1997, the procedure has to be slightly adjusted. The first differences $\Delta c_{i,t}$ and $\Delta y_{i,t}$ cease to be available and it has to be relied on second seasonal differences $\Delta_2 c_{i,t} = c_{i,t} - c_{i,t-2}$ and $\Delta_2 y_{i,t} = y_{i,t} - y_{i,t-2}$. However, this change does not cause major problems to the model. Appendix A3 contains more details on the estimation process and shows how to ensure identification of the parameters with biennial data.

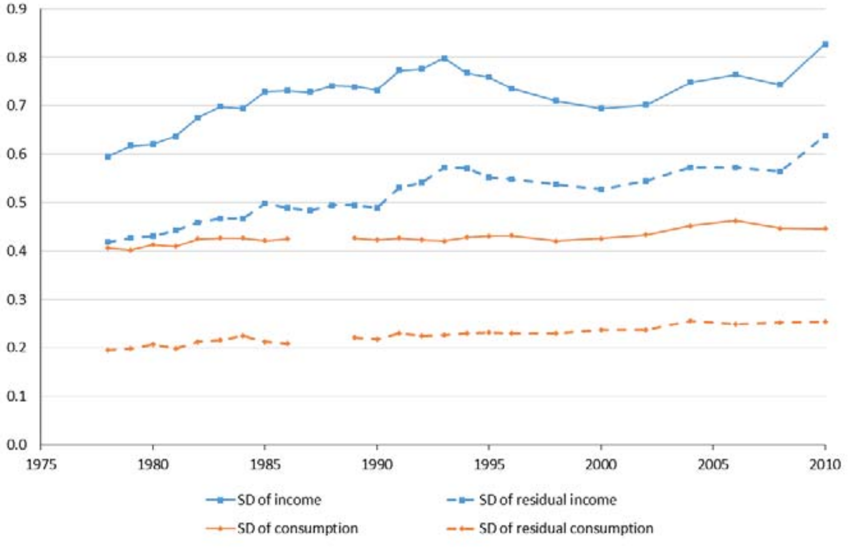
4 Results

Figure 1 shows the development of income and consumption inequality over time where inequality is measured as the standard deviation of log-variables. In line with much of the literature income and consumption inequality are both increasing over time, but inequality of imputed consumption increases far less than the inequality of household net income. This is not only true for overall income and consumption inequality represented by solid lines, but also for their residual measures, i.e. the share of inequality that cannot be explained by observable characteristics, which are depicted by the dotted lines.

To generate income and consumption residuals an OLS-regression is run that controls for a set of observable characteristics including the number of persons in the household, employment status and education of head (and spouse if present), state of residence, ethnicity,

[2] Since there is no economic intuition why $Var\left(\xi_{i,t}\right)$ should vary over time, this parameter is treated as stationary to speed up computation.

Figure 1: Inequality of household income and consumption, overall and residual



sex and age of the household head and presence of other income earners in the household. The effect of most right-hand side variables is allowed to vary with the respective year. R^2 for the income regression is 0.570, for the consumption regression it is 0.746 showing that a larger amount of household consumption can be explained by observable characteristics.

It can also be seen in Figure 1 that the increase in overall income inequality is nearly fully driven by an increase in residual income inequality. Overall income inequality increases by 0.232 between 1978 and 2010 whereas residual income inequality increases by 0.220. Residual consumption inequality even increases faster than overall inequality. Overall consumption inequality rises by 0.040 while residual inequality rises by 0.060 over the same time span. That is why the analysis in the following is concentrated on the residual measures of income and consumption that are the driving forces of the increase in overall income and consumption inequality in the last three decades.

4.1 Minimum Distance Results

Figure 2 depicts the development of $Var\left(\zeta_{i,t}\right)$ and $Var\left(\epsilon_{i,t}\right)$ at the aggregate level over time as estimated by the minimum distance procedure (additional estimation results and standard errors can also be found in Table A3 in the Appendix). It can be seen at first glance that the variance of transitory innovations which affect income only in the respective year is always larger in magnitude compared to the variance of permanent innovations whose effect on household income is lasting. Moreover, the transitory variance is increasing steadily over time whereas it is harder to detect a general trend for the variance of permanent innovations. The latter increases during the first half of the 1980s but decreases again to the initial level by the end of the 1980s. In the 1990s the permanent variance increases substantially, but then drops below the starting level in 1998 and stays there until the end of the sample period. In total, the variance of permanent innovations is more or less flat or even slightly decreasing so that on average the difference in the levels of $Var\left(\zeta_{i,t}\right)$ and $Var\left(\epsilon_{i,t}\right)$ is increasing over time. Additionally, the variance of transitory innovations ends with a jump upward in 2010. This could be a consequence of the financial crisis which substantially increased (temporary) unemployment.

Figure 2: Variance of permanent and transitory innovations over time

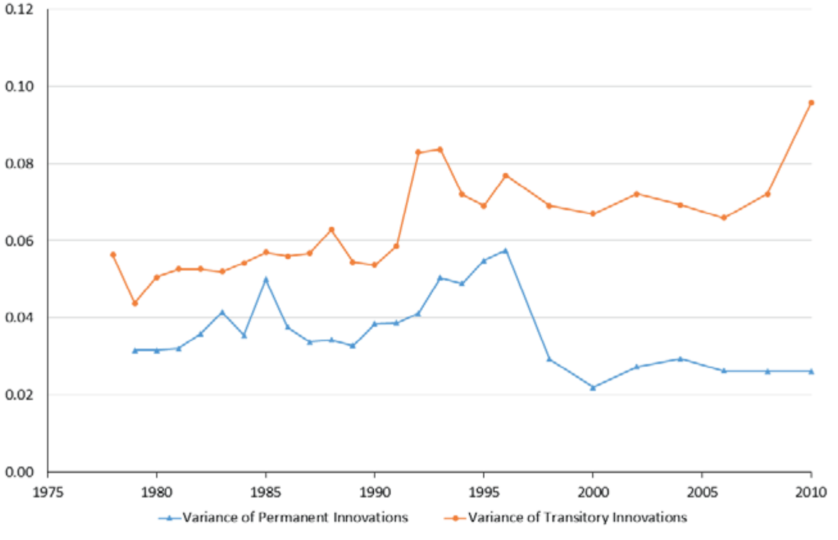


Figure 2 gives a first hint why consumption inequality does not increase as strongly as income inequality does. The composition of overall income inequality changes over time and the transitory part becomes more and more dominant. In case transitory shocks are easier to insure than permanent shocks, consumption will not be affected as much by an increase in the size of the former compared to a potential increase in permanent shocks by the same magnitude. However, this conclusion depends on the size of the insurance parameters. If these were constant over time, a steady increase in transitory inequality and a constant permanent inequality should lead to a steady increase in consumption inequality. If the insurance parameters also changed over time, the development of consumption inequality could be different.

The development of the parameters of insurance against transitory shocks (ψ_t) and permanent shocks (ϕ_t) is shown in Figure 3 that also depicts 95% confidence bands for both parameters. To ease computation both parameters have not been estimated separately for every single year, but time groups of around 4 years have been used (1979-1982, 1983-1986, 1990-1993, 1994-1998, 1999-2002, 2003-2006, 2007-2010). These groups should suffice to determine general trends over time. Estimates and standard errors for ϕ_t and ψ_t are also listed in Table 2.

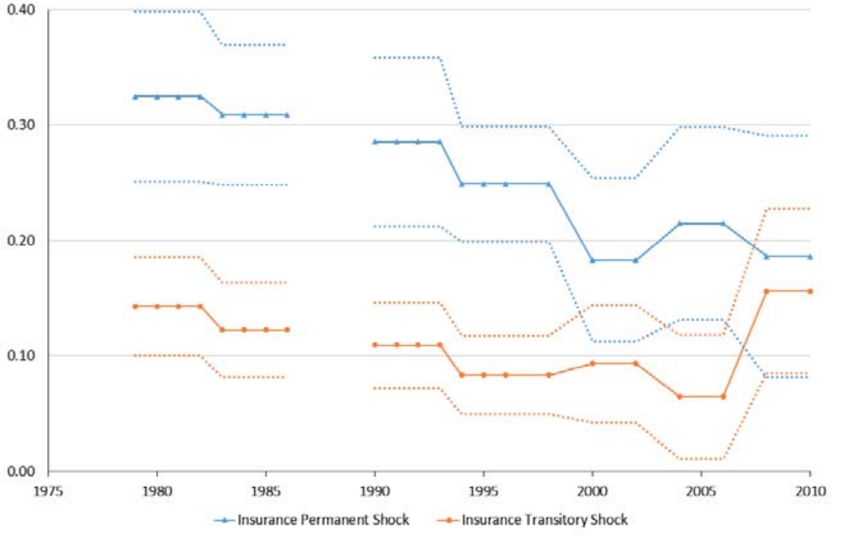
Table 2: Insurance parameters (standard errors in parantheses)

	79/82	83/86	90/93	94/98	99/02	03/06	07/10
ϕ_t	0.3247	0.3088	0.2854	0.2489	0.1829	0.2146	0.1860
	(0.0375)	(0.0311)	(0.0374)	(0.0255)	(0.0362)	(0.0425)	(0.0533)
ψ_t	0.1428	0.1223	0.1092	0.0834	0.0934	0.0644	0.1560
	(0.0217)	(0.0211)	(0.0191)	(0.0173)	(0.0260)	(0.0273)	(0.0361)

Figure 3 reveals some remarkable results: First, as expected above and also found by other studies, the insurance parameter ϕ_t always lies above ψ_t. This means that transitory shocks to income affect consumption to a lesser degree than permanent shocks to income in all years. Second, the insurance parameters are not constant over time at all. Both ϕ_t and ψ_t decrease steadily during the 1980s, the 1990s and most of the 2000s signaling an improvement in the insurance of income shocks. Estimated coefficients in the later periods lie outside the confidence bands of the initial period for both insurance parameters. Thus, the improvement in insurance is also statistically significant. The improvement is stronger in absolute terms for the insurance against permanent shocks whose parameter decreases from 0.3247 to 0.1860 between 1978 and 2010. The respective parameter for the insurance against transitory shocks declines from 0.1428 to 0.0644 between 1978 and 2006. Third, insurance against transitory income shocks jumps up drastically in the final period (2007-2010) that includes the financial crisis. This jump, though, indicates that the parameters also react to business cycle fluctuations.

Finally, although the degree of insurance against shocks is changing substantially over the course of the three decades, the level of ϕ_t and ψ_t between 1978 and 1993 is relatively stable. Both series are decreasing, but the differences in parameters are relatively small compared to the change over the whole sample period and they are not statistically significant. Hence, these results do not contradict the observations of Blundell et al. (2008) who find constant parameters in their study that covered the years between 1979 and 1992. Their result breaks

Figure 3: Development of insurance parameters over time with 95-percent confidence intervals



down, though, when observations for the years 1994-2010 are added.

4.2 Robustness

The robustness of the findings can be assessed by varying (i) the kind of variables used as well as (ii) the observations included in the sample. The first set of robustness tests tries to figure out by how much the results depend upon the imputed consumption variable. Two kinds of alternative consumption variables are logical candidates: the actual values for total household consumption of food that are available for the whole sample and the actual values for household nondurable consumption which are available from 1998 on.

Using only food consumption allows to check whether the improvement in consumption insurance can also be seen for another consumption variable that is available for the whole sample period. This route has already been followed by Hall & Mishkin (1982) in a first pioneering empirical contribution on the effect of permanent and transitory income changes on consumption. However, an important difference to nondurable consumption arises that has already been pointed out by Blundell et al. (2008): the food consumption variable only captures a part of total nondurable consumption expenditure and the fraction of food consumption to total consumption can change over time. The estimated parameters for consumption insurance will then be a product of the true parameter and the elasticity of food consumption with respect to total consumption expenditure (e.g. $\widehat{\phi_t} = \beta_t \phi_t$, where β_t is the mentioned elasticity). If household incomes and the standard of living increase on average since the 1980s, β_t will decreases over time. Then, the estimated parameters $\widehat{\phi}_t$ and $\widehat{\psi}_t$ will automatically decrease as well even if the true insurance parameters are constant. Combining a decreasing β_t and decreasing insurance parameters yields a stronger decrease then in the benchmark case. Thus, it can be expected that estimated insurance parameters will decrease more strongly over time when food consumption is used instead of the imputed variable for nondurable consumption. Moreover, since $\beta_t < 1$, the parameters found should also be smaller.

Figure 4: Robustness test - Insurance parameters when food consumption is used as variable of interest

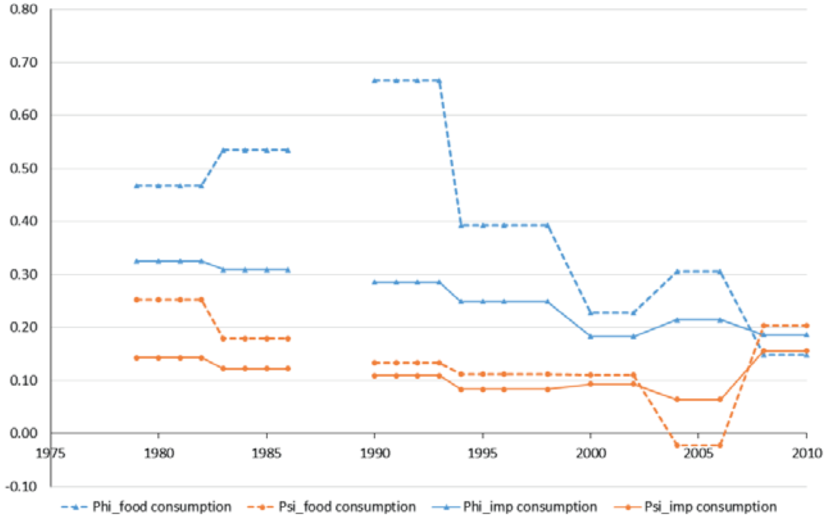
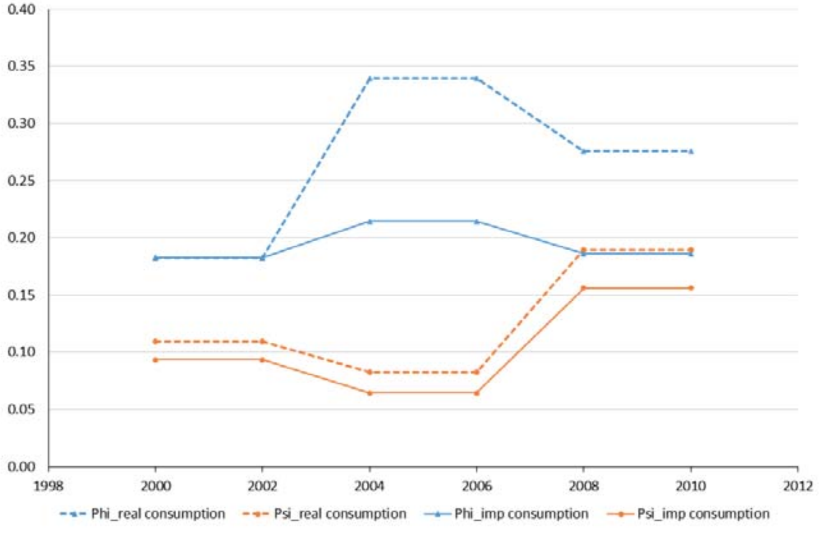


Figure 4 shows estimated parameters for the aggregate sample when only food consumption is used. The qualitative development is similar to the benchmark case and quantitatively the parameters show larger decreases as has been expected before. However, the estimated parameters are not smaller than in the benchmark case, but mostly larger. This counterintuitive finding can be explained by the nature of the imputed consumption variable: As the benchmark consumption variable has been imputed, its variance is smaller in level compared to comparative actual data. Since one of the main identifying equations for the insurance parameters is

$$Var\left(\Delta c_{i,t}\right) = \phi_t^2 Var\left(\zeta_{i,t}\right) + \psi_t^2 Var\left(\epsilon_{i,t}\right) + Var\left(\xi_{i,t}\right) + Var\left(u_{i,t}^c\right) + Var\left(u_{i,t-1}^c\right),$$

it is clear that $\widehat{\phi}_t$ and $\widehat{\psi}_t$ will increase when $Var\left(\Delta c_{i,t}\right)$ increases and the income shock variances are held constant. This effect can explain why the level of the insurance parameters is higher when actual data for food consumption is used. However, the finding that the insurance parameters decrease more strongly over time in Figure 4 is exactly in line with expectations and backs the benchmark results.

Figure 5: Robustness test - Insurance parameters when actual nondurable consumption is used for the years 1998-2010



The actual values for nondurable consumption that are available for the period 1998-2010 can be used to assess whether the estimated insurance coefficients for this period differ from the benchmark results. Again, the variances for the actual consumption variable are larger compared to imputed data. Hence, the level of the estimated coefficients should be higher compared to the benchmark case, but the change over time should be comparable. Figure 5 depicts the development of ϕ_t and ψ_t for the periods 2000/02, 2004/06, 2008/10. For ψ_t the results are in line with expectations: the level of insurance against transitory shocks is a bit lower (i.e. the estimated parameters are higher), but the change over time is parallel to the benchmark results. For ϕ_t the result is at least qualitatively similar to the benchmark case: the insurance parameter first increases in 2004 and then decreases back in 2008. Quantitatively, though, the level of ϕ_t in the initial period (2000/02) seems to be a bit too low to be in line with the benchmark case. This deviation could be due to problems related to the initial period and the fact that only a short spell of observations is available for the actual consumption variable compared to the imputed data. In general, this robustness test also shows that the level of consumption insurance varies over time and thereby supports the previous findings.

I have also experimented with generating a new imputed consumption variable by using the additional consumption expenditure variables that the PSID added in 2005 (clothing, vacation and entertainment). This broadens the "net consumption" variable $N_{i,t}$, but requires to drop all observations from the years 1999-2003 to rerun the imputation procedure. The results are not shown here since the outcome of the imputation regression and the results for ϕ_t and ψ_t do not differ much from the benchmark case.

The second set of robustness tests varies the amount of observations included in the dataset to check whether the results depend on sample selection. Many comparable studies exclude the SEO-subsample of the PSID that was introduced to collect information on poorer households in the US. Figure A1 in the Appendix compares estimated insurance parameters for the benchmark and for a sample that excludes the SEO-subsample. Moreover, the benchmark sample requires 4 observations on income for a household to be included in the dataset. This leads to a dataset consisting of a relatively high number of households with only few observations. To control that the results are not driven by this arbitrary choice, Figure A2 shows the benchmark case and a sample that excludes households with less than 30 observations on income and consumption. I have also estimated similar figures for a threshold of 35 and 40 observations, but the results were very similar. Figures A1 and A2 show that variations in sample selection do not lead to substantially different insurance parameters. In general, the key findings of the paper seem to be very robust with respect to changes in the consumption variable and the sample selection.

5 Discussion

Figure 2 has shown that transitory shocks to household net income increase in size over time while permanent shocks are roughly constant. Figure 3, though, made clear that a changing composition of income shocks - from less insurable permanent to transitory shocks - is not enough to explain the development of consumption inequality in the last three decades. The ability of households to insure their consumption expenditures against shocks to income changes as well. Both, insurance against transitory and against permanent shocks improve significantly over time (at least until 2006). By combining the evidence from the two figures it becomes obvious why the dramatic increase in income inequality is not mirrored in consumption inequality. The two explanations offered in the literature are present at the same time: Transitory shocks become more important as suggested by Blundell et al. (2008), but the insurability of all income shocks also improves over time as suggested by Krueger & Perri (2006).

The robustness tests have revealed that the exact level of consumption insurance is very sensitive to the kind of consumption variable used. Imputed variables have a lower variance compared to actual data and therefore yield smaller insurance coefficients. Hence, it is surprising that the quantities estimated by Blundell et al. (2008) who also use an imputed consumption variable, have received so much attention (e.g. Kaplan & Violante 2010). This paper has focused rather on the trends of consumption insurance parameters over time which seem to be relatively robust to changes in the consumption variable. Contrary to the findings of Blundell et al., both insurance coefficients fluctuate substantially over the course of the thirty years studied.

An important discussion in the literature on income and consumption inequality is how to distinguish consumption insurance from superior information. If an individual knows that her income will decrease at some point in the future, she is able to accumulate savings before so that at the point in time the income shock hits, it has no effect on the consumption level. Usually, the econometrician will not have information on foreseen shocks and will classify the lack of a response to the shock as a sign for well-functioning consumption insurance mechanisms. This point has been made very clear by Primiceri & van Rens (2009) who argue that most of the permanent income shocks are foreseen since consumption inequality does not react appropriately to increasing permanent income shocks. However, they also cannot identify information and insurance separately as data on income expectations are needed for that. Kaufmann & Pistaferri (2009) use an Italian dataset that contains subjective income expectations to distinguish foreseen and unforeseen shocks from each other. They can show that a huge part of transitory and about a third of permanent income shocks are indeed foreseen. For a detailed overview of the insurance vs. information-debate see Meghir & Pistaferri (2011).

Can the results of this paper potentially be explained by superior information? Due to the limitations of the data used, I have not been able to differentiate between insurance and information. Thus, it is important to keep in mind that everything that is identified as consumption insurance could also be a sign of superior information. The level of insurance found for both kind of shocks is relatively high and a part of what is identified as "insurance" is probably information. However, concerning the development over time it is hard to argue that there is a trend in superior information in the last thirty years. It seems implausible that households have significantly more information on their stochastic income processes today compared to the 1980s or that the income shocks itself have become more predictable. Thus, the main finding of the paper that consumption insurance improves over time, cannot be explained by superior information.

Krueger & Perri (2006) hypothesized that an expansion of credit supply is the reason for the trend in consumption inequality. The empirical framework of this paper, though, did not allow to identify the exact channel through which consumption insurance has improved.

Hence, we are not able to verify the credit supply-hypothesis within our analysis, but there is one finding which is very well in line with it: the insurance against transitory shocks worsens dramatically in the period 2007-2010. This result can be explained by a drying up of credit induced by the financial crisis that makes it harder to buffer temporary income losses. Thus, this piece of empirical evidence supports the credit supply-hypothesis which offers the best explanation for the increase in consumption insurance so far. Future research, however, should be concerned with ways to identify the precise channels that affect consumption insurance parameters.

6 Conclusion

This paper has analyzed the joint development of income and consumption inequality in the US over the last 30 years. By using a new method for imputing consumption into PSID data developed by Attanasio & Pistaferri (2014) a panel dataset that contains income and consumption was created. This data has been used to analyze by how much consumption reacts to transitory and permanent shocks to income and whether this reaction has changed over the course of the past three decades. Many other papers before have found that consumption inequality does not seem to react very much to increases in income inequality and the reasons for this phenomenon are still debated. While Krueger & Perri (2006) believe that the ability of households to insure consumption against income shocks has improved due to a better credit supply, Blundell et al. (2008) argue that it is the changing composition of income shocks that determines consumption inequality.

The results of this paper show that inequality of imputed consumption is also characterized by a relatively small increase since the 1980s whereas income inequality is steadily rising over the same time span. The minimum distance estimation of the model parameters shows that both explanations contribute to the development of consumption inequality. On the one hand the composition of shocks to income changes over time. The variance of transitory shocks increases constantly while the variance of permanent shocks is rather flat. On the other hand the capability of households to insure their consumption against income shocks improves over time. Both the ability to buffer permanent and transitory shocks increases so that shocks to household income are to a lesser extent transmitted into consumption. Taken together, these two developments explain why consumption inequality is relatively flat even in face of increasing income inequality.

Appendix A1: Sample Selection and Variable Construction

After merging all observations of the PSID waves between 1978 and 2011, the Latino and the Immigrant subsample is dropped. Since they are both only available for some years of the sample, its inclusion would change the sample composition for these years substantially and make comparisons to other years more difficult. Especially, the imputation procedure could yield erroneous results if the Immigrant sample was not left out in the regression and the results were projected onto years before 1996.

As the main focus of the work is to identify the households response to income shocks stemming from labor market risk, households with a head who is not in working age (i.e. where the head is younger than 25 or older than 65) are excluded from the sample. Moreover, observations are dropped when income or consumption values are drastic outliers. Income outliers are those values where household net income increases by more than 500% or by less than -80% compared to the period before. Observations where household net income, household food expenditure or the sum of all expenditure variables that were added in 1999 (household utilities, gasoline, car maintenance, transportation, health, education and child care) is less than 100$ a year are also identified as outliers and excluded. Households with a missing value for the educational attainment of the household head are dropped as well.

Finally, households that have less than four observations on household net income are excluded. This is done to prevent increases in income variation stemming from households that are only part of the dataset for one or two periods and for which there are no observations on continuous first differences. The decision to exclude those with exactly less than four observations is, of course, somewhat arbitrary. The reason is that four periods (i.e. three first differences) suffice to identify the key parameters of the study. However, robustness tests are conducted to see by how much the results depend on this decision. Table A1 summarizes how many observations are dropped in the process of sample selection.

Table A1: Number of total observations in the dataset by step of sample selection

Reason for exclusion	# dropped	# remain
Initial Sample		205,860
Latino & Immigrant Subsample	16,272	189,588
Head not in Working Age	40,652	148,936
Income & Consumption Outliers	4,852	144,084
No information on head's education	3,497	140,587
Too few income observations	9,895	130,692

One of the two central variables of the paper is household net income. To construct an accurate measure of annual household net income all top-coded income values in the PSID are set to missing. All income values that report weekly or monthly earnings are adjusted to annual values. Some of these adjusted values are unrealistically high so that they have to be dropped from the sample. Household net income is then defined as total family money minus the federal and state tax payments of head and wife and the federal tax payments of other family unit members.

Since federal tax payments are only available in the PSID up to 1991 and state taxes are not available at all, the variables have to be simulated. The simulation is realized with the help of NBER TaxSim (Feenberg & Coutts 1993). Unfortunately, it is not possible to simulate tax payments for other family unit members since there is to little information available in the PSID. Thus, federal tax payments of other family unit members can only be deducted up to 1991. However, as this variable is usually very small or zero, household net income should not be affected too much by this change in definition.

Household net income and consumption expenditure are then deflated using the corresponding Consumer Price Index (CPI) variables taken from the U.S. Bureau of Labor Statistics and, thus, are referring to prices of the year 1979. All variables are also equivalence weighted by the modified OECD scale to control for differences in family composition. All income variables of the PSID are referring to the previous year. I follow most of the literature by also treating consumption expenditure variables as referring to the previous year. Hence, yearly income data is available for the years 1978-1996 and biennial data for the years 1998, 2000, 2002, 2004, 2006, 2008 and 2010. (Imputed) Consumption data is available for the same years except for 1987 and 1988 where food consumption variables have not been collected.

Appendix A2: Results of the Imputation Regression

Imputation quality of the procedure by Attanasio & Pistaferri (2014) depends on the validity of the regression equation (1). I have tried to mirror the imputation procedure of Attanasio & Pistaferri as close as possible. However, due to slight differences in sample selection and the treatments of outliers and top-coded income values minor differences in the regression results can be expected. Table A2 summarizes and compares the regression results of Attanasio & Pistaferri and of this paper.

From a qualitative perspective nearly all point estimates of the two regressions share the same algebraic sign. The only important difference is that the ethnicity dummies (*White*, *Black*) are not statistically significant in the A&P-sample, but in my sample. The dummy *Black* has a negative impact on net consumption in this paper while it is very small, but

Table A2: Comparison of regression results

	A&P (2014)		This paper	
Hours worked/100	0.0057	***	0.0073	***
Family size	0.1250	***	0.0878	***
Education dummy 1	-0.4343	***	-0.4657	***
Education dummy 2	-0.3252	***	-0.3369	***
Education dummy 3	-0.1815	***	-0.1676	***
White	0.0373		0.0559	**
Black	0.0059		-0.0553	**
CPI	0.0088	***	0.0060	***
CPI_Food at home	-0.0092	***	-0.0069	***
CPI_Food away	-0.0057	***	-0.0052	***
CPI_Rent	0.0086	***	0.0077	***
Food consumption	0.0859	***	0.0629	***
Food consumption^2	-0.0026	***	-0.0015	***
Food consumption^3	0.0254	***	0.0142	***
Home owner dummy	0.3271	***	0.2047	***
N	26,815		28,846	
R^2	0.5058		0.5392	

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

positive in the sample of Attanasio & Pistaferri. Quantitatively, most of the values are rather close to the results of Attanasio & Pistaferri. Besides the ethnicity dummies only the variable for the family size and the dummy for home ownership show larger discrepancies. However, the comparison shows that there are no systematic differences between the regression results of Attanasio & Pistaferri (2014) and the results of this paper. The robustness test using the actual consumption values for the years 1998-2010 also shows that the imputed values are reasonably close to the true values (see Figure 5).

Appendix A3: Minimum Distance Estimation

Treatment of biennial data

For the years after 1996 yearly observations cease to be available and the dataset changes to a biennial frequency. Thus, first differences cannot be constructed and I have to rely on second (seasonal) differences. The second difference of residual income is defined as follows:

$$\begin{aligned} \Delta_2 y_{i,t} &= y_{i,t} - y_{i,t-2} = P_{i,t} + \epsilon_{i,t} - (P_{i,t-2} + \epsilon_{i,t-2}) \\ &= \zeta_{i,t} + \zeta_{i,t-1} + \epsilon_{i,t} - \epsilon_{i,t-2}. \end{aligned}$$

Using this definition it is again possible to set up covariance relationships. Note that due to the biennial structure $Cov\left(\Delta_2 y_{i,t}, \Delta_2 y_{i,t+s}\right)$ is only available for $s = 0, 2, 4, ...$ The structure of the resulting covariances is very similar to the case of yearly data:

$$\begin{aligned} Var\left(\Delta_2 y_{i,t}\right) &= Var\left(\zeta_{i,t}\right) + Var\left(\zeta_{i,t-1}\right) + Var\left(\epsilon_{i,t}\right) + Var\left(\epsilon_{i,t-2}\right), \\ Cov\left(\Delta_2 y_{i,t}, \Delta_2 y_{i,t+2}\right) &= -Var\left(\epsilon_{i,t}\right), \\ Cov\left(\Delta_2 y_{i,t}, \Delta_2 y_{i,t+s}\right) &= 0 \quad \forall s > 2. \end{aligned}$$

Unfortunately, the single yearly variance of permanent innovations $Var\left(\zeta_{i,t}\right)$ cannot be identified, only the sum of the variance in year t and the previous year $(Var\left(\zeta_{i,t}\right) + Var\left(\zeta_{i,t-1}\right))$ can be determined. Thus, I have to assume that the variances are equal for the respective years. The variance of growth in residual income therefore becomes:

$$Var\left(\Delta_2 y_{i,t}\right) = 2Var\left(\zeta_{i,t}\right) + Var\left(\epsilon_{i,t}\right) + Var\left(\epsilon_{i,t-2}\right)$$

Hence, the depicted variances in the years after 1996 always represent arithmetic means of two years. However, this should suffice to identify a broad trend in the permanent component.

The second difference of residual consumption is defined as

$$\begin{aligned} \Delta_2 c_{i,t} &= c_{i,t} - c_{i,t-2} = (c_{i,t} - c_{i,t-1}) + (c_{i,t-1} - c_{i,t-2}) = \Delta c_{i,t} + \Delta c_{i,t-1} \\ &= \phi_{i,t}\zeta_{i,t} + \phi_{i,t-1}\zeta_{i,t-1} + \psi_{i,t}\epsilon_{i,t} + \psi_{i,t-1}\epsilon_{i,t-1} + \xi_{i,t} + \xi_{i,t-1} + u_t^c - u_{t-2}^c. \end{aligned}$$

Assuming that the insurance coefficients do not differ for t and $t-1$, the following expression results:

$$\Delta_2 c_{i,t} = \phi_{i,t}\left(\zeta_{i,t} + \zeta_{i,t-1}\right) + \psi_{i,t}\left(\epsilon_{i,t} + \epsilon_{i,t-1}\right) + \xi_{i,t} + \xi_{i,t-1} + u_t^c - u_{t-2}^c.$$

For the variances and covariances this implies (again assuming also that $Var(\zeta_t) = Var(\zeta_{t-1})$, $Var(\epsilon_t) = Var(\epsilon_{t-1})$ and $Var(\xi_t) = Var(\xi_{t-1})$):

$$\begin{aligned} Var(\Delta_2 c_{i,t}) &= \phi_t^2 (2Var(\zeta_t)) + \psi_t^2 (2Var(\epsilon_t)) + 2Var(\xi_t) + Var(u_t^c) + Var(u_{t-2}^c), \\ Cov(\Delta_2 c_{i,t}, \Delta_2 c_{i,t+2}) &= -Var(u_t^c), \\ Cov(\Delta_2 c_{i,t}, \Delta_2 c_{i,t+s}) &= 0 \qquad \forall s > 2. \end{aligned}$$

as well as

$$\begin{aligned} Cov(\Delta_2 c_{i,t}, \Delta_2 y_{i,t}) &= \phi_t (2Var(\zeta_t)) + \psi_t Var(\epsilon_t), \\ Cov(\Delta_2 c_{i,t}, \Delta_2 y_{i,t+2}) &= -\psi_t Var(\epsilon_t), \\ Cov(\Delta_2 c_{i,t}, \Delta_2 y_{i,t+s}) &= 0 \qquad \forall s > 2 \text{ and } s < 0. \end{aligned}$$

The structure of the covariance relationships is analogous to the case of annual data. Thus, all important parameters can also be identified when only biennial data is available.

Estimation Procedure

For every household in the dataset the following column-vector of observations containing the first and second differences of residual consumption and income is collected:

$$x_i = \begin{array}{c} (\Delta c_{i,1979}, \Delta c_{i,1980}, ..., \Delta c_{i,1986}, \Delta c_{i,1990}, ..., \Delta c_{i,1996}, \Delta_2 c_{i,1998}, ..., \Delta_2 c_{i,2010}, \\ \Delta y_{i,1979}, \Delta y_{i,1980}, ..., \Delta y_{i,1996}, \Delta_2 y_{i,1998}, ..., \Delta_2 y_{i,2010})' \end{array}$$

For households that are present in the dataset throughout the whole sample period without missing observations, this yields 22 observations on consumption growth and 25 observations on income growth. Hence, $dim(x_i) = 47$. Missing values of income or consumption growth are set equal to zero. To control for missing values another vector d_i of the same dimension is constructed for every household in the dataset. d_i contains ones in years where the observations are non-missing and zeros for years with missing values, respectively.

With the help of x_i and d_i a symmetric 47x47-matrix M can be set up:

$$M = \left(\sum_i x_i x_i'\right) \bigotimes \left(\sum_i d_i d_i'\right)$$

where $\bigotimes$ represents an elementwise division. M contains estimates of all variances, covariances and autocovariances of income and consumption growth:

$$M = \left(\begin{array}{cccc} Var\left(\Delta c_{i,1979}\right) & & & \\ Cov\left(\Delta c_{i,1979}, \Delta c_{i,1980}\right) & & & \\ ... & & & \\ Cov\left(\Delta c_{i,1979}, \Delta_2 c_{i,2008}\right) & ... & Var\left(\Delta_2 c_{i,2008}\right) & \\ Cov\left(\Delta c_{i,1979}, \Delta_2 c_{i,2010}\right) & ... & Cov\left(\Delta_2 c_{i,2008}, \Delta_2 c_{i,2010}\right) & Var\left(\Delta_2 c_{i,2010}\right) \\ Cov\left(\Delta c_{i,1979}, \Delta y_{i,1979}\right) & ... & Cov\left(\Delta_2 c_{i,2008}, \Delta y_{i,1979}\right) & Cov\left(\Delta_2 c_{i,2010}, \Delta y_{i,1979}\right) \\ Cov\left(\Delta c_{i,1979}, \Delta y_{i,1980}\right) & ... & Cov\left(\Delta_2 c_{i,2008}, \Delta y_{i,1980}\right) & Cov\left(\Delta_2 c_{i,2010}, \Delta y_{i,1980}\right) \\ ... & & ... & ... \\ Cov\left(\Delta c_{i,1979}, \Delta_2 y_{i,2008}\right) & ... & Cov\left(\Delta_2 c_{i,2008}, \Delta_2 y_{i,2008}\right) & Cov\left(\Delta_2 c_{i,2010}, \Delta_2 y_{i,2008}\right) \\ Cov\left(\Delta c_{i,1979}, \Delta_2 y_{i,2010}\right) & ... & Cov\left(\Delta_2 c_{i,2008}, \Delta_2 y_{i,2010}\right) & Cov\left(\Delta_2 c_{i,2010}, \Delta_2 y_{i,2010}\right) \end{array} \right. ...$$

$$... \left. \begin{array}{ccccc} Var\left(\Delta y_{i,1979}\right) & & & & \\ Cov\left(\Delta y_{i,1979}, \Delta y_{i,1980}\right) & Var\left(\Delta y_{i,1980}\right) & & & \\ ... & ... & & & \\ Cov\left(\Delta y_{i,1979}, \Delta_2 y_{i,2008}\right) & Cov\left(\Delta y_{i,1980}, \Delta_2 y_{i,2008}\right) & ... & Var\left(\Delta_2 y_{i,2008}\right) & \\ Cov\left(\Delta y_{i,1979}, \Delta_2 y_{i,2010}\right) & Cov\left(\Delta y_{i,1980}, \Delta_2 y_{i,2010}\right) & ... & Cov\left(\Delta_2 y_{i,2008}, \Delta_2 y_{i,2010}\right) & Var\left(\Delta_2 y_{i,2010}\right) \end{array} \right)$$

For the minimum distance estimation the matrix M is half-vectorized according to $m = vech\left(M\right)$ so that the resulting vector m is of dimension $\frac{47*48}{2} = 1128$ and contains all unique second moments of the data.

$$m = \left(\begin{array}{c} Var\left(\Delta c_{i,1979}\right) \\ Cov\left(\Delta c_{i,1979}, \Delta c_{i,1980}\right) \\ ... \\ Cov\left(\Delta c_{i,1979}, \Delta_2 y_{i,2010}\right) \\ ... \\ Var\left(\Delta y_{i,1979}\right) \\ Cov\left(\Delta y_{i,1979}, \Delta y_{i,1980}\right) \\ ... \\ Var\left(\Delta_2 y_{i,2010}\right) \end{array} \right).$$

The estimation procedure minimizes the distance between the empirical second moments in the vector m and a vector $f(\theta)$ that contains the predictions of the parametric income model for these moments. The vector $f(\theta)$ is a function of the vector θ that contains the parameters of the income model that we seek to estimate (ϕ_t, ψ_t, $Var\left(\zeta_{i,t}\right)$, $Var\left(\epsilon_{i,t}\right)$, $Var\left(u_{i,t}^c\right)$, $Var\left(\xi_i\right)$):

$$f(\theta) = \begin{pmatrix} \phi_{1979}^2 Var\left(\zeta_{i,1979}\right) + \psi_{1979}^2 Var\left(\epsilon_{i,1979}\right) + Var\left(\xi_i\right) + Var\left(u_{i,1979}^c\right) + Var\left(u_{i,1978}^c\right) \\ -Var\left(u_{1979}^c\right) \\ ... \\ 0 \\ ... \\ Var\left(\zeta_{1979}\right) + Var\left(\epsilon_{1979}\right) + Var\left(\epsilon_{1978}\right) \\ -Var\left(\epsilon_{1979}\right) \\ ... \\ 2Var\left(\zeta_{2009/10}\right) + Var\left(\epsilon_{2010}\right) + Var\left(\epsilon_{2008}\right) \end{pmatrix}.$$

Equally weighted minimum distance (EWMD) estimates the parameter vector θ by

$$\min_{\theta} \left(m - f(\theta)\right)' \left(m - f(\theta)\right).$$

Standard errors for θ can be estimated as shown by Chamberlain (1984):

$$\widehat{Var\left(\widehat{\theta}\right)} = \left(G'G\right)^{-1} G'VG \left(G'G\right)^{-1}$$

where $G = \frac{\partial f(\theta)}{\partial \theta'}|_{\theta=\widehat{\theta}}$ and V is the variance-covariance matrix of m.

Estimation Results

Table A3: Additional results of the minimum distance estimation - standard errors in parantheses

$Var\,(\zeta_t)$				$Var\,(\epsilon_t)$			
		92	0.0412	$t = 78$	0.0563	92	0.0829
			(0.0063)		(0.0063)		(0.0067)
		93	0.0504	79	0.0438	93	0.0836
			(0.0071)		(0.0039)		(0.0063)
$t = 79/80$	0.0316	94	0.0490	80	0.0505	94	0.0720
	(0.0041)		(0.0069)		(0.0043)		(0.0065)
81	0.0321	95	0.0549	81	0.0527	95	0.0691
	(0.0043)		(0.0075)		(0.0045)		(0.0058)
82	0.0359	96	0.0576	82	0.0526	96	0.0769
	(0.0046)		(0.0066)		(0.0041)		(0.0062)
83	0.0415	97/98	0.0293	83	0.0520	98	0.0691
	(0.0055)		(0.0031)		(0.0043)		(0.0055)
84	0.0355	99/00	0.0221	84	0.0543	00	0.0670
	(0.0046)		(0.0030)		(0.0047)		(0.0053)
85	0.0501	01/02	0.0274	85	0.0570	02	0.0722
	(0.0055)		(0.0032)		(0.0047)		(0.0053)
86	0.0376	03/04	0.0295	86	0.0560	04	0.0693
	(0.0047)		(0.0057)		(0.0047)		(0.0053)
87	0.0338	05/06	0.0263	87	0.0568	06	0.0660
	(0.0055)		(0.0032)		(0.0049)		(0.0056)
88	0.0343	07/10	0.0262	88	0.0629	08	0.0722
	(0.0056)		(0.0033)		(0.0051)		(0.0060)
89	0.0328			89	0.0544	10	0.0958
	(0.0052)				(0.0050)		(0.0092)
90	0.0384			90	0.0537		
	(0.0053)				(0.0049)		
91	0.0387	$Var\,(\xi_t)$	0.0038	91	0.0586		
	(0.0054)		(0.0004)		(0.0054)		

Appendix A4: Robustness

Figure A1: Robustness test: Insurance parameters when SEO-subsample is dropped

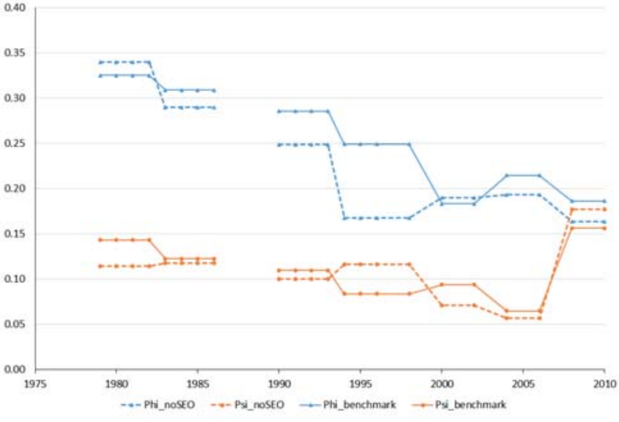
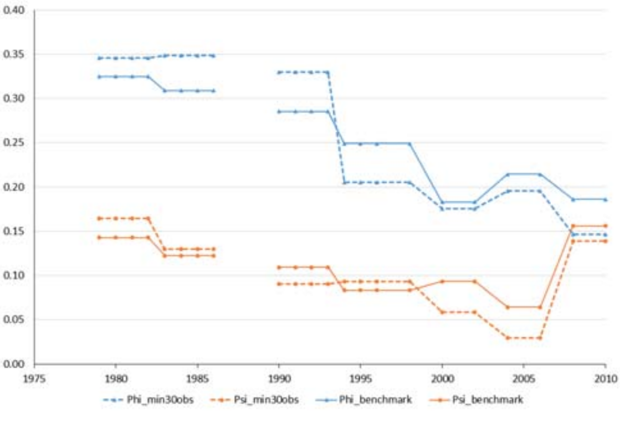


Figure A2: Robustness test: Insurance parameters when 30 observations per household are required



References

Abowd, J. M. & Card, D. (1989). On the Covariance Structure of Earnings and Hours Changes, *Econometrica* **57**(2): 411–445.

Aguiar, M. & Bils, M. (2013). Has Consumption Inequality Mirrored Income Inequality? Working Paper.

Attanasio, O., Hurst, E. & Pistaferri, L. (2013). The Evolution of Income, Consumption, and Leisure Inequality in the US, 1980-2010, *Improving the Measurement of Consumer Expenditures*, University of Chicago Press.

Attanasio, O. & Pistaferri, L. (2014). Consumption Inequality Over the Last Half Century. Some Evidence using the New PSID Consumption Measure, *American Economic Review Papers & Proceedings* **104**(5): 122–126.

Blundell, R., Pistaferri, L. & Preston, I. (2008). Consumption Inequality and Partial Insurance, *American Economic Review* **98**(5): 1887–1921.

Blundell, R., Pistaferri, L. & Saporta-Eksten, I. (2012). Consumption inequality and family labor supply. National Bureau of Economic Research Working Paper 18445.

Chamberlain, G. (1984). Panel data, *Handbook of econometrics*, Vol. 2, Elsevier, pp. 1247–1318.

Feenberg, D. & Coutts, E. (1993). An introduction to the TAXSIM model, *Journal of Policy Analysis and Management* **12**(1): 189–194.

Fisher, J. D., Johnson, D. S. & Smeeding, T. M. (2013). Measuring the Trends in Inequality of Individuals and Families: Income and Consumption, *American Economic Review Papers & Proceedings* **103**(3): pp. 184–188.

Hall, R. E. & Mishkin, F. S. (1982). The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households, *Econometrica* **50**(2): pp. 461–481.

Heathcote, J., Perri, F. & Violante, G. L. (2010). Unequal we stand: An empirical analysis of economic inequality in the United States, 1967-2006, *Review of Economic Dynamics* **13**(1): 15 – 51.

Kaplan, G. & Violante, G. L. (2010). How Much Consumption Insurance Beyond Self-Insurance?, *American Economic Journal: Macroeconomics* **2**(4): 53–87.

Kaufmann, K. & Pistaferri, L. (2009). Disentangling insurance and information in intertemporal consumption choices, *American Economic Review Papers & Proceedings* **99**(2): 387–392.

Krueger, D. & Perri, F. (2006). Does income inequality lead to consumption inequality? Evidence and theory, *The Review of Economic Studies* **73**(1): 163–193.

Meghir, C. & Pistaferri, L. (2011). Earnings, Consumption and Life Cycle Choices, *in* D. Card & O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 4, Part B, Chapter 9, Elsevier, pp. 773 – 854.

Meyer, B. D. & Sullivan, J. X. (2013a). Consumption and Income Inequality and the Great Recession, *American Economic Review Papers & Proceedings* **103**(3): pp. 178–183.

Meyer, B. D. & Sullivan, J. X. (2013b). Consumption and Income Inequality in the U.S. Since the 1960s. Working Paper.

Primiceri, G. E. & van Rens, T. (2009). Heterogeneous life-cycle profiles, income risk and consumption inequality, *Journal of Monetary Economics* **56**(1): 20 – 39.