OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 81, 6 (2019) 0305–9049
doi: 10.1111/obes.12310

Testing the Empirical Relevance of the 'Saving for a Rainy Day' Hypothesis in US Metro Areas*

André Kallåk Anundsen† and Ragnar Nymoen‡

†*Housing Lab, Oslo Metropolitan University, Oslo, Norway*
(e-mail: andre-kallak.anundsen@oslomet.no)
‡*Department of Economics, University of Oslo, Oslo, Norway*
(e-mail: ragnar.nymoen@econ.uio.no)

Abstract

The joint implication of the consumption Euler equation and cointegration between income and consumption is that savings predict future income declines, the 'saving for a rainy day' hypothesis. The empirical relevance of this hypothesis plays a key role in discussions of fiscal policy multipliers, and it holds under the null that the permanent income hypothesis is true. We find little support for this hypothesis using time series data for the 100 largest US Metropolitan Statistical Areas for the period 1980q1–2015q4. Our approach is to test for cointegration and weak exogeneity between income and consumption, and by exploring the direction of Granger causality between the two time series. We find that income more often predicts consumption and saving than the converse. We also give evidence that house price changes played a role in US income and consumption dynamics, before, during and after the Great Recession.

I. Introduction

Consumer expenditure is by far the largest component of spending in the US economy, and in most other countries as well. The study of saving and consumption dynamics is therefore of great importance for both economic policy analysis and economic forecasting.

JEL Classification numbers: C22, C32, C51, C52, E21, E62.
*We are thankful to the editor, Anindya Banerjee, and an anonymous reviewer for comments that have contributed to improve the paper considerably. Large parts of this paper were written while Anundsen was working at Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. Earlier versions of the paper was presented at the conference 'Econometric Modelling in a Rapidly Changing World' in Oxford, September 2014, at the 2014 European Winter Meeting of the Econometric Society in Madrid, December 2014, at the 37th Annual Meeting of the Norwegian Economic Association in Bergen, January 2015, at the 23rd Symposium of the Society for Nonlinear Dynamics and Econometrics in Oslo, March 2015, and at seminars and workshops in the Central Bank of Hungary, Norges Bank and Statistics Norway. We are grateful to the participants at these events for their comments. We thank Farooq Akram, Bruce Hansen, Veronica Harrington, Tord Krogh and Asbjørn Rødseth for comments and discussions. All results may be reproduced using an Ox-code that will be made available on the authors' websites.

It is well known that the rational expectations permanent income hypothesis (PIH hereafter) due to Hall (1978) is consistent with unit-root non-stationarity of income and stationarity of saving, see e.g. Muellbauer and Lattimore (1995, Ch. 3.2). When combined with the famous theoretical result of Hall (1978), stating that consumption follows a first-order Markov-process, we obtain the implication that Granger causation runs from lagged saving to current income and not from saving to consumption. In this paper, we explore the empirical relevance of these theoretical conjectures by testing the direction of Granger causality between consumption and income using quarterly time series data for the 100 largest US Metropolitan Statistical Areas (MSAs) over the period 1980q1–2015q4. The alternative hypothesis is that consumption is not well approximated by the Euler equation. Instead consumption adjusts (partly) to predictable income changes, and savings is thereby affected dynamically.

A common ground is represented by the idea that the savings rate may be a stationary variable, even though there are stochastic trends in the time series of both income and consumption. This common ground allows the analysis to be held within a vector autoregressive (VAR) framework. To account for stochastic trends in income and consumption, we apply econometric methods that are robust to non-stationarity. Under the hypothesis that the statistical relationship between consumption and income describes the PIH, a fall in saving anticipates a future increase in income and a rise in saving anticipates future income declines (see Campbell (1987)).[1] This also explains why the result has been dubbed the 'saving for a rainy day' hypothesis, cf. Attanasio (1999).

In his seminal paper, Campbell (1987) referred to (Granger) causation running from the savings rate to income growth – and not the other (Keynesian) way around – as the weak implication of the permanent income hypothesis. Empirically, using aggregate US data for the period 1953–84, Campbell found that the implication of the PIH for the direction of Granger causality preserved even if other implications of the PIH fared less well empirically.[2] In particular, the Euler equation found little support on aggregate data, see Flavin (1981), Campbell and Mankiw (1989) among others.

The conclusion that the PIH is at best partly correct, and that it needs to be supplemented by several factors to account for the many features of consumption dynamics that we are trying to understand, is well known (see e.g. Romer, 2006; Carroll, 2009; Jappelli and Pistaferri, 2010; Attanasio and Weber, 2010). Nevertheless, the PIH continues to be one of the core elements of modern macroeconomics (see e.g. Ljungqvist and Sargent, 2004, p. 3), it is the centre-piece of macroeconomic DSGE models, and it is essential to the iterative forward solution of these models, cf. Muellbauer (2016).

One reason why macroeconomists hold on to the principles of PIH, despite the empirical weakness of the consumption Euler equation, may be because it is thought that PIH nevertheless gets the system properties right: Income and consumption is cointegrated and income is the most important equilibrating variable, and savings can therefore hold some predictive power about future income. The objective of this paper is to contribute to the system assessment of the PIH.

[1] Campbell showed this for an infinitely lived consumer with quadratic utility function, equal and constant subjective discount rates and no credit constraints.

[2] Campbell (1987, p. 1267).

We test the empirical validity of the PIH restrictions on consumption-income VARs for 100 US MSAs. Since the restrictions implied by the alternative economic interpretation of the statistical system (the consumption function view) are testable within the same system, evidence for that view is also presented. The data from the MSAs represent aggregated outcomes, so our contribution is macroeconomic. However, the regional variation between the MSAs can be correlated with factors that influence consumption, for example distribution of income and wealth and unemployment. Hence, there is little reason to expect that the estimated VARs deliver identical cointegration parameters and adjustment coefficients. Some heterogeneity in the results is reasonable to expect, due to differences in the data generating processes. On the other hand, if the 100 empirical models show a clustering of results for the main parameters of interest, results are even more interesting, since the data is heterogenous. This also gives a panel data interpretation of our 100 separately estimated VARs, that is, as a fully heterogenous panel data investigation. We also estimate a panel data model with homogenous derivative parameters. Again, the results can be mutually supportive if estimation results for the parameters give the same main picture.

In addition, we estimate the models both on a sample covering only the Great Moderation period (1980q1–2007q4) and a sample including the Great Recession and the recovery (1980q1–2015q4). Studying both samples allow us to investigate the robustness of our results, in particular whether there is a change in the direction and significance of the link between income and consumption, which is the heart of the 'rainy day hypothesis'. Further, the empirical models we use allow us to shed light on the role of house prices for consumption dynamics before, during and after the Great Recession.

The empirical results from the Great Moderation sample strongly suggest that income is Granger causing consumption, while there is little support for Granger causation running in the other direction. This result is opposite to the early findings in Campbell (1987), but also nuances the results in Lettau and Ludvigson (2004), who find that their lagged equilibrium correction term, which includes income and wealth, explain little of the rate of change of consumption. Including the financial crisis period in our sample strengthens these result, and our findings therefore lead us to reject the 'saving for a rainy day' hypothesis.

We also show that house prices played a role in US income and consumption dynamics before, during and after the Great Recession. Moreover, our results suggest a strengthened role of house prices in affecting consumption dynamics after the financial crisis. This suggests that US consumers who saw their retirement funds saved up in the housing market completely wiped out during the housing bust increased their saving to compensate for this loss. These results are robust to controlling for stock prices and credit growth. In contrast to house prices, the link between consumption dynamics and credit growth is invariant to the extension of the data set to include the Great Recession and the period thereafter.

The paper proceeds as follows. In the next section, we outline the implied (and testable) VAR parameter restrictions of the PIH, and we discuss how we will proceed to explore the empirical relevance of these theoretical conjectures. In section IV, we present the data sets that are used in the econometric analyses. Results from the MSA-specific analyses over the Great Moderation are discussed in section V. In the same section, we explore how our main conclusions are affected by extending the data set to include the financial crisis period. In section VI, we analyse whether our results are robust to controlling for credit growth, the evolution of stock prices and alternative measures of consumption. The same section explores the robustness of our results to applying a panel data approach and to adopting an aggregate time series approach. We also provide evidence at the aggregate level that is congruent with the MSA evidence. The final section concludes the paper.

II. Related literature

Our paper follows the recent literature in empirical macroeconomic research which studies macroeconomic questions using disaggregate data (see e.g. Autor, Dorn and Hanson, 2013; Mian and Sufi, 2014; Beraja, Hurst and Ospina, 2016; Martin and Philippon, 2017). Another prominent example in this field is Nakamura and Steinsson (2014), who estimate a regional fiscal multiplier using state-level data. They argue that the regional data gives a powerful diagnostic for distinguishing between alternative macro models, and their estimated multiplier is on the high side of what is typically found using aggregate data. Based on these results, they suggest that researchers should give more weight on models where demand shocks have a larger effect on output. As pointed out in Nakamura and Steinsson (2018), the use of disaggregate data may be particularly useful to discriminate between different theory models. They may, however, be less suited for answering questions about aggregate effects, since it is not clear how regional responses translates into aggregate responses. Our paper is confined to asking whether the 'saving for a rainy day' hypothesis is a valid theoretical view of how the world works, and therefore the number of data points we get by taking a disaggregate approach seems useful to distinguish between different theoretical views.

The findings in this paper relate to the discussion about the role of expansionary fiscal policy during the jobs and incomes crisis that followed the financial crisis, cf. DeLong and Summers (2012), Eberly (2014) and Stiglitz (2014). The size of the fiscal multiplier depends on several premises, i.e. 'idle resources', the degree of import leakage and the marginal propensity to consume. With large numbers of unemployed, but employable, workers and a large domestic economy (small import leakage), the first two factors indicate that there has been a sizeable fiscal multiplier in the US over the last five to six years. However, if increased income ends up as private saving because the marginal propensity to consume is close to zero, the fiscal policy multiplier may nevertheless be very small.

We also contribute to the literature attempting to explain the puzzle that household saving declined during the Great Moderation and rose again during the Great Recession. A branch of this literature suggests the easing of credit conditions as an explanation, see e.g. Parker (2000) and Aron *et al.* (2012). Further, Guerrieri and Lorenzoni (2011), Eggertsson and Krugman (2012) and Hall (2011) find that the tightening of credit standards in the period succeeding the Great Recession can explain the sharp increase in the savings rate. An alternative explanation was highlighted in an earlier contribution by Carroll (1992), who suggested precautionary saving as an explanation for why savings rates tend to increase in recession periods. A more recent study by Alan, Crossley and Low (2012) reaches a similar conclusion. The evolution of the savings rate can also be affected by changes in households' net worth. Lettau and Ludvigson (2004) and Lettau and Ludvigson (2013) found only weak wealth effects. On the other hand, Mian, Rao and Sufi (2013) estimated a sizeable marginal propensity to consume out of housing net worth using US zip code level data for the 2006–09 period. Carroll, Slacalek and Sommer (2012) also investigated the relative importance of credit conditions, precautionary saving and the wealth channel in explaining US savings rate dynamics. While their results suggest that all three channels are important, they find that the largest contributor to the recent increase in the savings rate is the drop in household wealth. Our results are consistent with the findings in Carroll *et al.* (2012) and Mian *et al.* (2013).

III. The 'saving for a rainy day' hypothesis tested on cointegrated VARs for MSA data

Campbell (1987) derived the implication that, under the null hypothesis of the PIH, saving should encapsulate the superior information of the agent to the econometrician, meaning that lagged saving should Granger-cause income in a bivariate VAR. This is also consistent with Hall's consumption Euler-equation. Hence, consumption should not be Granger-caused by lagged income, or lagged saving, in the VAR. These implications contradicted the economic interpretation of empirical consumption functions of the Keynesian type, where the existing disequilibrium in saving was a predictor of next period's consumption change, Davidson *et al.* (1978).

A cointegrated VAR approach

Under the assumption that both consumption and income are integrated of order one, $I(1)$, cointegration represents a common ground between the consumption function approach, and the permanent-income/life-cycle theories (Eitrheim, Jansen and Nymoen, 2002).[3]

To test for cointegration between consumption and income, and to explore the direction of equilibrium correction and Granger causality, we develop MSA-specific econometric models. Our main reference is a VAR(p_j) model of the following form:

$$\boldsymbol{y}_{j,t} = \sum_{s=1}^{p_j} \boldsymbol{A}_{j,s}\boldsymbol{y}_{j,t-s} + \boldsymbol{\Phi}_j\boldsymbol{D}_{j,t} + \varepsilon_{j,t} \tag{1}$$

where the index j represents MSA unit. The vector $\boldsymbol{y}_{j,t}$ comprises real consumption and real disposable income. Deterministic terms (linear trend and a constant) are collected in the vector $\boldsymbol{D}_{j,t}$. House price changes and the real interest rate are also collected in $\boldsymbol{D}_{j,t}$. The disturbances are assumed to follow a multivariate normal distribution, with expectation $\boldsymbol{0}_{2\times1}$ and covariance matrix $\boldsymbol{\Sigma}_j$, i.e. $\varepsilon_{j,t} \sim MVN(\boldsymbol{0}_{2\times1}, \boldsymbol{\Sigma}_j)$.

For all areas, we start with a lag length of 5, i.e. $p_j = 5$. Then, we select the lag length (between 1 and 5) that minimizes the Akaike Information Criterium (AIC).[4] Conditional on the optimal lag truncation, p_j^*, we consider (1) on vector equilibrium correction (VECM) form. We follow the recommendation of Johansen (1995) and Harbo *et al.* (1998) and restrict a deterministic trend to enter the cointegration space in order to achieve a similar

[3] A detailed exposition of the framework with the use of a first order VAR is given in Online Appendix A.

[4] Other lag selection criteria also exists, and some researchers prefer the Schwarz Information Criterion (SIC). The main difference between the two criteria is that SIC punishes overparameterization relatively more than AIC, which gives more weight to fit (the likelihood). All of our main findings are robust to using SIC instead of AIC. Detailed results are available upon request.

test for cointegration rank.[5] Letting $\tilde{\boldsymbol{y}}_{j,t} = (\boldsymbol{y}'_{j,t}, trend_j)'$, the VECM representation of the VAR model takes the following form:

$$\Delta \boldsymbol{y}_{j,t} = \boldsymbol{\Pi}_j \tilde{\boldsymbol{y}}_{j,t-1} + \sum_{s=1}^{p_j^*-1} \boldsymbol{\Gamma}_{j,s} \Delta \boldsymbol{y}_{j,t-s} + \tilde{\boldsymbol{\Phi}}_j \tilde{\boldsymbol{D}}_{j,t} + \varepsilon_{j,t} \tag{2}$$

where $\tilde{\boldsymbol{D}}_{j,t}$ contains the constant term, the real interest rate and house price changes. All coefficient matrices are redefined conformably.

To determine the rank of the matrix $\boldsymbol{\Pi}_j$, we use the trace test of Johansen (1988). The rank of $\boldsymbol{\Pi}_j$ corresponds to the number of independent linear combinations between the variables in $\tilde{\boldsymbol{y}}_{j,t}$ that are stationary, i.e. the number of cointegrating relationships. When $\boldsymbol{\Pi}_j$ has reduced rank, we can write $\boldsymbol{\Pi}_j = \boldsymbol{\alpha}_j \boldsymbol{\beta}'_j$, where $\boldsymbol{\beta}_j$ is a $(l_j + 1) \times r_j$ matrix and $\boldsymbol{\alpha}_j$ is a $l_j \times r_j$ matrix corresponding to the long-run coefficients and loading factors (adjustment coefficients) respectively. The rank of $\boldsymbol{\Pi}_j$ is denoted by r_j, while $l_j + 1$ refers to the number of endogenous variables (plus the deterministic trend, which is restricted to lie in the space spanned by $\boldsymbol{\alpha}_j$). In all areas, l_j is equal to 2 (real consumption and real disposable income).

Conditional on non-zero rank, we can estimate the parameters in the cointegration space. In particular, our approach allows us to explore heterogeneities in both long-run income elasticities and the speed of adjustment parameters. Moreover, cointegration implies that there is Granger causality in at least one direction (Granger, 1986). To formally explore the direction of causality, in the Granger sense, consider the reduced rank representation of the VECM. Before exploring the direction of GNC, we test and impose the restriction that $\beta_{trend,j} = 0$. Hence, the VECM used for testing the 'rainy day hypothesis' can be written as:

$$\begin{aligned} \begin{pmatrix} \Delta c_{j,t} \\ \Delta y_{j,t} \end{pmatrix} &= \begin{pmatrix} \alpha_{c_j} \\ \alpha_{y_j} \end{pmatrix} \left(c_{j,t-1} - \beta_{y,j} y_{j,t-1} \right) \\ &\quad + \sum_{s=1}^{p_j^*-1} \begin{pmatrix} \Gamma_{11,j,s} & \Gamma_{12,j,s} \\ \Gamma_{21,j,s} & \Gamma_{22,j,s} \end{pmatrix} \begin{pmatrix} \Delta c_{j,t-s} \\ \Delta y_{j,t-s} \end{pmatrix} + \tilde{\boldsymbol{\Phi}}_j \tilde{\boldsymbol{D}}_{j,t} + \varepsilon_{j,t} \end{aligned} \tag{3}$$

where we have normalized the first coefficient in the cointegration space with respect to consumption.

The PIH holds that consumption growth is not Granger-caused by the lagged income, hence $\alpha_{c_j} = \Gamma_{12,j,s} = 0 \; \forall \; s$. Given $\alpha_{c_j} = 0$, cointegration implies that $0 < \alpha_{y_j} < 1$, since – as we know from the Engle–Granger representation theorem (Engle and Granger, 1987) – cointegration implies equilibrium correction, and *vice versa*.

Underlying the consumption function approach is the idea that consumption is equilibrium correcting, i.e. $-1 < \alpha_{c_j} < 0$. Given that this requirement is fulfilled, there are two possibilities for the coefficient α_{y_j}: (i) $0 < \alpha_{y_j} < 1$ or (ii) $\alpha_{y_j} = 0$. The first case is consistent with hours worked etc. being *demand determined* and that $y_{j,t}$ adjusts to past disequilibria. In econometric terms, there is mutual (Granger) causation between income and consumption, see Engle, Hendry and Richard (1983). The second possibility implies that income

[5] An alternative to our approach is to exclude the trend from the outset. All our results are robust to excluding the trend from the outset. Detailed results are available upon request.

is *supply determined*. In the context of the VAR, the restriction that $\alpha_{y_j} = 0$ implies that income is weakly exogenous with respect to the long-run income elasticity, β_{y_j}, see e.g. Johansen (1992). Moreover, there should be one-way Granger causation from income to consumption, so income is also strongly exogenous. To test for GNC from consumption to income in area j, we test the joint hypothesis that $\alpha_{y_j} = \Gamma_{21,j,s} = 0 \, \forall \, s$.

Allowing for MSA-specific structural breaks

When building the MSA-specific econometric models, we also make use of the *impulse indicator saturation* (IIS) algorithm, which is an integrated part of the *Autometrics* algorithm implemented within OxMetrics (see Doornik 2009; Hendry and Doornik, 2009) to allow for MSA-specific structural breaks.

The IIS algorithm includes an impulse dummy for each observation in the information set. More precisely, this implies that the baseline VAR in (1) can be modified to:

$$\boldsymbol{y}_{j,t} = \sum_{s=1}^{p_j} \boldsymbol{A}_{j,s}\boldsymbol{y}_{j,t-s} + \boldsymbol{\Phi}_j \boldsymbol{D}_{j,t} + \boldsymbol{\Psi}_j \boldsymbol{I}_t + \varepsilon_{j,t} \quad t = t_j, \ldots, T \tag{4}$$

where $\boldsymbol{I}_t$ is a $(T+1-t_j) \times (T+1-t_j)$ matrix of impulse dummies. Since this entails that there are more variables than observations, the model is estimated in blocks to determine which indicators are significant (see Hendry, Johansen and Santos, 2008; Johansen and Nielsen, 2009). If we let the retained indicators for area j be collected in the $(T+1-t_j) \times Q_j$ matrix $\tilde{\boldsymbol{I}}_{j,t}$, with $Q_j < (T+1-t_j)$, the IIS robust reparameterization of the VAR takes the following form:

$$\Delta \boldsymbol{y}_{j,t} = \boldsymbol{\Pi}_j \tilde{\boldsymbol{y}}_{j,t-1} + \sum_{s=1}^{p_j^*-1} \boldsymbol{\Gamma}_{j,s} \Delta \boldsymbol{y}_{j,t-s} + \tilde{\boldsymbol{\Phi}}_j \tilde{\boldsymbol{D}}_{j,t} + \tilde{\boldsymbol{\Psi}}_j \tilde{\boldsymbol{I}}_{j,t} + \varepsilon_{j,t} \tag{5}$$

After having estimated (4) employing Autometrics, we follow the same steps as those described in the previous section, i.e. we test down the lag length using the AIC, determine the rank of the matrix $\boldsymbol{\Pi}_j$, and conduct tests for both weak exogeneity and Granger non-causality. Thus, the estimates and tests obtained in this case can be seen as being robustified to MSA-specific structural breaks (Johansen and Nielsen, 2009).

Applying the IIS algorithm, an average of $\alpha^{IIS} \times (T+1-t_j)$ indicators will be retained by chance, where α^{IIS} denotes a pre-specified significance level used for the selection of indicators. When applying the IIS algorithm to the VAR model of area j, we set the significance level to 0.1%. As expected from the documented properties of the algorithm (Castle, Doornik and Hendry, 2012), very few indicators are picked up in the MSA analysis.

IV. Data

We have collected quarterly time series data at both the national level and for the 100 largest MSAs in the US. For most of the areas, the data set spans the period 1980q1–2015q4 ($T = 144$).

The MSAs included in our MSA data set cover all but four of the 50 US states and are spread out in different geographical regions. To ease the exposition, we shall follow the Census Bureau and divide the US into four major regions (West, Midwest, South and East) when discussing some of our results.[6] With reference to those regions, 25 of the MSAs included in our sample belong to the West, 20 to the East, 30 to the South and 25 to the Midwest.

The income data, $y_{j,t}$, measure personal disposable income in billions of USD. Disaggregate consumption data are only available for a relatively short time period at the MSA level. For that reason, we use data on retail sales in billions of USD as a proxy for consumption, $c_{j,t}$. That said, as pointed out by Sorensen and Luengo-Prado (2008), the correlation between aggregate US retail sales and non-durable consumption is very high. Thus, in the absence of data on MSA level consumption, we take this to be a relatively good proxy. Similar data have been used in e.g. Case, Quigley and Shiller (2012), Sorensen and Luengo-Prado (2008) and Dejuan, Seater and Wirjanto (2004), who all consider state level consumption in the US. We follow Case *et al.* (2012) and use the retail sales data supplied by Moody's (formerly supplied by Regional Financial Associates). That said, as shown in a separate robustness exercise in section 'Weak exogeneity and Granger non-causality', all our results are robust to using consumption data on the shorter sample. House price data are collected from the FHFA, and all series are deflated by the corresponding MSA level CPI measure, which has also been collected from Moody's. Finally, MSA-specific real interest rates are constructed by subtracting the MSA level CPI inflation from the nominal 3-month T-bill. In the empirical analysis, all variables, except the real interest rate, are included in log form.

The discussion in section III is based on the premise that the time series for income and consumption contain unit roots. To investigate the empirical relevance of this assumption, we have tested the order of integration of the data series using standard ADF tests (Dickey and Fuller, 1979) for each of the areas. In particular, we started with a lag length of 5, including a deterministic trend in the ADF regressions. Then, the optimal lag truncation was chosen by a sequence of t-tests. The average order of integration is close to one for both series.[7] We have also performed tests for unit roots in consumption and income using the ADF-GLS test suggested by Elliot, Rothenberg and Stock (1996). Results for consumption were almost unchanged, while there were more rejections than before for income, but still rejected in a large number of areas. Hence, the overall picture regarding the order of integration of the individual series remains the same. Based on these results, we continue our analysis under the modelling assumption that both series are integrated of order one.

V. Econometric results

Tests for slope homogeneity

While we have opted for an MSA-by-MSA analysis to allow for full heterogeneity in parameters, and in order to explore any regional differences in results, there are some

[6] The estimation and testing are, however, carried out separately for each MSA.

[7] Detailed results from the unit root tests at the MSA level are available upon request.

TABLE 1

Tests for slope homogeneity

Region	*Great moderation sample*	*Full sample*
West	0.0000	0.0000
East	0.2545	0.1298
South	0.0002	0.0000
Midwest	0.0000	0.0000
All	0.0000	0.0000

Notes: This table reports p-values from likelihood ratio tests for whether there is any information loss from imposing equal slope coefficients, i.e., testing whether there is any information loss from considering a fixed effects estimator relative to the mean group estimator. The table sorts the MSA by the four census regions; West, Northeast, South, Midwest. Results are reported for both the Great Moderation period and for the full sample.

advantages with using a standard panel estimator instead. First, conditional on the validity of the homogeneity assumption, it improves estimation efficiency. Second, in many cases, it may be the only feasible estimator, since the time dimension for each cross-sectional unit often is limited. However, the potential drawback of this method is equally obvious: it only allows the intercept to vary across units in the panel, while imposing the rather strict assumptions that slope coefficients are the same.

As stressed by Pesaran and Smith (1995), Im, Pesaran and Shin (2003), Pesaran, Shin and Smith (1999) and Phillips and Moon (2000), there are several instances where the usual panel assumption of homogenous slope coefficients (as imposed in the FE model) does not apply. To formally test whether slope homoegeneity is a valid assumption, we calculate the likelihood of the restricted model (FE) and test the validity of the imposed homogeneity restrictions against the unrestricted model (MG) by use of a likelihood ratio test. Since our analysis is carried out on both the Great Moderation sample (1980q1–2007q4) and a sample that includes the Great Recession (1980q1–2015q4), we test for slope homogeneity on both samples. Table 1 summarizes p-values from these tests across Census regions. It is clear that the assumption of homogeneity is rejected. This holds true in both samples. We therefore continue our analysis by estimating MSA-specific models when testing the empirical relevance of the 'saving for a rainy day' hypothesis.

Cointegration results for the MSA data set

In this section, we present the results from the MSA-specific econometric analysis using data for the Great Moderation (1980q1–2007q4). While detailed results for the individual MSAs are presented in Tables B.1–B.4 in section B of the Online Appendix, Table 2 reports a summary of the average results across each of the four census regions. In the first step of our estimation approach, the IIS algorithm picks up a little more than one dummy on average (confer the last row in the first column of Table 2). Based on AIC, we find the average optimal lag truncation to be just below 4, and the hypothesis of co-trending is supported for a majority of the areas (66%) when we use a 1% significance level.

As is also evident from Table 2, we find clear evidence in a majority of the areas that the residuals are well behaved, i.e. there are no signs of autocorrelation, heteroscedasticity

TABLE 2

Averages and percentages of key model features for Great Moderation sample (1980q1–2007q4), ordered by census region

Region	Dummies (avg.)	p^* (avg.)	Rank($\mathbf{\Pi}$) (avg.)	Auto. (%)	Norm. (%)	Hetero. (%)	$\beta_{trend} = 0$ (%)
West	1.36	3.16	0.48	100.00	100.00	76.00	76.00
East	0.90	3.75	0.95	100.00	100.00	100.00	70.00
Midwest	0.48	4.28	1.08	96.00	100.00	100.00	44.00
South	1.43	4.13	0.97	96.67	96.67	86.67	73.33
All	1.07	3.85	0.87	98.00	99.00	90.00	66.00

Notes: Columns 2–4 report the average number of dummies, Dummies (avg.), included in the econometric models within each of the four major regions, as well as the average optimal lag truncation, p^* (avg.) and average number of cointegrating relationships, Rank($\mathbf{\Pi}$). Columns 5–7 report the percentage number of times where we do not reject absence of autocorrelation, non-normality and heteroscedasticity. The final column displays the percentage number of areas where we find support for co-trending, i.e., $\beta_{trend} = 0$. The final row in each column reports the same figures for all the MSAs covered by the sample (all areas). Detailed results for the individual MSAs are reported in Tables B.1–B.4 in section B of the Online Appendix.

nor departures from normality. It is also evident that the average rank is very close to one, which is also what we will impose for the rest of the analysis.

Imposing the reduced rank restriction and normalizing the cointegrating vector with respect to consumption ($\beta_{c,j} = 1\ \forall j$), we obtain estimates of the long-run income elasticity. Detailed results of the individual MSAs are reported in Tables B.5–B.8 in section B of the Online Appendix. The same appendix contains a figure plotting the point estimates for the long-run income elasticity for all of the areas included in our sample (Figure B.1). Averages and medians across Census regions are, however, reported in Table 3.

The average long-run income elasticity across all areas is 0.84, and the standard error of this mean group estimate is 0.03, see the second and fourth column in Table 3.[8] Note that the hypotheses about the adjustment coefficients and the direction of Granger causality between income and consumption (our parameters of interest) do not require that $\beta_y = 1$. For this reason, we continue the analysis with as few restrictions as possible.

Our results imply that $\alpha_{c,j} < 0$ for a majority of the areas, which is consistent with a consumption function interpretation, and at odds with the 'rainy day' hypothesis. In the cases where a positive point estimate is obtained, we cannot reject the hypothesis that the parameter is equal to zero. Figure B.2 in the Online Appendix plots the distribution of the point estimate of the two adjustment parameters across the 100 MSAs, and we find that the average estimated α_{c_j} is -0.12, see column 5 in Table 3.[9] The estimated α_{y_j} is found to be positive in a majority of the cases.[10] The average estimate of α_{y_j} is 0.025, which is substantially lower than the absolute value of the average α_{c_j} estimates. Hence, the results thus far suggest mixed support for the 'saving for a rainy day' hypothesis as an empirically relevant description of US consumption behaviour.

[8] In calculating the mean group estimates, we have excluded the outliers where the estimated income elasticity was negative or above 2. This is only the case for 7 areas, meaning that mean group estimates are based on the remaining 93 MSAs.

[9] Detailed results for each MSA can be found in Tables B.5–B.8 in section B of the Online Appendix.

[10] In the few cases where the point estimate is negative, weak exogeneity is not rejected by the data.

TABLE 3

Summary of cointegration results for Great Moderation sample (1980q1–2007q4)

Region	$\hat{\beta}_y^{IIS}$			$\hat{\alpha}_c^{IIS}$			$\hat{\alpha}_y^{IIS}$		
	Mean	Median	Standard error	Mean	Median	Standard error	Mean	Median	Standard error
West	0.8786	0.8663	0.0367	−0.1132	−0.1036	0.0138	0.0295	0.0397	0.0077
East	0.7345	0.7386	0.1621	−0.1150	−0.1178	0.0134	0.0245	0.0025	0.0125
Midwest	0.8322	0.7960	0.0453	−0.1227	−0.1268	0.0179	0.0166	0.0204	0.0100
South	0.8750	0.8728	0.0281	−0.1170	−0.1101	0.0127	0.0282	0.0204	0.0072
All	0.8397	0.8406	0.0253	−0.1171	−0.1101	0.0072	0.0250	0.0209	0.0045

Notes: The table reports the average long-run income elasticities of consumption ($\hat{\beta}_y^{IIS}$), the adjustment parameter in the consumption equation ($\hat{\alpha}_c^{IIS}$) and the adjustment parameter in the income equation ($\hat{\alpha}_y^{IIS}$), grouped by census region. The table also reports the median and the standard error for each of these coefficients. The final row in each column reports the same figures for all the MSAs covered by the sample (all areas). Detailed results for the individual MSAs can be found in Table B.5–B.8 in Appendix B.

Weak exogeneity and Granger non-causality

Using the optimal lag truncations of the VAR models, as found in the previous section, together with the estimated cointegrating vectors, we derive the vector equilibrium correction representation of the CVAR models (confer (3)). The VECM for each area is estimated by FIML, and Table 4 summarizes the main results regarding tests for both weak exogeneity and Granger non-causality.[11] We use two approaches to test for GNC. The first is a standard approach, where we test within the framework of the VECM. The other approach, uses the level-based test of Toda and Yamamoto (1995). An advantage of this test is that one does not have to take a stand on the cointegrating relationship(s). Although this alternative test is mostly useful in the case where the research interest is not the cointegrating relationship itself (see discussion in Toda and Yamamoto, 1995), we also include results of this test for robustness.

As is evident by inspecting the fourth and fifth column in Table 4, weak exogeneity of consumption with respect to the cointegrating vector is rejected in a majority of the cases (84%), while weak exogeneity of income is rejected only for 33% of the MSAs. These findings are at odds with the weak implication of the PIH, and further support for this claim is provided by the results in the sixth and seventh column in Table 4, where we report the percentage number of times where we find evidence that income is Granger-causing consumption (87%) and the percentage number of times where we find evidence that consumption is Granger-causing income (56%). Similar results are obtained for the level-based test (see the second and third column in Table 4). The Granger causality tests for house prices suggest that house prices Granger-cause consumption in 60% of the areas, indicating that house prices may be important for consumption dynamics in some MSAs. The same figure for income is around 38%.

[11] Note that the reported results are based on the MSAs where the estimated long-run income elasticity was 'meaningful' – defined as $0 < \beta_{y_j} < 2$. This holds for all, except 7 areas, meaning that tests for weak exogneity and GNC are based on 93 MSAs.

TABLE 4

Tests for weak exogeneity and Granger non-causality for Great Moderation sample (1980q1–2007q4)

Region	*VAR in levels:*		*VECM:*					
	$y \underset{GC}{\rightarrow} c$	$c \underset{GC}{\rightarrow} y$	$\alpha_c \neq 0$	$\alpha_y \neq 0$	$y \underset{GC}{\rightarrow} c$	$c \underset{GC}{\rightarrow} y$	$ph \underset{GC}{\rightarrow} c$	$ph \underset{GC}{\rightarrow} y$
West	87.50	83.33	79.17	37.50	91.67	75.00	45.83	37.50
East	70.59	70.59	94.12	29.41	82.35	47.06	76.47	29.41
Midwest	47.83	47.83	86.96	26.09	86.96	34.78	56.52	43.48
South	86.21	82.76	79.31	37.93	86.21	62.07	65.52	37.93
All	74.19	72.04	83.87	33.33	87.10	55.91	60.22	37.63

Notes: Column 4–7 report the percentage number of times where weak exogeneity of consumption is rejected ($\alpha_c \neq 0$) and the percentage number of times where weak exogeneity of income ($\alpha_y \neq 0$) is rejected, as well as the percentage number of times where we find that income Granger-causes consumption ($y \underset{GC}{\rightarrow} c$) and vice versa ($c \underset{GC}{\rightarrow} y$). The final two columns report the percentage number of times where we find that house prices Granger-cause consumption ($ph \underset{GC}{\rightarrow} c$) and income ($ph \underset{GC}{\rightarrow} y$).

Including the financial crisis period

We have seen that the 'saving for a rainy day' hypothesis receives mixed support over the Great Moderation period. The empirical evidence is clearly supportive to the interpretation that consumption represents the main equilibrium correction mechanism.

In this section, we check if the assessment changes when we extend the sample to include the financial crisis period and the ensuing income and job crisis, i.e. the sample now covers the period from 1980q1 to 2015q4.

The estimated income coefficients (the β_{y_j}'s) are robust to extending the sample. Furthermore, the adjustment parameter in the consumption function is negative in most areas and resembles the results from the Great Moderation sample. The distribution of the adjustment parameter in the income equation is also similar to the Great Moderation sample.[12]

Overall, the inclusion of the financial crisis period in the estimation sample does not alter our main conclusions. If anything, our results are strengthened when the sample is extended.

To formally explore how the inclusion of the financial crisis period affects the tests for weak exogeneity and Granger non-causality, Table 5 reports average results across the four major census regions.[13]

There are several interesting observations in Table 5. First, the average number of dummies retained by the IIS algorithm (confer the final column) increases slightly compared to the Great Moderation sample. Second, the main results regarding weak exogeneity and Granger causality are maintained – in fact results are further strengthened when the financial crisis period is included, i.e., the rejection of the weak implication of the PIH is stronger when we include the financial crisis period. Finally, the evidence that house prices Granger-cause consumption is also stronger than we found for the Great Moderation sample. This is consistent with the view that the fall in house prices during the subprime crisis led to increased saving by US consumers to counteract the negative impact on their accu-

[12] Detailed results are reported in section C of the Online Appendix. Also for this extended sample do we find that the null cannot be rejected for the cases where the point estimates suggest $\alpha_c > 0$ and $\alpha_y > 0$.

[13] Detailed results for the individual MSAs are available upon request.

TABLE 5

Tests for weak exogeneity and Granger non-causality for full sample (1980q1–2015q4)

Region	*VAR in levels:*		*VECM:*					
	$y \underset{GC}{\rightarrow} c$	$c \underset{GC}{\rightarrow} y$	$\alpha_c \neq 0$	$\alpha_y \neq 0$	$y \underset{GC}{\rightarrow} c$	$c \underset{GC}{\rightarrow} y$	$ph \underset{GC}{\rightarrow} c$	$ph \underset{GC}{\rightarrow} y$
West	100.00	100.00	91.30	8.70	86.96	69.57	69.57	56.52
East	65.00	85.00	100.00	5.00	95.00	55.00	90.00	35.00
Midwest	92.00	88.00	100.00	12.00	96.00	44.00	76.00	60.00
South	100.00	96.55	89.66	44.83	96.55	68.97	79.31	51.72
All	90.72	92.78	94.85	19.59	93.81	59.79	78.35	51.55

Notes: Columns 4–7 report the percentage number of times where weak exogeneity of consumption ($\alpha_c \neq 0$) is rejected and the percentage number of times where weak exogeneity of income ($\alpha_y \neq 0$) is rejected, as well as the percentage number of times where we find that income Granger-causes consumption ($y \underset{GC}{\rightarrow} c$) and vice versa ($c \underset{GC}{\rightarrow} y$). Columns 8–9 report the percentage number of times where we find that house prices Granger-cause consumption ($ph \underset{GC}{\rightarrow} c$) and income ($ph \underset{GC}{\rightarrow} y$).

mulated wealth of the housing crash, i.e. that there are sizeable housing wealth effects on consumption, see also Carroll *et al.* (2012) and Mian *et al.* (2013).[14]

VI. Robustness and extensions

Panel estimation

Although the hypothesis of slope homogeneity is rejected empirically, we have also tested the robustness of our results to applying a standard panel estimator. First, we tested for unit roots in the panel using the test suggested by Breitung (2001) and Breitung and Das (2005), allowing for cross-sectional dependence. Results strongly suggest that both series are I(1), corroborating our MSA-by-MSA results. Second, we tested for cointegration using the Westerlund (2005) approach. We partitioned the panel in the four different Census regions. In all cases we included MSA-fixed effects. The models were estimated both on the Great Moderation sample and the full sample. Results are very similar to those obtained from the MSA-by-MSA analysis and the null of no cointegration is clearly rejected for all regions, see Table D.1 in section D of the Online Appendix.

To explore how our results on WE and GNC are affected by applying a panel approach, we formulated a panel error-correction model based on the cointegration results and tested for WE and GNC. The main conclusion that income Granger causes consumption is maintained in the panel analysis, i.e., the rejection of the rainy day hypothesis carries over to the panel. Detailed results are shown in Table D.2 and Table D.3 in section D of the Online Appendix.

[14] Household expectations might have played a similar role during the Great Depression period. Romer (1990) argues that households' perception of future income uncertainty increased significantly after the crash in the stock market in 1929, which led to the postponement of durable goods purchases.

Macro evidence

Macro time series of private income and consumption have features similar to the typical MSA series in that there are clear signs of both unit-root non-stationarity and intermittent structural breaks. To investigate if aggregate cointegration evidence and Granger causality is congruent with the picture that emerged from the analysis of the MSA-data, we redid the analysis using aggregate data for consumption and income.

Consistent with the average MSA results, we found evidence of cointegration. The zero restriction on the trend coefficient in the cointegration relationship is not rejected, which is similar to the analysis of the MSA data, where the co-trending restriction was accepted in about 70% of the MSAs.

The results strongly indicate that equilibrium correction is just as significant in consumption as it is in income. The macro results are robust to using data on total retail sales instead of personal consumption expenditures. In general, the aggregate analysis is in line with the MSA evidence, and a detailed description of the aggregate analysis is given in section E of the Online Appendix.

Using consumption data for shorter sample

Disaggregate consumption data are available for a much shorter time period than retail sales – starting only in 1997q4. That said, we have tested the robustness of our results to using consumption data. Results on cointegration rank remains robust, and the income coefficient moves closer to one. Our results still suggest that we reject weak exogeneity of consumption, and the results strongly favour that income Granger causes consumption. Hence, our key result is maintained in this setting. This holds true both for the Great Moderation sample and the full sample. These results are reported in section F of the Online Appendix.

Controlling for asset prices

Our results suggest that house prices play an important role for consumption dynamics and that the link between house prices and consumption has become stronger after the recent financial crisis. Housing wealth is one the main components of household wealth, which might directly influence their consumption and saving decisions. For instance, Carroll, Dynan and Krane (2003) do not find precautionary responses when they exclude home equity from household wealth. To analyse the link between wealth and consumption in a bit more detail, we consider a version of our model, where we extend the model to also include the real S&P500 index to control for financial wealth.

Our finding of a strong link between house prices and consumption, which has increased after the financial crisis, is maintained also in this case. Furthermore, we also establish a link between stock prices and consumption, but the evidence for this link is weaker than for the link between house prices and consumption. The finding that income Granger causes consumption is maintained in this extension. Detailed results are reported in section G of the Online Appendix.

Controlling for credit growth

In the presence of liquidity constraints, the sensitivity of consumption to income may decrease in strength, see e.g. Ludvigson (1999). For this reason, we investigate the robustness of our results to controlling for credit growth in the empirical analysis. MSA-level data for credit growth are only available from 1990q1, which restricts our sample somewhat. The inclusion of credit growth (and therefore also the reduction of the sample) has no material impact on our results and we still find Granger causality from income to consumption in a majority of the MSAs. Detailed results are reported in section H of the Online Appendix.[15]

Also in the case where we control for credit growth, the role of house prices for consumption dynamics is found to increase after the financial crisis. A link between consumption growth and credit is established in about 30% of the areas and this link has been relatively stable across the two samples.

VII. Conclusion

This paper has tested the so-called weak implication of the permanent income hypothesis, which entails that consumption growth does not respond to deviations from a long-run relationship between income and consumption, using metro level data for the US. The statistical implication of this is that consumption is weakly exogenous with respect to any long-run cointegrating relationship that exists between income and consumption. Our econometric analysis on the Great Moderation sample (1980q1–2007q4) gives little support for this hypothesis, and indicate that consumption responds to deviations from the long-run cointegrating relationship between income and consumption in a majority of the areas. Including the financial crisis period in the estimation sample, this result is strengthened, and the same is true for the results from the aggregate time series.

The VAR models that we use for testing include lagged growth rates in real house prices. In the MSA models we find significant effects of these conditioning variables, first on the 1980q1–2007q4 sample and even stronger effects when the financial crisis and Great Recession is included. On both samples, the overall direction of the effect is that lagged house price changes are positively related to consumption growth. The macro model corroborated the existence of such a relationship. Our finding therefore suggests that the large declines in housing equity in the aftermath of the subprime crash have strongly dampened consumer spending in the US. A similar conclusion is reached by Aron *et al.* (2012), Carroll *et al.* (2012) and Mian *et al.* (2013).

Final Manuscript Received: April 2019

References

Alan, S., Crossley, T. and Low, H. (2012). *Saving on a Rainy Day, Borrowing for a Rainy Day*, IFS Working Papers No. 11.

Aron, J., Duca, J. V., Muellbauer, J., Murata, K. and Murphy, A. (2012). 'Credit, housing collateral and consumption: Evidence from the UK, Japan and the US', *Review of Income and Wealth*, Vol. 58, pp. 397–423.

[15] Since credit data are only available from 1990q1, we also redid the original analysis on a similar sample and results are robust to considering this alternative sample period.

Attanasio, O. (1999). 'Consumption', in Taylor J. and Woodford M. (eds). *Handbook in Macroeconomics*, Vol. 1B, Chapter 11, Amsterdam: Elsevier, pp. 741–812.
Attanasio, O. P. and Weber, G. (2010). 'Consumption and saving: Models of intertemporal allocation and their implications for public policy', *Journal of Economic Literature*, Vol. 48, pp. 693–751.
Autor, D. H., Dorn, D. and Hanson, G. H. (2013). 'The China syndrome: Local labor market effects of import competition in the united states', *American Economic Review*, Vol. 103, pp. 2121–2168.
Beraja, M., Hurst, E. and Ospina, J. (2016). *The Aggregate Implications of Regional Business Cycles*, NBER Working Paper No. 21956.
Breitung, J. (2000). 'The local power of some unit root tests for panel data', in Baltagi B. H., Fomby T. B. and Hill R. C. (eds). *Non-Stationary Panels, Panel Cointegration and Dynamic Panels, Volume 15 of Advances in Econometrics*, Bingley, UK: Emerald Group Publishing Limited, pp. 161–177.
Breitung, J. and Das, S. (2005). 'Panel unit root tests under cross sectional dependence', *Statistica Neerlandica*, Vol. 59, pp. 414–433.
Campbell, J. Y. (1987). 'Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis', *Econometrica*, Vol. 55, pp. 1249–1273.
Campbell, J. Y. and Mankiw, N. G. (1989). 'Consumption, income and interest rates: Reinterpreting the time series evidence', in Blanchard O. J. and Fisher S. (eds). *NBER Macroeconomics Annual 1989*, Cambridge, Mass: MIT Press, pp. 185–216.
Carroll, C. (1992). 'The buffer-stock theory of saving: Some macroeconomic evidence', *Brookings Papers on Economic Activity*, Vol. 23, pp. 61–156.
Carroll, C. (2009). 'Precautionary saving and the marginal propensity to consume out of permanent income', *Journal of Monetary Economics*, Vol. 56, pp. 780–790.
Carroll, C., Dynan, K. and Krane, S. (2003). 'Unemployment risk and precautionary wealth: Evidence from households' balance sheets', *The Review of Economics and Statistics*, Vol. 85, pp. 586–604.
Carroll, C., Slacalek, J. and Sommer, M. (2012). *Dissecting Saving Dynamics: Measuring Credit, Wealth and Precautionary Effects*, IMF Working Paper No. 12/219.
Case, K. E., Quigley, J. M. and Shiller, R. J. (2012). 'Wealth effects revisited 1975–2012', *Critical Finance Review*, Vol. 2, pp. 101–128.
Castle, J. L., Doornik, J. A. and Hendry, D. F. (2012). 'Model selection when there are multiple breaks', *Journal of Econometrics*, Vol. 169, pp. 239–246.
Davidson, J. E. H., Hendry, D. F., Srba, F. and Yeo, S. (1978). 'Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom', *Economic Journal*, Vol. 88, pp. 661–692.
Dejuan, J. P., Seater, J. J. and Wirjanto, T. S. (2004). 'A direct test of the permanent income hypothesis with an application to the U.S. states', *Journal of Money, Credit and Banking*, Vol. 36, pp. 1091–1103.
DeLong, J. and Summers, L. (2012). 'Fiscal policy in a depressed economy', *Brookings Papers on Economic Activity*, Vol. 2012, pp. 233–274.
Dickey, D. A. and Fuller, W. A. (1979). 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association*, Vol. 74, pp. 427–431.
Doornik, J. (2009). 'Autometrics', in Castle J. L. and Shephard N. (eds). *The Methodology and Practice of Econometrics*, Oxford: Oxford University Press. pp. 88–121.
Eberly, J. (2014). 'Defining the reemerging role of fiscal policy', in Akerlof G., Blanchard O., Roemer D. and Stiglitz J. (eds.), *What Have We Learned? Macroeconomic Policy after the Crisis*, Chapter 14. Cambridge MA: MIT Press. pp. 179–182.
Eggertsson, G. B. and Krugman, P. (2012). 'Debt, deleveraging, and the liquidity trap: A Fisher-Minsky-Koo approach', *Quarterly Journal of Economics*, Vol. 127, pp. 1469–1513.
Eitrheim, Ø., Jansen, E. S. and Nymoen, R. (2002). 'Progress from forecast failure: The Norwegian consumption function', *Econometrics Journal*, Vol. 5, pp. 40–64.
Elliot, G., Rothenberg, T. J. and Stock, J. H. (1996). 'Tests for an autoregressive unit root', *Econometrica*, Vol. 64, pp. 813–836.
Engle, R. F. and Granger, C. W. J. (1987). 'Co-integration and error correction: Representation, estimation and testing', *Econometrica*, Vol. 55, pp. 251–276.
Engle, R. F., Hendry, D. F. and Richard, J.-F. (1983). 'Exogeneity', *Econometrica*, Vol. 51, pp. 277–304.

Flavin, M. A. (1981). ‘The adjustment of consumption to changing expectations about future income’, *Journal of Political Economy*, Vol. 89, pp. 974–1009.

Granger, C. W. J. (1986). ‘Developments in the study of cointegrated economic variables’, *Oxford Bulletin of Economics and Statistics*, Vol. 48, pp. 213–228.

Guerrieri, V. and Lorenzoni, G. (2011). *Credit Crises, Precautionary Savings, and the Liquidity Trap*. NBER Working Papers No. 17583, National Bureau of Economic Research, Inc.

Hall, R. E. (1978). ‘Stochastic implications of the life cycle permanent income hypothesis’, *Journal of Political Economy*, Vol. 86, pp. 971–987.

Hall, R. E. (2011). ‘The long slump’, *American Economic Review*, Vol. 101, pp. 461–469.

Harbo, I., Johansen, S., Nielsen, B. and Rahbek, A. (1998). ‘Asymptotic inference on cointegrating rank in partial systems’, *Journal of Business and Economic Statistics*, Vol. 16, pp. 388–399.

Hendry, D. F. and Doornik, J. (2009). *Empirical Econometric Modelling Using PcGive: Volume I*, Timberlake Consultant Press, London.

Hendry, D. F., Johansen, S. and Santos, C. (2008). ‘Automatic selection of indicators in a fully saturated regression’, *Computational Statistics*, Vol. 33, pp. 317–335.

Im, K. S., Pesaran, M. H. and Shin, Y. (2003). ‘Testing for unit roots in heterogeneous panels’, *Journal of Econometrics*, Vol. 115, pp. 53–74.

Jappelli, T. and Pistaferri, L. (2010). ‘The consumption response to income changes’, *Annual Review of Economics*, Vol. 2, pp. 479–506.

Johansen, S. (1988). ‘Statistical analysis of cointegration vectors’, *Journal of Economic Dynamics and Control*, Vol. 12, pp. 231–254.

Johansen, S. (1992). ‘Cointegration in partial systems and the efficiency of single-equation analysis’, *Journal of Econometrics*, Vol. 52, pp. 389–402.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

Johansen, S. and Nielsen, B. (2009). ‘An analysis of the indicator saturation estimator as a robust regression estimator’, in Castle J. L. and Shepard N. (eds). *The Methodology and Practice of Econometrics*, Oxford: Oxford University Press. pp. 1–36.

Lettau, M. and Ludvigson, S. (2004). ‘Understanding trend and cycle in asset values: Re-evaluating the wealth effects on consumption’, *The American Economic Review*, Vol. 94, pp. 276–299.

Lettau, M. and Ludvigson, S. (2013). ‘Shocks and crashes’, *NBER Macroeconomics Annual*, Vol. 28, pp. 293–354.

Ljungqvist, L. and Sargent, T. (2004). *Recursive Macroeconoimic Theory*, 2nd edn, MIT Press, Cambridge MA.

Ludvigson, S. (1999). ‘Consumption and credit: A model of time-varying liquidity constraints’, *Review of Economics and Statistics*, Vol. 81, pp. 434–447.

Martin, P. and Philippon, T. (2017). ‘Inspecting the mechanism: Leverage and the great recession in the eurozone’, *American Economic Review*, Vol. 107, pp. 1904–1937.

Mian, A. and Sufi, A. (2014). ‘What explains the 2007–2009 drop in employment? *Econometrica*, Vol. 82, pp. 2197–2223.

Mian, A., Rao, K. and Sufi, A. (2013). ‘Household balance sheets, consumption, and the economic slump’, *Quarterly Journal of Economics*, Vol. 128, pp. 1687–1726.

Muellbauer, J. (2016). *Macroeconomics and Consumption*, CEPR Working Paper No. 11588.

Muellbauer, J. and Lattimore, R. (1995). ‘The consumption function: A theoretical and emprical overview’, in M.H. P. and Wickens M. (eds). *Handbook of Applied Economterics*, Volume 1: Macroeconomics, Chapter 5, Oxford: Blackwell, pp. 221–311.

Nakamura, E. and Steinsson, J. (2014). ‘Fiscal stimulus in a monetary union: Evidence from US regions’, *American Economic Review*, Vol. 104, pp. 753–792.

Nakamura, E. and Steinsson, J. (2018). ‘Identification in macroeconomics’, *Journal of Economic Perspectives*, Vol. 32, pp. 59–86.

Parker, J. A. (2000). ‘Spendthrift in America? On two decades of decline in the U.S. saving rate’, in Bernanke B. S. and Rotemberg J. J. (eds). NBER Macroeconomics Annual 1999, vol. 14, Cambridge, MA: MIT Press. pp. 317–387.

Pesaran, M. H. and Smith, R. (1995). 'Estimating long-run relationships from dynamic heterogeneous panels', *Journal of Econometrics*, Vol. 68, pp. 79–113.

Pesaran, M. H., Shin, Y. and Smith, R. P. (1999). 'Pooled mean group estimation of dynamic heterogeneous panels', *Journal of the American Statistical Association*, Vol. 94, pp. 621–634.

Phillips, P. C. B. and Moon, H. R. (2000). 'Nonstationary panel data analysis: An overview of some recent developments', *Econometric Reviews*, Vol. 19, pp. 263–286.

Romer, C. D. (1990). 'The great crash and the onset of the great depression', *Quarterly Journal of Economics*, Vol. 105, pp. 597–624.

Romer, D. (2006). *Advanced Macroeconomics*, 3rd edn, New York: McGraw-Hill.

Sorensen, B. E. and Luengo-Prado, M. J. (2008). 'What can explain excess smoothness and sensitivity of state-level consumption? *Review of Economics and Statistics*, Vol. 90, 65–88.

Stiglitz, J. E. (2014). 'The lessons from the North Atlantic crisis for economic theory and policy', in Akerlof G., Blanchard O., Roemer D. and Stiglitz J. (eds). *What Have We Learned? Macroeconomic Policy after the Crisis.*, Chapter 29. Cambridge, MA: MIT Press, pp. 335–347.

Toda, H. Y. and Yamamoto, T. (1995). 'Statistical inference in vector autoregressions with possibly integrated processes', *Journal of Econometrics*, Vol. 66, pp. 225–250.

Westerlund, J. (2005). 'New simple tests for panel cointegration', *Econometric Reviews*, Vol. 24, pp. 297–316.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix A. Testing alternative consumption theories within a cointegrated VAR with first order dynamics.
Appendix B. Detailed econometric results by MSA (reported by Census region), Great Moderation sample.
Appendix C. Additional results from full sample analysis.
Appendix D. Results from standard panel analysis.
Appendix E. Macro evidence.
Appendix F. Results from using consumption data for shorter sample.
Appendix G. Controlling for growth in stock prices.
Appendix H. Controlling for credit growth.