The Decay of *cay*

Moritz Dauber & Jochen Lawrenz*

August, 2023

Abstract

We revisit the ability of different versions of the consumption-wealth ratio (*cay*) to predict stock market returns and show that forecasting power has declined over at least the last decade up to the point that it is neither in-sample, out-of-sample, nor economically significant. We uncover that the loss in predictability goes along with a structural shift in the underlying cointegrating relationship. Over the past decades, the development of asset wealth is increasingly detached from consumption, which makes it unlikely that a predictor derived from the representative agent's intertemporal budget constraint can capture stock market behavior.

JEL classification: E21, E44, G10

Keywords: consumption-wealth ratio, intertemporal budget constraint, cointegration, predictive regression

*Moritz Dauber, Department of Banking & Finance, Innsbruck University, A-6020 Innsbruck, Tel.:+43.512.507.73108, email: moritz.dauber@uibk.ac.at, Corresponding author

Jochen Lawrenz, Department of Banking & Finance, Innsbruck University, A-6020 Innsbruck, Tel.:+43.512.507.73110, email: jochen.lawrenz@uibk.ac.at

1 Introduction

A basic tenet of neoclassical financial economics is that the representative agent maximizes utility over consumption under some budget constraint. Following this logic and building on ideas in Campbell and Mankiw (1989), Lettau and Ludvigson (2001) exploit the economic relations of the intertemporal budget constraint between wealth, consumption and expected returns to develop the consumption-wealth ratio (*cay*) as a predictor for stock market excess returns in the time-series. However, although being theoretically appealing, it has received ambiguous empirical support, where conflicting results are largely due to the stance, the researchers take with respect to in-sample and out-of-sample tests.

In this paper, we revisit the forecasting ability of *cay* and offer three contributions: *First*, we conduct comprehensive in-sample, out-of-sample, as well as economic significance tests on U.S. post-war data and find a remarkable and sustained decline of the predictive power of *cay* during (at least) the last decade. From the perspective of most recent data, *cay* appears to have lost even its in-sample ability to predict stock market returns. We conclude that, independent of how to assess predictability, *cay* does not reliably predict stock market returns. *Second*, we propose two improvements for the construction of *cay*: On the one hand, we de-filter the consumption time series to recover richer time variation, and on the other hand, we derive *cay* from the data for the top-10% richest households in an attempt to better identify the marginal investor. We find some modest improvements in forecasting power from our extensions, but document still the same pattern of declining predictability in most recent times. *Third*, our analysis uncovers that the declining predictive power of *cay* goes along with a structural shift in its presumed cointegrating relationship. We find that the coefficient on wealth in the cointegrating vector shows a sustained decline over at least the last three decades, both in terms of value and significance up to the point that it ends up insignificant from the perspective of most recent data. Furthermore, the econometric evidence to reject the null of no cointegration has weakened substantially, so that there is currently no reason to believe that a stable cointegrating relation exists in the first place. We argue that these structural shifts are largely responsible for the *decay* in the forecasting ability of *cay* and explain more generally the strong discrepancy between the in-sample (IS) and (strict) out-of-sample (OOS) performance of *cay*.

On the more general conceptual level, our recent evidence suggests that the theoretically motivated belief that consumption and wealth is cointegrated because the representative agent optimally adjusts consumption to future wealth expectations appears not well supported. In contrast, in particular during the last decades, the development of wealth appears increasingly detached from consumption data. In this respect, our results confirm and complement the findings of Lustig et al. (2013), who estimated that only a tiny fraction of total wealth is due to stock market wealth and therefore concluded that the time variation in wealth is largely disconnected to aggregate stock market behavior.

Even more than twenty years after its formulation in Lettau and Ludvigson (2001), the assessment of the predictive ability of *cay* still tends to be somewhat ambiguous. While Guo (2006), Guo et al. (2013), Hahn and Lee (2006), Kalotay et al. (2007), Della Corte et al. (2010), and more recently Ren et al. (2014) and Lettau and Ludvigson (2019) confirm the strong and robust forecasting power of *cay*, Avramov (2002) and later Welch and Goyal (2008) argue that the construction of *cay* suffers from a 'look-ahead bias' when the cointegrating vector is estimated using the full sample and pointed towards the weak out-of-sample performance when the test is carried out under the assumption that only available information is used for prediction. Brennan and Xia (2005) address the same 'look-ahead bias' and argue that a cointegrating relationship between wealth and a simple time trend would generate similar (or even better) forecasting results. The conflicting results are due to different conceptions of how to conduct OOS tests for the residual of a cointegrated relationship, and more generally due to the substantial model uncertainty for cointegration test as pointed out by e.g., Koop et al. (2008). While Avramov (2002), Welch and Goyal (2008), and Brennan and Xia (2005) favor the perspective of a real-time decision-maker and only allow for contemporaneous available information, Lettau and Ludvigson (2001) and more explicitly Lettau and Ludvigson (2005b) do reply to this criticism by making the strong case, that using only partial information to estimate the equilibrium cointegrating vector is inappropriate because it neglects information that would have been available to the decision-maker if the underlying theory is correct.

A second line of disagreement in the assessment of the forecasting power comes from the debate between in-sample and out-of-sample tests. While it is generally believed that OOS results are more reliable evidence for predictability, e.g., Inoue and Kilian (2005) are often credited for showing that OOS is not necessarily superior and has rather low power to detect predictability. So, given its strong in-sample evidence, weak OOS performance is not necessarily a reason to dismiss the forecasting ability of *cay*. However, it is remarkable that *cay* tends to display an unusually large discrepancy between IS and (strict) OOS performance as documented by e.g. Welch and Goyal (2008).

Finally, a third, well-established approach in the literature is to assess the economic performance of forecasts by estimating the utility gain of a mean-variance investor in the spirit of Fleming et al. (2001) and Campbell and Thompson (2008). In this sense, Andrade et al. (2006) and Della Corte et al. (2010) provide evidence which shows that, despite its poor statistical OOS results, *cay* performs well in economic terms.

In our analysis, we take a comprehensive look at the performance of *cay* from all three perspectives with U.S. post-war data from 1952:Q1 up to 2019:Q4. We deliberately end our sample for baseline results before the Covid-19 episode and provide separate analyses which extend up to 2022:Q4. Our main finding is to show that independent of the way how you look at the assessment of *cay*, it shows a remarkable decline over at least the last decade to end up insignificant from the perspective of IS, OOS and economic performance. Figure 1 illustrates a preview on our main point. In the lower panel, we show (in-sample) predictive regression coefficients and utility gains (in bp) for the original *cay*-measure. Beginning with at least 2010 both performance metrics decline to end up insignificant from the perspective of 2019:Q4, which is in line with the recent evidence of Goyal et al. (2021). In the upper

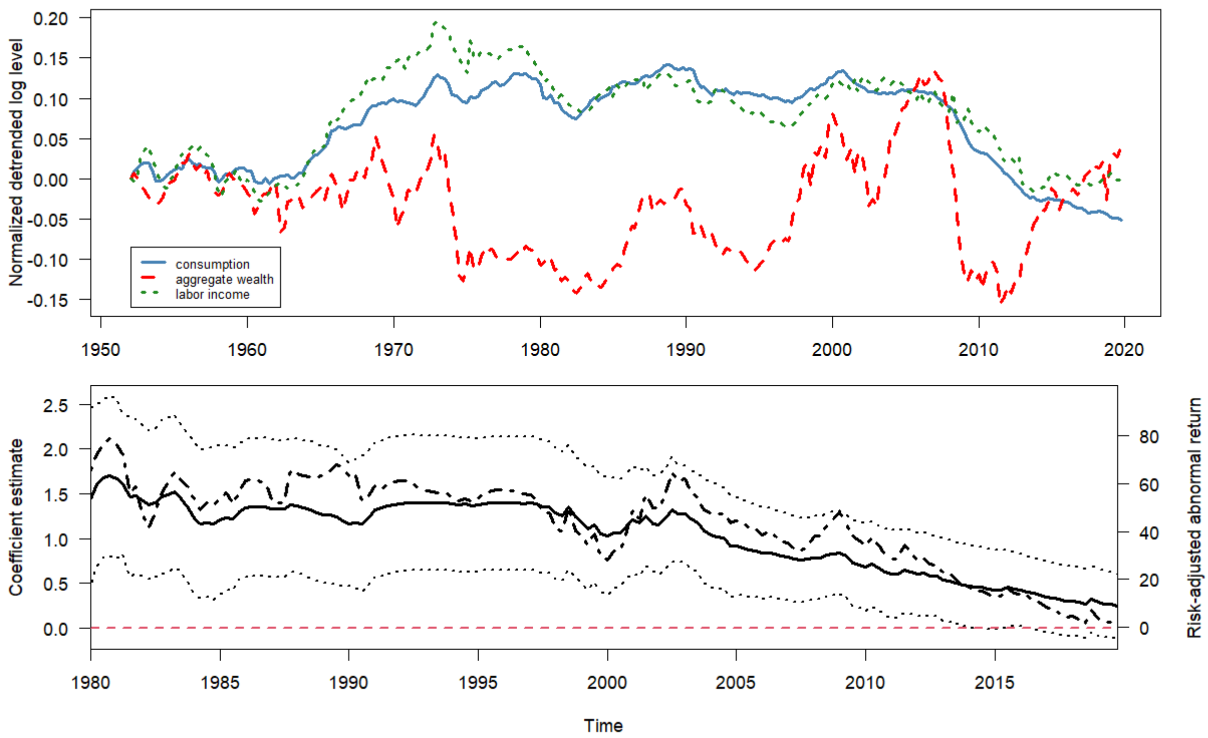


Figure 1: The upper panel shows consumption (blue), aggregate wealth (red) and labor income (green) in normalized detrended log levels. This means that the log series are detrended assuming a simple linear time trend and afterwards are normalized to start at zero.
The lower panel shows the coefficient estimate of cay_t (solid line) from quarterly forecasts of excess returns on the CRSP NYSE/NYSE MKT/NASDAQ/Arca Value-Weighted Market Index, accompanied by its 90%-confidence band as the dotted lines. The dashed line represents the risk-adjusted abnormal return of the $\mathcal{CAY}$-strategy against the $\mathcal{AVE}$-strategy. Fore more details see Section 4. The sample period is 1952:1-2019:4.

panel, we display the input data to the construction of *cay* which is the log of consumption, income and asset wealth after removing the deterministic time trend. We do observe a strong comovement between consumption and income over the entire sample. In contrast, asset wealth displays a strongly detached development from consumption and income, which is particularly pronounced since around 2000, and manifests itself as a structural shift in the cointegrating relationship where the coefficient on asset wealth declines steadily. Taking the claim of Lettau and Ludvigson (2005b) at face value to use the most comprehensive (i.e. most recent) data set for the estimation of the equilibrium leaves us with a time series for *cay* that would have never displayed forecasting power neither in- nor out-of-sample.

Given that *cay* has a rigorous theoretical foundation, the current disappointing empirical evidence also casts doubt if it is a sufficient description of agents' behavior. Recall that *cay* is supposed to be predictive due to the optimizing behavior of the representative agent who adjusts consumption to expected wealth changes. In an attempt to still reconcile theory with observations, we propose two novel approaches to construct *cay*. Our first extension is guided towards better identifying the representative agent. Against the background of significant wealth and income inequality, it is debatable if market aggregates are able to represent the behavior of the representative investor. Thus, inspired by recent work of Lettau et al. (2019), who propose capital share risk as a surprisingly successful proxy for the pricing kernel, we construct a measure of *cay* from the data of the 10% richest households. Given that rich households tend to hold a disproportionally large fraction of stock market wealth relative to their consumption, we hypothesize that *cay* should be more predictive if the theory is appropriate.[1]

The second extension builds upon the insights from Savov (2011) and Kroencke (2017). Both contributions address the well-known equity premium puzzle, which can be restated as the fact that the consumption time series displays too little variation. While Savov (2011) uses garbage data to recover richer time variation, Kroencke (2017) removes the filtering procedure which is implicit in the construction of the official data release. Since Savov (2011) as well as Kroencke (2017) show that their procedure allows to better fit risk premia, we hypothesize that their approach may also benefit the forecasting ability of *cay*.

We find that unfiltering *cay* does not improve or even worsens the predictive performance. The time variation which can be recovered from the procedure has too high frequency as to be able to favorably impact the forecasts. With respect to the top-10% *cay*, we do find some modest improvements. While being almost identical to the original *cay*-measure up to around 2000, we find a slightly better predictive performance post-2000, yielding some evidence that the increasing wealth inequality does impact the forecasting ability. However, although the decline in *cay* appears less pronounced, we still find the same decaying pattern up to the point that the top-10% *cay* is (in-sample) only significant at the the 10%-level from most recent data.

Our assessment of the predictive power of *cay* is clearly in contrast to the original work of Lettau and Ludvigson (2001) as well as their update in Lettau and Ludvigson (2019). It is also in contrast to the earlier works of Hahn and Lee (2006), Della Corte et al. (2010) and

[1] Lustig et al. (2013) estimate that in the market *aggregate*, stock market wealth represents only 1% of total wealth. It is reasonable to expect that among rich households, this fraction will be substantially larger.

others, while being in line with the most recent assessment of Goyal et al. (2021)[2]. Therefore, it is instructive to analyze in more detail the origin of the discrepancy of earlier results to our current assessment. This analysis focuses on the stability of the cointegrating relationship and shows that – as mentioned above – sustained shifts have taken place, in particular with respect to the coefficients of the cointegrating vector. A related aspect is the role of a deterministic time trend in the cointegration equation. Brennan and Xia (2005) and later Hahn and Lee (2006) did already point towards the potentially critical role of an omitted time trend. In particular, Hahn and Lee (2006) have shown that the performance of *cay* is biased upwards when the cointegrating relationship is restricted to have no time trend. Although substantially smaller, their results still find forecasting power for *cay* when estimated with a time trend. Our findings are even stronger, in that *cay* is actually uninformative over the entire sample when estimated with a time trend. However, we point out that from the perspective of most recent data, the inclusion of a time trend is econometrically no longer justified and rather confirms the original assumption of Lettau and Ludvigson (2001). In this sense, we document some irony of history. Up until around 2010, the original *cay* measure displayed stable forecasting power, but was subject to the potential criticism that it earned its predictability due to the omitted (but statistically significant) time trend. Since 2010, the data does support the assumption of omitting a time trend, but *cay* no longer displays forecasting power.

A further implication of the sustained shift in the cointegrating vector is the fact that *cay* tends to be always biased downwards in the most recent observations, which follows from the fact that the fixed coefficient estimates are estimated over the full observations and will therefore overweigh asset wealth in the latest periods. This finding is able to explain the strong discrepancy between the in-sample and OOS performance results of *cay*.

Finally, we document that the inclusion of the Covid-19 episode, i.e., the extension of the data sample to 2022:Q4 would be further evidence for the instability of the cointegrating relationship, but which is hard to reconcile with the idea of *cay*, since consumption dropped due to pandemic-induced shutdowns, while income soared due to massive government transfer payments.

We complement our main analysis with comprehensive robustness checks. We test various cointegration regression approaches. We use PCE consumption data instead of the more common NDS data. We implement the improvement suggested by Sousa (2010) to disentangle financial wealth, which they labelled *cday*. We follow the suggestion of Guo (2006) to augment the predictive regression by stock market volatility to avoid an omitted variable bias. We also take up the argument by Lettau and Ludvigson (2005a) and Lettau and Ludvigson (2005b) to avoid estimating the cointegrating relationship and rather conduct a multivariate regression. None of these tests change our main conclusions that the last decades have seen a sustained *decay of cay*.

[2] Goyal et al. (2021) assess a large number of predictors and do not analyze or discuss their negative evidence for *cay* in any detail.

The remainder of the article is structured as follows: Section 2 recaps the economic theory behind the consumption-wealth ratio and the estimation of *cay*. Section 3 continues in describing our data and the three different specifications of *cay* before presenting the results from the estimation of the cointegrating relationship(s). Section 4 documents the decay of *cay* from an in-sample, out-of-sample and economic perspective. In Section 5, we address the structural shift in the cointegrating relationship including the role of a deterministic time trend as well as the impact of the Covid-19 pandemic on *cay*. Section 6 provides various robustness tests before Section 7 concludes.

2 The consumption-wealth ratio

2.1 Theoretical background

Campbell and Mankiw (1989) assume a representative agent economy in which the consumer's intertemporal budget constraint is given by

$$W_{t+1} = R_{w,t+1} \left(W_t - C_t \right) \tag{1}$$

where W_t denotes aggregate wealth, C_t denotes aggregate consumption and $R_{w,t+1}$ is the return on aggregate wealth between periods t and $t+1$. Under the assumption that the consumption-wealth ratio is stationary, Campbell and Mankiw (1989) show that via a first-order Taylor series expansion

$$\Delta w_{t-1} \approx k + r_{w,t+1} + (1 - 1/\rho)(c_t - w_t) \tag{2}$$

where Δ denotes the difference operator, c_t log consumption, w_t log wealth, $r_{w,t} = \log(1 + R_{w,t})$ the log return on aggregate wealth, k a constant and ρ the steady-state value of invested wealth to total wealth, i.e., $(W - C)/W$. Using the fact that

$$\Delta w_{t+1} = \Delta c_{t+1} + (c_t - w_t) - (c_{t+1} - w_{t+1}),$$

transforms (2) into

$$c_t - w_t \approx \rho(r_{w,t+1} - \Delta c_{t+1}) + \rho(c_{t+1} - w_{t+1}) + \rho k. \tag{3}$$

After omitting the unimportant constant k, solving (3) forward and imposing the transversality condition that $\lim_{i\to\infty} \rho(c_{t+i} - w_{t+i}) = 0$ together with taking expectations on both sides of the equation, Campbell and Mankiw (1989) derive

$$c_t - w_t = E_t \sum_{i=1}^{\infty} \rho^i (r_{w,t+i} - \Delta c_{t+i}). \tag{4}$$

This means, the consumption-wealth ratio is a function of future expected returns and future expected consumption growth and must thus forecast either one of them or both. This crucial linkage allows us to infer expectations on future returns from the consumption-wealth ratio using observable consumption data.

A direct observation of the consumption-wealth ratio, however, is impossible due to the inability to observe the component of human capital, H_t, in total wealth. We therefore follow the approach by Lettau and Ludvigson (2001) of using log labor income, y_t, as a proxy for the nonstationary component of log human capital, h_t. This means we first approximate $W_t = A_t + H_t$ via

$$w_t \approx \alpha_a a_t + (1 - \alpha_a) h_t \tag{5}$$

where a_t denotes log asset wealth and α_a represents the average share of asset holdings in total wealth, i.e., A/W. We then replace h_t with $\kappa + y_t + z_t$ where κ denotes an unimportant constant and z_t is a mean zero stationary random variable. Together with the approximation

$$r_{w,t} \approx \alpha_a r_{a,t} + (1 - \alpha_a) r_{h,t} \tag{6}$$

for log returns by Campbell (1996), this transforms (4) into

$$\begin{aligned} c_t - \alpha_a a_t - (1 - \alpha_a) y_t = E_t \sum_{i=1}^{\infty} \rho^i \left([\alpha_a r_{a,t+i} + (1 - \alpha_a) r_{h,t+i}] - \Delta c_{t+i} \right) \\ + (1 - \alpha_a) z_t. \end{aligned} \tag{7}$$

Since all terms on the right-hand side are assumed to be stationary, the left-hand side must also be stationary and thus implying that c_t, a_t and y_t must be cointegrated. Hence, the left-hand side gives us the deviation in the common trend between consumption, aggregate wealth, and labor income, which, according to Lettau and Ludvigson (2001), defines *cay* as

$$cay_t := c_t - \alpha_a a_t - (1 - \alpha_a) y_t. \tag{8}$$

From (7) we get that as long as expected future returns on human capital, $r_{h,t+i}$, and consumption growth, Δc_{t+i}, are not too variable, or at least highly correlated with expected returns on assets, cay_t, as a proxy for the consumption-wealth ratio, $c_t - w_t$, should be a good predictor for market expectations on future asset returns, $r_{a,t+i}$.

Since α_a is unobservable, the next step will be to estimate the cointegrating relationship between consumption, aggregate wealth and labor income, in order to get a proxy for *cay*.

2.2 Estimation of the cointegrating relationship

Under the assumption of the presence of a (single) cointegrating relationship between given consumption, wealth and labor income series, our goal is to estimate the cointegrating parameters, i.e., to obtain an estimate for α_a from equation (8).

Our baseline method is Stock and Watson's (1993) dynamic least squares (DLS) technique, also applied by Lettau and Ludvigson (2001), that adds leads and lags of the first differences in aggregate wealth and labor income to an otherwise normal ordinary least squares (OLS) regression of consumption on aggregate wealth and labor income. We follow Lettau and Ludvigson (2001) in taking eight leads and lags into account but we also find that the exact number, between one and eight, does not impact the results substantially. Thus, we estimate

$$c_t = \alpha + \beta_a a_t + \beta_y y_t + \sum_{i=-8}^{8} b_{1,i} \Delta a_{t-i} + \sum_{i=-8}^{8} b_{2,i} \Delta y_{t-i} + \varepsilon_t \tag{9}$$

where Δ again denotes the difference operator.

The main advantage of the DLS specification over the simple OLS specification is that it eliminates the effects of regressor endogeneity. Despite that, Lettau and Ludvigson (2001) point out that the estimates of β_a and β_y will be consistent, even though the error term ε_t will typically be correlated with a_t and y_t. This comes from the fact that OLS estimates of cointegrating parameters converge to the true parameters at a faster rate than in normal estimations (Stock (1987)). Despite using the DLS estimates throughout the paper, we also report the cointegrating parameters and forecasts based on them for other estimation methods, next to OLS including Park's (1992) canonical cointegration regression (CCR) and Phillips and Hansen's (1990) fully modified estimator (FME), in Appendix A.1.

With the estimated parameters $\widehat{\beta}_a$ and $\widehat{\beta}_y$ from (9), we then define

$$\widehat{cay}_t := c_t - \widehat{\beta}_a a_t - \widehat{\beta}_y y_t, \tag{10}$$

as the 'residual' from the cointegrating regression. It is the estimated trend deviation in the long-run relationship between consumption, aggregate wealth and labor income and by that a proxy for the consumption-wealth ratio.

Our baseline measure of consumption is nondurables and services excluding clothing and footwear (NDS) from the NIPA tables by the BEA. The idea behind using NDS instead of total personal consumption expenditures (PCE) is, that the former more accurately captures the flow of consumption, whereas expenditures on durable goods that are contained in PCE are rather to be seen as additions and replacements to existing stock. Lettau and Ludvigson (2001) follow Blinder et al. (1985) and Galí (1990) in assuming that total consumption, i.e., PCE, is a constant multiple of NDS. This allows them to assume a constant scale factor between log total consumption and log NDS, leading to a valid cointegrating relationship between NDS consumption, aggregate wealth and labor income. While we follow this approach and use NDS as our measure of consumption throughout the paper, it is noteworthy that in a revised note, Lettau and Ludvigson (2019) show that the assumption of PCE being a constant multiple of NDS is no longer valid. In lack of any other suitable measure for the flow of consumption, they conclude that under certain assumptions, PCE has replaced NDS as the more appropriate consumption measure. We address this issue in Section 6 and show that our main results are not substantially impacted by what measure of consumption is used.

2.3 Data

Before estimating $\widehat{cay}_t$, we briefly describe our data sources. A detailed description of how each series is derived can be found in Appendix A.2. Unless otherwise noted, we use quarterly samples spanning the period of 1952:1-2019:4.

All data on consumption and income are taken from the NIPA tables by the Bureau of Economic Analysis (BEA). The wealth series come from the Financial Accounts publication by the Board of Governors of the Federal Reserve System (FED).

The data on income and wealth inequality come from the World Inequality Database (WID). Following Lettau et al. (2019), our measure for the labor share is the labor share of the nonfarm business sector by the Bureau of Labor Statistics (BLS).

The macroeconomic uncertainty measure, used to unfilter the NIPA consumption as described in Kroencke (2017), is taken from Sydney Ludvigson's website. Garbage is the municipal solid waste (MSW) series by the Environmental Protection Agency (EPA) as in Savov (2011).

Asset returns are calculated with the CRSP NYSE/NYSE MKT/NASDAQ/Arca Value-Weighted Market Index provided by the Center for Research in Security Prices (CRSP). We refer to this index as the market portfolio. Interest rates are collected from the Board of Governors of the Federal Reserve System's H.15 Selected Interest Rates publication. Our proxy for the risk-free rate is the 3-Month Treasury Bill Secondary Market Rate.

3 Alternative specifications of *cay*

The previous section outlined the general theory underlying *cay* and we will present empirical results in Section 4. Before getting there, this section will introduce two methodological improvements over the baseline specification of Lettau and Ludvigson (2001). First, we discuss how to construct *cay* from data on the wealthiest households, and second, we de-filter the consumption time series to reconstruct richer time variation.

3.1 Consumption-wealth ratio for the 10% richest household

The basic rationale behind *cay* is derived from the representative agent's intertemporal budget constraint. In the standard empirical implementation, *cay* is estimated from *total* consumption flow and aggregate wealth. However, in contrast to consumption, the distribution of asset wealth is strongly unequal due to limited capital market participation. Lettau et al. (2019), e.g., report that while the raw stock market participation rate has increased from 31.7% in 1989 to a more or less constant level of around 50% since the end of the 1990s, the wealth-weighted participation rate remains at far lower levels of only 20.2% most recently in 2013. Even more importantly, they show that the top 10% of households along the wealth distribution own more than 70% of stock wealth among stockowners, and even around 85% among all households. Thus, a substantial part of total consumption is contributed from households that do not own any sizeable financial wealth and therefore their intertemporal budget constraint should not materially be affected by expectations over future capital market returns.
Thus, guided by recent work of Lettau et al. (2019), who propose capital share risk as a surprisingly successful proxy for the pricing kernel, we construct a measure of *cay* from the data of the 10% richest households by regressing the income share of the top 10% wealthiest households onto the capital share, in order to account for only systematic risk in the variation of the income share as some of it across different percentile groups may be idiosyncratic and can be diversified away. We apply the same procedure to the wealth share of the top 10% wealthiest households. Our original series for the income and wealth share are on an annual basis and are converted to a quarterly frequency by a simple linear interpolation between yearly fourth quarter observations. The capital share series is quarterly.

For our sample from 1952:1-2019:4, the two regressions of the income and wealth share on the capital share are described by

$$\frac{X_t^{\text{top10}}}{X_t} = \alpha_X + \beta_X \, KS_t + \varepsilon_t, \qquad \text{with } X \in \{A, Y\}. \tag{11}$$

A_t^{top10}/A_t and Y_t^{top10}/Y_t denote the wealth and income share of the top 10% wealthiest households, respectively, and KS_t is the capital share. The capital share is defined as $1 - LS_t$ where LS_t is the labor share.

The coefficients for both specifications of (11) can be found in Table 1.

In a second step, we then multiply our consumption and labor income series in levels with the fitted values from the income share regression on the capital share, and the aggregate wealth series in levels with the fitted values from the wealth share regression on the capital share. While this should give us some reasonable proxies for the respective series for the top 10% wealthiest

$X \in \{A, Y\}$	Y_t^{top10}/Y_t		A_t^{top10}/A_t	
$\widehat{\alpha}_X$	0.012	(0.731)	0.534***	(30.019)
$\widehat{\beta}_X$	1.147***	(23.700)	0.434***	(8.016)
R^2	0.675		0.192	
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 1: Coefficient estimates and R^2 statistics of the income and wealth share regressed on the capital share. The normal OLS t-statistics are reported in parentheses. The sample period is 1952:1-2019:4.

households, we have to mention that especially multiplying the consumption and income series with the same fitted values has to be treated with care. A richer data foundation in the future is therefore likely to generate stronger results. All of this is done in an attempt to account for the fact that the wealthier households finance their consumption primarily through assets and not like the majority of households through labor income, implying that the marginal investor's consumption, asset wealth and labor income should predict future asset returns more accurately than their aggregated counterparts. We will refer to this as the top-10% specification and denote the corresponding *cay* by cay^{top10}.

3.2 Consumption-wealth ratio from unfiltering consumption data

Our second alternative for constructing *cay* is inspired by the observations of Savov (2011) and Kroencke (2017) that consumption data derived from the NIPA tables by the BEA displays comparatively low time variation. Savov (2011) as well as Kroencke (2017) focus on the explanation of the classical equity premium puzzle of Mehra and Prescott (1985), which can be restated as the problem that consumption growth volatility is too low for reasonable risk aversion parameters. Savov (2011) resolves the 'puzzle' by recovering richer time variation from data on municipal solid waste (MSW), or simply garbage. In contrast, Kroencke (2017) argues that the BEA employs a filtering process to their raw consumption data before publishing the NIPA tables. He shows how to undo the filtering process in order to recover the original, more volatile consumption series. We closely follow his approach and undo the filtering by the BEA with the help of the uncertainty measure of Jurado et al. (2015). We will refer to this as the unfiltered specification and the corresponding version of *cay* is denoted by cay^{unfil}. In this case, our sample is still quarterly but starts not until 1960:3. We spare the details of the filtering model here as they are identical to the method described by Kroencke (2017), except that we specifically use what he calls 'JLN unfilter', based on the uncertainty measure of Jurado et al. (2015), because it is as model-free as possible and uses a wide range of macroeconomic indicators. Despite the fact that this may not precisely reflect consumption uncertainty, we still consider it to be advantageous over the use of a GARCH-model as the main model of consumption volatility.

3.3 Estimation results

We now estimate the cointegration regression (9) discussed in Section 2 for the baseline and our two alternative *cay* specifications and report the results in Table 2. Note that due to data availability, cay^{unfil} can only be estimated from 1960:3 onwards.

	1952:1-2019:4		1960:3-2019:4		
	baseline	top-10%	baseline	top-10%	unfiltered
$\widehat{\beta}_a$	0.035	0.097***	-0.064	0.082	-0.099
	(0.987)	(2.812)	(-1.041)	(1.385)	(-1.559)
$\widehat{\beta}_y$	0.906***	0.855***	1.053***	0.873***	1.097***
	(23.950)	(25.486)	(12.746)	(13.049)	(12.753)
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 2: Cointegrating parameter estimates for our baseline, top-10% and unfiltered specification for different sample periods specified in the first line of the table. Newey and West (1987) corrected t-statistics appear in parentheses.

When comparing the cointegrating parameters among the different specifications, we see that in the sample from 1952:1 (left columns), the coefficient estimates for the baseline and the top-10% specification are close to each other, although it is noteworthy that the coefficient for aggregate wealth ($\widehat{\beta}_a$) is significant in the latter compared to a t-statistic below one for the baseline case. The comparison with cay^{unfil} can only be done within the shorter sample from 1960:3 (right columns) and here we actually find that $\widehat{\beta}_a$ is not significant in any of the three specifications.

With the cointegrating parameters from Table 2, we can now estimate the corresponding *cay*s according to (10). Summary statistics, including the log excess return of the market portfolio, $r_t - r_{f,t}$, where r_t denotes the log return of the market portfolio and $r_{f,t}$ denotes the log risk-free rate, can be found in Table 3.

Figure 2 shows all three *cay*s for their respective full sample and in units of standard deviation, together with log excess returns. Overall the three different *cay*s behave rather similar, especially the one for the baseline and the top-10% specification. The unfiltered *cay* is somewhat more different but which is mainly due to the different sample period. In general, we see that in the past, negative developments in *cay* often foreshadowed negative excess returns, e.g., the oil crisis in the 1970s or the bear market after the burst of the dotcom bubble. Other events, such as the 1987 stock market crash, but also the financial crisis of 2007/08, however, were not accompanied by a foregone decline in *cay*. More recently, *cay* has entered an ongoing negative development since the financial crisis, albeit mostly positive returns during the last decade.

Especially the negative development of *cay* during the last decade is in sharp contrast to the positive stock market development during that time span. This observation is already a preview

	$\widehat{cay}_t$	$\widehat{cay}_t^{\text{top10}}$	$\widehat{cay}_t^{\text{unfil}}$	$r_t - r_{f,t}$
Panel A: Univariate summary statistics				
Mean	0.612	0.281	0.329	0.015
Standard deviation	0.023	0.022	0.030	0.083
Autocorrelation	0.946	0.931	0.914	0.065
Panel B: Correlation matrix				
$\widehat{cay}_t$	1.000	0.956	0.819	0.047
$\widehat{cay}_t^{\text{top10}}$		1.000	0.726	0.055
$\widehat{cay}_t^{\text{unfil}}$			1.000	0.068
$r_t - r_{f,t}$				1.000

Table 3: Summary statistics for $\widehat{cay}_t$, $\widehat{cay}_t^{\text{top10}}$, $\widehat{cay}_t^{\text{unfil}}$ and quarterly excess returns. The sample period is 1952:1-2019:4, except for $\widehat{cay}_t^{\text{unfil}}$ it is 1960:3-2019:4.

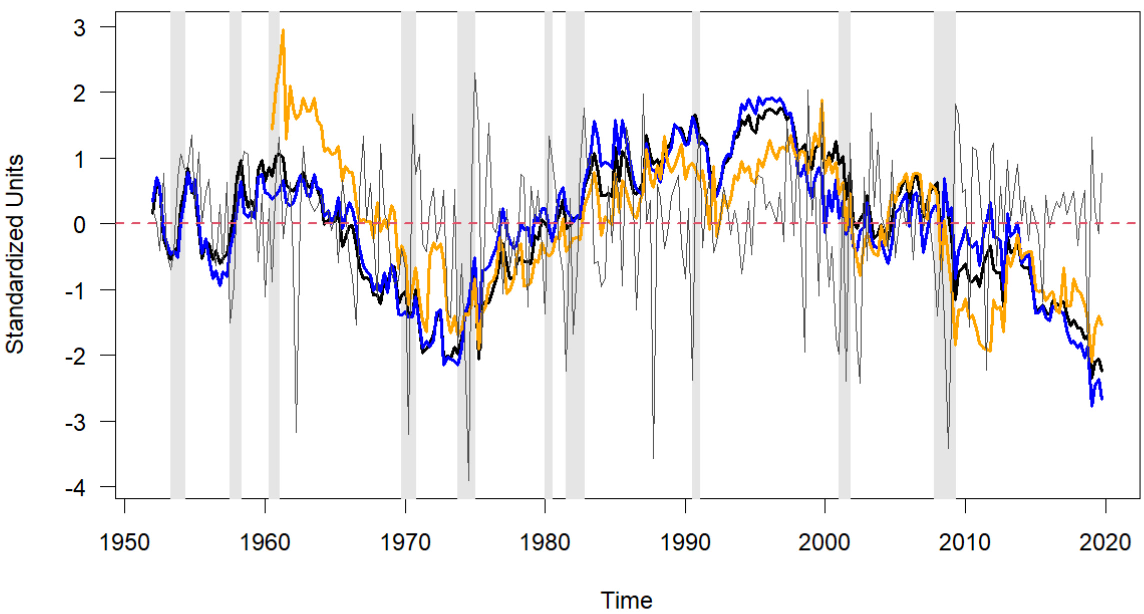


Figure 2: The plot shows $\widehat{cay}_t$ (black), $\widehat{cay}_t^{\text{top10}}$ (blue) and $\widehat{cay}_t^{\text{unfil}}$ (orange). The gray curve represents log excess returns, $r_t - r_{f,t}$. The gray shaded areas mark NBER recessions. The sample period is 1952:1-2019:4, except for cay_t^{unfil} it is 1960:3-2019:4.

on the weaker forecasting ability, which we will investigate in closer detail in the upcoming section.

4 Empirical forecasting results

4.1 Predictive regressions

After having estimated the cointegrating parameters of our various *cay*s, we are now able to run predictive regressions of H-period ahead excess returns on *cay*, i.e. we estimate by OLS the model,

$$\tilde{r}_{t,H} = \alpha + \gamma\,\widehat{cay}_t + \varepsilon_{t,H} \tag{12}$$

where $\tilde{r}_{t,H} = r_{t+1} - r_{f,t+1} + \cdots + r_{t+H} - r_{f,t+H}$. In Table 4, we report the results for horizons up to five years. Even though we already employ the commonly used Newey and West (1987) standard

	Dependent variable: $\tilde{r}_{t,H}$						
	$H = 1$	$H = 2$	$H = 4$	$H = 8$	$H = 12$	$H = 16$	$H = 20$
$\widehat{cay}_t$	0.240	0.467	0.813	1.537	2.152	2.970	3.520
	(1.079)	(1.060)	(0.692)	(0.703)	(0.662)	(0.965)	(0.991)
	[0.001]	[0.004]	[0.008]	[0.019]	[0.030]	[0.050]	[0.055]
$\widehat{cay}_t^{\text{top10}}$	0.398*	0.777*	1.482	2.889	3.817	4.968**	5.592*
	(1.762)	(1.744)	(1.290)	(1.517)	(1.629)	(2.286)	(1.914)
	[0.007]	[0.015]	[0.031]	[0.066]	[0.086]	[0.123]	[0.121]
$\widehat{cay}_t^{\text{unfil}}$	0.023	-0.012	-0.245	-0.594	-0.599	-0.550	-0.875
	(0.127)	(-0.036)	(-0.304)	(-0.459)	(.0.307)	(-0.258)	(.0.394)
	[−0.004]	[−0.004]	[−0.002]	[0.002]	[0.000]	[−0.001]	[0.002]
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 4: The table reports forecasting regression estimates for $\widehat{cay}_t$, $\widehat{cay}_t^{\text{top10}}$ and $\widehat{cay}_t^{\text{unfil}}$. Forecasts span different horizons from one quarter up to five years. Newey and West (1987) corrected t-statistics appear in parentheses and adjusted R^2 statistics in square brackets. The sample period is 1952:1-2019:4, except for $\widehat{cay}_t^{\text{unfil}}$ it is 1960:3-2019:4.

errors, we emphasize the importance of being careful with long-horizon predictions, as recently elaborated in Kostakis et al. (2023). To avoid any overlapping problems and circumvent this issue, we will focus on one-quarter ahead forecasts in the remainder of this paper unless stated otherwise. In principal, (7) also allows for the possibility of *cay* to predict future consumption growth. In Appendix A.3, we present evidence that this is still not the case even though the predictability of asset returns has weakened over time.

The results displayed in Table 4 are striking. At the 1%-significance level, we find none of the *cay* versions over any horizon to be significant predictors. The comparatively best predictor turns out to be $\widehat{cay}_t^{\text{top10}}$, for which we at least get significance at the 5% or 10% level at horizons of four and five years. The adjusted R^2 statistics are also well below one percent for a one quarter horizon. These results are in sharp contrast to findings in prior literature. For example, Lettau and Ludvigson (2001) reported a coefficient estimate of 2.165 with a t-statistic of over three and an adjusted R^2 of 9% for a sample period from 1952:4-1998:3. Other papers (e.g., Guo (2006), Hahn and Lee (2006), Della Corte et al. (2010), Sousa (2010), Guo et al. (2013)) confirmed these strongly significant findings for different sample periods. Lettau and Ludvigson (2001) also concluded that the impact of a one-standard deviation rise in $\widehat{cay}_t$ is economically large with expected excess returns going up by about 220 basis points, roughly a nine percent increase at an annual rate. With a current standard deviation of 0.023, we only get an increase of about 56 basis points for $\widehat{cay}_t$, which is not even significant.

Taken together, the results lead us to conclude that the predictive ability of *cay* has fundamentally weakened over the last roughly 15 years. We investigate this apparent decay more closely from an in-sample, out-of-sample and economic perspective in the following subsections.

4.2 The decay in-sample

Whenever we deal with variables as predictors for asset returns, an important question is whether the forecasts are performed in-sample or out-of-sample. At first sight, one might think that using only past values of *cay* to predict future excess returns is a true out-of-sample forecast. However, Lettau and Ludvigson (2001) themselves, as well as others (e.g., Brennan and Xia (2005)), have pointed out that for the estimation of the cointegrating parameters used to construct $\widehat{cay}_t$, the use of the full sample introduces somewhat of a 'look-ahead bias'. Therefore, the forecasts we presented above are rather to be labeled as in-sample forecasts. Nonetheless, we can also re-estimate $\widehat{cay}_t$, and its alterations, each quarter in order to sequentially update the forecast regression and to see how the predictive ability varies over time. We do so beginning no earlier than in the first quarter of 1985 to have a sufficiently large sample from which to estimate the first set of cointegrating parameters.

The results of this exercise are presented in Figure 3. The upper plot shows the coefficient estimates for our various *cay*s together with their 90%-confidence intervals. For the top-10% and unfiltered specification, it took a little longer to reach significant levels but from the late 1980s on, all three *cay*s have been strong predictors for next quarter's excess returns. They stayed more or less on their respective levels for roughly 15 years, before peaking in 2002:3 and starting to decay afterwards. For the baseline and unfiltered specification, this decay is somewhat more pronounced than for the top-10% specification, but in general all three exhibit a significant decline towards the levels we already observed in Table 4. While the unfiltered forecasts reached insignificant levels already around 2010, it took until 2015 for the baseline forecasts to do the same.

The bottom plot of Figure 3 shows the economic impact in terms of basis points (left axis, solid lines) given a one standard deviation increase in *cay* as well as the corresponding adjusted R^2 statistics (right axis, dashed lines). Both measures exhibit the same decay as the coefficient

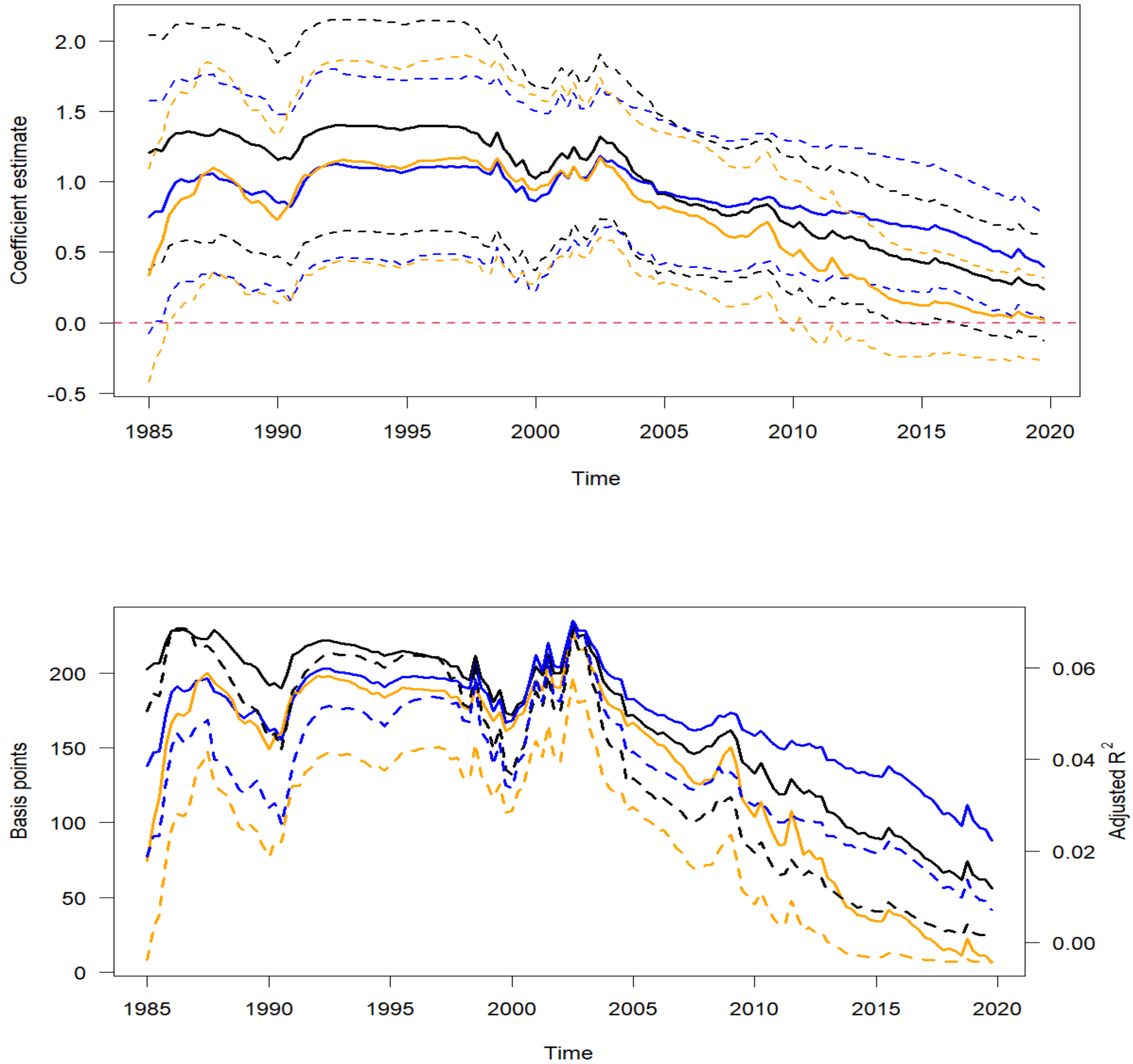


Figure 3: Both plots show results from quarterly re-estimated one-quarter ahead forecast regressions of excess returns on $\widehat{cay}_t$ (black), $\widehat{cay}_t^{\text{top10}}$ (blue) and $\widehat{cay}_t^{\text{unfil}}$ (orange). The upper plot represents the coefficient estimates as the solid lines accompanied by their respective 90%-confidence band based on Newey and West (1987) standard errors as the dashed lines. The bottom plot shows the impact of a one-standard deviation increase in *cay* on one-quarter ahead excess returns in basis points as the solid lines on the left axis and the adjusted R^2 statistics on the right axis. The sample period is 1952:1-2019:4, except for cay_t^{unfil} it is 1960:3-2019:4.

estimate itself. From levels of well above 200 basis points in the early 2000s, the impact has weakened to 56, 88 and 7 basis points for our three specifications, respectively. In a very similar manner, the R^2 statistics have declined from more than 6% to levels below 1% already seen in Table 4. The difference to the 9% that Lettau and Ludvigson (2001) found can be attributed to revisions in the underlying series from the BEA.

Altogether, Figure 3 makes clearly visible what we already suspected when comparing our results to previous findings from others. The decay of *cay* has been an ongoing development for at least the last decade.

4.3 The decay out-of-sample

After having seen the decay of *cay* in-sample above, we now turn our attention towards changes in the predictive power of *cay* out-of-sample. In their seminal paper, Welch and Goyal (2008) assess the forecasting abilities of excess returns for various variables, including *cay*. While their models make annual predictions, we stick with our quarterly frequency. Aside from that obvious difference, we follow their approach in measuring the in-sample (IS) and out-of-sample (OOS) performance against two simple mean models. For the IS performance, we take the full sample mean from Table 3 as a prediction, whereas for the OOS performance, we compute prevailing historical means based on the data available up to that point in time. An intuitive way to then evaluate the performance of our *cay*-models against their respective mean models, is given by taking the difference of the sum of squared errors (SSE) over time. For a negative difference between the mean model and the *cay*-model, we have a worse performance of the *cay*-model and vice versa. At the same time an increase in this difference means a better performance of the *cay*-model in that specific quarter.

We allow a 15-year training period for the out-of-sample forecasts and report results in Figure 4. We can confirm the findings of Welch and Goyal (2008) that it was mainly the oil shock that accounts for most of the positive in-sample performance of *cay*. Aside from that, we also find that the OOS performance of *cay* is clearly worse than the prevailing historical mean model. What is more important for us, however, is that we can again clearly see the decay of *cay* setting in in the early 2000s for all three specifications, further worsening their OOS forecasting ability and bringing them to new lows by the end of 2019. In line with our IS findings, we again see that the top-10% specification performs better than the baseline and unfiltered specification. Nevertheless, it still cannot match its IS counterpart and has lost its advantage over the mean model during its recent decay. It is thus not only a feature of the IS estimation that the forecasting power of *cay* has ceased recently, but also of its OOS evidence. Together, this means that from a statistical point of view, the relevance of *cay* has decreased substantially to non-significant levels recently. Apart from that, one can ask if there is still some economic significance left that could be exploited in terms of a trading strategy based on *cay*-forecasts. This is precisely what we will investigate next.

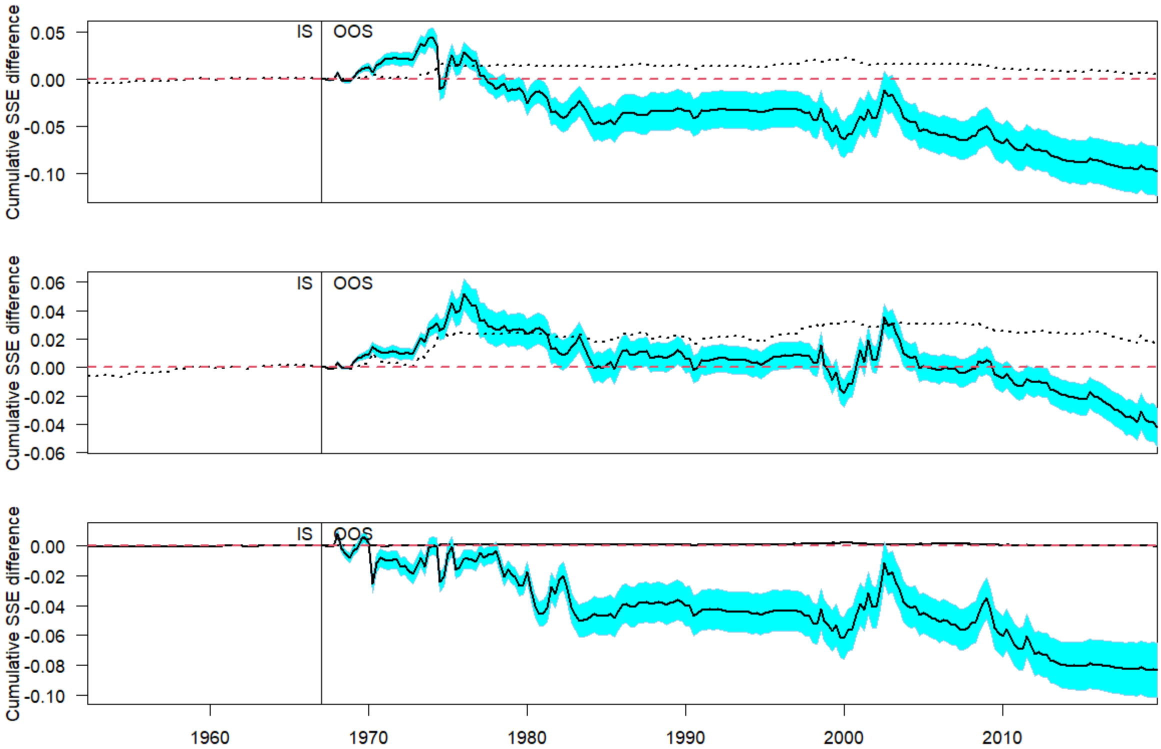


Figure 4: The plots show the OOS performance of our three *cay* specifications. In the top we have the baseline, in the middle the top-10%, and in the bottom the unfiltered specification. The solid lines represent the difference in sum of squared errors between a prevailing mean model and OOS forecasts based on an recursively updated *cay*. The blue shaded area marks the 95% confidence band based on OOS-t critical values from McCracken (2007). The dotted line is the respective IS counterpart representing the difference in sum of squared errors between a full sample mean model and IS forecasts estimated using the full available sample at that time. The sample periods are as in Table 2.

4.4 The decay in economic significance

After having seen the undeniable decay of *cay* in statistical terms, we now turn our attention towards a more economic assessment of *cay* and ask whether a trading strategy based on *cay*-forecasts would have made an investor better off over time than simple mean-investing.

For that, we mainly follow Della Corte et al. (2010) in assuming a mean-variance investor who rebuilds her portfolio on a quarterly basis. This means she can allocate her wealth between a risk-free asset, here represented by the interest rate of 3-month U.S. T-bills, and a risky asset, represented by our market portfolio based on the CRSP NYSE/NYSE MKT/NASDAQ/Arca Value-Weighted Market Index. She then puts a weight of

$$w_t = \frac{1}{\lambda} \frac{E_t[r_{t+1} - r_{f,t+1}]}{Var_t[r_{t+1} - r_{f,t+1}]} \tag{13}$$

on the risky asset and a weight of $1 - w_t$ on the risk-free asset. We follow Campbell and Thompson (2008) in assuming a relative risk aversion (RRA) coefficient, λ, of three. Further, we do not allow for short-selling of either asset and thus restrict the weights to be in the interval $[0, 1]$. In cases for which $w_t > 1$, we set $w_t = 1$ and for $w_t < 0$ we set $w_t = 0$, respectively. In particular, this means that for a negative expected excess return in the next period, i.e., $E_t[r_{t+1} - r_{f,t+1}] < 0$, our investor will invest only in the risk-free asset.

We compare two trading strategies. The benchmark is labelled average-strategy ($\mathcal{AVE}$) where forecasts for the next quarter are based on a prevailing historical mean. Our *cay*-strategy ($\mathcal{CAY}$) uses data available up to the respective quarter to build a one-quarter ahead forecast based on the estimated trend deviation measure $\widehat{cay}_t$.

We then evaluate both strategies using two different performance measures. The first one is the the manipulation-proof risk-adjusted abnormal return, θ, from Goetzmann et al. (2007), that essentially measures a portfolio's excess premium after adjusting for risk. It is defined as

$$\theta = \frac{1}{1-\lambda}\left[\log\left(\frac{1}{T}\sum_{t=0}^{T-1}\left(\frac{R^{\mathcal{CAY}}_{p,t+1}}{R_{f,t+1}}\right)^{1-\lambda}\right) - \log\left(\frac{1}{T}\sum_{t=0}^{T-1}\left(\frac{R^{\mathcal{AVE}}_{p,t+1}}{R_{f,t+1}}\right)^{1-\lambda}\right)\right]$$

where $R_{p,t+1} = R_{f,t+1} + w_t(r_{t+1} - r_{f,t+1})$ is a portfolio's gross return and $R_{f,t} = 1 + r_{f,t}$ is the gross risk-free rate. A positive value of θ then indicates that the $\mathcal{CAY}$-strategy would have been the better choice for an investor after adjusting for risk.

As a second metric, we also assess the performance fee, ϕ, which can be interpreted as the maximum performance fee a risk-averse investor would be willing to pay to switch from the $\mathcal{AVE}$ to the $\mathcal{CAY}$-strategy. For that, we assume a quadratic utility of the form

$$U(R_p) = U\left(\{R_{p,t+1}\}_{t=0}^{t=T-1}\right) = \frac{1}{T}\sum_{t=0}^{T-1}\left(R_{p,t+1} - \frac{\lambda}{2(1+\lambda)}R^2_{p,t+1}\right),$$

following West et al. (1993) and Fleming et al. (2001). The performance fee, ϕ, is then defined via the equation

$$U\left(R_p^{\mathcal{CAY}} - \phi\right) = U\left(R_p^{\mathcal{AVE}}\right). \tag{14}$$

Hence, just like for θ, a positive value of ϕ indicates a better performance of the $\mathcal{CAY}$-strategy.

Since the developments of θ and ϕ over time are almost identical, we limit ourselves to showing the results for θ in Figure 5, while the results for ϕ can be found in Appendix A.4. The baseline and top-10% specification were basically on the same level between 40 and 80 basis points for a long time, before starting to decay down to almost zero by the end of 2019. The unfiltered specification also exhibits a decline after the financial crisis, but having never exceeded 40 basis points, this decay appears less pronounced.

Figure 6 shows the cumulative performance of the three portfolios for having invested 1 U.S.-Dollar at the end of 1966. Until the financial crisis, all three $\mathcal{CAY}$-portfolios outperformed

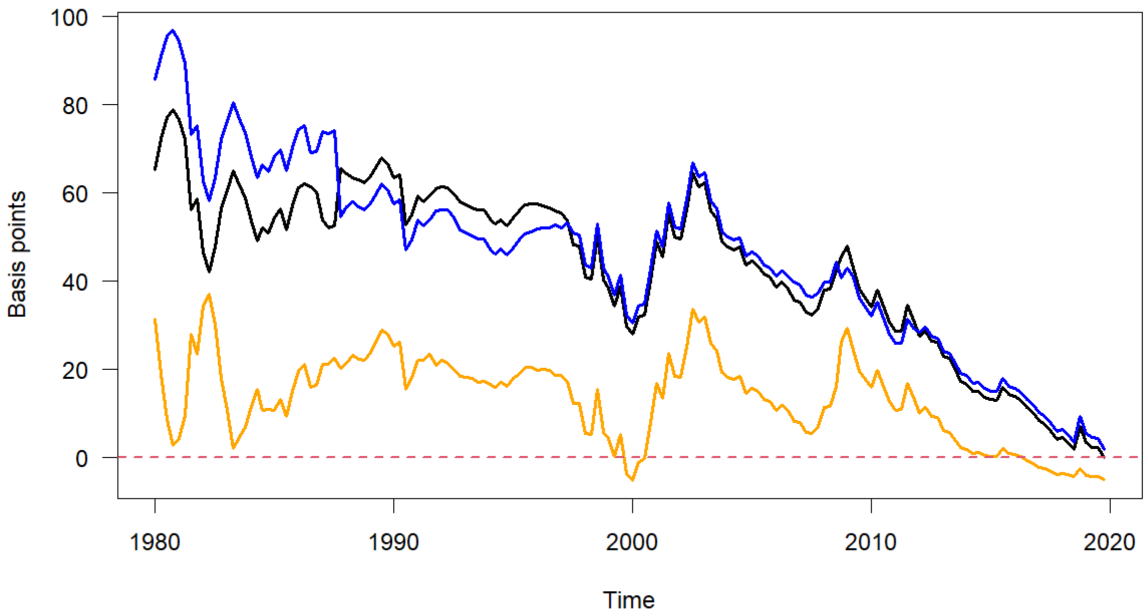


Figure 5: The plot shows the risk-adjusted abnormal return, θ, for the baseline (black), top-10% (blue) and unfiltered (orange) specification. The sample period is 1952:1-2019:4, except for $\widehat{cay}_t^{\text{unfil}}$ it is 1960:3-2019:4.

the $\mathcal{AVE}$-strategy (dashed line). Since then the baseline and top-10% specification were almost always invested in the risk-free asset due to the negative development of *cay* exhibited in Figure 2. At the same time, the $\mathcal{AVE}$-strategy invested roughly about two-thirds into the risky asset and by that closed the gap to the $\mathcal{CAY}$-portfolios. When comparing them to the market itself, both the baseline and top-10% specification performed remarkably well until the financial crisis, in fact outperforming the market for most times. But again, since the financial crisis this has changed dramatically leading to a market portfolio value more than three times as high as those of the $\mathcal{CAY}$-portfolios.

Figure 6 also illustrates that we do not only see the decay of *cay* in the decline of the performance measures θ and ϕ but also directly in the recent stagnating portfolio value of the $\mathcal{CAY}$-strategy when compared to the $\mathcal{AVE}$-strategy and even more so against the market portfolio. We therefore conclude that the decay of *cay* also holds up when assessing the economic significance of *cay* and that it is not merely a feature of statistical considerations.

5 A structural shift in the cointegration relation

Having documented the substantial decay of *cay* in the previous part, this section sheds light on the origin of the declining forecasting power. Recall again, that *cay* is expected to be a viable

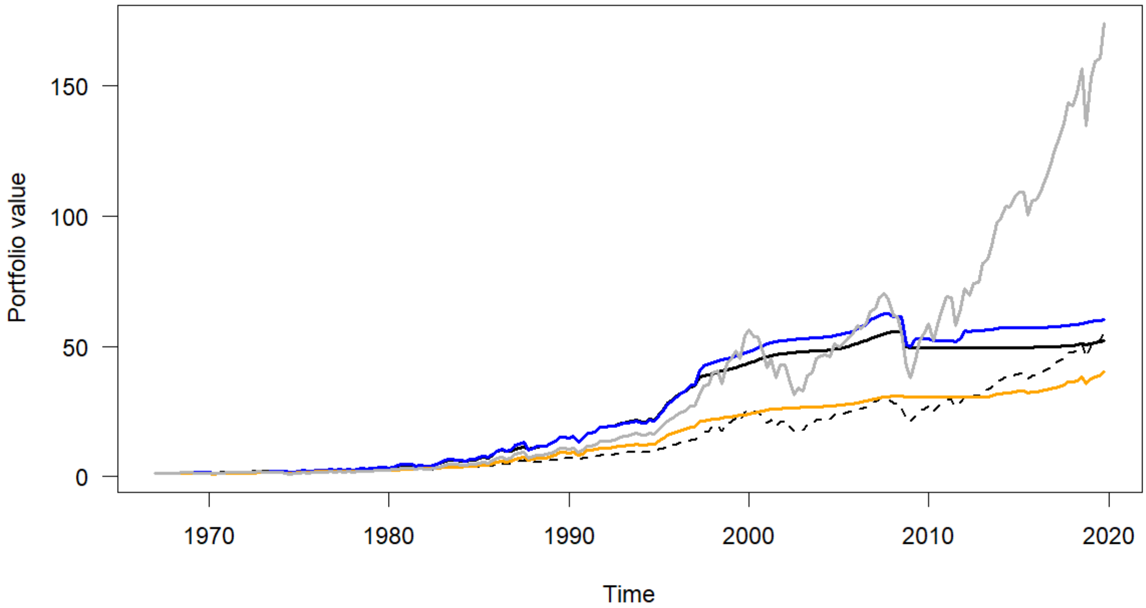


Figure 6: Portfolio values of the baseline (black), top-10% (blue) and unfiltered (orange) $\mathcal{CAY}$-strategy when investing 1 U.S.-dollar at the end of 1966. The gray line represents an investment in the market portfolio only, whereas the dashed line resembles the $\mathcal{AVE}$-strategy.

predictor for stock market returns since it is derived as the residual from the (single) cointegration equation between consumption, aggregate wealth and labor income. The assumption of a stable equilibrium relationship may be justified from the rationale of economic theory as invoked by Lettau and Ludvigson (2001), who argue that the parameters in *cay* are steady state wealth shares which are perfectly known to an agent in equilibrium. Our previous section on results followed this reasoning in the sense that we employed *cay* without questioning its properties as stationary cointegrating residual. However, the existence of a stable equilibrium relationship is eventually an empirical question, which we investigate in more detail in this section. Testing cointegration relations requires a number of modelling choices on the part of the econometrician, such as the existence of trends, the number of lags and allowance of intercepts. We first conduct different cointegration tests and analyze the results of the Vector Error Correction Model.

5.1 Stability of the cointegrating coefficients

In their reply to Brennan and Xia's (2005) criticism of a 'look-ahead bias' incorporated into *cay*, Lettau and Ludvigson (2005b) emphasize the importance of using the full available sample to estimate the cointegrating coefficients and not just a subsample that contains only information available up to a certain point in time. We take Lettau and Ludvigson (2005b) at face value and thus estimate $\widehat{\beta}_a$ and $\widehat{\beta}_y$ from equation (10) using the full available sample from 1952:1-2019:4 and report results in Panel A of Table 5.

Panel A: Cointegrating parameters				
	this paper	LL2001	HL2006	DC2010
Sample	52:1-19:4	52:4-98:3	52:4-02:4	1946-2006
$\widehat{\beta}_a$	0.035	0.310***	0.275***	0.274***
	(0.987)	(11.700)	(27.500)	(11.417)
$\widehat{\beta}_y$	0.906***	0.590***	0.616***	0.684***
	(23.950)	(23.920)	(61.600)	(28.500)
Note:				*p<0.1; **p<0.05; ***p<0.01

Panel B: Phillips and Ouliaris (1990) cointegration tests						
	1952:1-1998:3			1952:1-2019:4		
	Statistic	Lags	p-value	Statistic	Lags	p-value
	-2.780	1	0.3462	-1.443	1	0.9084

Panel C: Vector Error Correction Model						
	1952:1-1998:3			1952:1-2019:4		
	Δc_t	Δa_t	Δy_t	Δc_t	Δa_t	Δy_t
Δc_{t-1}	0.211**	0.096	0.486***	0.334***	0.238	0.626***
	(2.602)	(0.370)	(3.139)	(5.327)	(0.905)	(4.737)
Δa_{t-1}	0.071***	0.240***	0.081*	0.053***	0.164***	0.052
	(2.968)	(3.137)	(1.769)	(3.548)	(2.599)	(1.627)
Δy_{t-1}	0.064	-0.154	0.021	0.044	-0.314	-0.312**
	(1.395)	(-1.049)	(0.240)	(1.382)	(-1.007)	(-1.980)
$\widehat{cay}_{t-1}$	-0.020	0.167**	0.020	0.004	0.020	0.102
	(-0.931)	(2.482)	(0.504)	(0.373)	(0.458)	(1.451)
$\bar{R}^2$	0.171	0.065	0.096	0.221	0.020	0.102
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5: The table reports the coefficient estimates of vector error correction models (VECM) for two different sample periods specified in the first line. Ordinary t-statistics appear in parentheses.

Column one reports our results (already shown in Table 2) and contrast them with the corresponding results in Lettau and Ludvigson (2001), Hahn and Lee (2006), and Della Corte et al. (2010). Strikingly, our most recent results differ remarkably from previous findings.[3] While $\widehat{\beta}_y$ shows similar levels of significance, the coefficient itself is larger. Vice versa, $\widehat{\beta}_a$ is smaller by an order of magnitude and entirely lost its strong significance (Lettau and Ludvigson (2001) report t-stats of more than eleven). Thus, the cointegrating regression leaves us with the

[3]Note, that by restricting our data to the same samples as in these papers, we are able to replicate their results.

conclusion that asset wealth is not able to significantly explain the variation of consumption. Furthermore, standard cointegration tests (e.g., Engle and Granger (1987); Phillips and Ouliaris (1990); Johansen (1988, 1991)) that take no cointegration as the null hypothesis fail to reject the null in neither the full sample nor different subsamples, as can be seen in Panel B of Table 5 and in more detail in Appendix A.5.[4] We are thus not able to replicate the test results from Lettau and Ludvigson (2001).[5] Interestingly, in the latest assessment of *cay* in Lettau and Ludvigson (2019), they only report results from not being able to reject the null of cointegration. Thereby following the argument of e.g. Ogaki and Park (1997), that taking the null of no cointegration may have low power.

To directly assess the stationarity, we conduct Phillips and Ouliaris (1990) residual based cointegration tests of the constructed *cay* measure and report results in Panel B of Table 5. We are far from being able to reject the null of a unit root and thus cannot confirm that *cay* appears to be stationary.

Finally, to assess the short-term dynamics of the cointegrating relationship, we report the results of estimating the Vector Error Correction Model with data up to 1998:3 (left side) and up to the most recent observations (2019:4, right side) in Panel C of Table 5. The coefficient of interest is the impact of $\widehat{cay}_{t-1}$ on Δa_t. Replicating the results of Lettau and Ludvigson (2001), it is strongly significant with a value of 0.167 in the sample until 1998:3 and shows transitory movements in asset wealth rather than consumption or labor income. However, $\widehat{cay}_{t-1}$ is far from significant (t-stat of 0.458) from the perspective of most recent data. Consistent with the results from the cointegrating regression in Panel A, it leads to the conclusion that *cay* does not (or, no longer) predict subsequent (transitory) changes in asset wealth.

To elaborate further on the structural shift, we re-estimate the cointegrating relationship in an expanding window for every quarter beginning in 1967 and report results for $\widehat{\beta}_{a,t}$ and $\widehat{\beta}_{y,t}$ over time in Figure 7. We find stable values for $\widehat{\beta}_{a,t}$ of roughly 0.275 from the mid 1980s to the late 1990s, from which point on it steadily declined. Since the cointegrating parameters should add up to approximately one, $\widehat{\beta}_{y,t}$ shows the mirror behavior with an increase towards values close to one from the perspective of most recent data. This result is essentially the econometric confirmation of what can be observed from the development of the original data as illustrated in the upper panel of Figure 1. After detrending the consumption, asset wealth and labor income series, we can observe that, beginning with the late 1990s, asset wealth behaves increasingly

[4] As pointed out in the introduction to this section, cointegration tests are subject to different modelling choices. We stick to the specification, where we allow a linear trend in the data (but none in the cointegration relation) and an intercept in the cointegration relation. We discuss the particular role of the deterministic time trend in the cointegration equation in more detail in the later subsection 5.3. The Johansen (1988, 1991) cointegration test can usually be chosen among five alternative specifications. In Appendix A.6, we report results for the number of cointegrating relationships within all five specifications for comparison. In line with Koop et al. (2008), we find substantial differences among them, which further illustrates the model uncertainty.

[5] Our findings are in line with Rudd and Whelan (2006).

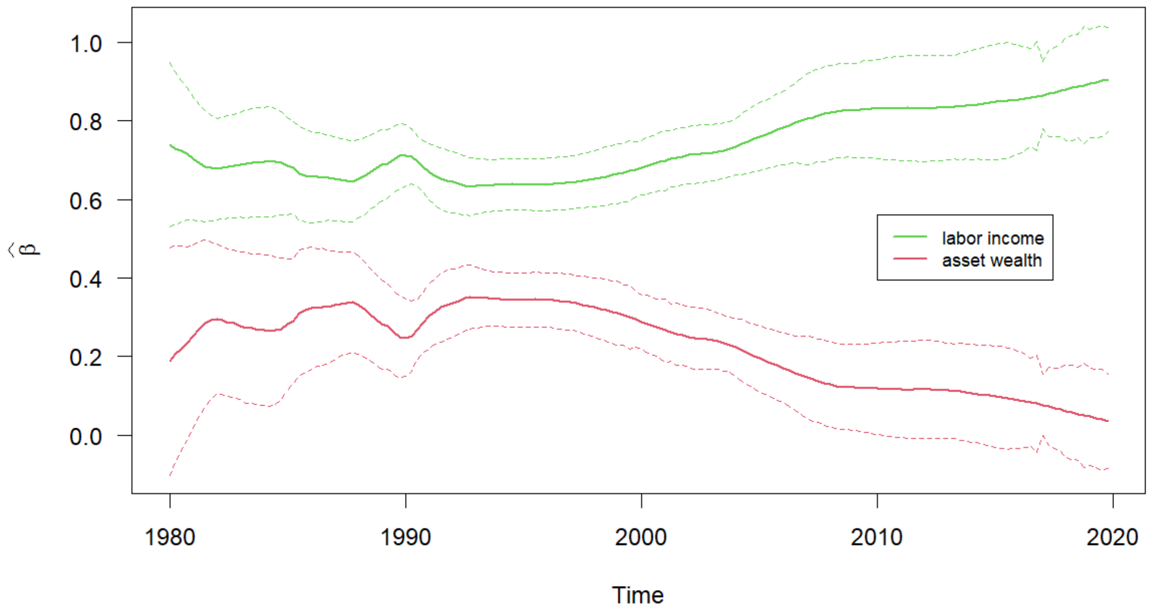


Figure 7: Cointegrating parameters for asset wealth (red) and labor income (green) quarterly re-estimated. The dashed lines represent the corresponding confidence bands of $\pm$ 2 standard errors. The sample period is 1952:1-2019:4.

detached from its alleged long-run relationship with consumption and labor income.

This result may raise several issues. First, the sustained shift in parameters opens the question of allowing for a deterministic time trend in the cointegrating relationship. Second, the observed shift may be driven by changes in the composition of the consumption measure. The latter point has been put forward in Lettau and Ludvigson (2019), who propose to switch from NDS to PCE as consumption series. We address both issues in subsequent sections, but already note that it will not affect our main results.

A third issue refers to which cointegrating relation shall be taken as the relevant one. From the pure econometrician's perspective (and in line with the argumentation of Lettau and Ludvigson, 2005b), parameter estimates from the most comprehensive data sample should be used. From economic theory consideration, one may argue however, that some equilibrium relationship has to exist and maybe the Lettau and Ludvigson (2001) parameters do best represent it. We therefore ask, how holding fixed a set of parameters influences the predictive ability of *cay* over time. In contrast to our previous analysis of Section 4.2, where we re-estimated the cointegrating parameters sequentially in the expanding window, we will now fix it – being fully aware that it is subject to the 'look-ahead bias' discussion of e.g. Brennan and Xia (2005). We use two parameter sets: The most recent one (i.e. parameter $\widehat{\beta}_{a,}$ and $\widehat{\beta}_y$ from column one in Table 5), and the one from Lettau and Ludvigson (2001) (i.e. column two in Table 5). We report the coefficients from the predictive regression of (12) in both cases in Figure 8. As known from

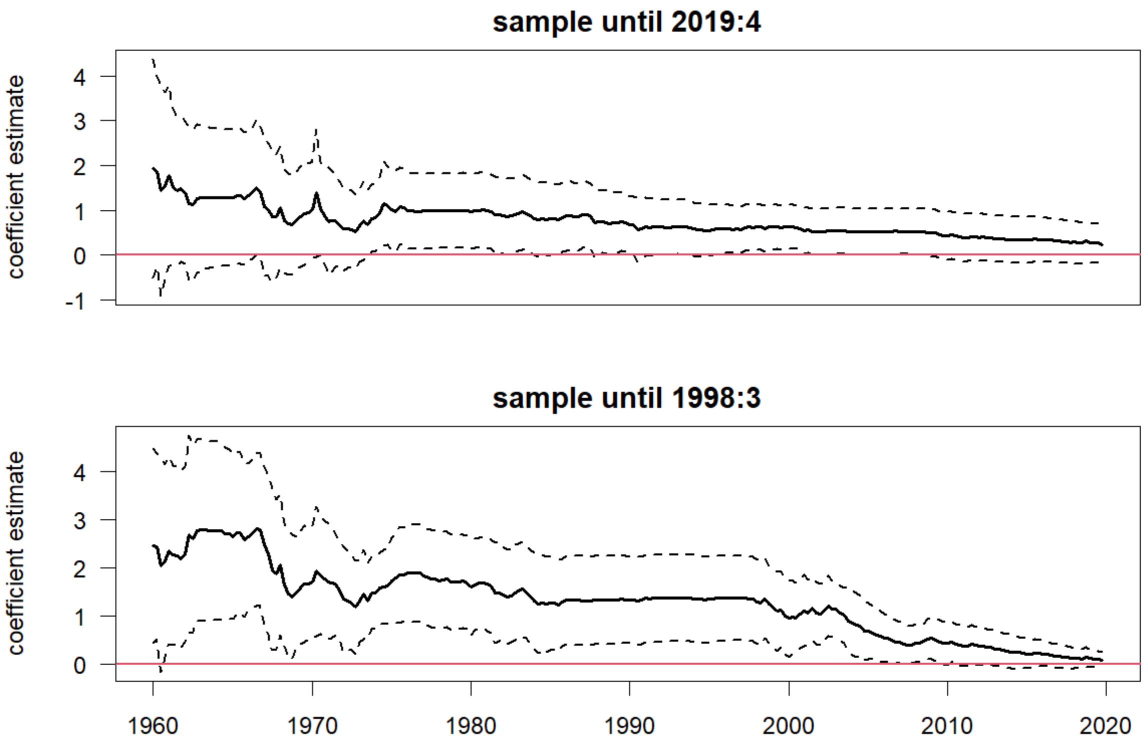


Figure 8: Both plots show the period wise re-estimated coefficient estimate of $\widehat{cay}_t$ in quarterly forecasts of excess returns accompanied by its 95% confidence band based on Newey and West (1987)-standard errors as the dashed lines. The upper plot represents the results using the cointegrating parameters estimated using the full sample from 1952:1-2019:4, while the bottom plot uses the cointegrating parameters from the sample 1952:1-1998:3.

Lettau and Ludvigson (2001), their parameter set worked well for (in-sample) prediction until 2000. However, the lower panel of Figure 8 shows that even by arguing that this parameter set is the right one (and therefore holding it fixed), we find the same decay in forecasting power. Vice versa, by using the most recent parameter set, we find that *cay* would have *never* worked. The upper panel indeed shows that the predictive coefficient has been insignificant for the entire sample and would not have shown predictability even in the period prior to 1998.

5.2 Most recent *cay* tends to be negative – Explaining poor OOS performance

The shift in the cointegration coefficients casts serious doubt on the forecasting power of *cay* in general. In particular, it can also explain the surprisingly large difference between in-sample and out-of-sample performance. Recall from Figure 3 that *cay* appeared to work well in-sample until 2000, but that in the same period, OOS results were very poor and even worse thereafter (see Figure 4). As we documented in the previous section, the coefficient on wealth $\widehat{\beta}_a$

shows a declining pattern in an expanding window. So, for the estimation of *cay* as the residual $\widehat{cay}_t = c_t - \widehat{\beta}_a a_t - \widehat{\beta}_y y_t$ at each endpoint in the expanding window, it will tend to overestimate the impact of wealth and therefore tend to be biased to negative numbers. To investigate this reasoning empirically, we re-estimate *cay* in the fourth quarter of each year from 1966 to 2019 and retain values only for the last three most recent years. We further subdivide the sample in three groups with the first spanning the first 18 years from 1966-1983, the second spanning the time period from 1984-2001, and the last spanning the time period from 2002-2019. Thus, each group contains a total of 216 observations. Results are plotted in Figure 9 which shows the distribution of the sequentially most recent estimates of *cay* (in standardized units). We observe

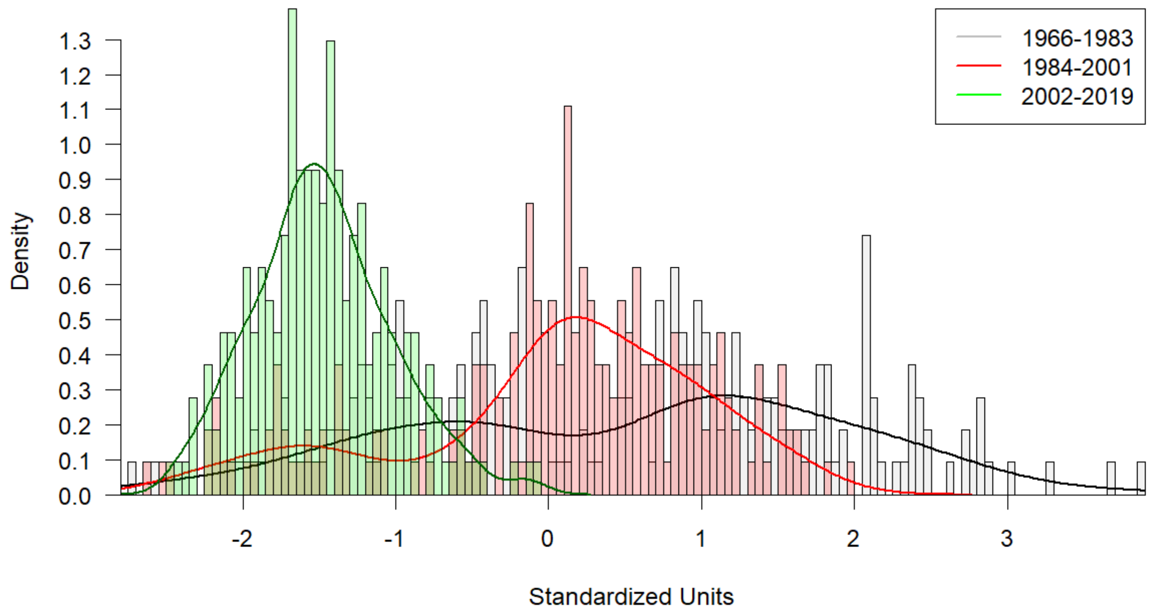


Figure 9: We re-estimate $\widehat{cay}_t$ every fourth quarter from 1966:4 to 2019:4. We then plot histograms of the last three years of each $\widehat{cay}_t$ for three equally sized subsamples of 18 estimations. The solid line indicates the respective kernel density estimate assuming a normal distribution. In gray we have the estimations from 1966-1983, red represents the years 1984-2001 and green indicates the time period from 2002-2019. The samples always start in 1952:1.

that in the first and second sub-sample, the latest *cay*-estimates were roughly equally likely to be positive or negative. However, when looking at the third sub-sample (i.e. from 2002 onward), we find a distribution that is highly shifted to negative values – in fact, *cay* was almost always negative and on average more than one standard deviation away from the historical mean. This result confirms the above reasoning, that the parameter shift will bias *cay* to the downside.

Note that according to the underlying theory, a low consumption-to-wealth ratio (i.e. negative *cay*) would signal that the representative agent anticipates low returns on wealth in the future

(which is why she reduces current consumption). Therefore we expect a positive coefficient on *cay* in the predictive regression. In the evaluation of the OOS performance, the most recent prediction of the market return will be compared to the realized return. If *cay* is biased downwards, also the predicted return will be biased downwards and therefore deliver poor results against the realized returns.
In economic terms, a negative *cay* and therefore a low market return forecast, will imply that the asset allocation of the representative agent should be tilted away from the risky asset. This is actually confirmed in Figure 6, which has already shown that from 2000 onwards, the *cay* portfolio was almost always invested in the risk-free asset and therefore underperformed dramatically.

5.3 The role of a deterministic time trend

In this subsection, we revisit the possibility that the structural shift in the cointegrating coefficients could be addressed through allowing for a deterministic time trend. This is ultimately a question of model choice regarding the cointegrating relationship, an issue extensively discussed in Koop et al. (2008) which we pick up in Appendix A.6. The relevance of an omitted time trend in the cointegrating relationship was first brought to attention by Brennan and Xia (2005), who argue that a mechanistic variable, which they call *tay*, that simply replaces consumption with calendar time in the estimation of the cointegrating relationship, predicts asset returns at least as good as *cay*.[6] A more detailed analysis of the time trend is provided by Hahn and Lee (2006), who show that that the restricted version of *cay* (i.e. without time trend) from Lettau and Ludvigson (2001) is nothing but the sum of a bias component and the unrestricted *cay* estimated with a deterministic time trend. Their main finding is that the time trend takes away a substantial part of the forecasting ability of *cay*, or in other terms, that the unrestricted *cay* (with trend) does show substantially less forecasting power.

To investigate the role of the time trend, we follow Hahn and Lee (2006) and augment equation (9) by calendar time t and derive *cay* from $\widehat{cay}_t = c_t - \hat{\pi}\, t - \hat{\beta}_a a_t - \hat{\beta}_y y_t$, where $\hat{\pi}$ is the estimated coefficient on time t. We estimate the cointegration regression sequentially in an expanding window and report the time variation of the coefficients $\hat{\pi}_t$, $\hat{\beta}_{a,t}$, and $\hat{\beta}_{y,t}$ in Figure 10. We find that the deterministic time trend was strongly significant until 2010 (with t-values of more than 6), but sharply dropped thereafter to become insignificant from the perspective of most recent data.[7] The inclusion of the time trend has the remarkable effect that the coefficient

[6] In their strong reply to Brennan and Xia (2005), Lettau and Ludvigson (2005b) not only argue that from a theoretical standpoint, a meaningful long-run equilibrium between consumption, aggregate wealth and labor income does not allow for a deterministic trend. They also claim that *tay* contains more economic content than admitted by Brennan and Xia (2005), by the mere fact that *tay* is nothing but a proxy for *cay* due to a large part of the variability of consumption being governed by a simple deterministic time trend. A forward-looking consumer anticipates this relationship which brings it perfectly in line with the model assumptions.

[7] As of 2019:4, $\hat{\pi} = 0.024$ with a t-stat of 0.303. Estimation by DLS.

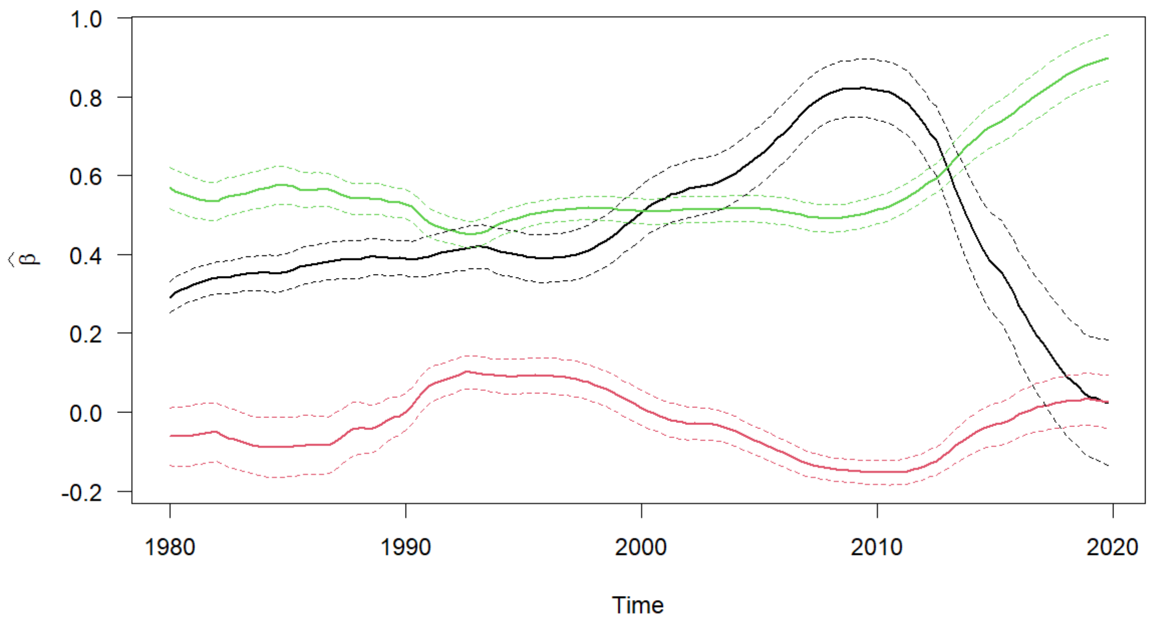


Figure 10: Cointegrating parameters for time trend (black), asset wealth (red) and labor income (green) quarterly re-estimated. The dashed lines represent the corresponding confidence bands of ± 2 standard errors. The sample period is 1952:1-2019:4.

on wealth, $\hat{\beta}_{a,t}$, would have never been significantly different from zero, which in turn leads to the result that *cay*, estimated from the unrestricted model, would have never displayed forecasting power. This result reinforces the findings of Hahn and Lee (2006) and their conclusion that the good (in-sample) forecasting performance of *cay* was (at least partly) due to omitting a time trend. However, note that the situation changes dramatically after 2010, where the inclusion of the time trend is no longer warranted from an econometric point of view. Furthermore, note that the decline of the time trend goes along with a substantial increase in $\hat{\beta}_{y,t}$, i.e. the coefficient on income after 2010. So, the time trend apparently absorbs the impact of wealth prior to 2010, but we still observe a similar structural shift as already documented above in Figure 8. In sum, the analysis of a time trend reveals some irony of history. Prior to 2010, *cay* showed good (in-sample) forecasting power, but was subject to the potential criticism that it earned its predictability due to the omitted (but statistically significant) time trend. Since 2010, the data does support the theoretically motivated assumption of omitting a time trend, but even the restricted *cay* no longer displays forecasting power.

5.4 The Covid-19 pandemic episode

We deliberately ended our sample in 2019:4 and excluded observations that were impacted by the Covid-19 pandemic, hitting the U.S. in the first quarter of 2020. In this section, we illustrate the impact of the pandemic-affected observations on *cay* and its predictive ability. Recall again that a low *cay* is assumed to predict low future returns, since the representative agent reduces consumption in anticipation of lower growth rates for wealth. It is therefore instructive to first look at the development of the consumption and income time series during the Covid-19 episode, whose quarterly growth rates are shown in the upper two panels of Figure 11. It is more than obvious that the pandemic-affected quarters display extraordinary behavior. In 2020:1 and 2020:2, consumption growth dropped at unprecedented magnitude while income spiked upwards in 2020:2 and 2021:1. A decline in consumption (i.e. the numerator) together with an increase in income (i.e. the denominator) propelled the consumption-wealth ratio eight standard deviations below its historic mean (see third panel in Figure 11). Within the interpretation of the theory, it would be a signal that the representative agent has massively reduced consumption because she had anticipated extraordinarily low future returns. While it is true that the pandemic induced a negative shock on growth expectations, a closer inspection of the episode clearly shows that the decline in consumption is primarily due to government-imposed shutdowns, while the huge spike in income can be explained by the substantial transfer payments that were made in the context of various rescue packages. Therefore, although the sign would be correct, it is hard to argue that during the Covid-19 episode *cay* is able to forecast future returns on the basis of its theoretical basis. Consumption declined rather due to the inability to spend than due to anticipated low growth rates.

In sum, observations that are affected by the Covid-19 pandemic are subject to circumstances (economy shutdown, transfer payments) that are not captured by the theoretical model assumptions. Furthermore, by their extraordinary order of magnitude, these observations massively affect the proper estimation of $\widehat{cay}_t$. Unsurprisingly, this leads to a complete disappearance of any stock return predictability. Therefore, we limit ourselves to the sample ending in the fourth quarter of 2019. Moreover, when addressing the decay of *cay*, we always refer to the fundamental changes in the underlying cointegrating relationship and the consequential decline in predictability within this sample.

6 Robustness

6.1 Avoiding a two-step approach

Confronted with the criticism that the construction of *cay* as cointegrating residual involves the alleged 'look-ahead bias', Lettau and Ludvigson (2005b) argue that the same results can also be obtained by avoiding this two-step approach and directly estimating a multivariate regression

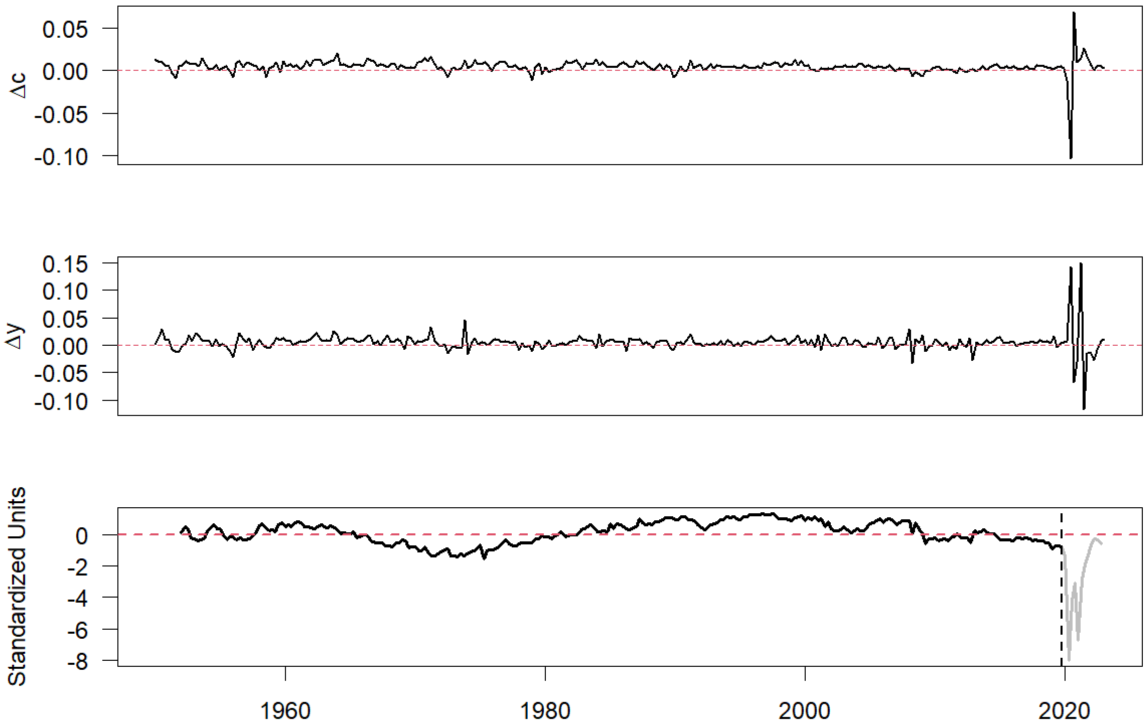


Figure 11: The two plots in the top show the (log) growth rates of consumption and labor income, respectively. The bottom plot shows $\widehat{cay}_t$ in standardized units. The sample period is 1952:1-2022:4.

of excess returns on consumption, aggregate wealth and labor income. This procedure is valid under the assumption that the series are cointegrated and ensures that there are no future observations in use. Taking for the moment an agnostic stance on the underlying assumption of cointegration, we investigate whether the decay of *cay*, i.e. the disappearance of the forecasting ability of *cay* is also found from the direct multivariate predictive regression. In line with Lettau and Ludvigson (2005b), we estimate

$$\tilde{r}_{t,H} = \alpha + \gamma_c\, c_t + \gamma_a\, a_t + \gamma_y\, y_t + \varepsilon_{t,H} \tag{15}$$

and report results for the full sample in Table 6 for horizons up to $H = 20$ quarters. Clearly, there are no significant results for any of the three regressors over any horizon. Moreover, the coefficient estimates for consumption in the first row are especially for the shorter horizons very close to the ones for $\widehat{cay}_t$ from Table 4. This is exactly what we would asymptotically expect in large samples. Taken together, the results in Table 6 document that from most recent data, future returns cannot be predicted over any horizon from the direct multivariate regression.

To illustrate the temporal development, we plot the recursively re-estimated coefficient estimates γ_c, γ_a, γ_y of consumption, aggregate wealth and labor income respectively from quarterly

	Dependent variable: $\tilde{r}_{t,H}$						
	$H = 1$	$H = 2$	$H = 4$	$H = 8$	$H = 12$	$H = 16$	$H = 20$
c_t	0.238	0.468	0.848	1.689	2.455	3.446	4.320
	(1.060)	(1.024)	(0.708)	(0.741)	(0.696)	(0.735)	(1.009)
a_t	-0.015	-0.054	-0.158	-0.315	-0.426	-0.544	-0.765
	(-0.323)	(-0.597)	(-0.570)	(-0.469)	(-0.483)	(-0.398)	(-0.850)
y_t	-0.211	-0.388	-0.637	-1.277	-1.887	-2.696	-3.275
	(-0.926)	(-0.870)	(-0.538)	(-0.596)	(-0.604)	(-0.756)	(-0.930)
$\bar{R}^2$	-0.006	-0.002	0.007	0.027	0.041	0.064	0.077
Note:					*p<0.1; **p<0.05; ***p<0.01		

Table 6: The table reports regression estimates for c_t, a_t and y_t and adjusted R^2 statistics from forecasts of H-period ahead excess returns spanning different horizons from one quarter up to five years. Newey and West (1987) corrected t-statistics appear in parenthesis. The sample period is 1952:1-2019:4.

forecast regressions in Figure 12.

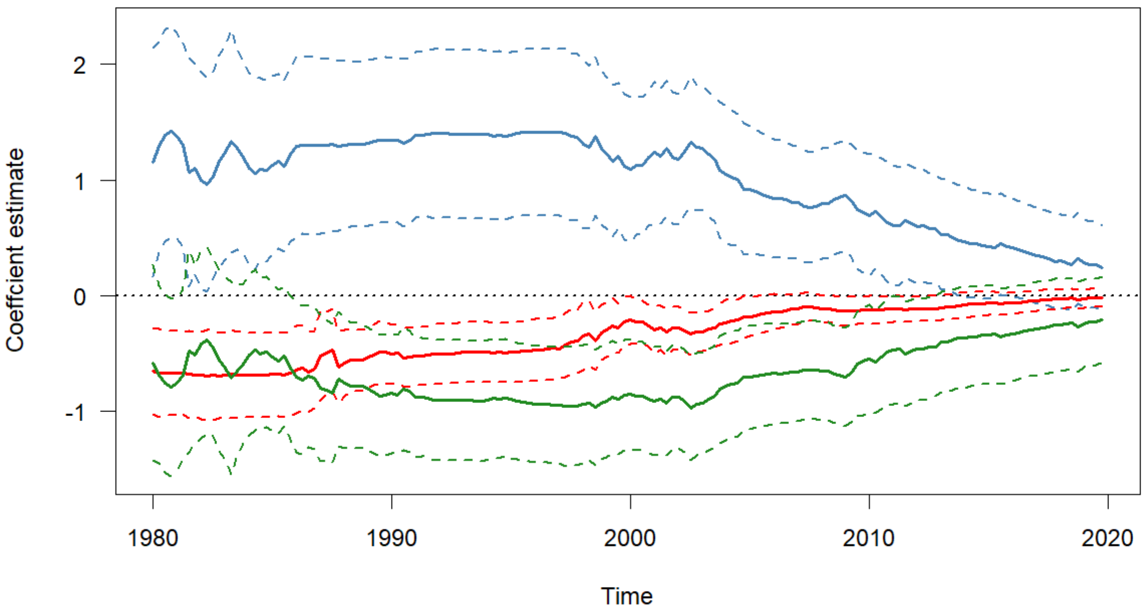


Figure 12: The plot shows the coefficient estimates of consumption (blue), aggregate wealth (red) and labor income (green) γ_c, γ_a, γ_y accompanied by their 90%-confidence bands based on Newey and West (1987)-corrected standard errors. The sample period is 1952:1-2019:4.

We find the by now familiar pattern, that until around 2005, coefficient estimates are significant and roughly stable. From then on all three coefficients tend to approach zero and are insignificant from around 2010. The coefficient on consumption γ_c (blue line) shows an almost identical pattern as the coefficient on $\widehat{cay}_t$ itself as shown in Figure 3. This result is in line with asymptotic theory which implies that both approaches should be the same in large samples. These results confirm that the decay of *cay* is not simply an artefact of a two-step estimation procedure but directly translates to the underlying series that form *cay* through their cointegrating relationship.

6.2 Using PCE as consumption flow measure

As mentioned in Section 3 already, in a revision to their original paper on *cay*, Lettau and Ludvigson (2019) argue that in light of a continuously declining ratio of NDS over PCE over the last four decades, it is meanwhile more appropriate to consider PCE as the measure of consumption in the estimation of $\widehat{cay}_t$. So far, we purposely used NDS to track the historical development of *cay* and its forecasting ability, and to be comparable to prior literature. As a robustness check, we will repeat (part of) our analysis by using PCE consumption instead of NDS. To avoid any confusion, we will denote the consumption measure in use by a superscript from here on out. The estimated cointegrating parameters using PCE within the full sample turn out to be

$$\widehat{cay}_t^{\mathrm{PCE}} = c_t^{\mathrm{PCE}} - \underset{(7.67)}{0.232}\ a_t - \underset{(24.22)}{0.788}\ y_t, \tag{16}$$

with t-stats in brackets below coefficient estimates. Using $\widehat{cay}_t^{\mathrm{PCE}}$ as predictor, we summarize multiple horizon (in-sample) predictive regression results in Table 7. Overall, results based on

	Dependent variable: $\tilde{r}_{t,H}$						
	$H = 1$	$H = 2$	$H = 4$	$H = 8$	$H = 12$	$H = 16$	$H = 20$
$\widehat{cay}_t^{\mathrm{PCE}}$	0.438	0.937*	2.017	4.202*	5.710**	6.935***	7.759***
	(1.617)	(1.855)	(1.480)	(1.830)	(2.116)	(3.462)	(3.303)
	[0.007]	[0.019]	[0.047]	[0.114]	[0.156]	[0.193]	[0.188]
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 7: The table reports forecasting regression estimates for $\widehat{cay}_t^{\mathrm{PCE}}$. Forecasts span different horizons from one quarter up to five years. Newey and West (1987) corrected t-statistics appear in parenthesis and adjusted R^2 statistics in square brackets. The sample period is 1952:1-2019:4.

PCE are better than with NDS data. The cointegrating coefficient on wealth is larger and significant although showing also a decline in the last ten years. Returns up to $H = 8$ quarters are not reliably predictable from $\widehat{cay}_t^{\mathrm{PCE}}$, but become significant for longer horizons. While being better, we still find a comparable decay in the forecasting performance as illustrated in Figure 13, which plots the coefficient estimate over time in the predictive regression for our baseline, top-10% and unfiltered specification, each estimated using PCE instead of NDS. It is the exact analogue of Figure 3. For the baseline and the unfiltered specification the picture is almost the same. After a long-lasting period in which both the estimate with its economic impact, as well as the adjusted R^2 statistics, stayed more or less stable on a certain level, they began to decay around 2002/03. This observation is almost identical to the case where NDS is the consumption measure. The only notable difference occurs in the case of the top-10% specification. Here, the decay is still present but less pronounced. For example, the coefficient estimate of the full sample is still significant and also the economic impact is much higher at 126 basis points, compared to only 87 in the baseline specification. This is interesting, since until the early 2000s, both were almost exactly at the same level. The fact that the top-10% specification yields much stronger forecasting results is again a consequence of the very different behavior of consumption and labor income on the one side and aggregate wealth on the other especially after the financial crisis. It shows that adjusting the consumption-wealth ratio such that it captures the marginal investor's behavior is a meaningful improvement.

The general finding that *cay* behaves comparatively better but otherwise similar for PCE instead of NDS as the consumption measure, and most importantly that it exhibits the above documented decay, also holds up when considering its out-of-sample as well as its economic significance. All of this taken together leads us to conclude that the decay of *cay* is not merely a result of a measurement issue with consumption.

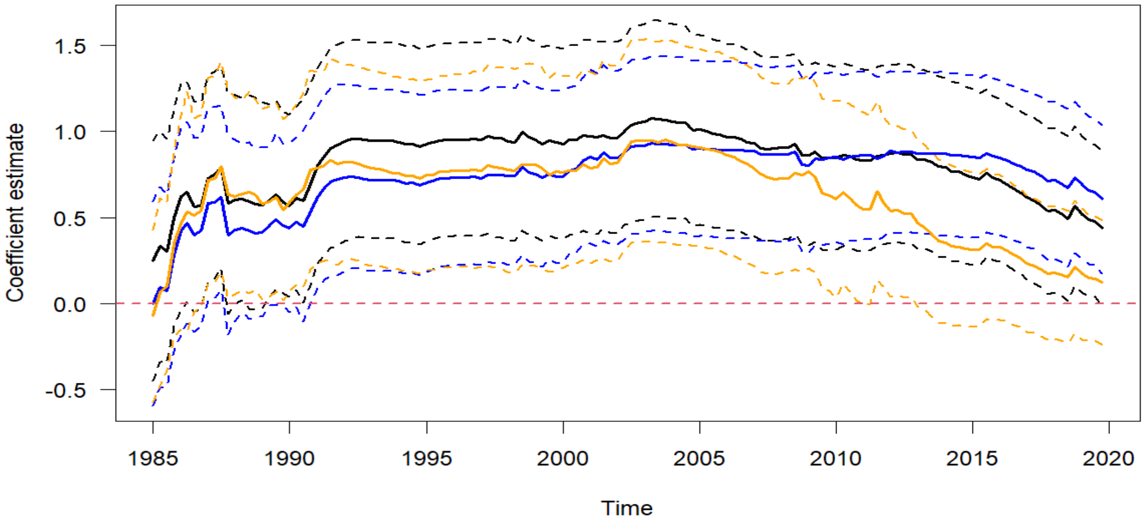
6.3 Disaggregating asset wealth

Sousa (2010) proposes an improved version of *cay*, called *cday*, where aggregate (asset) wealth is separated (or disaggregated) into financial wealth and housing wealth. He argues that it is mainly financial wealth that captures transitory movements in the long-run cointegrating relationship between consumption, asset wealth and labor income, and shows that over a sample period of 1975:1-2008:4, *cday* slightly outperformed *cay* in terms of forecasting ability in the U.S. The overall methodology in estimating *cday* is analogous to the one for *cay* from Lettau and Ludvigson (2001) and only splits asset wealth, A_t, up into the sum of financial wealth, F_t, and housing wealth, U_t. We follow that approach and use four leads and lags of the first differences of financial wealth, housing wealth, and labor income in the DLS specification of *cday*. The cointegrating parameter estimates for $\widehat{cday}_t$ obtained in this way are

$$\widehat{cday}_t^{\text{NDS}} = c_t^{\text{NDS}} + \underset{(-1.03)}{0.019}\ f_t - \underset{(6.09)}{0.093}\ u_t - \underset{(29.96)}{0.856}\ y_t,$$

$$\widehat{cday}_t^{\text{PCE}} = c_t^{\text{PCE}} - \underset{(5.61)}{0.109}\ f_t - \underset{(8.32)}{0.117}\ u_t - \underset{(29.70)}{0.800}\ y_t,$$

for our sample period 1952:1-2019:4. A comparison with the results from Sousa (2010) with



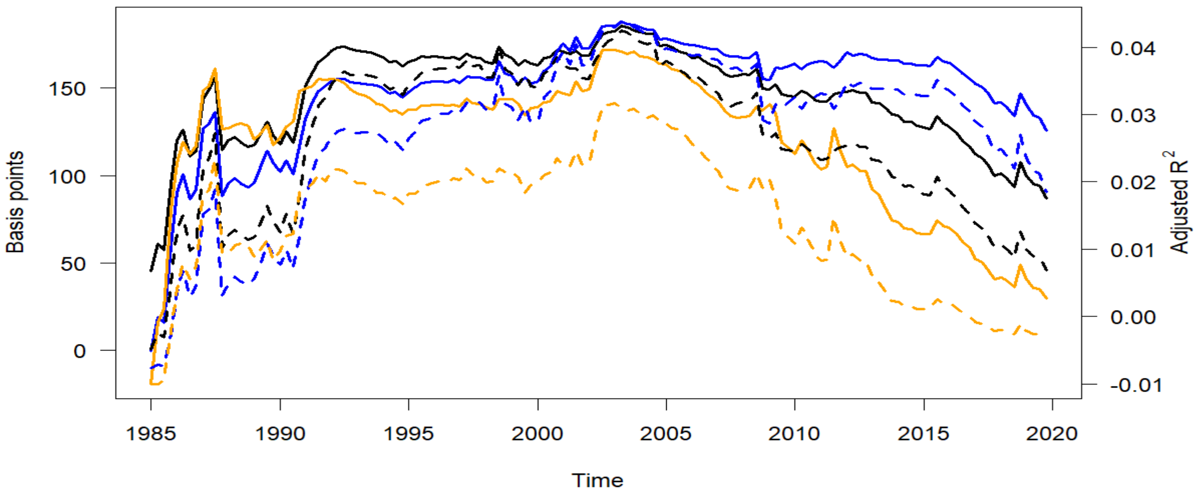


Figure 13: Both plots show results from quarterly re-estimated one-quarter ahead forecast regressions of excess returns on $\widehat{cay}_t$ (black), $\widehat{cay}_t^{\text{top10}}$ (blue) and $\widehat{cay}_t^{\text{unfil}}$ (orange) with PCE as the consumption measure. The upper plot represents the coefficient estimates as the solid lines accompanied by their respective 90%-confidence band based on Newey and West (1987) standard errors as the dashed lines. The bottom plot shows the impact of a one-standard deviation increase in *cay* on one-quarter ahead excess returns in basis points as the solid lines on the left axis and the adjusted R^2 statistics on the right axis. The sample period is 1952:1-2019:4, except for cay_t^{unfil} it is 1960:3-2019:4.

coefficients of 0.16, 0.02 and 1.02, however, is somewhat delicate due to the very different sample period.

In general, *cday* behaves similarly to *cay*, although the higher sensitivity of *cday* towards changes in housing wealth became distinctly visible during the course of the financial crisis of 2007/08 when housing wealth was deviating massively from its long-term trend with consumption, financial wealth and labor income. Figure 14 shows $\widehat{cday}_t$ over our full sample period in comparison to $\widehat{cay}_t$ for both (baseline) consumption measures. The results of our forecast regressions using *cday* are presented in Table 8.

	Dependent variable: $\tilde{r}_{t,H}$						
	$H=1$	$H=2$	$H=4$	$H=8$	$H=12$	$H=16$	$H=20$
$\widehat{cday}_t^{\text{NDS}}$	0.285	0.560	0.994	1.883	2.477	3.102	2.996
	(1.225)	(1.138)	(0.805)	(0.732)	(0.571)	(0.605)	(0.586)
	[0.001]	[0.005]	[0.010]	[0.022]	[0.030]	[0.041]	[0.030]
$\widehat{cday}_t^{\text{PCE}}$	0.471	1.023*	2.244*	4.728**	6.318**	7.464**	7.682**
	(1.630)	(1.751)	(1.686)	(2.137)	(2.381)	(2.513)	(2.293)
	[0.008]	[0.021]	[0.053]	[0.127]	[0.166]	[0.194]	[0.162]
Note:				*p<0.1; **p<0.05; ***p<0.01			

Table 8: The table reports forecasting regression estimates for $\widehat{cday}_t^{\text{NDS}}$ and $\widehat{cday}_t^{\text{PCE}}$. Forecasts span different horizons from one quarter up to five years. Newey and West (1987) corrected t-statistics appear in parenthesis and adjusted R^2 statistics in square brackets. The sample period is 1952:1-2019:4.

We see that the coefficient estimates and significance levels are very much comparable to the results for $\widehat{cay}_t$. The same holds when assessing the in-sample and out-of-sample power of *cday*. This shows that even though splitting up asset wealth into financial and housing wealth does slightly improve the forecasting ability of *cay*, it cannot eliminate nor explain the decay of $c(d)ay$.

6.4 Including volatility into the forecasting regression

In his paper, Guo (2006) shows that the forecasts of quarterly-ahead excess returns based solely on *cay* suffer from an omitted variable problem. He includes aggregate stock market volatility, σ_m^2, into the forecast regression to get much stronger results for both the coefficient estimate of *cay* as well as the explained variation in terms of R^2 statistics.[8] He mainly attributes this finding to the fact that while *cay* and σ_m^2 are negatively correlated to one another, they are both positively correlated with future stock returns. He furthermore shows that including the stochastically detrended risk-free rate, labelled *rrel*, into the forecast regressions yields significant estimates for all three variables and leads to an even higher R^2 statistic of 16.3% for the full sample of 1952:3-2002:4. These findings are much stronger when taking only the first half of the sample into account. Especially the finding that *rrel* has some further explanatory

[8]Similarly strong results are found in the later work of Guo et al. (2013).

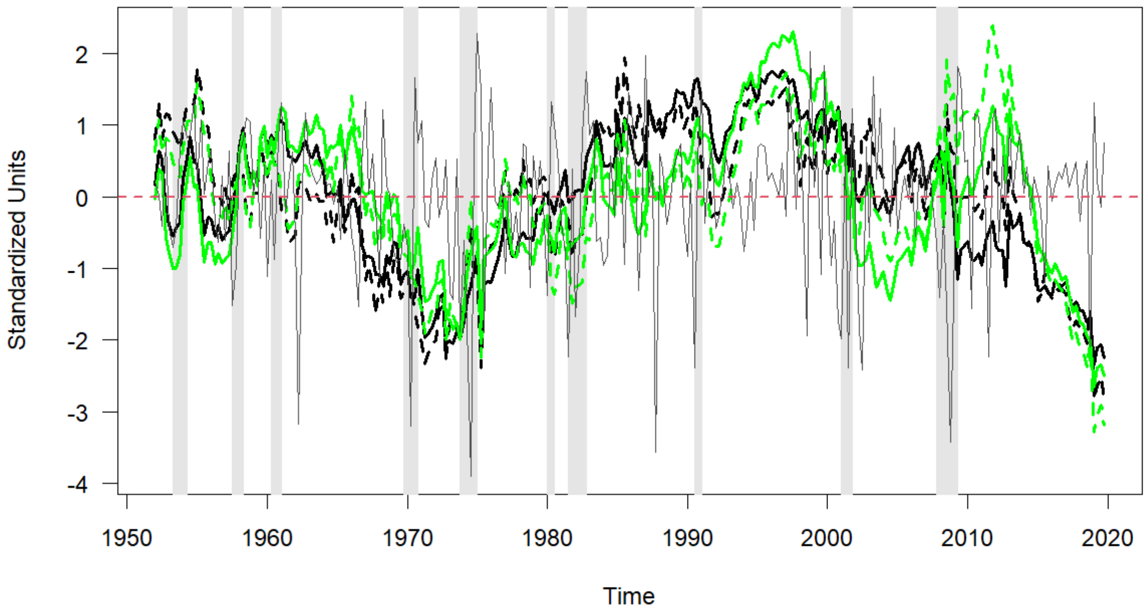


Figure 14: The plot shows $\widehat{cay}_t^{\mathrm{NDS}}$ and $\widehat{cday}_t^{\mathrm{NDS}}$ in black and green as the solid lines. The dashed lines represent $\widehat{cay}_t^{\mathrm{PCE}}$ and $\widehat{cday}_t^{\mathrm{PCE}}$, respectively. The sample period is 1952:1-2019:4.

power does not hold in the second half of the sample, whereas the omitted variable problem due to excluding σ_m^2 seems to still exist.

We address the omitted variables problem and follow Guo's (2006) approach by adding the variance of our market portfolio to the forecast regressions. We compute σ_m^2 by taking the sum of the squared daily log returns of the CRSP NYSE/NYSE MKT/NASDAQ/Arca Value-Weighted Market Index for each quarter. The results for the full sample are presented in Table 9. We also experimented with including the relative bill rate, i.e., $rrel$, into the forecasts but did not find any significant results for $rrel$ itself, or any sizeable improvement in R^2 statistics. Therefore, we decided to only incorporate σ_m^2 into our forecasts for the remainder of this section.[9]

In principal, the results including stock market volatility into the forecast regressions are just as bad as the ones for only using *cay* as a predictor variable. This at first sight contrary finding to Guo (2006), leads us to believe that the decay of *cay* is still present even when adding stock market volatility to our forecast regressions. Indeed, when again re-estimating the quarterly

[9]Different from Guo (2006), we did not adjust downward observations with an unusually high amplitude of realized stock market variance. The main reason for that is, next to the somewhat arbitrary decision which observations should be adjusted downward, that our findings are simply not sensitive to downward adjustment. By that we mean that we indeed do get statistically stronger results when adjusting downward, but our main interest lies in the decay of *cay* which is given in almost the same way, independent of downward adjustment. We therefore spare ourselves the decision which observations should be adjusted downward.

	Dependent variable: $\tilde{r}_{t,H}$						
	$H = 1$	$H = 2$	$H = 4$	$H = 8$	$H = 12$	$H = 16$	$H = 20$
$\widehat{cay}_t$	0.238	0.453	0.794	1.502	2.137	2.949	3.481
	(1.072)	(1.009)	(0.670)	(0.654)	(0.654)	(0.945)	(0.973)
$\sigma^2_{m,t}$	0.178	1.216	1.886	2.978	1.875	2.866	5.249**
	(0.215)	(1.478)	(1.529)	(1.557)	(0.854)	(1.368)	(2.518)
$\bar{R}^2$	-0.003	0.009	0.015	0.031	0.031	0.056	0.077
Note:					*p<0.1; **p<0.05; ***p<0.01		

Table 9: The table reports forecasting regression estimates for $\widehat{cay}_t$ and $\sigma^2_{m,t}$. Forecasts span different horizons from one quarter up to five years. Newey and West (1987) corrected t-statistics appear in parenthesis and adjusted R^2 statistics in square brackets. The sample period is 1952:1-2019:4.

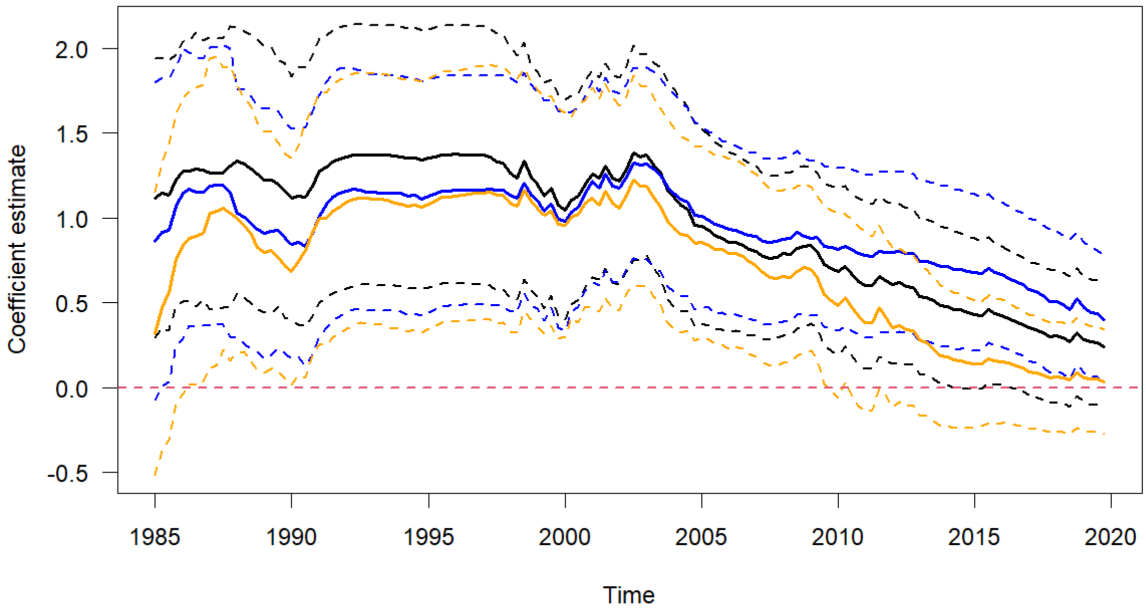


Figure 15: The plot shows the coefficient estimates of $\widehat{cay}_t$ (black), $\widehat{cay}_t^{\text{top10}}$ (blue) and $\widehat{cay}_t^{\text{unfil}}$ (orange), accompanied by their 90%-confidence bands as the dashed lines, from quarterly forecast regressions of excess returns that also include the realized stock market variance $\sigma^2_{m,t}$. The sample period is 1952:1-2019:4, except for cay_t^{unfil} it is 1960:3-2019:4.

forecasts step by step, we get the already familiar picture of a strongly significant coefficient estimate in the past that has declined sharply over the last roughly 15 years, which is illustrated in Figure 15. We can thus conclude that the decay of *cay* is not only robust against changes in the methodology in the estimation of $\widehat{cay}_t$ as seen above, but also not affected by an omitted variable problem in the forecast regression itself.

7 Conclusion

In this paper we show that the ability of the consumption-wealth ratio *cay* to predict future market excess returns has dramatically declined over more than the last ten years both in-sample, out-of-sample as well as in terms of economic performance. In an attempt to improve the basic construction of *cay*, we propose two alternative derivations: First, we take up the argumentation put forward in e.g. Lettau et al. (2019), that due to wealth inequality, the representative agent may be better captured by the richest households. Therefore, we determine consumption and wealth data for the top-10% of richest households. Second, we make use of an unfiltering process of the consumption data in order to recover richer time variation in the spirit of Kroencke (2017). Neither approach is able to substantially alter our main conclusion. Although in particular the top-10% implementation is able to yield some improvement in the forecasting ability, we still find a pronounced decline in predictability over the last 10-15 years. We conduct extensive robustness checks, but neither the avoidance of estimating the cointegrating residual, nor the decomposition into financial and housing wealth, nor the inclusion of potentially omitted variables changes our results qualitatively. Arguably the best improvement is found for estimating *cay* from PCE instead of NDS consumption data and adjusting for the top-10% households, for which at least for longer horizons predictive regressions still deliver significant (in-sample) results. However, although significance levels are better, we do observe the same declining pattern.

We identify the source of this decay to be a structural shift in the cointegrating relationship that governs the comovement between consumption and wealth. In particular asset wealth appears to be increasingly detached from its long-run relation with consumption and labor income over at least the last two decades. This behavior implies that the cointegrating parameters of asset wealth and labor income are increasingly drifting apart. Among others, we show that this structural shift drives the poor OOS performance as well as the failure to generate utility gains over the past two decades. We do not take a stance on whether this behavior results from a long-lasting deviation from a previously reached equilibrium or convergence towards another equilibrium, which may be a question for future research. The bottom line of our contribution is the comprehensive documentation of the recent *decay of cay* as a meaningful predictor of asset returns, which casts doubt if a predictor derived from the representative agent's intertemporal budget constraint can meaningfully predict future stock market behavior.

References

Andrade, S. C., Babenko, I., and Tserlukevich, Y. (2006). Market Timing with cay. *The Journal of Portfolio Management*, 32(2):70–80.

Avramov, D. (2002). Stock Return Predictability and Model Uncertainty. *Journal of Financial Economics*, 64(3):423–458.

Blinder, A. S., Deaton, A., Hall, R. E., and Hubbard, R. G. (1985). The Time Series Consumption Function Revisited. *Brookings Papers on Economic Activity*, 1985(2):465.

Brennan, M. J. and Xia, Y. (2005). Tay's as good as cay. *Finance Research Letters*, 2(1):1–14.

Campbell, J. Y. (1996). Understanding Risk and Return. *Journal of Political Economy*, 104(2):298–345.

Campbell, J. Y. and Mankiw, N. G. (1989). Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence. *NBER Macroeconomics Annual*, 4:185–216.

Campbell, J. Y. and Thompson, S. B. (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *The Review of Financial Studies*, 21(4):1509–1531.

Della Corte, P., Sarno, L., and Valente, G. (2010). A Century of Equity Premium Predictability and the Consumption–Wealth Ratio: An International Perspective. *Journal of Empirical Finance*, 17(3):313–331.

Engle, R. F. and Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2):251.

Fleming, J., Kirby, C., and Ostdiek, B. (2001). The Economic Value of Volatility Timing. *The Journal of Finance*, 56(1):329–352.

Galí, J. (1990). Finite Horizons, Life-Cycle Savings, and Time-Series Evidence on Consumption. *Journal of Monetary Economics*, 26(3):433–452.

Goetzmann, W., Ingersoll, J., Spiegel, M., and Welch, I. (2007). Portfolio Performance Manipulation and Manipulation-Proof Performance Measures. *The Review of Financial Studies*, 20(5):1503–1546.

Goyal, A., Welch, I., and Zafirov, A. (2021). A Comprehensive 2021 Look at the Empirical Performance of Equity Premium Prediction II. *SSRN Electronic Journal*.

Guo, H. (2006). On the Out–of–Sample Predictability of Stock Market Returns. *The Journal of Business*, 79(2):645–670.

Guo, H., Wang, Z., and Yang, J. (2013). Time-Varying Risk-Return Trade-off in the Stock Market. *Journal of Money, Credit and Banking*, 45(4):623–650.

Hahn, J. and Lee, H. (2006). Interpreting the Predictive Power of the Consumption–Wealth Ratio. *Journal of Empirical Finance*, 13(2):183–202.

Inoue, A. and Kilian, L. (2005). In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use? *Econometric Reviews*, 23(4):371–402.

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.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring Uncertainty. *American Economic Review*, 105(3):1177–1216.

Kalotay, E., Gray, P., and Sin, S. (2007). Consumer Expectations and Short-Horizon Return Predictability. *Journal of Banking & Finance*, 31(10):3102–3124.

Koop, G., Potter, S. M., and Strachan, R. W. (2008). Re-Examining the Consumption–Wealth Relationship: The Role of Model Uncertainty. *Journal of Money, Credit and Banking*, 40(2-3):341–367.

Kostakis, A., Magdalinos, T., and Stamatogiannis, M. P. (2023). Taking stock of long-horizon predictability tests: Are factor returns predictable? *Journal of Econometrics*, page 105380.

Kroencke, T. A. (2017). Asset Pricing without Garbage. *The Journal of Finance*, 72(1):47–98.

Lettau, M. and Ludvigson, S. C. (2001). Consumption, Aggregate Wealth and Expected Stock Returns. *The Journal of Finance*, 56(3):815–849.

Lettau, M. and Ludvigson, S. C. (2005a). Expected Returns and Expected Dividend Growth. *Journal of Financial Economics*, 76(3):583–626.

Lettau, M. and Ludvigson, S. C. (2005b). Tay's as good as cay: Reply. *Finance Research Letters*, 2(1):15–22.

Lettau, M. and Ludvigson, S. C. (2019). Using PCE Consumption in cay_t.

Lettau, M., Ludvigson, S. C., and Ma, S. (2019). Capital Share Risk in U.S. Asset Pricing. *The Journal of Finance*, 74(4):1753–1792.

Lustig, H., van Nieuwerburgh, S., and Verdelhan, A. (2013). The Wealth-Consumption Ratio. *The Review of Asset Pricing Studies*, 3(1):38–94.

McCracken, M. W. (2007). Asymptotics for Out of Sample Tests of Granger Causality. *Journal of Econometrics*, 140(2):719–752.

Mehra, R. and Prescott, E. C. (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics*, 15(2):145–161.

Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3):703.

Ogaki, M. and Park, J. Y. (1997). A Cointegration Approach to Estimating Preference Parameters. *Journal of Econometrics*, 82(1):107–134.

Park, J. Y. (1992). Canonical Cointegrating Regressions. *Econometrica*, 60(1):119.

Phillips, P. C. B. and Hansen, B. E. (1990). Statistical Inference in Instrumental Variables Regression with I(1) Processes. *The Review of Economic Studies*, 57(1):99.

Phillips, P. C. B. and Ouliaris, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica*, 58(1):165.

Ren, Y., Yuan, Y., and Zhang, Y. (2014). Human Capital, Household Capital and Asset Returns. *Journal of Banking & Finance*, 42:11–22.

Rudd, J. and Whelan, K. (2006). Empirical Proxies for the Consumption–Wealth Ratio. *Review of Economic Dynamics*, 9(1):34–51.

Savov, A. (2011). Asset Pricing with Garbage. *The Journal of Finance*, 66(1):177–201.

Sousa, R. M. (2010). Consumption, (Dis) Aggregate Wealth and Asset Returns. *Journal of Empirical Finance*, 17(4):606–622.

Stock, J. H. (1987). Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. *Econometrica*, 55(5):1035.

Stock, J. H. and Watson, M. W. (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica*, 61(4):783.

Welch, I. and Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *The Review of Financial Studies*, 21(4):1455–1508.

West, K. D., Edison, H. J., and Cho, D. (1993). A Utility-Based Comparison of Some Models of Exchange Rate Volatility. *Journal of International Economics*, 35(1-2):23–45.

A Appendix

A.1 Other cointegration estimates

Next to the DLS method used throughout the paper when estimating the cointegrating parameters of the long-run relationship between consumption, aggregate wealth and labor income, there are also other approaches of getting estimates for said cointegration relation. Table 10 presents these for three other methods, namely simple OLS, Park's (1992) canonical cointegrating regression (CCR) and Phillips and Hansen's (1990) fully modified estimator (FME).

	OLS	CCR	FME
Panel A: baseline specification			
$\widehat{\beta}_a$	0.036	0.028	0.028
	(1.102)	(0.872)	(0.876)
$\widehat{\beta}_y$	0.902***	0.910***	0.910***
	(25.908)	(23.786)	(23.785)
Panel B: top-10% specification			
$\widehat{\beta}_a$	0.073***	0.071**	0.071**
	(2.609)	(2.223)	(2.235)
$\widehat{\beta}_y$	0.876***	0.878***	0.878***
	(33.485)	(26.570)	(26.561)
Panel C: unfiltered specification			
$\widehat{\beta}_a$	0.027	0.016	0.017
	(0.790)	(0.413)	(0.418)
$\widehat{\beta}_y$	0.908***	0.920***	0.920***
	(23.098)	(18.234)	(18.092)
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 10: Cointegrating parameters for standard OLS, Park's (1992) canonical cointegrating regression and Phillips and Hansen's (1990) fully modified estimation. Newey and West (1987)-corrected t-statistics appear in parentheses.

We see that all methods produce similar cointegrating parameters and by that also closely related forecasts, due to the fact that $\widehat{cay}_t$ is being estimated using almost the same parameters. Table 12 reports the results of forecasts based on the baseline specification using the full sample period.

	Dependent variable: $\tilde{r}_{t,H}$						
	$H = 1$	$H = 2$	$H = 4$	$H = 8$	$H = 12$	$H = 16$	$H = 20$
$\widehat{cay}_t^{\mathrm{OLS}}$	0.285	0.560	0.994	1.883	2.477	3.102	2.996
	(1.225)	(1.138)	(0.805)	(0.732)	(0.571)	(0.605)	(0.586)
	[0.001]	[0.005]	[0.010]	[0.022]	[0.030]	[0.041]	[0.030]
$\widehat{cay}_t^{\mathrm{CCR}}$	0.471	1.023*	2.244*	4.728**	6.318**	7.464**	7.682**
	(1.630)	(1.751)	(1.686)	(2.137)	(2.381)	(2.513)	(2.293)
	[0.008]	[0.021]	[0.053]	[0.127]	[0.166]	[0.194]	[0.162]
$\widehat{cay}_t^{\mathrm{FME}}$	0.471	1.023*	2.244*	4.728**	6.318**	7.464**	7.682**
	(1.630)	(1.751)	(1.686)	(2.137)	(2.381)	(2.513)	(2.293)
	[0.008]	[0.021]	[0.053]	[0.127]	[0.166]	[0.194]	[0.162]
Note:					*p<0.1; **p<0.05; ***p<0.01		

Table 11: The table reports forecasting regression estimates for $\widehat{cay}_t$ being estimated with OLS, CCR and FME instead of DLS. Forecasts span different horizons from one quarter up to five years. Newey and West (1987) corrected t-statistics appear in parenthesis and adjusted R^2 statistics in square brackets. The sample period is 1952:1-2019:4.

A.2 Data

Consumption. Consumption is defined in two different ways. One being nondurables plus services minus clothing and footwear (NDS). The other is total personal consumption expenditures (PCE). Data are quarterly, seasonally adjusted at an annual rate and measured in billions of dollars. Series comprise the period 1947:1–2022:4. The source is the U.S. Bureau of Economic Analysis, NIPA Table 2.3.5, lines 1, 8, 10, 13.
Each series is separately deflated to 2012-dollars with its respective deflator, see below, before NDS is computed as the sum of nondurables and services minus clothing and footwear. Series are transformed in per capita terms and put in the logarithmic form.

Aggregate wealth. Aggregate wealth is defined as the net worth of households and nonprofit organizations. Data are quarterly, end of period, not seasonally adjusted. Series comprises the period 1951:4–2022:4. The source is the Board of Governors of the Federal Reserve System, Financial Accounts, Table B.101, line 40.

Financial wealth. Financial wealth is defined as financial assets (debt securities and loans, corporate equities, mutual fund shares, deposits, life insurance reserves, pension entitlements, miscellaneous assets, equity in noncorporate business, and grants and trade receivables) minus financial liabilities (debt security and loans, trade payables, and deferred and unpaid life insurance premiums). Data are quarterly, end of period, not seasonally adjusted. Series comprise the period 1951:4–2022:4. The source is the Board of Governors of the Federal Reserve System, Financial Accounts, Table B.101, lines 9 and 30.

Housing wealth. Housing wealth is defined as the value of real estate held by households minus home mortgages. Data are quarterly, end of period, not seasonally adjusted. Series comprise the period 1951:4–2022:4. The source is the Board of Governors of the Federal Reserve System, Financial Accounts, Table B.101, lines 4 and 33.

Labor income. Labor income is defined as the sum of wages and salaries (line 3), personal current transfer receipts (line 16), and employer contributions for employee pension and insurance funds (line 7) minus employee contributions for government social insurance, and taxes. Employee contributions for government social insurance are defined as personal contributions for government social insurance (line 25) minus employer contributions for government social insurance (line 8). Taxes are defined as [(wages and salaries (line 3))/(wages and salaries (line3) + proprietor's income with inventory valuation and capital consumption adjustments (line 9) + rental income of persons with capital consumption adjustment (line 12) + personal dividend income (line 15) + personal interest income (line 14))] $*$ (personal current taxes (line 26)). Data are quarterly, seasonally adjusted at annual rates and measured in billions of dollars. Series comprise the period 1947:1–2022:4. The source is the U.S. Bureau of Economic Analysis, NIPA Table 2.1.

The aggregate wealth, financial wealth, housing wealth and labor income series are all deflated to 2012-dollars using the PCE chain-type price deflator before being transformed into per capita terms and put in the logarithmic form.
In the following, we will describe the population series as well as the price deflators used for that in more detail.

Population. Population is defined as the population series by the BEA, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/B230RC0Q173SBEA, April 5, 2023. Data are quarterly, not seasonally adjusted and comprise the period 1947:1-2022:4.

Price deflator. The chain-type price deflators for personal consumption expenditures, nondurables, services and clothing and footwear are quarterly, seasonally adjusted and in 2012-dollars. The source is the U.S. Bureau of Economic Analysis, NIPA Table 2.3.4, lines 1, 8, 10 and 13.

Inflation rate. The inflation rate is computed from the PCE chain-type price deflator. Series comprises the period 1947:2–2022:4.

Interest rate ("Risk-free" rate). The "risk-free" rate is defined as the 3-month U.S. Treasury bills real interest rate per quarter. Original data are monthly not seasonally adjusted nominal rates in percent per annum and are converted to a quarterly frequency by computing the simple arithmetic average of three consecutive months and applying the discount method. The real interest rates are computed as the difference between nominal interest rates and the inflation rate. The source of the 3-month U.S. Treasury bill secondary market rate is the H.15 publication of the Board of Governors of the Federal Reserve System and it comprises the period 1934:1-2022:4.

Asset returns. Asset returns are computed using the CRSP NYSE/NYSE MKT/NASDAQ/Arca Value-Weighted Market Index. It measures market performance assuming a full reinvestment of all distributions. The original series is daily. Quarterly returns are computed using the data from the last business day of each quarter. The series comprises the period 1926:1-2022:4.

A.3 Forecasting consumption growth

The theory centered around equation (4) also allows for the consumption-wealth ratio to predict future consumption growth rather than asset returns. This translates into the possibility of *cay* to predict Δc_{t+1} according to (7). To illustrate that this is still not the case, even though the predictability of asset returns has also vanished, we display the results from H-period ahead consumption growth forecasts on $\widehat{cay}_t$ for both NDS and PCE as the consumption measure in Table 12.

	Dependent variable: $\Delta c_{t+1} + \cdots + \Delta c_{t+H}$						
	$H = 1$	$H = 2$	$H = 4$	$H = 8$	$H = 12$	$H = 16$	$H = 20$
$\widehat{cay}_t^{\text{NDS}}$	0.006	0.014	0.028	0.042	0.067	0.104	0.123
	(0.322)	(0.296)	(0.220)	(0.192)	(0.264)	(0.385)	(0.402)
	[−0.003]	[−0.002]	[−0.001]	[−0.002]	[−0.001]	[0.000]	[0.000]
$\widehat{cay}_t^{\text{PCE}}$	-0.030	-0.036	-0.069	-0.122	-0.190	-0.269	-0.401
	(-1.010)	(-0.423)	(-0.318)	(-0.341)	(-0.517)	(-0.527)	(-0.789)
	[0.004]	[0.000]	[0.001]	[0.002]	[0.005]	[0.009]	[0.017]
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 12: The table reports coefficient estimates for $\widehat{cay}_t$ and $\widehat{cday}_t$ using both NDS and PCE as the consumption measure from forecasts of H-period ahead consumption growth. The forecasts span different horizons from one quarter up to five years. Newey and West (1987) corrected t-statistics appear in parenthesis and adjusted R^2 statistics in square brackets. The sample period is 1952:1-2019:4.

A.4 Performance fee ϕ

Next to the risk-adjusted abnormal return, θ, we also considered the performance fee, ϕ, when assessing the economic advantage of a *cay*-based trading strategy in comparison to a simple historical average strategy. The level of ϕ over time is displayed in Figure 16.

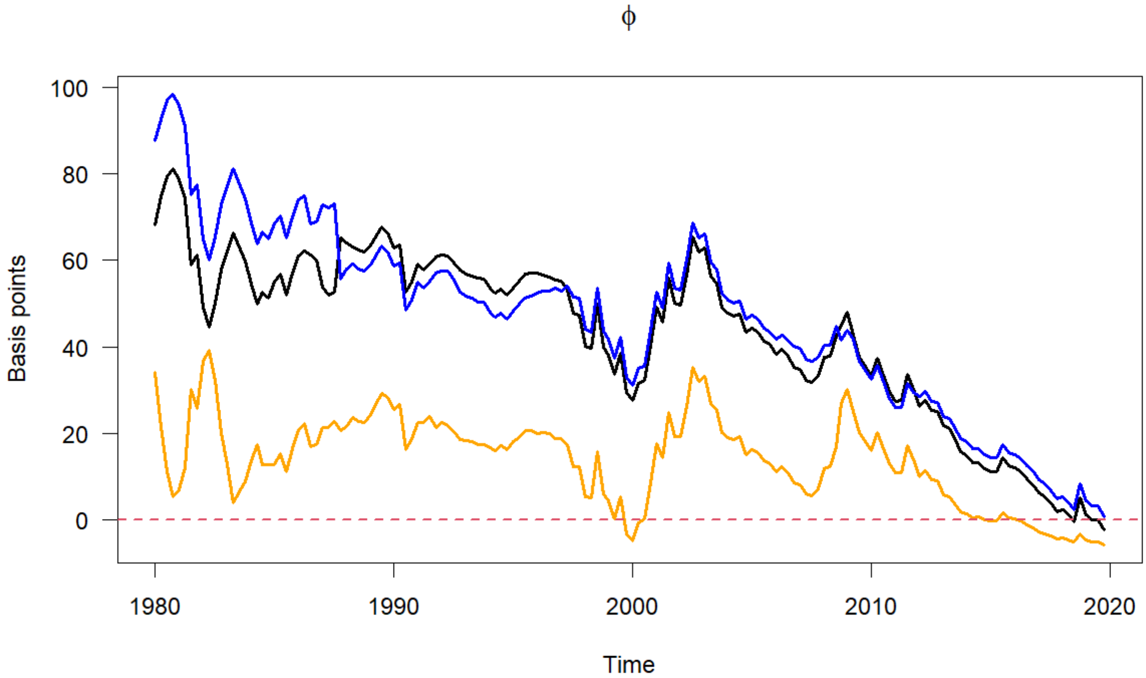


Figure 16: The plot shows the performance fee, ϕ, for the baseline (black), top-10% (blue) and unfiltered (orange) specification. The sample period is 1952:1-2019:4, except for $\widehat{cay}_t^{\text{unfil}}$ it is 1960:3-2019:4.

A.5 Cointegration test results

As mentioned in Section 5, the whole theory of why consumption, aggregate wealth and labor income should be cointegrated, has its weaknesses when being tested empirically. Below we report test results from Engle and Granger (1987) as well as Phillips and Ouliaris (1990) tests with various lag specifications. Note that we already provided the Phillips and Ouliaris (1990) results for one lag in Panel B of Table 5. Furthermore, we also conduct the Johansen (1988, 1991) procedure that not only tests for stationarity in the residuals of a single cointegrating equation as the other two procedures do but also for the actual number of cointegrating relationships. The underlying consumption, aggregate wealth and labor income series are all $I(1)$ and the Johansen (1988, 1991) tests are therefore valid.

Lag = 1	Lag = 2	Lag = 3	Lag = 4	5% CV	10% CV
Panel A: Engle and Granger (1987) cointegration tests					
-1.466	-1.438	-1.175	-1.032	-3.741	-3.452
Panel B: Phillips and Ouliaris (1990) cointegration tests					
-1.443	-1.469	-1.293	-1.244	-3.768	-3.449

Table 13: The table reports the test statistics from Engle and Granger (1987) and Phillips and Ouliaris (1990) tests for cointegration. The sample period is 1952:1-2019:4.

L-Max		Trace		$H_0 = r$
Test Statistic	90% CV	Test Statistic	90% CV	r
Panel A: Lag in VAR Model equal to 1				
12.99	18.89	19.71	27.07	0
6.21	12.30	6.72	13.43	1
0.51	2.71	0.51	2.71	2
Panel B: Lag in VAR Model equal to 2				
12.74	18.89	19.89	27.07	0
6.80	12.30	7.16	13.43	1
0.35	2.71	0.35	2.71	2
Panel C: Lag in VAR Model equal to 3				
13.70	18.89	18.59	27.07	0
4.41	12.30	4.88	13.43	1
0.48	2.71	0.48	2.71	2
Panel D: Lag in VAR Model equal to 4				
20.16	18.89	24.60	27.07	0
4.21	12.30	4.44	13.43	1
0.23	2.71	0.23	2.71	2

Table 14: The table reports Johansen (1988, 1991) test results including one to four lags and assuming a linear trend in the underlying data but not in the cointegrating relationship itself.

A.6 Model uncertainty in cointegration tests

As mentioned in the main body, there are several issues regarding the right model choice for the cointegrating relationship. On the one hand, theory tells us that a specification with a linear trend in the underlying data and only an intercept in the cointegrating relationship is the most appropriate choice. For that reason, Lettau and Ludvigson (2001) assume this specification which we follow suit. On the other hand, Koop et al. (2008) show that several model specification are consistent with the theory, and examine the substantial amount of model uncertainty among them. Therefore, we also report results on model choice in Table 16 below. The five model specifications therein are in line with the codes in Koop et al. (2008) and are defined in Table 15.

Code	intercept	trend	trend in data
$d=1$	yes	yes	quadratic
$d=2$	yes	yes	linear
$d=3$	yes	no	linear
$d=4$	yes	no	none
$d=5$	no	no	none

Table 15: The table reports the assumptions on the cointegrating relationship specification by linking them to the codes used in Koop et al. (2008).

The standard assumption by Lettau and Ludvigson (2001) that we also adopted in this paper is given by $d=3$. In Subsection 5.3 we also discussed the case of $d=2$ in more detail.

Lags	Test	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$
Panel A: Sample period from 1952:1-2019:4						
1	L-Max	0	0	0	1	1
	Trace	0	0	0	1	1
2	L-Max	0	0	0	1	1
	Trace	0	0	0	1	1
3	L-Max	0	0	0	1	1
	Trace	0	0	0	1	1
4	L-Max	0	0	1	1	1
	Trace	0	0	0	1	1
Panel B: Sample period from 1952:1-1998:3						
1	L-Max	0	0	0	1	1
	Trace	0	0	0	1	3
2	L-Max	0	0	0	1	3
	Trace	0	0	0	1	3
3	L-Max	0	0	0	1	2
	Trace	0	0	0	1	2
4	L-Max	0	0	0	1	2
	Trace	0	0	0	1	2

Table 16: The table reports the number of cointegrating relationships using the Johansen (1988, 1991) procedure with 90% critical values. The five model specifications follow the codes in Table 15.

What the results in Table 16 again show us is that we have almost no evidence of a cointegrating relationship under our standard assumptions.[10] This is in line with the results reported in Table 14. Interestingly, this also holds in the subsample until 1998:3. Clear evidence for only one single cointegrating relationship is only given in the case where the underlying data are not assumed to have a trend, i.e., the cases $d = 4$ and $d = 5$. This however clearly goes against the actual observed consumption, asset wealth and labor income series which is why we do not consider these specifications as suited for our model.

[10] Except for the L-Max statistic with four lags there is no evidence of cointegration.