

Gardberg, Malin; Pozzi, Lorenzo

**Working Paper**

## Consumption and wealth in the long run: an integrated unobserved component approach

Tinbergen Institute Discussion Paper, No. TI 2018-046/VI

**Provided in Cooperation with:**

Tinbergen Institute, Amsterdam and Rotterdam

*Suggested Citation:* Gardberg, Malin; Pozzi, Lorenzo (2018) : Consumption and wealth in the long run: an integrated unobserved component approach, Tinbergen Institute Discussion Paper, No. TI 2018-046/VI, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/185565>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

TI 2018-046/VI  
Tinbergen Institute Discussion Paper



# Consumption and wealth in the long run: an integrated unobserved component approach

**Revision: September 2018**

*Malin Gardberg<sup>1,2</sup>*

*Lorenzo (L.C.G.) Pozzi<sup>1</sup>*

<sup>1</sup>Erasmus University Rotterdam

<sup>2</sup>Research Institute of Industrial Economics (IFN)

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: [discussionpapers@tinbergen.nl](mailto:discussionpapers@tinbergen.nl)

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam  
Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900

# Consumption and wealth in the long run: an integrated unobserved component approach\*

Malin Gardberg<sup>1</sup> and Lorenzo Pozzi<sup>†2</sup>

<sup>1</sup>*Erasmus University Rotterdam & Tinbergen Institute & Research Institute of Industrial Economics*

<sup>2</sup>*Erasmus University Rotterdam & Tinbergen Institute*

September 2018

## Abstract

The ratio of consumption to total household wealth (i.e., tangible assets plus unobserved human wealth) is commonly calculated from the estimation of a log-linear version of the household intertemporal budget constraint as a cointegrating relationship between consumption, assets and earnings (i.e., the variable "cay"). The evidence in favor of a stable cointegrating relationship between these variables in the US is weak however. This paper follows an alternative empirical approach using an unobserved component model applied to US data over the period 1951Q4-2016Q4. The regression of consumption on assets and earnings is augmented with an unobserved stochastic trend, i.e., an integrated component. The results strongly support the presence of such an unobserved component in the consumption equation. We provide evidence that this component is related to financial liberalization which, by relaxing liquidity constraints of consumers, has permanently increased the consumption-to-wealth ratio over the sample period. We calculate an alternative "cay" variable, i.e., the stationary part of the consumption-to-wealth ratio, and find that its predictive ability for future (excess) stock returns is comparable to that of the standard "cay" variable.

**JEL Classification:** E21, C32, C11

**Keywords:** consumption, wealth, cay, financial liberalization, stationarity, excess returns, unobserved component, Bayesian

---

\*For constructive comments and suggestions, we thank Gerdie Everaert and participants of the following seminars and conferences: econometrics and economics seminars at the Erasmus School of Economics (Rotterdam, Netherlands), the 2018 workshop on financial econometrics and empirical modeling of financial markets at Kiel Institute for the World Economy (Kiel, Germany), the 2018 International Association for Applied Econometrics conference (Montreal, Canada), and the 2018 European summer meeting of the Econometric Society (Cologne, Germany).

<sup>†</sup>Corresponding author at Department of Economics, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. Email: [pozzi@ese.eur.nl](mailto:pozzi@ese.eur.nl). Website: <http://people.few.eur.nl/pozzi>.

# 1 Introduction

The study of the long run aggregate relationship between consumption and total household wealth (i.e., the sum of tangible assets and unobserved human wealth) serves a twofold purpose. First, it is central to the estimation of long run elasticities of consumption to wealth and its components. These elasticities are useful, among other things, to evaluate and predict the implications of wealth changes for economic growth. Second, it allows for the estimation of the unobserved ratio of consumption to total wealth. According to the aggregate intertemporal household budget constraint, the consumption-to-wealth ratio contains information about future variables, in particular future consumption changes and future returns to wealth.

The contemporaneous literature on the long run aggregate relationship between consumption and wealth originates from Lettau and Ludvigson (2001, 2004) who propose a cointegration approach to proxy the unobserved consumption to total wealth ratio. To this end, they regress log consumption on log assets and log earnings (as a proxy for human wealth) using US data and argue in favor of a stable cointegration relationship between these variables. From the stationary regression residuals, they calculate a proxy for the consumption to total wealth ratio (i.e, the variable "cay") and find that it has strong predictive power for (excess) stock returns. Rudd and Whelan (2006) however find no cointegration in the US upon adjusting the data to make it more consistent with the underlying intertemporal household budget constraint, i.e., they argue that "cay" is not stationary.<sup>1</sup> Their finding suggests that results obtained from the cointegration approach to measure wealth effects on consumption and to predict stock returns using the estimated consumption-wealth ratio "cay" are spurious.<sup>2</sup> In a recent contribution, Bianchi et al. (2017) acknowledge these issues and - after adjusting the data along the lines of Rudd and Whelan (2006) and failing to find stationarity for the standard "cay" measure - propose a new proxy for the consumption-wealth ratio obtained from a regression that includes a two-state Markov switching intercept. This allows for regime shifts in the consumption to wealth ratio which they link to the US Federal Reserve's monetary policy. They argue that the consumption-wealth ratio obtained from this approach is stationary and has stronger predictive ability for (excess) stock returns when compared to the traditional "cay" variable.

This paper contributes to the literature by estimating the long run consumption-wealth relationship using an alternative empirical approach. In particular, we propose an unobserved component model applied to US data over the period 1951Q4-2016Q4. The regression equation of log consumption on log

---

<sup>1</sup>As far as international evidence is concerned, Slacalek (2004) fails to find evidence in favor of a stable cointegrating relationship between consumption, earnings and assets in a sample of 26 countries.

<sup>2</sup>Given considerations like absence of cointegration and instability of the cointegrating vector, Carroll et al. (2011) argue against cointegration methods to measure wealth effects and instead propose an indirect approach based on the consumption Euler equation and stickiness in consumption growth.

assets and log earnings is augmented with an integrated unobserved component (see e.g., Harvey et al., 1986; Canarella et al., 1990; Sarantis and Stewart, 2001; Planas et al., 2007; Everaert, 2010, for the inclusion of such a component in regressions conducted in a different context). This inclusion allows us to reliably estimate the long run relationship between consumption and total household wealth, even though consumption, assets and earnings are not cointegrated. Our unobserved component model is estimated using Bayesian state space methods with model selection along the lines of Frühwirth-Schnatter and Wagner (2010). This allows for the calculation of the posterior probability that an integrated unobserved component is present in the regression equation. Posterior distributions are calculated for the elasticities of consumption to assets and earnings and these elasticities are compared to elasticities obtained for models without an integrated unobserved component. An alternative "cay" variable - i.e., the stationary part of the consumption to wealth ratio - is calculated from this set-up and its persistence and predictive ability for (excess) stock returns are evaluated and compared to that of the standard "cay" variable.

The paper contains the following results. Using standard frequentist cointegration tests, we confirm that the evidence in favour of a stable cointegrating relationship between consumption, earnings and assets in the US is very weak. Our Bayesian unobserved component approach with model selection strongly supports the presence of an unobserved integrated component in the consumption equation. We argue that a model with an integrated unobserved component added to the regression is preferred over alternative options to deal with the non-stationarity of the consumption-to-wealth ratio such as the modelling of regime switches in its mean. With respect to the estimates of the elasticities of consumption to assets and earnings we find, in line with the previous literature, that asset wealth has a much smaller impact on consumption than earnings (as a proxy for human wealth), and that non-stock asset wealth (i.e., housing wealth) is more important for consumption than stock assets (see e.g., Davis and Palumbo, 2001; Carroll et al., 2011). Importantly, we find that the estimated elasticities for assets (and their components) tend to be overestimated when we wrongfully exclude the unobserved integrated component from the model. Hence, the failure to appropriately account for the non-stationarity in the consumption equation leads to elasticity estimates that are misleading. We further provide an interpretation for the unobserved integrated component and test this interpretation. In particular, we argue that financial liberalization has, by relaxing liquidity constraints of consumers, permanently increased the consumption-to-wealth ratio over the sample period. Finally, we evaluate both the in-sample and out-of-sample predictive ability for excess stock returns of our new "cay" variable, i.e., the stationary part of the consumption to wealth ratio. In-sample, we find that its predictive ability for future excess stock returns, while diminished compared to that of the traditional "cay" variable, is statistically and economically significant and in line with the univariate predictability results typically reported in the literature for excess stock returns using other

predictors (see e.g., Ferson et al., 2003, for an overview). Out-of-sample, we find that its predictive ability for future excess stock returns is comparable to that of the standard "cay" variable.

The structure of the paper is as follows. Section 2 derives the estimable equations that relate consumption to earnings and asset wealth (and its components). Section 3 provides a brief description of the data while Section 4 provides evidence on cointegration between consumption, assets and earnings based on standard frequentist cointegration tests. Section 5 presents and discusses our unobserved component approach and the results obtained using this method. A theoretical interpretation of the estimated integrated unobserved component is presented and tested and the predictive ability for excess stock returns of our alternative "cay" is discussed. Finally, Section 6 concludes.

## 2 Theoretical framework

This section derives two estimable equations that relate consumption to earnings and assets and to earnings and asset components (i.e., stock wealth and non-stock wealth). The derivations are largely based on the intertemporal budget constraint framework considered initially by Campbell and Mankiw (1989) and then explored further by Lettau and Ludvigson (2001, 2004). Note that linearization constants are ignored throughout the section.

The per period budget constraint of a representative consumer is given by,

$$\frac{A_{t+1}}{1+r_{t+1}} = A_t + Y_t - C_t \quad (1)$$

where  $A_t$  are real assets,  $Y_t$  is real disposable labor income (earnings),  $C_t$  is real consumption and  $r_t$  is the real rate of return. Alternatively, write

$$\frac{W_{t+1}}{1+r_{t+1}} = W_t - C_t \quad (2)$$

where  $W_t \equiv A_t + H_t$  is total wealth and  $H_t$  is human wealth.

To reconcile both eqs.(1) and (2), note that

$$H_t = Y_t + \frac{H_{t+1}}{1+r_{t+1}} \quad (3)$$

which is consistent with the definition of human wealth as the presented discounted value of future earnings.<sup>3</sup>

After log-linearizing eq.(2) and solving the resulting log-linearized per period constraint forward, we

---

<sup>3</sup>For ease of exposition, we do not decompose the return on total wealth into a return on human wealth and a return on asset wealth hence effectively assuming these returns are equal. The framework can be easily extended to incorporate this distinction but this does not offer additional insight.

can write

$$c_t - w_t = E_t \sum_{j=1}^{\infty} \rho^j (r_{t+j} - \Delta c_{t+j}) + \lim_{k \rightarrow \infty} \rho^k E_t (c_{t+k} - w_{t+k}) \quad (4)$$

where  $c_t$  is the log of real consumption  $C_t$ ,  $w_t$  is the log of total real wealth  $W_t$ ,  $E_t$  is the expectation operator conditional on period  $t$  information and  $\rho$  is the discount factor ( $0 < \rho < 1$ ).<sup>4</sup> The transversality condition  $\lim_{k \rightarrow \infty} \rho^k E_t (c_{t+k} - w_{t+k}) = 0$  is imposed so that

$$c_t = w_t + n_t \quad (5)$$

where  $n_t = E_t \sum_{j=1}^{\infty} \rho^j (r_{t+j} - \Delta c_{t+j})$ .

We then log-linearize  $W_t \equiv A_t + H_t$  to obtain  $w_t = \alpha a_t + \beta h_t$  so that eq.(5) becomes

$$c_t = \alpha a_t + \beta h_t + n_t \quad (6)$$

where  $a_t$  is the log of assets  $A_t$ ,  $h_t$  is the log of human wealth  $H_t$ ,  $\alpha$  is the elasticity of consumption with respect to assets and  $\beta$  is the elasticity of consumption with respect to human wealth.<sup>5</sup>

Since log human wealth  $h_t$  is unobserved, we log-linearize eq.(3), solve the resulting expression forward and impose a transversality condition to obtain

$$h_t = y_t + m_t \quad (7)$$

where  $m_t = E_t \sum_{j=1}^{\infty} \bar{\rho}^j (\Delta y_{t+j} - r_{t+j})$  with discount factor  $\bar{\rho}$  ( $0 < \bar{\rho} < 1$ ).<sup>6</sup>

Substituting eq.(7) into eq.(6) then gives the baseline equation,

$$c_t = \alpha a_t + \beta y_t + z_t \quad (8)$$

with  $z_t = n_t + \beta m_t$ .

If we decompose assets  $A_t$  into stock wealth  $A_t^s$  (financial wealth) and non-stock wealth  $A_t^{ns}$  (mainly housing wealth), we have  $W_t \equiv A_t^s + A_t^{ns} + H_t$  and we can derive the extended equation,

$$c_t = \alpha^s a_t^s + \alpha^{ns} a_t^{ns} + \beta y_t + z_t \quad (9)$$

where  $a_t^s$  is the log of stock wealth  $A_t^s$ ,  $a_t^{ns}$  is the log of non-stock wealth  $A_t^{ns}$ ,  $\alpha^s$  is the elasticity of consumption with respect to stock wealth and  $\alpha^{ns}$  is the elasticity of consumption with respect to non-stock wealth.<sup>7</sup>

<sup>4</sup>In particular,  $\rho = 1 - \frac{C}{W}$  where  $C$  and  $W$  are the steady state values of  $C_t$  and  $W_t$  around which we linearize.

<sup>5</sup>We note that  $\alpha = \frac{A}{W}$  and  $\beta = \frac{H}{W}$  where  $A$ ,  $H$  and  $W$  are the steady state values of  $A_t$ ,  $H_t$  and  $W_t$  around which we linearize.

<sup>6</sup>In particular,  $\bar{\rho} = 1 - \frac{Y}{H}$  where  $Y$  and  $H$  are the steady state values of  $Y_t$  and  $H_t$  around which we linearize.

<sup>7</sup>We note that  $\alpha^s = \frac{A^s}{W}$  and  $\alpha^{ns} = \frac{A^{ns}}{W}$  where  $A^s$ ,  $A^{ns}$  and  $W$  are the steady state values of  $A_t^s$ ,  $A_t^{ns}$  and  $W_t$  around which we linearize.



This set-up allows for the estimation of the elasticities of consumption to assets and to asset components, but also for the construction of a proxy for the consumption to total wealth ratio  $c_t - w_t$ , i.e., the estimated variable  $z_t$  - which in the case of eq.(8) - is commonly denoted by "cay". The ways to estimate these equations are discussed in detail in Sections 4 and 5 after a description of the data is first given in the next section.

### 3 Data

For the estimation of eqs.(8) and (9) we use quarterly US data over the period 1951Q4 – 2016Q4. A detailed description of the data can be found in Appendix A. Our dataset is constructed identically as the dataset considered by Lettau and Ludvigson (2015). For  $c_t$  we use the log of real per capita total personal consumption expenditures.<sup>8</sup> For  $y_t$  we use the log of real per capita disposable labor income. For  $a_t$  we use the log of real per capita household total net worth which includes stock and non-stock financial assets, housing wealth and durable goods.<sup>9</sup> For  $a_t^s$  we use the log of real per capita stock market assets while for  $a_t^{ns}$  we use the log of real per capita non-stock assets (i.e., real estate, non-stock financial assets, durable goods) minus liabilities. The asset variables  $a_t$ ,  $a_t^s$  and  $a_t^{ns}$  are all measured at the end of the period.

### 4 Cointegration

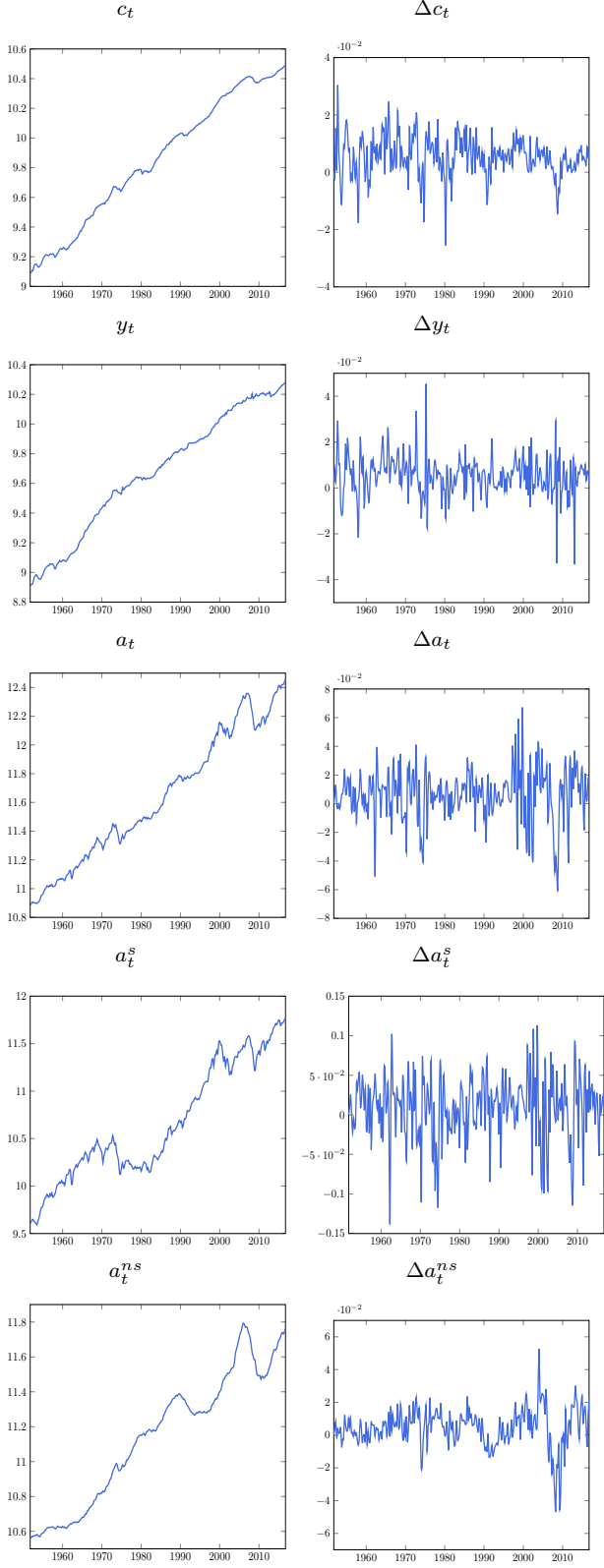
This section investigates whether regression eqs.(8) and (9) are cointegrating relationships. First, we briefly discuss the integratedness of the variables included in these equations. Next, we conduct a standard frequentist cointegration analysis and discuss its results.

---

<sup>8</sup>Lettau and Ludvigson (2001, 2004) originally used expenditures on nondurable goods and services as a proxy for consumption but they recently switched to the use of total personal consumption expenditures (see Lettau and Ludvigson, 2015).

<sup>9</sup>Rudd and Whelan (2006) argue in favor of excluding durable goods from assets as  $c_t$  already includes expenditures on durable goods. However, as durable goods are not fully consumed in the period of purchase, we follow Lettau and Ludvigson (2004, 2015) and do not exclude them from net asset wealth  $a_t$ . Our results are robust to using a definition of wealth that excludes durable goods however. These results are not reported but are available from the authors upon request.

**Figure 1:** Real per capita personal consumption expenditures ( $c$ ), disposable labor income ( $y$ ), net assets ( $a$ ), net stock wealth ( $a^s$ ) and net non-stock wealth ( $a^{n.s}$ ) - US data 1951Q4 – 2016Q4.



## 4.1 Integrated variables

Figure 1 presents the variables  $c_t$ ,  $y_t$ ,  $a_t$ ,  $a_t^s$  and  $a_t^{ns}$  (logs) and  $\Delta c_t$ ,  $\Delta y_t$ ,  $\Delta a_t$ ,  $\Delta a_t^s$  and  $\Delta a_t^{ns}$  (growth rates) used in our analysis. From the figure, we note that the variables  $c_t$ ,  $y_t$ ,  $a_t$ ,  $a_t^s$  and  $a_t^{ns}$  are stochastically trended. In particular, unreported (augmented) Dickey-Fuller tests suggest that they are integrated of order one -  $I(1)$  - and become integrated of order zero -  $I(0)$  - upon first differencing, i.e., implying that the growth rates  $\Delta c_t$ ,  $\Delta y_t$ ,  $\Delta a_t$ ,  $\Delta a_t^s$  and  $\Delta a_t^{ns}$  are stationary. With respect to the stationarity of these growth rates, it should be noted however that Dickey-Fuller tests may fail to detect a unit root if it underlies a volatile high frequency component. As such, growth rates may *appear* stationary rather than be stationary. We argue below that, due to financial liberalization, aggregate consumption growth  $\Delta c_t$  contains a slow moving low frequency stochastic trend underneath its volatile high frequency component.

## 4.2 Cointegration

If the regressions given by eqs.(8) and (9) are believed to consist of integrated  $I(1)$  variables, then the standard approach to estimate the elasticities  $\alpha$ ,  $\beta$ ,  $\alpha^s$  and  $\alpha^{ns}$  is to estimate these equations as cointegrating regressions (see Engle and Granger, 1987). For this approach to be valid, the term  $z_t$  in both equations should be stationary, i.e., the variables  $c_t$ ,  $y_t$  and  $a_t$  in eq.(8) and the variables  $c_t$ ,  $y_t$ ,  $a_t^s$  and  $a_t^{ns}$  in eq.(9) must be cointegrated. In Section 4.2.1, we use standard frequentist cointegration tests to find out whether  $z_t$  is stationary. In Section 4.2.2, we provide potential reasons why  $z_t$  is found to be non-stationary. In Section 4.2.3, we discuss the implications of a non-stationary  $z_t$ .

### 4.2.1 Testing whether $z_t$ is stationary

**Table 1:** Engle-Granger cointegration test between  $c$ ,  $a$  and  $y$  and between  $c$ ,  $a^s$ ,  $a^{ns}$  and  $y$

$c, a, y$		$c, a^s, a^{ns}, y$	
1951Q4 – 2016Q4	1951Q4 – 2007Q3	1951Q4 – 2016Q4	1951Q4 – 2007Q3
Dickey-Fuller t-statistic			
-2.32	-2.57	-2.36	-2.47
5% and 10% critical values (3 variables)		5% and 10% critical values (4 variables)	
5%	-3.74	5%	-4.10
10%	-3.45	10%	-3.81

Notes: The null hypothesis tested is the no cointegration hypothesis. A constant is included in the cointegrating equation. The 5% and 10% critical values are taken from MacKinnon (2010). The number of lags included in the augmented Dickey-Fuller regression is based on the Bayesian information criterion.

To check whether  $z_t$  is stationary or not, we conduct a standard Engle-Granger cointegration test (see Engle and Granger, 1987) and report the results in Table 1. From the table, we note that there is very little evidence in favor of cointegration between the variables  $c_t$ ,  $y_t$  and  $a_t$  or between the variables  $c_t$ ,  $y_t$ ,  $a_t^s$  and  $a_t^{ns}$ . This is the case for the full sample period (i.e., the period 1951Q4 – 2016Q4) as well as for the sample that excludes the years from the Great Recession onward (i.e., the period 1951Q4 – 2007Q3). We have also conducted a number of alternative cointegration tests, the results of which are reported in Appendix B. These tests also strongly suggest that there is no cointegration between the variables considered. As we argue in the appendix, the failure of finding cointegration does not seem to originate from statistical issues - i.e., the lack of power to reject a false null hypothesis - but rather suggests that there is in fact a unit root present in  $z_t$ .

#### 4.2.2 Reasons why $z_t$ is non-stationary

There are different potential reasons as to why eqs.(8) and (9) are not cointegrating regressions. First, the stationarity of  $z_t$  requires the stationarity of the term  $n_t$  in eq.(5). Since the result in eq.(5) is obtained after imposing a transversality condition, a violation of this condition renders  $n_t$  in eq.(5) non-stationary. Additionally, the log-linearization conducted to derive eq.(5) pushes potentially non-stationary higher-order terms in  $n_t$ . The same arguments can be evoked when considering the stationarity of the term  $m_t$  in eq.(7), which also is a component of  $z_t$ . Second, the model presented in Section 2 assumes stability in the steady state ratios of the model variables and hence in the resulting elasticities. This assumption may not hold in practice.<sup>10</sup> Structural instability in the cointegrating equation may lead to an  $I(1)$  component in  $z_t$ .<sup>11</sup> Third, from the definition of the terms  $n_t$  in eq.(5) and  $m_t$  in eq.(7) these terms are stationary only if the variables  $r_t$ ,  $\Delta c_t$ , and  $\Delta y_t$  in the model are stationary. As argued by Bianchi et al. (2017), regime shifts in US monetary policy may - through  $r_t$  - have caused shifts in the mean of the consumption to wealth ratio, rendering it non-stationary. In this paper, we instead focus on aggregate consumption growth  $\Delta c_t$ . In what follows we argue that financial liberalization occurring over the sample period has rendered  $\Delta c_t$  non-stationary, i.e., underneath the volatile high frequency component which is apparent from Figure 1, there is a slow moving low frequency stochastic trend.

<sup>10</sup>For instance, different forms of heterogeneity at the micro level may imply structural instability in the aggregate relationship estimated using time series data (see Cooper, 2016, and references therein).

<sup>11</sup>To see this, note that if the true relationship between  $c_t$ ,  $a_t$  and  $y_t$  is given by  $c_t = \alpha_t a_t + \beta_t y_t + v_t$  where the time-varying parameters  $\alpha_t$  and  $\beta_t$  are given by  $\alpha_t = \alpha + \tilde{\alpha}_t$  and  $\beta_t = \beta + \tilde{\beta}_t$  - i.e., consisting of a constant and a time-varying component where the latter can be stationary or integrated - then estimating  $c_t = \alpha a_t + \beta y_t + z_t$  causes  $z_t$  to have an  $I(1)$  component since  $z_t = v_t + \tilde{\alpha}_t a_t + \tilde{\beta}_t y_t$ .

### 4.2.3 Implications of a non-stationary $z_t$

The estimation of the relationship given by eq.(8) or eq.(9) as a cointegrating relationship using standard methods like static OLS or dynamic OLS is spurious if  $z_t$  is non-stationary (see Granger and Newbold, 1974; Phillips, 1986). Highly significant OLS estimates could be obtained for the elasticities  $\alpha$  and  $\beta$  with high  $R^2$ 's even if  $c_t$  and the regressors  $a_t$  and  $y_t$  are completely independent merely because these variables are all stochastically trended. While it is unlikely that the variables in eqs.(8) and (9) are completely independent, the estimates for the elasticities obtained via standard cointegration analysis are nonetheless unreliable if  $z_t$  contains a unit root. Additionally, the non-stationarity of the variable  $z_t$  may invalidate its use as a predictor for other variables, in particular (excess) stock returns (see e.g., Rudd and Whelan, 2006). These issues are tackled in the next section where we discuss an alternative empirical approach based on an unobserved component model.

## 5 Unobserved component model

We can reliably estimate the long-run relationship between  $c_t$ ,  $a_t$  and  $y_t$  in eq.(8) or between  $c_t$ ,  $a_t^s$ ,  $a_t^{ns}$  and  $y_t$  in eq.(9) - even when these variables are not cointegrated - using an unobserved component or state space framework (see Harvey, 1989; Durbin and Koopman, 2001). In such a set-up, the omitted  $I(1)$  component in  $z_t$  can be explicitly added to the regression equation as an unobserved component and estimated together with the parameters - among which are the elasticities  $\alpha$ ,  $\alpha^s$ ,  $\alpha^{ns}$  and  $\beta$  - of the model (see e.g., Harvey et al., 1986; Canarella et al., 1990; Planas et al., 2007; Everaert, 2010). Section 5.1 presents the specification of the unobserved component model. Section 5.2 discusses the methodology used to estimate the model. Section 5.3 presents the main estimation results.

### 5.1 Empirical specification

#### 5.1.1 Set-up

We can write eqs.(8) and (9) in general form as,

$$c_t = x_t \phi + z_t \tag{10}$$

where  $x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$  and  $\phi = \begin{bmatrix} \alpha & \beta \end{bmatrix}'$  for eq.(8) or  $x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$  and  $\phi = \begin{bmatrix} \alpha^s & \alpha^{ns} & \beta \end{bmatrix}'$  for eq.(9).

The unobserved variable  $z_t$  is modelled as the sum of an  $I(1)$  non-stationary component  $\mu_t$  and an  $I(0)$  stationary component  $v_t$ . As such, we have,

$$z_t = \mu_t + v_t \tag{11}$$

The non-stationary component  $\mu_t$  is modelled as a random walk process, i.e.,

$$\mu_t = \mu_{t-1} + \eta_t \quad (12)$$

where  $\eta_t \sim iid\mathcal{N}(0, \sigma_\eta^2)$ .

Following the literature where dynamic OLS is typically applied to the estimation of regression equations between consumption, earnings and assets, the stationary component  $v_t$  is modelled as consisting of an error term  $\varepsilon_t$  and lags, leads and contemporaneous values of the first difference of the regressors  $x_t$ , i.e.,

$$v_t = \sum_{j=-p}^p \Delta x_{t+j} \rho_j + \varepsilon_t \quad (13)$$

where  $\varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$ .<sup>12</sup>

### 5.1.2 Discussion

The presence of an integrated unobserved component  $\mu_t$  in eq.(10) may have multiple causes.<sup>13</sup> As noted in Section 4.2.2 above, the reason for failing to find cointegration between the variables  $c_t$ ,  $y_t$  and  $a_t$  - or between the variables  $c_t$ ,  $y_t$ ,  $a_t^s$  and  $a_t^{ns}$  - may be due to a violation of the transversality condition, the neglect of potentially integrated higher-order terms in the conducted log-linearizations or structural instability in the steady state ratios of the model variables. And while Bianchi et al. (2017) argue that monetary policy shifts render the returns on wealth non-stationary, this paper focusses instead on financial liberalization and its impact on aggregate consumption growth (see Section 5.4). Of course, objections can be raised to giving an integrated variable or unit root interpretation to some of these failures of finding cointegration. For instance, violations of the transversality condition may be short-lived (i.e., bubbles tend to burst). However, for other failures this concern seems unwarranted. For example, the neglected higher-order terms in the log-linearizations are functions of the levels of the variables in the model and are therefore unquestionably integrated variables. Additionally, even if one objects to the principle of there being an integrated component in this context, this component may actually provide the best way to approximate the non-stationarity of  $z_t$  *within a given sample period*. Indeed, as we detail in Appendix C, using an alternative and more general modelling approach for the unobserved component suggests that the non-stationarity of  $z_t$  is best modelled through an integrated variable. More specifically, we find that the data supports a regression model with an integrated unobserved component included to capture non-stationarity rather than a regression model containing a Markov switching intercept along the lines

<sup>12</sup>The results reported in this section change very little if, additionally, lags of the first difference of the dependent variable  $c_t$  are added to the component  $v_t$ .

<sup>13</sup>Instead of an "unobserved component", the variable  $\mu_t$  could alternatively be denoted as a "time-varying intercept" without changing the implications of the results found in the paper.

of Bianchi et al. (2017). Before discussing the results of the estimation of the unobserved component model, the next section first elaborates on the methodology used.

## 5.2 Methodology

We first discuss the advantages of Bayesian estimation. Then, we present the Bayesian model selection approach to determine whether or not a non-stationary unobserved component is present in the regression. Finally, we discuss our choice of parameter priors. The general outline and technical details of the Gibbs sampling algorithm together with a convergence analysis of the sampler are provided in Appendix D.

### 5.2.1 Bayesian estimation

Using a standard classical approach to state space estimation, the Kalman filter could be applied to estimate the unobserved component and calculate the likelihood, where the latter would be maximized with respect to the model parameters using an iterative numerical procedure.<sup>14</sup> This paper instead follows a Bayesian approach. In particular, we use a Gibbs sampling approach which is a Markov Chain Monte Carlo (MCMC) method used to simulate draws from the intractable joint posterior distribution of the parameters and the unobserved state using only tractable conditional distributions. Our Bayesian approach has a number of advantages. First, as discussed in the next section, model selection allows for the calculation of the posterior probability that an integrated unobserved component is present in the regression, i.e., we calculate the probability of a model with integrated unobserved component versus a model without integrated unobserved component. Second, model averaging is possible so that we may calculate parameter estimates - in particular, elasticities - averaged across both models. In this way, valid parameter estimates can be obtained for any value of the posterior probability that a non-stationary unobserved component is present in the regression. This stands in contrast to the standard classical cointegration approach where valid parameter estimates can only be obtained under stationarity, i.e., when there is no integrated unobserved component in the regression error term. Third, the Gibbs sampler provides small sample posterior distributions of the parameters making it possible to avoid the use of asymptotic approximations to parameter distributions.<sup>15</sup>

---

<sup>14</sup>Under stationarity assumptions, the obtained maximum likelihood estimates are consistent and asymptotically Gaussian provided that the model is identified (see e.g, Hamilton, 1994). Chang et al. (2009) show that these asymptotic properties can be extended to state space models with integrated observed and latent variables.

<sup>15</sup>Further advantages compared to the standard classical maximum likelihood (ML) estimation approach to state space models are (i) the fact that Gibbs sampling is computationally easier to implement than ML estimation and, as such, does not suffer from the numerical optimization problems inherent to ML estimation (see Kim and Kim, 2011) and (ii) the fact that both the parameters and the unobserved state are treated as random variables in a Bayesian setting whereas the traditional ML approach treats the ML estimates of the parameters as if they are the true values hence neglecting parameter uncertainty when conducting inference on the unobserved state (see Kim and Nelson, 1999).

### 5.2.2 Model selection

We test whether to include or exclude the integrated unobserved component to the regression equation using the stochastic model selection approach for Bayesian state space models as developed by Frühwirth-Schnatter and Wagner (2010). Testing whether or not an integrated unobserved component  $\mu_t$  is present in eq.(10) amounts to testing  $\sigma_\eta^2 > 0$  against  $\sigma_\eta^2 = 0$ . This is a non-regular testing problem from a classical viewpoint as the null hypothesis lies on the boundary of parameter space. Our Bayesian approach is convenient to deal with this problem. In a Bayesian setting each of the two potential models - i.e. the model for  $c_t$  that includes the integrated unobserved component and the model for  $c_t$  that excludes the integrated unobserved component - is assigned a prior probability, and the posterior probability of each model is then calculated conditional on the data. Frühwirth-Schnatter and Wagner (2010) extend the Bayesian model selection approach of George and McCulloch (1993) to state space models. Following their approach, we consider a non-centered parameterization of the unobserved component  $\mu_t$ . A binary stochastic indicator is then assigned to the integrated unobserved component so that it is either included in or excluded from eq.(10). This indicator is sampled together with the model parameters in our Gibbs sampling algorithm.

#### Non-centered specification

We can rewrite eq.(12) as

$$\mu_t = \mu + \sigma_\eta \mu_t^* \tag{14}$$

$$\mu_t^* = \mu_{t-1}^* + \eta_t^* \qquad \mu_0^* = 0 \qquad \eta_t^* \sim iid\mathcal{N}(0, 1) \tag{15}$$

where  $\mu$  is the initial value of  $\mu_t$  if  $\mu_t$  is time-varying ( $\sigma_\eta^2 > 0$ ) and it is the constant value of  $\mu_t$  when  $\mu_t$  is constant ( $\sigma_\eta^2 = 0$ ). Crucially, the non-centered specification is not identified as the signs of  $\sigma_\eta$  - i.e., the square root of the variance  $\sigma_\eta^2$  - and  $\mu_t^*$  in eq.(14) can be changed without changing their product. As a result of this non-identification, the likelihood is symmetric around 0 along the  $\sigma_\eta$  dimension. When  $\sigma_\eta^2 > 0$ , the likelihood is bimodal with modes  $-\sigma_\eta$  and  $\sigma_\eta$ . When  $\sigma_\eta^2 = 0$ , the likelihood is unimodal around zero. Allowing for non-identification of  $\sigma_\eta$  therefore provides useful information on whether  $\sigma_\eta^2 > 0$ , and this non-identification can be exploited when drawing  $\sigma_\eta$  and  $\mu_t^*$  in the Gibbs algorithm through the use of a random sign switch, i.e., when sampling we multiply both by  $-1$  with probability 0.5 and leave both unaltered with probability 0.5.



### **Selection of the integrated unobserved component**

The non-centered parameterization is useful for model selection as the transformed component  $\mu_t^*$ , in contrast to  $\mu_t$ , does not degenerate to a static component if the innovation variance equals zero. This means that if the variance  $\sigma_\eta^2 = 0$ , then  $\sigma_\eta = 0$  in eq.(14) and the time-varying part  $\mu_t^*$  of the unobserved component  $\mu_t$  drops from the model. Hence, in the non-centered parameterization the presence or absence of a non-stationary unobserved component can be expressed as a standard variable selection problem. In particular, we write

$$\mu_t = \mu + \iota\sigma_\eta\mu_t^* \tag{16}$$

where  $\iota$  is a binary inclusion indicator which is either 0 or 1. If  $\iota = 1$ , there is an integrated unobserved component,  $\mu$  is the initial value of  $\mu_t$  and  $\sigma_\eta$  is estimated from the data. If, on the other hand,  $\iota = 0$ , there is no integrated unobserved component,  $\mu_t$  becomes constant as  $\mu_t = \mu$  and  $\sigma_\eta$  is set to 0. The binary indicator  $\iota$  is sampled together with the other parameters so that from its posterior distribution we can calculate the posterior inclusion probability of the integrated unobserved component in the regression.

#### **5.2.3 Parameter priors**

Our Bayesian estimation approach requires choosing prior distributions for the model parameters.

##### **Prior for the binary indicator $\iota$**

For the binary inclusion indicator, we assume a Bernoulli prior distribution with probability  $p_0$ . When calculating the posterior probability of the unobserved integrated component in the regression, we set  $p_0 = 0.5$  but also report results for  $p_0 = 0.25$  and  $p_0 = 0.75$ .

##### **Prior for the innovation variance $\sigma_\eta^2$ of the unobserved component**

Following Frühwirth-Schnatter and Wagner (2010), the standard inverse gamma (IG) prior for the variance of the innovation to the unobserved component is replaced by a Gaussian prior centered at zero for the square root of the variance. The reason for this is that when using the standard IG prior distribution for variance parameters, the choice of the shape and scale hyperparameters that define this distribution has a strong influence on the posterior distribution when the true value of the variance is close to zero. More specifically, as the IG distribution does not have probability mass at zero, using it as a prior distribution tends to push the posterior density away from zero. This is of particular importance when estimating the variance  $\sigma_\eta^2$  of the innovation to the unobserved component  $\mu_t$  as we want to decide whether or not to include this component in the regression. As can be noted from eqs.(10), (11) and

(14), since  $\sigma_\eta$  is a regression coefficient in the consumption equation, a further important advantage of the non-centered parameterization is therefore that it allows us to replace the standard IG prior on the variance parameter  $\sigma_\eta^2$  by a Gaussian prior centered at zero on  $\sigma_\eta$ . Centering the prior distribution at zero makes sense as, for both  $\sigma_\eta^2 = 0$  and  $\sigma_\eta^2 > 0$ ,  $\sigma_\eta$  is symmetric around zero. Frühwirth-Schnatter and Wagner (2010) show that the posterior density of  $\sigma_\eta$  is much less sensitive to the hyperparameters of the Gaussian distribution and is not pushed away from zero when  $\sigma_\eta^2 = 0$ . In Table 2 we report the prior distributions assumed for the model parameters. As can be seen in the table, we use a Gaussian prior distribution for  $\sigma_\eta$  with mean zero and a variance equal to 0.10. The variance is chosen such that the prior distribution has support over a sufficiently large range of relevant parameter values.

### Other priors

As can be seen in Table 2, we also use Gaussian prior distributions centered at zero for the elasticities  $\phi$ , for the constant  $\mu$  of the integrated unobserved component and for the coefficients  $\rho$  on the contemporaneous values and leads and lags of the first differences of the regressors. Again, the variances of these prior distributions are chosen such that the distributions have support over a sufficiently large range of relevant parameter values, i.e., a unit variance is used for the prior distributions of  $\phi$ ,  $\mu$  and  $\rho$ . Finally, for the variance  $\sigma_\varepsilon^2$  of the error term  $\varepsilon$  we use an IG prior distribution with a relatively high belief equal to 0.01 and a relatively low strength equal to 0.01 (see Bauwens et al., 2000, for details on prior beliefs and strengths).<sup>16</sup> By giving a relatively high prior value to the parameter  $\sigma_\varepsilon^2$ , we give a relatively high weight to the stationary component in the regression  $v_t$  compared to the integrated component  $\mu_t$ . In general, all priors are rather flat so that the results reported in the next section are driven mostly by the data and are not very sensitive to the priors chosen.

---

<sup>16</sup>Note that the strength 0.01 times the sample size  $T$  can be interpreted as the number of fictional observations used to construct the prior belief.

**Table 2:** Prior distributions regression parameters

Gaussian priors $\mathcal{N}(b_0, V_0)$				Percentiles	
		mean ( $b_0$ )	variance ( $V_0$ )	5%	95%
Elasticities on regressors $x_t$	$\phi$	0.00	1.00	-1.64	1.64
Constant $I(1)$ component	$\mu$	0.00	1.00	-1.64	1.64
Square root variance $I(1)$ component	$\sigma_\eta$	0.00	0.10	-0.52	0.52
Coeff. on lags/leads of $\Delta x_t$ (DOLS terms)	$\rho$	0.00	1.00	-1.64	1.64
Inverse Gamma prior $IG(\nu_0 T, \nu_0 T \sigma_0^2)$				Percentiles	
		belief ( $\sigma_0^2$ )	strength ( $\nu_0$ )	5%	95%
Variance error term $\varepsilon_t$	$\sigma_\varepsilon^2$	0.01	0.01	0.004	0.12

Notes: The regression equation is  $c_t = x_t \phi + \mu_t + v_t$ . The random walk  $I(1)$  component is  $\mu_t = \mu + \iota \sigma_\eta \mu_t^*$  with  $\mu_t^* = \mu_{t-1}^* + \eta_t^*$ . The stationary  $I(0)$  component is  $v_t = \sum_{j=-p}^p \Delta x_{t+j} \rho_j + \varepsilon_t$ . With  $p = 6$  and data available over the period 1951Q4 – 2016Q4, the effective sample period is 1953Q3 – 2015Q2 and the effective sample size is  $T = 248$  (i.e., 261 observations minus 1 for first-differencing and minus 12 for constructing leads and lags).

### 5.3 Results

This section presents the results of the estimation of the unobserved component model. Table 3 presents the posterior probabilities that there is an integrated unobserved component present in the basic regression eq.(10). These are reported for different prior probabilities  $p_0$  and for the two regression specifications that we consider, i.e., for regressor vector  $x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$  (panel A) and regressor vector  $x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$  (panel B). The results provide strong evidence that there is an unobserved random walk component present in the regression, i.e., all posterior probabilities are found to be equal to 1.

**Table 3:** Posterior inclusion probabilities  $p(\iota = 1)$  of an integrated unobserved component (for different prior inclusion probabilities  $p_0$ )

	(A)	(B)
	$x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$	$x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$
$p_0 = 0.5$	1	1
$p_0 = 0.75$	1	1
$p_0 = 0.25$	1	1

Notes: The regression equation is  $c_t = x_t \phi + \mu_t + v_t$ . Reported is the posterior inclusion probability of the integrated unobserved component  $\mu_t$ . It is calculated as the average of the 10.000  $\iota$ 's with each  $\iota$  sampled in a Gibbs iteration. The prior distribution of the binary indicator  $\iota$  is Bernoulli with probability  $p_0$ . The effective sample period is 1953Q3 – 2015Q2.

In Table 4 we present the posterior means and 90% highest posterior density (HPD) intervals of the posterior distributions for the fixed parameters of the unobserved component model given by eqs.(10), (11), (13), (16) and (15) with the exception of the coefficients  $\rho_j$  in eq.(13) which are excluded due to space constraints. Again, we present results for regressor vector  $x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$  (panel A) and for regressor vector  $x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$  (panel B). Furthermore, we present results both with the binary indicator set to 1 and with the binary indicator set to 0. Setting  $\iota = 1$  is in line with the posterior inclusion probabilities equal to 1 as reported in Table 3, i.e., the integrated component is included in the model and estimated. Setting  $\iota = 0$  is in line with the models estimated in the existing literature, i.e., the non-stationarity of  $z_t$  in eq.(10) is typically not accounted for. From the table we note that when we neglect the non-stationary component in  $z_t$  (i.e. when  $\iota = 0$ ) the impact of log assets  $a_t$  on log consumption  $c_t$  is somewhat overestimated, i.e., in panel A the elasticity  $\alpha$  is found to be equal to 0.27 when  $\iota = 0$  and equal to 0.24 when  $\iota = 1$ . From panel B in the table we find that this difference can be attributed to the non-stock part of assets  $a^{ns}$  as  $\alpha^{ns}$  equals 0.25 when  $\iota = 0$  and only 0.18 when  $\iota = 1$  (while the estimates for  $\alpha^s$  are essentially the same when  $\iota = 0$  and when  $\iota = 1$ ). Panel B of the table further shows that non-stock asset wealth, which consists mainly of housing wealth, has a larger long run impact on consumption than stock wealth. This finding is in line with results reported in the literature (see e.g., Davis and Palumbo, 2001; Carroll et al., 2011). The standard deviation  $|\sigma_\eta|$  of the error term of the integrated random walk component  $\mu_t$  is estimated only when  $\iota = 1$  and is found to be larger than zero both in panels A and B. This result for  $|\sigma_\eta|$  suggests that there is important time-variation in  $\mu_t$  and corroborates the finding of a posterior inclusion probability  $p(\iota = 1)$  equal to 1 as reported in Table 3.

**Table 4:** Posterior distributions parameters of equation

$$c_t = x_t \phi + \mu_t + v_t$$

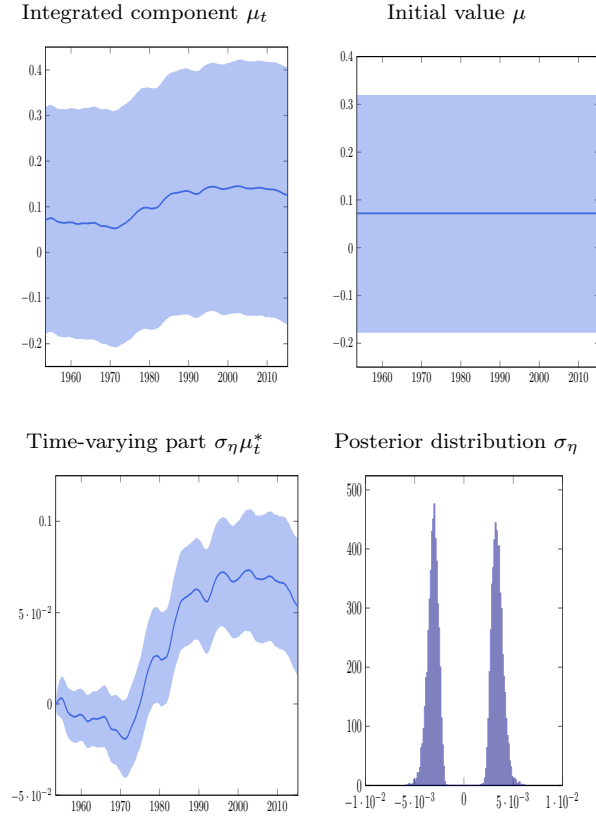
	(A)		(B)	
	$x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$		$x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$	
	(1)	(2)	(1)	(2)
	$\iota = 1$	$\iota = 0$	$\iota = 1$	$\iota = 0$
$\alpha$	0.2382	0.2728	-	-
	[0.1722,0.3041]	[0.2405,0.3050]	[-,-]	[-,-]
$\alpha^s$	-	-	0.0817	0.0881
	[-,-]	[-,-]	[0.0428,0.1243]	[0.0763,0.1000]
$\alpha^{ns}$	-	-	0.1783	0.2486
	[-,-]	[-,-]	[0.1175,0.2391]	[0.2078,0.2895]
$\beta$	0.7209	0.7575	0.7161	0.7115
	[0.6412,0.8004]	[0.7218,0.7934]	[0.6249,0.8045]	[0.6662,0.7576]
$\mu$	0.0718	-0.6491	0.0424	-0.7449
	[-0.1781,0.3196]	[-0.7226,-0.5753]	[-0.1800,0.2778]	[-0.8554,-0.6345]
$ \sigma_\eta $	0.0033	-	0.0035	-
	[0.0025,0.0043]	[-,-]	[0.0026,0.0045]	[-,-]
$\sigma_\varepsilon^2$	.00028	.00055	.00027	.00045
	[.00024,.00032]	[.00048,.00064]	[.00023,.00032]	[.00039,.00052]

Notes: Reported are the posterior mean with 90% HPD interval (in square brackets). In panel A of the table the coefficient vector is  $\phi = \begin{bmatrix} \alpha & \beta \end{bmatrix}'$  and in panel B it is  $\phi = \begin{bmatrix} \alpha^s & \alpha^{ns} & \beta \end{bmatrix}'$ . The random walk I(1) component is  $\mu_t = \mu + \iota \sigma_\eta \mu_t^*$  with  $\mu_t^* = \mu_{t-1}^* + \eta_t^*$ . The stationary I(0) component is  $v_t = \sum_{j=-p}^p \Delta x_{t+j} \rho_j + \varepsilon_t$ . The coefficients  $\rho_j$  are excluded from the table due to space constraints. With  $p = 6$  and data available over the period 1951Q4 – 2016Q4, the effective sample period is 1953Q3 – 2015Q2 and the effective sample size is  $T = 248$  (i.e., 261 observations minus 1 for first-differencing and minus 12 for constructing leads and lags).

Figure 2 presents the estimated random walk  $\mu_t$  and its components as obtained from the estimation of eq.(10) with  $x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$  and  $\iota = 1$ , i.e., as obtained from the estimates reported in Table 4 panel A with  $\iota = 1$ . Results obtained with  $x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$  and  $\iota = 1$  are very similar and therefore not reported. The upper left panel of the figure presents  $\mu_t$ . From the non-centered specification given by eq.(14), the initial value  $\mu$  of the random walk is presented in the upper right panel of the figure while the time-varying part  $\sigma_\eta \mu_t^*$  of the random walk is presented in the lower left panel. Based on these figures we confirm that there is considerable time-variation in  $\mu_t$ . The rather wide 90% HPD intervals around the posterior mean of  $\mu_t$  stem mainly from uncertainty surrounding the initial value  $\mu$  while the HPD intervals around the time-varying part  $\sigma_\eta \mu_t^*$  are much narrower. The posterior distribution of  $\sigma_\eta$  (i.e.,

the square root of the variance  $\sigma_\eta^2$ ) presented in the lower right panel of the figure is clearly bimodal and therefore again constitutes evidence that points toward time-variation in  $\mu_t$  and thus to the presence of a random walk  $\mu_t$  in eq.(10).

**Figure 2:** The integrated component  $\mu_t$ , its initial value  $\mu$  and time-varying part  $\sigma_\eta\mu_t^*$ , and the posterior distribution of the square root of its error variance  $\sigma_\eta$



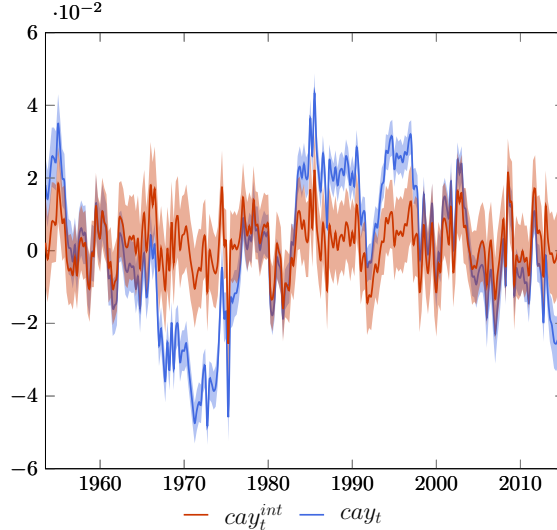
Note: Depicted in the first three figures are the posterior means with shaded areas denoting the 90% HPD interval. All figures are based on the results of Table 4 panel A with  $\iota = 1$ . The effective sample period is 1953Q3 – 2015Q2.

Figure 3 then presents the posterior means and 90% HPD intervals of the stationary component  $v_t$  both for  $\iota = 0$ , i.e., when no integrated unobserved component is included in the model, and for  $\iota = 1$ , i.e., when the non-stationary random walk  $\mu_t$  is included in the model and estimated. In the former case,  $v_t$  corresponds to the traditional proxy for the consumption to total wealth ratio "cay" estimated in the literature so we denote it by  $cay_t$ . We find that it is nearly identical to the standard "cay" as calculated and reported by Lettau and Ludvigson (2015). In the latter case,  $v_t$  constitutes a new "cay" obtained after controlling for an unobserved integrated component. We denote this series by  $cay_t^{int}$ . The figure reveals that  $cay_t^{int}$  is considerably less persistent (i.e., more stationary) than  $cay_t$  as the non-stationarity in  $z_t$  in eq.(10) is entirely soaked up by  $\mu_t$  in this case.<sup>17</sup> The variable  $cay_t^{int}$  can be interpreted as the

<sup>17</sup>Moreover, note from the figure that the HPD interval around  $cay_t^{int}$  is wider than that around  $cay_t$  which stems from the fact that the estimation of  $cay_t^{int}$  entails the estimation of both fixed parameters and a time-varying state - i.e., the

stationary or transitory component of the consumption-to-wealth ratio while the integrated unobserved component  $\mu_t$  constitutes the permanent component of the consumption-to-wealth ratio. In Section 5.5, we take a closer look at  $cay_t^{int}$  but first, in Section 5.4, we present an interpretation for the unobserved integrated component  $\mu_t$ .

**Figure 3:** The stationary component  $v_t$  for  $\iota = 1$  ( $cay_t^{int}$ ) and  $\iota = 0$  ( $cay_t$ )



Notes: Depicted are the posterior means. The shaded areas denote the 90% HPD interval.  $cay_t^{int}$  is calculated from the results of Table 4 panel A with  $\iota = 1$ .  $cay_t$  is calculated from the results of Table 4 panel A with  $\iota = 0$ . The effective sample period is 1953Q3 – 2015Q2.

## 5.4 Interpretation of the unobserved component $\mu_t$

In this section we first link the estimated integrated unobserved component found in the regression of consumption on assets and earnings to a proxy for financial liberalization. Second, we discuss and estimate the theoretical channel by which the unobserved component - interpreted as stemming from financial liberalization - affects the consumption-to-wealth ratio.

### 5.4.1 Linking the unobserved component to financial liberalization

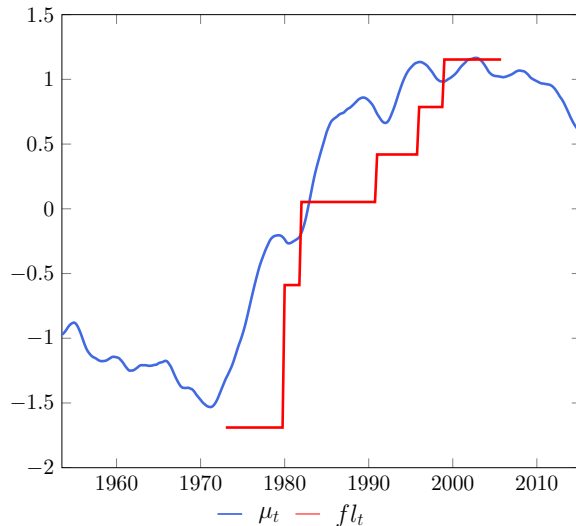
By relaxing liquidity constraints, financial liberalization may have increased the propensity to consume out of (total) wealth over time, and thereby raised the consumption to wealth ratio. Figure 4 shows the estimated unobserved component  $\mu_t$  and Abiad et al. (2008)'s proxy for financial liberalization in the US which we denote by  $fl_t$  and which is available over the period 1973Q1 – 2005Q4. Note that in the figure both variables are re-scaled so as to make them comparable graphically. The financial liberalization index of Abiad et al. (2008) is a mixture of indicators of financial development (credit controls and ceilings, interest rate liberalization, capital account transactions, securities market development, banking sector

---

unobserved component  $\mu_t$  - while the estimation of  $cay_t$  entails only the estimation of fixed parameters.

supervision). More details on this index are provided in Appendix A. We choose this particular proxy over other proxies because - since it is based on readings of laws and regulations - it is more likely to be driven by credit supply and less by demand.<sup>18</sup> From the figure, we note the similar trend in both measures. This suggests that financial liberalization may be - at least partially - responsible for the structural increase in the proxied consumption-to-wealth ratio and hence for the failure to find a stationary long-run relationship (i.e., cointegration) between consumption and wealth (where the latter is proxied by assets and earnings). We note that the structural increase in the (unobserved) US consumption to total wealth ratio due to financial liberalization is in line with the structural decrease in the (observed) US household saving to disposable income ratio as documented by Carroll et al. (2012). They note that this development can be largely attributed to financial liberalization.

**Figure 4:** The unobserved component  $\mu_t$  and Abiad et al. (2008)'s index of financial liberalization  $fl_t$



Notes: Depicted is the posterior mean of  $\mu_t$  re-scaled by subtracting its mean and dividing by its standard deviation.  $\mu_t$  is calculated from the results of Table 4 panel A with  $\iota = 1$ . Depicted is Abiad et al. (2008)'s index of financial liberalization  $fl_t$  re-scaled by subtracting its mean and dividing by its standard deviation. The effective sample period is 1953Q3 – 2015Q2 but  $fl_t$  is only available over the period 1973Q1 – 2005Q4.

A more formal approach however is needed to investigate the financial liberalization interpretation of  $\mu_t$ . To this end, we conduct Engle-Granger cointegration tests between consumption, assets, earnings and our proxy for financial liberalization. We note that  $c_t$ ,  $a_t$ ,  $y_t$  and  $\mu_t$  are cointegrated by construction since the unobserved component  $\mu_t$  is added to the regression equation for that specific purpose. Finding cointegration between  $c_t$ ,  $a_t$ ,  $y_t$  and  $fl_t$  however when - as noted in Section 4.2.1 - there is no cointegration between  $c_t$ ,  $a_t$  and  $y_t$  would provide strong support for the liberalization story. The results of the

<sup>18</sup>This in contrast to alternative measures such as household liabilities to disposable income or the credit easing accumulated (CEA) index - see Carroll et al. (2012) - which is strongly based on the household liabilities ratio. Upon using the CEA index instrumented by the Abiad et al. (2008) index, we find similar results as those reported in this section for the Abiad et al. (2008) index.



conducted tests are presented in Table 5. We test for cointegration between  $c_t$ ,  $a_t$ ,  $y_t$  and  $fl_t/fl_t^+$  where  $fl_t^+$  equals  $fl_t$  over the period 1973Q1 – 2005Q4 after which it is set to the last available value of  $fl_t$  for the remainder of the sample which is a value that indicates full financial liberalization. We also test for cointegration between  $c_t$ ,  $a_t^s$ ,  $a_t^{ns}$ ,  $y_t$  and  $fl_t/fl_t^+$ . From the table, we note that the null hypothesis of no cointegration is rejected at the 10% level for all the conducted regressions between the variables  $c_t$ ,  $a_t$ ,  $y_t$  and  $fl_t/fl_t^+$ . For the variables  $c_t$ ,  $a_t^s$ ,  $a_t^{ns}$ ,  $y_t$  and  $fl_t/fl_t^+$ , the null of no cointegration is rejected at the 10% level only over the sample period 1973Q1 – 2016Q4 (i.e., for the largest sample size which provides more power to the test). All in all, the results presented provide reasonable evidence that the lack of cointegration between consumption, assets and earnings and, therefore, the presence of an integrated unobserved component in the relationship between these variables - which implies a non-stationary consumption-to-wealth ratio  $c_t - w_t$  - can be attributed to financial liberalization. The theoretical channel by which this occurs is investigated in the next section.

**Table 5:** Engle-Granger cointegration test between  $c$ ,  $a$ ,  $y$  and  $fl$  and between  $c$ ,  $a^s$ ,  $a^{ns}$ ,  $y$  and  $fl$

$c, a, y, fl$			$c, a^s, a^{ns}, y, fl$		
73Q1 – 05Q4	73Q1 – 16Q4	73Q1 – 07Q3	73Q1 – 05Q4	73Q1 – 16Q4	73Q1 – 07Q3
Dickey-Fuller t-statistic					
-3.82	-4.09	-3.98	-3.81	-4.25	-3.99
5% and 10% critical values (4 variables)			5% and 10% critical values (5 variables)		
5%	-4.10		5%	-4.41	
10%	-3.81		10%	-4.13	

Notes: The null hypothesis tested is the no cointegration hypothesis. A constant is included in the cointegrating equation. The 5% and 10% critical values are taken from MacKinnon (2010). The number of lags included in the augmented Dickey-Fuller regression is based on the Bayesian information criterion. The sample 1973Q1 – 2005Q4 uses Abiad et al. (2008)'s index of financial liberalization  $fl_t$  which is only available over this period. The samples 1973Q1 – 2016Q4 and 1973Q1 – 2007Q3 use the financial liberalization variable  $fl_t^+$  which equals Abiad et al. (2008)'s index  $fl_t$  over the period 1973Q1 – 2005Q4 after which it is set to the last value  $fl_t$  takes for the remainder of the sample which is a value indicating full financial liberalization.

#### 5.4.2 Channel

The theoretical framework of Section 2 is based only on the budget constraint of a representative consumer. To incorporate financial liberalization into our model, we now also impose behavioral restrictions on the representative consumer. In particular, we consider a first-order condition that incorporates a potentially binding liquidity constraint (see e.g, Zeldes, 1989), i.e.,

$$E_t \left[ \delta R_{t+1} \frac{u'(C_{t+1})}{u'(C_t)} (1 + \lambda_t) \right] = 1 \quad (17)$$

where  $\delta$  is the consumer's discount factor,  $R_t = 1 + r_{t+1}$  is the gross real rate of return on assets,  $u'(C_t)$  is marginal utility of consumption with  $u(C_t)$  the utility function and  $C_t$  the period  $t$  consumption level, and  $\lambda_t$  is the (normalized) Lagrange multiplier associated with the period  $t$  liquidity constraint which is positive when the constraint is binding and zero when the constraint is not binding. From eq.(17), we can write  $\delta R_{t+1} \frac{u'(C_{t+1})}{u'(C_t)} (1 + \lambda_t) = 1 + e_{t+1}$  where  $e_{t+1}$  is an expectation error uncorrelated with period  $t$  information. We can then specify the utility function as being of the CRRA type so that  $u(C) = \frac{C^{1-\theta}}{1-\theta}$  with coefficient of relative risk aversion  $\theta > 0$  and write  $\delta R_{t+1} \left(\frac{C_{t+1}}{C_t}\right)^{-\theta} (1 + \lambda_t) = 1 + e_{t+1}$ . After taking logs of both sides of this expression and solving for the growth rate in consumption  $\Delta c_{t+1} = c_{t+1} - c_t = \ln(C_{t+1}) - \ln(C_t)$ , we obtain,

$$\Delta c_{t+1} = \frac{1}{\theta} \ln \delta + \frac{1}{\theta} \ln R_{t+1} + \frac{1}{\theta} \ln(1 + \lambda_t) - \frac{1}{\theta} (1 + e_{t+1}) \quad (18)$$

Financial liberalization can be expected to make liquidity constraints less binding, thereby reducing  $\lambda_t$  and leading to a lower future consumption growth rate  $\Delta c_{t+1}$ . If financial liberalization is (stochastically) trended, aggregate consumption growth is non-stationary. And if the trend in financial liberalization is upward, then consumption growth decreases permanently. If future consumption growth is permanently reduced, the intertemporal budget constraint then suggests that - ceteris paribus - the (log) current consumption-to-wealth ratio  $c_t - w_t$  can increase permanently. This should be clear from eq.(5), which is repeated here for convenience, i.e.,

$$c_t - w_t = E_t \sum_{j=1}^{\infty} \rho^j (r_{t+j} - \Delta c_{t+j}) \quad (19)$$

The proposed channel can be tested by conducting regressions of the present discounted value of future consumption growth  $\sum_{j=1}^h \rho^j \Delta c_{t+j}$  - with  $h$  the considered horizon - on the integrated unobserved component  $\mu_t$ , which is the permanent component of the consumption-to-wealth ratio  $c_t - w_t$ . The financial liberalization interpretation of  $\mu_t$  suggests that the impact of  $\mu_t$  should be negative, i.e., higher financial liberalization and therefore a higher  $\mu_t$  and  $c_t - w_t$  implies lower future aggregate consumption growth. In Table 6, we present the results of the estimation of this relationship for discount factors  $\rho$  equal to 1 and 0.99. Because  $\mu_t$  is non-stationary and  $\Delta c_t$  is expected to be non-stationary - whereby considering longer horizons exacerbates the problem as summing consumption growth rates increases their persistence by construction - our estimation method must be such that spurious results are avoided. As noted by Hamilton (1994, pages 561-562), this can be achieved by a Cochrane-Orcutt adjustment for first-order serial correlation in the residuals of the regression of  $\sum_{j=1}^h \rho^j \Delta c_{t+j}$  on  $\mu_t$  (which is asymptotically equivalent to first-differencing the regression equation). In the table we further report the results of a Cochrane-Orcutt regression of  $\sum_{j=1}^h \rho^j r_{t+j}$  on  $\mu_t$ . We do this because from eq.(19), it can be seen that

an integrated unobserved component  $\mu_t$  leading to a structural increase in the consumption to wealth ratio  $c_t - w_t$  could stem from a negative impact of that component on the present discounted value of aggregate consumption growth, but also from a positive impact of that component on the present discounted value of the returns on wealth. Since no data are available for the returns on total wealth, we use real stock market returns  $r_{s,t}$  as a proxy for  $r_t$  (see Appendix A for details). The results reported in the table show that there is indeed a negative impact of the unobserved component  $\mu_t$  on the present discounted value of future consumption growth, which is significant for all horizons considered with the exception of the one quarter horizon case. This result provides further support for the interpretation of the integrated unobserved component  $\mu_t$  as stemming from financial liberalization and acting through the channel described above. On the other hand, while the impact of  $\mu_t$  on the present discounted value of future returns on wealth (proxied by stock market returns) is as expected positive, it is never significant.

**Table 6:** Cochrane-Orcutt estimation of equation  $\sum_{j=1}^h \rho^j m_{t+j} = a + b\mu_t + \epsilon_{t+h}$   
(with  $\epsilon_{t+h} = \delta\epsilon_{t+h-1} + \varepsilon_{t+h}$ )

	$\rho = 1$					$\rho = 0.99$				
	Horizon $h$ (in quarters)					Horizon $h$ (in quarters)				
$m_t =$	1	16	32	48	64	1	16	32	48	64
$\Delta c_t$	-0.02	-1.24	-2.56	-1.54	-2.63	-0.02	-1.25	-2.32	-1.50	-2.16
	(-0.91)	(-2.21)	(-3.86)	(-1.98)	(-3.43)	(-0.92)	(-2.38)	(-3.85)	(-2.35)	(-3.41)
$r_{s,t}$	0.10	2.91	2.78	6.18	1.19	0.10	2.64	2.64	5.00	1.94
	(0.61)	(1.28)	(0.68)	(1.12)	(0.16)	(0.61)	(1.25)	(0.74)	(1.11)	(0.36)

Notes: The table reports maximum likelihood estimates of the parameter  $b$  with t-statistics between brackets. The t-statistics are calculated from standard errors based on outer gradient products.  $\Delta c_t$  is the growth rate in real per capita personal consumption expenditures and  $r_{s,t}$  is the real stock market return. For  $\mu_t$  the posterior mean of the integrated unobserved component is used which is obtained from estimating the unobserved component model discussed above. The effective sample period is different in every case depending on the horizon  $h$  considered.

## 5.5 Characteristics of the $cay^{int}$ variable

Following the discussion of the non-stationary component of the consumption-to-wealth ratio  $c_t - w_t$ , we now take a closer look at the stationary component of  $c_t - w_t$ , which we denoted in Section 5.3 by  $cay_t^{int}$ . The persistence and stationarity properties of the  $cay$  variables calculated from the unobserved component model - i.e.,  $cay_t^{int}$  which is calculated from the model with an unobserved integrated component included and estimated, and  $cay_t$  which is calculated from the standard model with no unobserved integrated component included - are quite different. The AR(1) coefficient obtained from a regression of each  $cay$  on its own lag and a constant shows that  $cay_t^{int}$  - with an AR(1) coefficient of less than 0.5 - is considerably

less persistent than the standard  $cay_t$  - which has an AR(1) coefficient equal to 0.91. We also conduct an (augmented) Dickey-Fuller unit root test on the standard  $cay_t$  variable, which shows that a unit root cannot be rejected using the standard critical values at both the 5% and even 10% levels of significance. This confirms the results reported above in Table 1 for the standard cointegration approach. The  $cay_t^{int}$  variable, on the other hand, is stationary by construction.

The findings concerning the persistence and stationarity of the variables  $cay_t^{int}$  and  $cay_t$  are important to evaluate their predictive ability for excess stock returns. According to the theory discussed in Section 2, the consumption to total wealth ratio is expected to have predictive power for the returns on wealth. The literature has in particular focussed on the ability of the "cay" proxy to predict future excess stock returns, as much of the variation in "cay" is attributed to the stock market component of assets  $a_t$  around the less variable consumption  $c_t$  and earnings  $y_t$  variables (see e.g, Lettau and Ludvigson, 2004). As noted by Ferson et al. (2003), if underlying *expected* excess returns are persistent then using highly persistent (i.e. non-stationary) regressors in typical forecasting models of excess returns will yield spurious results. Moreover, as noted by Rudd and Whelan (2006), this problem is likely more serious when longer horizon excess returns are used, as these are more persistent by construction. Since the variable  $cay_t^{int}$  is stationary by construction while the evidence reported above suggests that the standard  $cay_t$  variable is not stationary, it is useful to investigate the predictive power of  $cay_t^{int}$  for excess stock returns and compare it to the potentially spurious predictive ability of the  $cay_t$  variable. We consider both the in-sample and out-of-sample predictive ability for excess stock returns of our  $cay_t^{int}$  variable as calculated from a regression of  $c_t$  on  $a_t$  and  $y_t$  that contains an integrated unobserved component. Excess stock returns, denoted by  $r_{s,t}^e$ , are defined as the difference between the log real stock market return and the log real three-month Treasury bill rate. More details on this series and its calculation are provided in Appendix A. We note that the results for stock returns *not* in excess of a risk-free rate are very similar to those for excess stock returns. Hence, we follow the literature which has focussed especially on excess returns and report these results. The results for returns that are not in excess of a risk-free rate are unreported but available from the authors upon request.

Table 7 presents the results of in-sample forecast regressions at different horizons for excess stock returns using as predictors the  $cay_t^{int}$  and  $cay_t$  variables. Following the literature, we consider horizons of one, four, eight, twelve and sixteen quarters and we report the regression coefficient on each "cay" variable of the forecast regression for  $r_{s,t}^e$ , its corresponding Newey-West corrected t-statistic, and the adjusted  $R^2$  of the forecast regression. From the t-statistics and adjusted  $R^2$ 's reported in the table, we note that the standard  $cay_t$  variable has strong predictive power for excess stock returns and that this predictive ability increases considerably when longer horizons are considered. However, since the discussion of the previous

section casts doubt on the stationarity of the variable  $cay_t$ , these results may be spurious. It is therefore interesting to note that our variable  $cay_t^{int}$ , which is stationary by construction, still has considerable predictive power for excess stock returns. As can be seen when comparing the t-statistics and  $R^2$ 's in the table, the predictive power of  $cay_t^{int}$  is lower compared to that of  $cay_t$ , but it is nonetheless still important. We find t-statistics of the regression coefficient on  $cay_t^{int}$  in the forecast regression above 2.5 for horizons from four to sixteen quarters and  $R^2$ 's of the forecast regression as high as 10% and 12% at eight and twelve quarter horizons respectively. Ferson et al. (2003) argue that  $R^2$ 's of this magnitude can be considered economically significant. Moreover, the magnitudes of the  $R^2$ 's reported in Table 7 are in accordance with the univariate predictability results typically reported in the literature using other predictors for excess stock returns (see Ferson et al., 2003, for an overview).

**Table 7:** In-sample forecast regressions for excess stock returns

$$\sum_{j=1}^h r_{s,t+i}^e = a + bk_t + \varepsilon_{t+h}$$

$k_t =$	Horizon $h$ (in quarters)				
	1	4	8	12	16
$cay_t^{int}$	1.50 (1.80) [0.01]	5.93 (3.16) [0.06]	10.79 (4.49) [0.12]	10.93 (3.61) [0.10]	9.32 (2.68) [0.09]
$cay_t$	0.73 (2.71) [0.02]	2.91 (3.21) [0.10]	5.54 (4.37) [0.21]	7.20 (5.60) [0.27]	8.27 (6.91) [0.31]

Notes: The effective sample period is 1953Q3 – 2015Q2. The table reports OLS estimates of the parameter  $b$  with Newey-West corrected t-statistics between brackets and the adjusted  $R^2$  of the regression between square brackets.  $r_{s,t}^e$  is the stock excess return (see Appendix A).  $cay_t^{int}$  is the estimated stationary part of the log consumption to wealth ratio as obtained from the model containing an integrated unobserved component.  $cay_t$  is the estimated log consumption to wealth ratio as obtained from the standard model without an integrated unobserved component.

We next consider the the out-of-sample forecast performance of the  $cay_t^{int}$  variable for excess stock returns. In particular, our evaluation is based on root mean squared error (RMSE) ratios calculated from 60-quarter rolling forecasting regressions for excess stock returns, again over different horizons. Table 8 presents RMSE ratios which are calculated as the ratio of the RSME based on a forecast regression for  $cay_t^{int}$  over the RSME based on a forecast regression for an alternative predictor. As alternative predictors for the excess stock returns  $r_{s,t}^e$ , we consider the variable  $cay_t$ , the lagged excess stock return  $r_{s,t-1}^e$  (i.e., assuming an AR(1) process for excess returns) and a constant  $c$  (i.e., using the sample mean

of  $r_{s,t}^e$  as a predictor). A ratio below one means that the  $cay_t^{int}$  based forecast model performs better than the alternative forecast model and a ratio above one means that the  $cay_t^{int}$  based model performs worse. We also calculate the modified Diebold-Mariano (MDM) statistic that tests the null hypothesis that the mean squared error (MSE) of the forecasts obtained with  $cay_t^{int}$  is the same as the one obtained from the alternative forecast model considered. The p-value of this test is reported between square brackets. From the table we note that, based on the reported RMSE ratios, the  $cay_t^{int}$  variable performs worse than  $cay_t$  at longer horizons (i.e., at horizons of twelve and sixteen quarters) and better at shorter horizons (i.e., at horizons of four and eight quarters). The reported MDM statistics however suggest that - with the exception of the sixteen quarter horizon case - these differences are not significant. Hence, we can conclude that the predictive ability of  $cay_t^{int}$  for excess stock returns is largely comparable to that of  $cay_t$ . The table further shows that the RMSE ratios obtained from using  $cay_t^{int}$  as a predictor for excess stock returns versus lagged returns or a simple constant are lower than one at all horizons. This suggests out-of-sample prediction using  $cay_t^{int}$  which is superior to that which can be achieved by these naive models. However, upon looking at the p-values of the MDM statistics reported in the table, we find that these differences are seldom significant (i.e., only in the eight quarter horizon case).

**Table 8:** Out-of-sample forecast evaluations for excess stock returns  
( $RMSE_{cay_t^{int}}/RMSE_{k_t}$  measure)

$k_t =$	Horizon $h$ (in quarters)				
	1	4	8	12	16
$cay_t$	1.00	0.97	0.95	1.06	1.22
	[0.59]	[0.58]	[0.60]	[0.17]	[0.02]
$r_{s,t-1}^e$	0.98	0.98	0.96	0.95	0.95
	[0.88]	[0.11]	[0.03]	[0.21]	[0.20]
$c$	0.99	0.99	0.96	0.97	0.96
	[0.70]	[0.35]	[0.05]	[0.29]	[0.38]

Notes: The effective sample period is 1953Q3 – 2015Q2. The table reports root mean squared error (RMSE) ratios obtained from out-of-sample  $h$ -period ahead forecasts of the excess stock returns  $r_{s,t}^e$  using 60-quarter rolling subsamples. Forecasts are based on the predictive variables  $cay_t^{int}$  and  $k_t$  where  $k_t$  is either  $cay_t$ , the lagged excess return  $r_{s,t-1}^e$  or a constant  $c$ . RMSE ratios are calculated as the ratio of the RSME based on  $cay_t^{int}$  as a predictive variable over the RSME based on  $k_t$  as a predictive variable. A ratio below 1 means that the  $cay_t^{int}$  based forecast model performs better than the alternative forecast model and a ratio above 1 means that the  $cay_t^{int}$  based model performs worse. Between square brackets is the p-value of the modified Diebold-Mariano statistic that tests the null hypothesis that the mean squared error (MSE) of the forecasts obtained with  $cay_t^{int}$  and  $k_t$  are the same.  $cay_t^{int}$  is the estimated stationary part of the log consumption to wealth ratio as obtained from the model containing an integrated unobserved component.  $cay_t$  is the estimated log consumption to wealth ratio as obtained from the standard model without an integrated unobserved component.

## 6 Conclusions

This paper proposes an alternative empirical approach to study the long run aggregate relationship between household consumption and household wealth, where household wealth consists of tangible assets and unobserved human wealth. The evidence in favor of a stable cointegrating relationship between consumption, assets and earnings (as a proxy for human wealth) in the US is weak. Hence, the consumption to total wealth ratio (i.e., the variable "cay") estimated from such a relationship is non-stationary - rendering it inadequate to predict excess stock returns - while the elasticities of consumption to wealth estimated from this type of regression are unreliable.

The approach followed in this paper applies an unobserved component model to US data over the period 1951Q4 – 2016Q4 whereby the regression of consumption on assets and earnings is augmented with an integrated unobserved component. Our results strongly support the presence of an integrated unobserved component in the consumption equation. The residuals of this regression are stationary

because consumption, assets, earnings *and* the integrated unobserved component are cointegrated by construction. The elasticities of consumption to assets and earnings are positive and those estimated for assets (and their components) tend to be lower compared to the case where no integrated component is added to the regression. We interpret the integrated unobserved component as stemming from financial liberalization which, by relaxing liquidity constraints of consumers, has permanently increased the consumption-to-wealth ratio over the sample period. We calculate an alternative "cay" variable, i.e., the stationary part of the consumption-to-wealth ratio, which is much less persistent than the traditional "cay" variable. In-sample, we find that its predictive ability for future excess stock returns, while diminished compared to that of the traditional "cay" variable, is statistically and economically significant and in line with the univariate predictability results typically reported in the literature for excess stock returns using other predictors. Out-of-sample, we find that its predictive ability for future excess stock returns is comparable to that of the standard "cay" variable.

## References

- Abiad, A., Detragiache, E., and Tressel, T. (2008). A new database of financial reforms. Working paper 266, International Monetary Fund.
- Bauwens, L., Lubrano, M., and Richard, J.-F. (2000). *Bayesian inference in dynamic econometric models*. Oxford University Press.
- Bianchi, F., Lettau, M., and Ludvigson, S. (2017). Monetary policy and asset valuation. *Unpublished manuscript*.
- Campbell, J. and Mankiw, N. (1989). Consumption, income, and interest rates: reinterpreting the time series evidence. *NBER Macroeconomics Annual*, 4:185–216.
- Canarella, G., Pollard, S., and Lai, K. (1990). Cointegration between exchange rates and relative prices: another view. *European Economic Review*, 34(7):1303–1322.
- Carroll, C. D., Otsuka, M., and Slacalek, J. (2011). How large are housing and financial wealth effects? *Journal of Money, Credit and Banking*, 43(1):55–79.
- Carroll, C. D., Slacalek, J., and Sommer, M. (2012). Dissecting saving dynamics: measuring wealth, precautionary, and credit effects. *IMF Working Paper*, WP/12/219.
- Carter, C. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81:541–53.



- Chang, Y., Miller, J., and Park, J. (2009). Extracting a common stochastic trend: theory with some applications. *Journal of Econometrics*, 150(2):231–47.
- Cooper, D. (2016). Wealth effects and macroeconomic dynamics. *Journal of Economic Surveys*, 30(1):34–55.
- Davis, M. and Palumbo, M. (2001). A primer on the economics and time series econometrics of wealth effects. *Finance and Economics Discussion Series Federal Reserve Board Washington*, 2001-09.
- Durbin, J. and Koopman, S. (2001). *Time series analysis by state space methods*. Oxford University Press.
- Engle, R. and Granger, C. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica*, 55(2):251–76.
- Everaert, G. (2010). Estimation and inference in time series with omitted  $I(1)$  variables. *Journal of Time Series Econometrics*, 2(2):1–25.
- Ferson, W., Sarkissian, S., and Simin, T. (2003). Spurious regressions in financial economics? *Journal of Finance*, LVIII(4):1393–1413.
- Frühwirth-Schnatter, S. and Wagner, H. (2010). Stochastic model specification search for Gaussian and partial non-Gaussian state space models. *Journal of Econometrics*, 154(1):85–100.
- George, E. and McCulloch, R. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, 88:881–889.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Berger, J., Bernardo, J., Dawid, A., and Smith, A., editors, *Bayesian statistics*. Oxford University Press.
- Granger, C. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 35:143–159.
- Groen, J., Paap, R., and Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business and Economic Statistics*, 31(1):29–44.
- Hamilton, J. (1994). *Time series analysis*. Princeton.
- Harvey, A. (1989). *Forecasting structural time series models and the Kalman filter*. Cambridge University Press.

- Harvey, A., Henry, S., Peters, S., and Wren-Lewis, S. (1986). Stochastic trends in dynamic regression models: an application to the employment-output equation. *Economic Journal*, 96:975–985.
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3):231–254.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6):1551–1580.
- Kim, C.-J. and Nelson, C. (1999). *State-space models with regime switching. Classical and Gibbs-sampling approaches with applications*. MIT Press, Cambridge.
- Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH Models. *Review of Economic Studies*, 65(3):361–93.
- Kim, Y. and Kim, C.-J. (2011). Dealing with endogeneity in a time-varying parameter model: joint estimation and two-step estimation procedures. *Econometrics Journal*, 14:487–497.
- Lettau, M. and Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance*, LVI(3):815–849.
- Lettau, M. and Ludvigson, S. (2004). Understanding trend and cycle in asset values: re-evaluating the wealth effect on consumption. *American Economic Review*, 94(1):276–299.
- Lettau, M. and Ludvigson, S. (2015). Changes in the measurement of consumption in cay. *Report*.
- MacKinnon, J. (2010). Critical values for cointegration tests. *Working Paper Economics Department Queen's University*, 1227.
- McCulloch, R. and Tsay, R. (1993). Bayesian inference and prediction for mean and variance shifts in autoregressive time series. *Journal of the American Statistical Association*, 88(423):968–978.
- Newey, W. and West, K. (1987). A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708.
- Ogaki, M. and Park, J. Y. (1997). A cointegration approach to estimating preference parameters. *Journal of Econometrics*, 82(1):107–134.
- Park, J. Y. (1990). Testing for unit roots and cointegration by variable addition. *Advances in econometrics*, 8(2):107–133.

- Phillips, P. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, 33:311–340.
- Phillips, P. C. B. and Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica*, 58(1):165–193.
- Planas, C., Roeger, W., and Rossi, A. (2007). How much has labor taxation contributed to European structural unemployment? *Journal of Economic Dynamics and Control*, 31(4).
- Rudd, J. and Whelan, K. (2006). Empirical proxies for the consumption-wealth ratio. *Review of Economic Dynamics*, 9:34–51.
- Sarantis, N. and Stewart, C. (2001). Unobserved components in an error-correction model of consumption for Southern European countries. *Empirical Economics*, 26:391–405.
- Slacalek, J. (2004). International evidence on cointegration between consumption, income and wealth. *Unpublished manuscript*.
- Zeldes, S. (1989). Consumption and liquidity constraints: an empirical investigation. *Journal of Political Economy*, 97(2):305–346.

## Appendix A Data

Quarterly seasonally adjusted data for consumption, earnings (disposable labor income), population and the price deflator are collected from the National income and Product Accounts (NIPA) from the Bureau of Economic Analysis (BEA) at the U.S. Department of Commerce. The assets (wealth) data are collected from the Flow of Funds Accounts of the Board of Governors of the Federal Reserve System.

*Consumption* is measured as total personal consumption expenditures (line 1 of NIPA Table 2.3.5).

*Earnings* are defined as the sum of compensation for employees (line 2 of NIPA Table 2.1) plus personal current transfer receipts (line 16) minus contributions for domestic government social insurance (line 25) and minus personal labor taxes. Personal labor taxes are derived by first calculating the labor income fraction of total income, and subsequently using this ratio to back out the share of labor taxes from the total personal current taxes (line 26). The labor income to total income ratio is defined as the ratio of wages and salaries (line 3) to the sum of wages and salaries (line 3), proprietors' income (line 9), rental income (line 12) and personal income receipts on assets (line 13).

*Assets* are defined as the net worth of households and nonprofit organizations, measured at the end of the period. Stock market wealth is defined as households' and nonprofit organizations' holdings of

corporate equities, mutual fund shares, life insurance reserves, and private and public pension entitlements. Non-stock wealth is defined as non-stock assets minus liabilities. The non-stock assets consist of households' and nonprofit organizations' holdings of consumer durable goods, real estate, and non-stock financial assets (total currency and deposits including money market fund shares, debt securities, mortgages, proprietors' equity in non-corporate business and other assets). Liabilities include all loans, such as mortgages, consumer debt and other loans.

*Stock returns* are the returns (excluding dividends) of the value-weighted CRSP index from the Center for Research in Security Prices. The CRSP index is a broad stock market index including the NYSE, AMEX, NASDAQ and ARCA, and the data are collected from CRSP via Wharton Research Data Services (WRDS).

*Excess stock returns* are defined as the difference between the quarterly log real stock market return (as stated above) and the quarterly log real 3-month Treasury Bill return (i.e., the "risk free rate"). The 3-month Treasury Bill data is the secondary market rate, not seasonally adjusted, collected from the Federal Reserve Bank of St.Louis.

*Financial liberalization* is proxied by the Index of Financial Reform by Abiad et al. (2008). The annual index covers the period 1973-2005 and includes seven different dimensions of financial sector policy: credit controls and reserve requirements, interest rate controls, entry barriers, state ownership, policies on securities markets, banking regulations and restrictions on the capital account. Liberalization scores for each category are combined in a graded index which is normalized from zero to one.

All series except the financial liberalization proxy are deflated with the price index for total personal consumption expenditures (line 1 of NIPA Table 2.3.4) with base year 2009 = 100. All variables except the (excess) stock returns and the financial liberalization proxy are further expressed in per capita terms, with population data also collected from the NIPA (line 40 of Table 2.1).

## Appendix B Cointegration tests

**Table B-1:** A battery of frequentist tests for cointegration between  $c$ ,  $a$  and  $y$  and between  $c$ ,  $a^s$ ,  $a^{ns}$  and  $y$

Sample	$c, a, y$		$c, a^s, a^{ns}, y$	
	1951Q4 – 2016Q4	1951Q4 – 2007Q3	1951Q4 – 2016Q4	1951Q4 – 2007Q3
Engle-Granger	No	No	No	No
Phillips-Ouliaris	No	No	No	No
Johansen Trace 1 lag	No	No	No	No
Johansen Max Eig. 1 lag	No	No	No	No
Park's H(0,1) test	Yes	No	Yes	No
Park's H(0,3) test	No	No	No	No

Notes: The residual-based Engle-Granger and Phillips-Ouliaris cointegration tests and the Johansen Trace and Maximum Eigenvalue methodology test the null hypothesis of no cointegration, whereas Park's H( $q$ ,  $p$ ) added variable test, with time trends of powers  $q$  up to order  $p$ , evaluates the null hypothesis of cointegration. Conclusions are based on 5% significance levels.

In Table B-1, we present a summary of the results of a number of additional tests conducted to determine whether there is cointegration between the variables  $c$ ,  $a$  and  $y$  or between the variables  $c$ ,  $a^s$ ,  $a^{ns}$  and  $y$ . These tests by and large suggest that there is no cointegration between the variables  $c$ ,  $a$  and  $y$  nor between the variables  $c$ ,  $a^s$ ,  $a^{ns}$  and  $y$ . We note that some of these tests like the Engle and Granger (1987), Phillips and Ouliaris (1990), and Johansen (1988, 1991) tests have the absence of cointegration as the null hypothesis while others like Park (1990)'s added variable test have the presence of cointegration as the null hypothesis. According to Park (1990) and Ogaki and Park (1997), the inability to reject the null hypothesis of a unit root in the variable  $z_t$  in eqs.(8) and (9) may be due to a potential lack of power of cointegration tests that have the absence of cointegration as the null hypothesis. Park's (1990) added variable test has cointegration as the null hypothesis and - for the full sample period - cannot reject the null of cointegration when a linear deterministic time trend is added to mimic the potential stochastic trend in  $z_t$ , but does reject cointegration when the potential stochastic trend is mimicked in a more general and realistic way through the addition of a linear, quadratic and cubic time trend to the regression. This suggests that the failure of finding cointegration does not stem from statistical issues - i.e., the lack of power to reject a false null hypothesis - of standard cointegration tests. Rather, it suggests that there is in fact a unit root present in  $z_t$ .

## Appendix C A more general unobserved component model

We also consider and estimate a more general unobserved component model. Specifically, we model the unobserved component  $\mu_t$  as a weighted average of two random walks,  $\check{\mu}_t$  and  $\tilde{\mu}_t$ , with the weights given

by a regime switching variable  $S_t$ , i.e.,

$$\mu_t = (1 - S_t)\check{\mu}_t + S_t\tilde{\mu}_t \quad (\text{C-1})$$

The variable  $S_t$  is assumed to follow a first-order two-state Markov switching process (i.e., it takes on either the value of 0 or the value of 1). As in Section 5.2.2, we specify the random walks in non-centered form as,

$$\check{\mu}_t = \mu' + \sigma'_\eta \check{\mu}_t^* = \mu' \quad (\text{C-2})$$

and

$$\tilde{\mu}_t = \mu + \sigma_\eta \mu_t^* \quad (\text{C-3})$$

where  $\check{\mu}_t^*$  and  $\mu_t^*$  are standardized random walks given by  $\check{\mu}_t^* = \check{\mu}_{t-1}^* + \eta_t'$  with  $\check{\mu}_0^* = 0$  and  $\eta_t' \sim iid\mathcal{N}(0, 1)$ , respectively  $\mu_t^* = \mu_{t-1}^* + \eta_t^*$  with  $\mu_0^* = 0$  and  $\eta_t^* \sim iid\mathcal{N}(0, 1)$ . In eq.(C-2), we assume that  $\sigma'_\eta = 0$  so that  $\check{\mu}_t$  is constant and the unobserved component  $\mu_t$  is a constant  $\mu'$  in one regime and follows a standard random walk with initial value  $\mu$  in the other regime. Using eqs.(C-2) and (C-3) into eq.(C-1) then gives,

$$\mu_t = (1 - S_t)\mu' + S_t\mu + S_t\sigma_\eta\mu_t^* \quad (\text{C-4})$$

This specification nests a number of models. First, if  $\sigma_\eta = 0$ , the consumption equation does not contain an unobserved integrated process but includes a Markov switching intercept along the lines of Bianchi et al. (2017). Second, if  $\mu = \mu'$ , the unobserved component is given by  $\mu_t = \mu + S_t\sigma_\eta\mu_t^*$ . This model is similar in spirit to a mixture innovation model along the lines of McCulloch and Tsay (1993), i.e., if the variable  $S_t = 0$  then  $\mu_t = \mu$  and the unobserved component is constant while if  $S_t = 1$  then  $\mu_t = \mu + \sigma_\eta\mu_t^*$  and the unobserved component follows a random walk. Third, if  $S_t = 1 (\forall t)$ , the model collapses to the model discussed in Sections 5.1.1 and 5.2.2 of the main text, i.e., with an unobserved component given by  $\mu_t = \mu + \sigma_\eta\mu_t^*$ . Upon estimation of the state space model presented in the text with a more general specification for the unobserved component  $\mu_t$  as given by eq.(C-4), we do indeed estimate  $S_t = 1 (\forall t)$  while we find  $\sigma_\eta \neq 0$  so that this model collapses to the model presented in the main text of this paper. We do not present the estimation details nor the estimation results for this model but these are available from the authors upon request.

## Appendix D Gibbs sampler

This appendix, first, briefly presents the steps of the Gibbs sampler in Section D.1. Then, the technical details are discussed in Section D.2. Finally, a convergence analysis is provided in Section D.3.

## D.1 General outline

We collect the constant parameters in a vector  $\Gamma$ , i.e.,  $\Gamma = (\iota, \phi, \rho, \mu, \sigma_\eta, \sigma_\varepsilon^2)$ . The Gibbs approach allows us to simulate draws from the intractable joint posterior distribution of parameters  $\Gamma$  and state  $\mu^*$ , i.e.,  $f(\Gamma, \mu^* | \text{data})$ , using only tractable conditional distributions. In particular, given the prior distribution of the parameter vector  $f(\Gamma)$  and an initial draw for  $\mu^*$  taken from its prior distribution, the following steps are implemented:

1. Sample the constant parameters  $\Gamma$  conditional on the unobserved state  $\mu^*$  and the data
  - (a) Sample the binary indicator  $\iota$  marginalizing over the parameter  $\sigma_\eta$  for which variable selection is carried out (see Frühwirth-Schnatter and Wagner, 2010).
  - (b) If  $\iota = 1$ , sample the parameters  $\phi, \rho, \mu, \sigma_\eta, \sigma_\varepsilon^2$ . If  $\iota = 0$ , sample the parameters  $\phi, \rho, \mu$  and  $\sigma_\varepsilon^2$ . In the latter case, we set  $\sigma_\eta = 0$ .
2. Sample the unobserved state  $\mu^*$  conditional on the constant parameters  $\Gamma$  and the data. To this end, if  $\iota = 1$ , we use the multimove sampler for state space models of Carter and Kohn (1994) (see also Kim and Nelson, 1999). If  $\iota = 0$ , we draw  $\mu^*$  from its prior distribution. To exploit the non-identification of the non-centered specification discussed in Section 5.2.2, we conduct a random sign switch on  $\mu^*$  and  $\sigma_\eta$ , i.e., with probability 0.5 we multiply both by  $-1$  and with probability 0.5 we leave both unaltered.

These steps are iterated  $J$  times and in each iteration  $\Gamma$  and  $\mu^*$  are sampled. After a number of burn-in draws  $B$ , the sequence  $(B + 1, \dots, J)$  of draws of  $\Gamma$  and  $\mu^*$  approximates a sample from the posterior distributions of  $\Gamma$  and  $\mu^*$ . The results reported below are based on  $J = 20.000$  iterations with the first  $B = 10.000$  draws discarded as a burn-in sequence, i.e., the reported results are all based on posterior distributions constructed from  $J - B = 10.000$  draws. Note further that from the distribution of the binary indicator  $\iota$ , we calculate the posterior probability that there is an unobserved integrated component in regression eq.(10) as the fraction of  $\iota$ 's that are equal to 1 over the  $J - B$  draws of the Gibbs sampler.

## D.2 Details on the steps of the sampler

### D.2.1 Sample the constant parameters $\Gamma$

The parameters contained in  $\Gamma$  can be sampled from a standard regression model,

$$y = w^r \gamma^r + \chi \tag{D-1}$$

where  $y$  is a  $T \times 1$  vector containing  $T$  observations on the dependent variable,  $w$  is a  $T \times M$  matrix containing  $T$  observations of  $M$  predictor variables,  $\gamma$  is the  $M \times 1$  parameter vector and  $\chi$  is the  $T \times 1$  vector of error terms for which  $\chi \sim iid\mathcal{N}(0, \sigma_\chi^2 I_T)$ . If the binary indicators  $\kappa$  are equal to 1 then the restricted parameter vector  $\gamma^r$  and the corresponding restricted predictor matrix  $w^r$  are equal to  $\gamma$  respectively  $w$ . Otherwise, the restricted  $\gamma^r$  and  $w^r$  exclude those elements in  $w$  and  $\gamma$  for which the corresponding binary indicators  $\kappa$  are equal to 0. The prior distribution of  $\gamma^r$  is given by  $\gamma^r \sim \mathcal{N}(b_0^r, B_0^r \sigma_\chi^2)$  with  $b_0^r$  a  $M^r \times 1$  vector and  $B_0^r$  a  $M^r \times M^r$  matrix. The prior distribution of  $\sigma_\chi^2$  is given by  $\sigma_\chi^2 \sim \mathcal{IG}(s_0, S_0)$  with scalars  $s_0$  (shape) and  $S_0$  (scale). The posterior distributions (conditional on  $y$ ,  $w^r$ , and  $\kappa$ ) of  $\gamma^r$  and  $\sigma_\chi^2$  are then given by  $\gamma^r \sim \mathcal{N}(b^r, B^r \sigma_\chi^2)$  and  $\sigma_\chi^2 \sim \mathcal{IG}(s, S^r)$  with,

$$\begin{aligned} B^r &= [(w^r)'w^r + (B_0^r)^{-1}]^{-1} \\ b^r &= B^r [(w^r)'y + (B_0^r)^{-1}b_0^r] \\ s &= s_0 + T/2 \\ S^r &= S_0 + \frac{1}{2} [y'y + (b_0^r)'(B_0^r)^{-1}b_0^r - (b^r)'(B^r)^{-1}b^r] \end{aligned} \tag{D-2}$$

Following Frühwirth-Schnatter and Wagner (2010), we marginalize over the parameters  $\gamma$  when sampling  $\kappa$  and then draw  $\gamma^r$  conditional on  $\kappa$ . The posterior distribution of the binary indicators  $\kappa$  is obtained from Bayes' theorem as,

$$p(\kappa|y, w, \sigma_\chi^2) \propto p(y|\kappa, w, \sigma_\chi^2)p(\kappa) \tag{D-3}$$

where  $p(\kappa)$  is the prior distribution of  $\kappa$  and  $p(y|\kappa, w, \sigma_\chi^2)$  is the marginal likelihood of regression eq.(D-1) where the effect of the parameters  $\gamma$  has been integrated out. We refer to Frühwirth-Schnatter and Wagner (2010) (their eq.(25)) for the closed-form expression of the marginal likelihood for the regression model of eq.(D-1).

### Sample the binary indicator $\iota$

Our regression has one binary indicator  $\iota$ , so  $\kappa = \iota$ . We sample  $\iota$  by calculating the marginal likelihoods  $p(y|\iota = 1, w, \sigma_\chi^2)$  and  $p(y|\iota = 0, w, \sigma_\chi^2)$  (see Frühwirth-Schnatter and Wagner, 2010, for the correct expressions). Upon combining the marginal likelihoods with the Bernoulli prior distributions of the binary indicators  $p(\iota = 1) = p_0$  and  $p(\iota = 0) = 1 - p_0$ , the posterior distributions  $p(\iota = 1|y, w, \sigma_\chi^2)$  and  $p(\iota = 0|y, w, \sigma_\chi^2)$  are obtained from which the probability  $prob(\iota = 1|y, w, \sigma_\chi^2) = \frac{p(\iota=1|y, w, \sigma_\chi^2)}{p(\iota=1|y, w, \sigma_\chi^2) + p(\iota=0|y, w, \sigma_\chi^2)}$  is calculated which is used to sample  $\iota$ , i.e., draw a random number  $r$  from a uniform distribution with support between 0 and 1 and set  $\iota = 1$  if  $r < prob(\cdot)$  and  $\iota = 0$  if  $r > prob(\cdot)$ .



## Sample the other parameters in $\Gamma$

We then sample the regression coefficients  $\phi$ ,  $\rho$ ,  $\mu$  and  $\sigma_\eta$  and the regression error variance  $\sigma_\varepsilon^2$  conditional on  $\iota$ , the data and the unobserved component  $\mu_t^*$ . The dependent variable is  $y = c$  where  $c$  is the  $T \times 1$  vector containing consumption  $c_t$  stacked over time while the error term is  $\chi = \varepsilon$  with  $\varepsilon$  containing  $\varepsilon_t$  stacked over time and where the variance is given by  $\sigma_\chi^2 = \sigma_\varepsilon^2$ . When  $\iota = 1$ , we have  $w^r = w = \begin{bmatrix} x & \Delta x_{-p} & \dots & \Delta x_{+p} & e & \mu^* \end{bmatrix}$  and  $\gamma^r = \gamma = \begin{bmatrix} \phi' & \rho'_{-p} & \dots & \rho'_{+p} & \mu & \sigma_\eta \end{bmatrix}'$  where  $e$  is a  $T \times 1$  vector of ones and  $\mu^*$  is a  $T \times 1$  vector containing  $\mu_t^*$  stacked over time. With  $x$  and every  $\Delta x_j$  (for  $j = -p \dots +p$ ) being  $T \times K$  matrices then  $\phi$  and every  $\rho_j$  are  $K \times 1$  vectors and we have  $M = K(2p+2) + 2$ . When  $\iota = 0$ , we have  $w^r = \begin{bmatrix} x & \Delta x_{-p} & \dots & \Delta x_{+p} & e \end{bmatrix}$  and  $\gamma^r = \begin{bmatrix} \phi' & \rho'_{-p} & \dots & \rho'_{+p} & \mu \end{bmatrix}'$ . In this case, we have  $M^r = K(2p+2) + 1$ . Once the matrices of eq.(D-1) are determined, the parameters  $\gamma^r$  and  $\sigma_\chi^2$  can be sampled from the Gaussian posterior distributions given above with the prior distributions as specified in Table 2 in the text.<sup>19</sup>

### D.2.2 Sample the unobserved state $\mu^*$

If  $\iota = 0$ , the unobserved component is drawn from its prior distribution. In particular,  $\mu_t^*$  is drawn from eq.(15), i.e., as a cumulative sum of standard normally distributed shocks  $\eta_t^*$  so  $\mu_t^* = \sum_{s=1}^t \eta_s^*$ . If  $\iota = 1$ , the unobserved component  $\mu_t^*$  is sampled conditional on the constant parameters and on the data using a state space approach. In particular, we use the forward-filtering backward-sampling approach discussed in detail in Kim and Nelson (1999) to sample the unobserved state. The general form of the state space model is given by,

$$Y_t = ZS_t + V_t, \quad V_t \sim iid\mathcal{N}(0, H), \quad (\text{D-4})$$

$$S_t = TS_{t-1} + KE_t, \quad E_t \sim iid\mathcal{N}(0, Q), \quad (\text{D-5})$$

$$S_0 \sim iid\mathcal{N}(s_0, P_0), \quad (\text{D-6})$$

(where  $t = 1, \dots, T$ ) with observation vector  $Y_t$  ( $n \times 1$ ), state vector  $S_t$  ( $n^s \times 1$ ), error vectors  $V_t$  ( $n \times 1$ ) and  $E_t$  ( $n^{ss} \times 1$  with  $n^{ss} \leq n^s$ ) that are assumed to be serially uncorrelated and independent of each other, and with the system matrices that are assumed to be known (conditioned upon) namely  $Z$  ( $n \times n^s$ ),  $T$  ( $n^s \times n^s$ ),  $K$  ( $n^s \times n^{ss}$ ),  $H$  ( $n \times n$ ),  $Q$  ( $n^{ss} \times n^{ss}$ ) and the mean  $s_0$  ( $n^s \times 1$ ) and variance  $P_0$  ( $n^s \times n^s$ ) of the initial state vector  $S_0$ . As eqs. (D-4)-(D-6) constitute a linear Gaussian state space model, the unknown state variables in  $S_t$  can be filtered using the standard Kalman filter. Sampling  $S = [S_1, \dots, S_T]$  from its

<sup>19</sup>From the specification of the prior distributions in Table 2, we note that  $s_0 = \nu_0 T = 0.01T$ ,  $S_0 = \nu_0 T \sigma_0^2 = 0.01 \times T \times 0.01$  and that  $b_0^r$  is a  $M^r \times 1$  vector of zeros. Further,  $B_0^r$  is an  $M^r \times M^r$  diagonal matrix containing as elements the variances 0.1 (for parameter  $\sigma_\eta$ ) or 1 (for all other regression parameters) - i.e., the variable  $V_0$  in Table 2 - divided by the prior belief for  $\sigma_\varepsilon^2$  - i.e., the variable  $\sigma_0^2$  in Table 2.

conditional distribution can then be done using the multimove Gibbs sampler of Carter and Kohn (1994). Given our state space system presented in eqs.(10), (11), (13), (14) and (15) we have  $n = n^s = n^{ss} = 1$ . The matrices are then given by  $Y_t = c_t - x_t\phi - \mu - \sum_{j=-p}^p \Delta x_{t+j}\rho_j$ ,  $Z = \sigma_\eta$ ,  $S_t = \mu_t^*$ ,  $V_t = \varepsilon_t$ ,  $H = \sigma_\varepsilon^2$ ,  $T = 1$ ,  $K = 1$ ,  $E_t = \eta_t^*$ ,  $Q = 1$ ,  $s_0 = \mu_0^* = 0$  and  $P_0 = 10^{-6}$ .

### D.3 Convergence analysis

We analyse the convergence of the MCMC sampler using the simulation inefficiency factors as proposed by Kim et al. (1998) and the convergence diagnostic of Geweke (1992) for equality of means across subsamples of draws from the Markov chain (see Groen et al., 2013, for a similar convergence analysis).

For each fixed parameter and for every point-in-time estimate of the unobserved component, we calculate the inefficiency factor as  $IF = 1 + 2 \sum_{l=1}^m \kappa(l, m) \hat{\theta}(l)$  where  $\hat{\theta}(l)$  is the estimated the  $l$ -th order autocorrelation of the chain of retained draws and  $\kappa(l, m)$  is the kernel used to weigh the autocorrelations. We use a Bartlett kernel with bandwidth  $m$ , i.e.,  $\kappa(l, m) = 1 - \frac{l}{m+1}$  where we set  $m$  equal to 4% of the retained sampler draws  $J - B = 10.000$  (see Section D.1 above). If we assume that  $n$  draws are sufficient to cover the posterior distribution in the ideal case where draws from the Markov chain are fully independent, then  $n \times IF$  provides an indication of the minimum number of draws that are necessary to cover the posterior distribution when the draws are not independent. Usually,  $n$  is set to 100. Then, for example, an inefficiency factor equal to 20 suggests that we need at least 2.000 draws from the sampler for a reasonably accurate analysis of the parameter of interest. Additionally, we also compute the p-values of the Geweke (1992) test which tests the null hypothesis of equality of the means of the first 20% and last 40% of the retained draws obtained from the sampler for each fixed parameter and for every point-in-time estimate of the unobserved component. The variances of the respective means are calculated using the Newey and West (1987) robust variance estimator using a Bartlett kernel with bandwidth equal to 4% of the respective sample sizes.

In Table D-1, we present the convergence analysis corresponding to the results reported in Table 4 in the text. The convergence results are reported for individual parameters or for groups of parameters. Groups are considered when the parameters can be meaningfully grouped which is the case for the  $k$  elasticity parameters in  $\phi$  (with  $k = 2$  or  $k = 3$  depending on whether  $x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$  or  $x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$ ), for the  $k \times (p+1)$  parameters  $\rho$  of the DOLS specification of the stationary component  $v_t$  (where, given  $p = 6$ , we have 26 or 39 parameters depending again on whether  $x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$  or  $x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$ ), and for the unobserved component  $\mu^*$  which is a state, i.e., a time series of length  $T = 248$ . In both tables, we report statistics of the distributions of the inefficiency factors for every parameter or parameter group, i.e., median, minimum, maximum, and - for the state  $\mu^*$  - the 5% and

10% quantiles. Obviously, these statistics are identical for the non-grouped parameters. The tables also report the rejection rates of the Geweke tests conducted both at the 5% and 10% levels of significance. These rates are equal to the number of rejections of the null hypothesis of the test per parameter group divided by the number of parameters in a parameter group. These rates can only be 0 or 1 for individual (non-grouped) parameters but can lie between 0 and 1 for the grouped parameters.

The calculated inefficiency factors suggest that the MCMC sampler performs well and that all parameters are well converged using our retained 10,000 draws. In fact, an accurate analysis could have been conducted with less than 10,000 draws. From the tables, we note that more draws are required when the integrated unobserved component (UC) is included in the model and estimated, i.e., for cases where  $\iota = 1$ , while the inefficiency factors are all close to 1 when the estimated model is a standard regression without unobserved component, i.e, when  $\iota = 0$ . When  $\iota = 1$ , the most draws are required to estimate the posterior distribution of the initial values  $\mu$  of the unobserved component. That this parameter is somewhat harder to estimate is not surprising and is also clear from inspection of the wide 90% HPD interval surrounding its posterior mean as for instance depicted in Figure 2 (for the case corresponding to the convergence results in Table D-1 with  $x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$  and  $\iota = 1$ ). Our findings for the inefficiency factors are corroborated by the results for the Geweke (1992) test for equality of means across subsamples of the retained draws. The rejection rates reported in the tables are, with few exceptions, equal to 0 and therefore strongly suggest that the means of the first 20% and last 40% of the retained draws are equal. In a few instances, high rejection rates are observed, in particular again for the parameter  $\mu$  and sometimes also for the elasticities  $\phi$ . We argue that these high rejection rates are due to the particular sample of draws and are not indicative of non-convergence as these rejection rates are not withheld when we rerun the sampler using another seed. Hence, in general, we can conclude that the convergence of the sampler for the retained number of draws is satisfactory.

**Table D-1:** Inefficiency factors and convergence diagnostics (results Table 4)

Regressors	UC	Parameters	Number	Inefficiency factors (Stats distribution)					Convergence (Rejection rates)	
				Median	Min	Max	5%	10%	5%	10%
$x_t = \begin{bmatrix} a_t & y_t \end{bmatrix}$	$\iota = 1$	$\phi$	2	16.15	9.12	23.18	-	-	0.00	0.00
		$\mu$	1	55.42	55.42	55.42	-	-	0.00	0.00
		$ \sigma_\eta $	1	28.68	28.68	28.68	-	-	0.00	0.00
		$\sigma_\varepsilon^2$	1	1.13	1.13	1.13	-	-	0.00	0.00
		$\rho$	26	1.44	1.02	1.86	-	-	0.00	0.08
		$\mu^*$	248	1.07	0.83	1.16	0.92	1.12	0.00	0.02
	$\iota = 0$	$\phi$	2	0.83	0.83	0.83	-	-	0.00	0.00
		$\mu$	1	0.78	0.78	0.78	-	-	0.00	0.00
		$ \sigma_\eta $	1	-	-	-	-	-	-	-
		$\sigma_\varepsilon^2$	1	1.03	1.03	1.03	-	-	0.00	0.00
		$\rho$	26	0.95	0.81	1.18	-	-	0.04	0.08
		$\mu^*$	248	-	-	-	-	-	-	-
$x_t = \begin{bmatrix} a_t^s & a_t^{ns} & y_t \end{bmatrix}$	$\iota = 1$	$\phi$	3	23.96	17.30	26.83	-	-	0.67	0.67
		$\mu$	1	27.34	27.34	27.34	-	-	1.00	1.00
		$ \sigma_\eta $	1	15.93	15.93	15.93	-	-	0.00	0.00
		$\sigma_\varepsilon^2$	1	1.12	1.12	1.12	-	-	0.00	0.00
		$\rho$	39	1.83	0.95	4.02	-	-	0.08	0.15
		$\mu^*$	248	0.93	0.84	1.14	0.88	1.07	0.00	0.00
	$\iota = 0$	$\phi$	3	1.09	1.05	1.17	-	-	0.00	0.00
		$\mu$	1	0.86	0.86	0.86	-	-	0.00	0.00
		$ \sigma_\eta $	1	-	-	-	-	-	-	-
		$\sigma_\varepsilon^2$	1	0.91	0.91	0.91	-	-	0.00	0.00
		$\rho$	39	0.94	0.80	1.25	-	-	0.05	0.10
		$\mu^*$	248	-	-	-	-	-	-	-

Notes: The convergence analysis corresponds to the results reported in Table 4. The statistics of the distribution of the inefficiency factors are presented in columns 5 to 9 for every parameter or group of parameters. These statistics are identical when parameters are considered individually as only one inefficiency factor is calculated in these cases. The inefficiency factors are calculated for every fixed parameter and for every point-in-time estimate of the unobserved component using a Bartlett kernel with bandwidth equal to 4% of the 10.000 retained sampler draws. The rejection rates of the Geweke (1992) test conducted at the 5% and 10% levels of significance are reported in columns 10 and 11. These rates are equal to the number of rejections of the null hypothesis of the test per parameter group divided by the number of parameters in a parameter group. These rates are either 1 or 0 for parameters that are considered individually. They are based on the p-value of the Geweke test of the hypothesis of equal means across the first 20% and last 40% of the 10.000 retained draws which is calculated for every fixed parameter and for every point-in-time estimate of the unobserved component. The variances of the respective means in the Geweke (1992) test are calculated with the Newey and West (1987) robust variance estimator using a Bartlett kernel with bandwidth equal to 4% of the respective sample sizes.