Consumption Disconnect Redux*

Alessandro Melone[†]

This version: June 28, 2022

Abstract

I document that consumption dynamics are characterized by a permanent level warranted by productivity expectations plus temporary deviations around this trend, termed the consumption gap. At horizons from one quarter to five years, I find that the consumption gap forecasts stock returns in- and out-of-sample, largely subsuming other popular predictors. Using the consumption gap to time the market, a mean-variance investor would achieve an annualized certainty equivalent return of 2.39%. In the cross–section of stocks, the consumption gap relates to the time–varying coefficients of a stochastic discount factor linear in the market, suggesting that temporary consumption fluctuations track the price of market risk, consistent with benchmark macro–finance theories.

Keywords: Consumption Levels, Return Predictability, Conditional Models, Time–Varying Market Price of Risk.
JEL codes: E21, E32, G11, G12.

---

*I am grateful to Carlo Favero, Bob Korajczyk, Andrea Tamoni, and Josef Zechner for helpful discussions and invaluable advice. I have benefited from numerous conversations with Ravi Jagannathan and Christian Wagner. Further, I would like to thank Lorenzo Bretscher, Larry Christiano, Max Croce, Daniele d'Arienzo, Ian Dew-Becker, Richard Evans, Jack Favilukis, Adlai Fisher, Lorenzo Garlappi, Stefano Giglio, Roberto Gomez-Cram, Andrei Gonçalves, Niels Gormsen, Ben Hébert, Philipp Illeditsch, Cosmin Ilut, Zhengyang Jiang, Tim Kroencke, Lars Lochstoer, Guido Lorenzoni, Semyon Malamud, Andy Neuhierl, Giorgio Ottonello, Dimitris Papanikolaou, Roberto Pinto, Sébastien Pouget, Otto Randl, Paola Sapienza, Clemens Sialm, Stefan Voigt, Michael Weber, Rüdiger Weber, Zhengyang Xu, Tony Zhang, and many seminar and conference participants for insightful comments. Part of this paper has been written while visiting Northwestern Kellogg; I thankfully acknowledge financial support from the OeAD (Marietta Blau Grant) during this time. All errors, of course, are my own.

[†]Fisher College of Business, The Ohio State University, 2100 Neil Avenue, Columbus, OH 43210. E-mail: alessandro.melone@wu.ac.at. Personal website: Homepage.

1 Introduction

A sound link between the time–varying market price of risk and fluctuations in consumption is a central prediction of several macro–finance theories, including Campbell and Cochrane (1999) and Chan and Kogan (2002). Yet, empirically, it has been challenging to validate such a prediction. Indeed, the puzzling empirical disconnect between time–series movements in stock returns and macroeconomic fundamentals is still a matter of debate (see, e.g., Lettau and Ludvigson, 2010).[1]

One approach to studying the relationship between consumption and return predictability is considered in Lettau and Ludvigson (2001a), who document that a proxy for the log consumption–wealth ratio—denoted *cay*—predicts stock returns. However, *cay* is a function of aggregate wealth, hence containing the level of stock prices. This fact suggests that the forecasting power documented can be generated by a mean–reversion of prices, as in the case of the dividend–price ratio. Furthermore, *cay* appears to capture temporary movements in wealth rather than consumption fluctuations (see, e.g., Brennan and Xia, 2005), thus leaving open the question of why return predictability cannot be linked to the real economy.

This paper proposes a novel approach to investigating the link between expected stock returns and time-series movements in consumption. I start from a key result in real business cycle models that permanent technological innovations lead to cointegrated consumption and output levels (e.g., King, Plosser, Stock, and Watson, 1991), consistent with a version of the life cycle–permanent income model in which individuals set consumption equal to their long–run productivity expectations (e.g., Blanchard, L'Huillier, and Lorenzoni, 2013).

[1] My work is also inspired by Fama (1991): "[. . .] we can hope for a coherent story that (1) relates the cross–section properties of expected returns to the variation of expected returns through time, and (2) relates the behavior of expected returns to the real economy [. . .]. We can, however, hope to know more about the links between expected returns and the macro–variables." This paper aims at establishing one of these links.

The presence of a stochastic trend in consumption *levels* has implications for testing the linkages between stock returns and consumption *growth* rates. This empirical analysis has traditionally been carried out by regressing stock returns on (current or past) consumption growth, i.e., by studying the first–difference of log consumption expenditures. However, when a macroeconomic time-series features both permanent (non–stationary) and transitory (stationary) components, simply first–differencing the series confounds the contribution of each component (see, e.g., Stock and Watson, 1988b). To address this issue, I provide a new permanent–transitory decomposition of consumption levels and study how different components relate to stock returns.

I document that consumption levels are characterized by the stochastic trend in real GDP in the long–run, but temporarily deviate from this permanent level warranted by productivity expectations. I denote the difference between consumption expenditures and the stochastic trend in real GDP as the consumption gap.[2] These transitory consumption deviations are the only component of aggregate consumption expenditures that relate to stock returns.

Specifically, I find that the consumption gap is a strong predictor of stock returns. Using US quarterly stock market data in the period 1967–2020, I find that for a drop in current consumption below the trend in real GDP of 1%, the expected market excess return over the next quarter (next year) increases by about 0.9% (3.6%). The R^2 from these predictive regressions is about 4% at a quarterly horizon, and it raises to 12% at one–year horizon. Stock market predictability holds out-of-sample, for different aggregation horizons, using alternative measures for the permanent component, and on international data. Also, a reduced–form VAR exercise confirms that the consumption gap has high marginal predictive power for market returns.

[2] I use the CBO potential output as a benchmark measure for the stochastic trend in real GDP.

What is the economic intuition of these findings? Situations in which consumption expenditures fall below their permanent component warranted by productivity, and the consumption gap is negative, correspond to bad times—e.g., the aftermath of the Great Financial Crisis and the COVID-19 pandemic—in which the marginal utility from current consumption is high and individuals ask for a larger compensation for bearing risks, thus leading to a higher market risk premium. The converse is true in good times, when consumption rises above the stochastic trend in real GDP, the consumption gap is positive, marginal utility is low, and individuals demand lower compensation for market risk. Therefore, the consumption gap should be negatively associated with future expected stock returns, consistent with what I find.

As macroeconomic data typically feature measurement errors and are revised, and the stochastic trend in real GDP is a latent process that needs to be estimated, a natural question is whether using only information available in real–time would affect the forecasting performance of the consumption gap. To address this concern, I replicate both in-sample and out-of-sample analyses using only data and parameter estimates that were known when the forecast was made. Using strictly information available in real–time, I find uniform evidence that the consumption gap predicts stock returns.

The informative content of the consumption gap to predict excess market returns broadly subsumes the forecasting power of other predictors. Specifically, I run a horse race of the consumption gap against other well–known return predictors, including alternative consumption-based predictors (e.g., *cay*, Lettau and Ludvigson, 2001a; cyclical consumption, Atanasov, Møller, and Priestley, 2020), and I find that the consumption gap captures most of the variation in expected excess market returns. In addition, the consumption gap predicts other characteristics-based systematic financial factors, including value (Fama and French, 1993) and the q-factors (Hou, Xue, and Zhang, 2015), suggesting that it captures a systematic asset pricing component.

This time series predictability is both statistically and economically significant. A mean-variance investor who allocates her wealth between equity and risk-free Treasury bills each quarter would be willing to pay an annualized fee of 2.39% to a portfolio manager who uses the consumption gap to predict aggregate market returns rather than the historical mean. This market timing strategy generates a significant positive alpha with respect to both unconditional and alternative conditional models. Having at disposal a parsimonious predictive model with a single predictor is particularly valuable in the age of big data in which the curse of dimensionality is a severe concern (e.g., Martin and Nagel, 2021).

Why does the consumption gap predict returns? Particularly, is predictability due to mispricing or time–varying market risk premium? My analysis supports the latter view for two reasons. First, the consumption gap does not contain the level of stock prices, thus removing any suspicion that returns are predictable due to a mechanical mean-reversion of prices. Second, the cross-sectional analysis shows that the consumption gap tracks the market price of risk over time.

Specifically, I estimate a conditional version of the Capital Asset Pricing Model (CAPM) in which the stochastic discount factor (SDF) is linear in the market portfolio and its coefficients are time-varying. The coefficients are a linear function of the consumption gap. The intuition for this specification is that the price of market risk is higher in bad times, when the consumption gap is negative and marginal utility is high, than it is in good times, when the consumption gap is positive and marginal utility is low. Conditioning the SDF coefficients using the information contained in the consumption gap leads to an unconditional two-factor model, in which the two factors are the excess market return and the excess market return scaled by the consumption gap. The market portfolio is the only fundamental factor. The scaled excess market can be interpreted as a managed portfolio which invests more on the market when the consumption gap is high, and vice versa. Since the consumption gap negatively predicts the stock market, the managed portfolio represents a hedge against time-varying equity premia as it performs well when the stock market is low and marginal utility is high. Thus, the risk premium associated with this managed portfolio should be negative. The cross-sectional analysis confirms this prediction.

Furthermore, the conditional version of the CAPM in which I use the consumption gap as a conditioning variable shows a significant ability at explaining cross-sectional variation in betas across different test assets when compared with other unconditional and conditional single-factor models. For the 25 Portfolios sorted on Size and Book-to-Market, using the consumption gap as a conditioning variable leads to an improvement in the cross-sectional R^2 of about 14% with respect to the unconditional CAPM and of about 8% with respect to the conditional CAPM of Lettau and Ludvigson (2001b). I find similar improvements for many test assets, including the large cross-section of test assets used, e.g., in Kelly, Kozak, and Giglio (2020) and Kozak, Nagel, and Santosh (2020).

Finally, I use the SDF estimates to construct an empirical SDF and study its dynamics. This estimated SDF fluctuates substantially over time, with an average annualized conditional variance of 0.9, a value comparable to theoretical counterparts used, e.g., in long-run risk or habit models. Moreover, I find that the SDF variance is descriptive about the state of the economy. Indeed, the variance of the SDF correlates with measures of economic activity, and large spikes in the SDF coincide with periods of bad economic conditions and financial distress—the 1970s energy crisis, the early 1980s recession, the Dot-com bubble, the

Great Financial Crisis 2007–2009, and the COVID-19 pandemic. Overall, these results lend support to macro-finance theories aimed at understanding cyclical variation in the price of market risk (e.g., Constantinides and Duffie, 1996; Campbell and Cochrane, 1999; Chan and Kogan, 2002; Bansal and Yaron, 2004; Barro, 2009; Gârleanu and Panageas, 2015; Barberis, Greenwood, Jin, and Shleifer, 2015).

Related Literature. This paper contributes to the empirical literature on conditional pricing models.[3] Especially, my analysis relates to work by Lettau and Ludvigson (2001a; 2001b). Lettau and Ludvigson (2001b), building on Lettau and Ludvigson (2001a), study several conditional versions of the CAPM and the consumption CAPM (CCAPM) using *cay*—a proxy for the consumption-to-wealth ratio—as a conditioning variable. I investigate the relationship between consumption and *permanent* productivity rather than *current* wealth. Indeed, using current wealth to detrend consumption partially cancels temporary fluctuations in consumption away when consumption and wealth are exposed to same aggregate shocks. I show that these temporary consumption fluctuations track the price of market risk over the business cycle.

This paper also relates to the empirical macro-finance literature that investigates the relationship between consumption and asset prices. Several relevant papers in this literature address the empirical failures of consumption-based asset pricing models by proposing alternative consumption *growth risk* measures and testing their cross-sectional pricing ability (e.g., Parker and Julliard, 2005; Jagannathan and Wang, 2007; Savov, 2011; Kroencke, 2017; Bandi and Tamoni, 2020; Liu and Matthies, 2021; Bryzgalova and Julliard, 2021). My ap-

[3]My paper draws important economic insights from Gibbons and Ferson (1985); Hansen and Richard (1987); Bollerslev, Engle, and Wooldridge (1988); Ferson and Schadt (1996); Ferson and Harvey (1999); Jagannathan and Wang (1996); Duffee (2005); Santos and Veronesi (2006); Bali (2008); Whelan (2008); Nagel and Singleton (2011); Roussanov (2014); Møller and Rangvid (2015); Hasler and Martineau (2020); Barroso, Boons, and Karehnke (2021); Gormsen and Jensen (2021).

proach is complementary to this literature. Indeed, I find that studying consumption *levels* uncovers a novel consumption-based return predictor.

2 Consumption Expenditures, Productivity Expectations, and the Consumption Gap

In this Section, I show how to construct the consumption gap and study its properties. Then, I discuss the potential origins of temporary fluctuations in aggregate consumption and their interpretation.

2.1 The Consumption Gap

Individuals form estimates of their ability to produce in the long run, and, according to this estimate, they decide how much to consume today. This estimate can be permanent income (Friedman, 1957), wealth (Modigliani, 1971), or long-run productivity (e.g., Blanchard, L'Huillier, and Lorenzoni, 2013).

Following Blanchard, L'Huillier, and Lorenzoni (2013), I assume that households set consumption equal to their long-run productivity expectations:

$$c_t = \lim_{j\to\infty} \mathrm{E}_t\left[y_{t+j} - j\mu\right] \tag{1}$$

where c_t is log aggregate consumption, y_t is log real gross domestic product (GDP), and $\mu \equiv \mathrm{E}[\Delta y_t]$. Notice that $\lim_{j\to\infty} \mathrm{E}_t\left[y_{t+j} - j\mu\right] = y_t + \sum_{j=1}^{\infty} \mathrm{E}_t\left[\Delta y_{t+j} - \mu\right]$: long-run productivity expectations are current real GDP plus all expected future real GDP growth.

Equivalently, long-run productivity expectations are the stochastic trend in real GDP.[4] I propose to employ the US Congressional Budget Office (CBO) potential output as a benchmark measure of long-term productivity expectations.[5] However, I show that both in-sample and out-of-sample results hold using alternative measures of the permanent component of consumption (see Section 3.3). The fact that predictability results are robust to alternative methods and variables used to estimate the permanent consumption component suggests that predictability stems from the informative content in temporary consumption expenditures fluctuations.

In the main analysis, I use the CBO potential output rather than alternative measures for the trend in real GDP for several reasons. First, it is based on the standard framework of the Solow growth model allowing for a clear accounting for the different sources of growth—labor, capital accumulation, and total factor productivity. Second, using as an alternative to compute the stochastic trend in real GDP a filtering method like the Hodrick-Prescott (HP) filter can introduce spurious dynamic relations that have no basis in the underlying data-generating process and generates inconsistent filtered values across the estimation sample (e.g., Hamilton, 2018). Then, many filtering methods do not benchmark their trends to any external measure of capacity. Therefore, these estimates cannot be interpreted as the level of output consistent with stable inflation. Finally, it is ready-to-use, extensively used in

[4]This is equivalent to assuming that consumption and real GDP share the same stochastic trend. The idea that output and consumption share the same stochastic trend, i.e., potential consumption equals potential output, is common in the real business cycle literature. Recently, this assumption has been used also in the consumption-based macro-finance model of Campbell, Pflueger, and Viceira (2020). In Appendix B, I report results for this cointegrating relationship.

[5]Potential output is the real GDP an economy would produce with a high rate of use of its capital and labor resources; it is a measure of the level of real GDP consistent with stable inflation. The US Congressional Budget Office (CBO) provides estimates of potential output together with its projections up to ten-years ahead. Data are available at https://fred.stlouisfed.org/series/GDPPOT; I report more details on its construction in Appendix A.1.

academic work and by policy makers, and freely available on FRED. One possible concern about CBO potential output is that it is subject to measurement errors and revisions that can potentially affect results in empirical investigations. I address this issue using also unrevised real-time data and find uniform evidence in favor of return predictability.

I formally test whether log potential GDP describes log aggregate consumption. To this end, I consider the following regression:

$$c_t = \alpha + \beta ypot_t + v_t \ , \tag{2}$$

where c_t and $ypot_t$ are respectively quarterly real per-capita log consumption NIPA expenditures (non-durable and services) and real per-capita log potential output. If equation (2) delivers stationary residuals, then there exists a long-run relation between consumption levels and potential output such that the two series cannot drift apart for an indefinite time. In this case, consumption expenditures and its permanent level warranted by productivity expectations have a special relationship: They are cointegrated.

I estimate equation (2) over the sample period from 1967Q1 to 2020Q4. To do so, I use the Dynamic OLS (DOLS) approach proposed by Stock and Watson (1993). The DOLS technique delivers optimal estimates of the parameters in a cointegrating regression by adding n leads and lags to the regression to eliminate asymptotically any bias due to serial correlation or endogeneity.[6] Thus, I estimate the following regression by OLS:

$$c_t = \alpha + \beta ypot_t + \sum_{j=-n}^{n} b_j \Delta ypot_{t-j} + v_t \ ,$$

where Δ denotes first difference. Following Lettau and Ludvigson (2001a), I set $n = 8$ and exclude the deterministic trend. I obtain the following estimates

$$c_t = -\underset{(-0.749,\ 0.099)}{0.325} + \underset{(0.941,\ 1.024)}{0.982} \cdot ypot_t \ ,$$

where in parenthesis I report the 95% confidence intervals.

A few comments are in order. First, as predicted by equation (1), I find that log consumption and log potential output have a cointegrating vector $(1, -1)$. Second, the null of non-stationarity for the residuals is uniformly rejected: the Phillips-Perron (PP) Unit Root Test delivers a p-value lower than 0.01 (PP test statistic $= -39.77$), and the Engle-Granger (EG) Cointegration Test delivers a t-stat of -4.164, which is far below the robust t-stat proposed by MacKinnon (2010) for testing cointegrating residuals. Also, the null of no-cointegration in the Phillips-Ouliaris (PO) Cointegration Test is rejected with a p-value lower than 0.01 (PO test statistics $= -40.68$).

Given this empirical evidence, I define the consumption gap as:

$$cgap_t = c_t - ypot_t \ . \tag{3}$$

[6] It is important to estimate cointegrating regression parameters via DOLS, because DOLS estimates are "superconsistent", namely the estimates converge to the true parameter values at a rate proportional to the sample size T rather than proportional to $\sqrt{T}$. This fact guarantees that estimates of α and β will be consistent despite the fact that the errors v_t will typically be correlated with $ypot_t$.

As recommended by Ferson, Sarkissian, and Simin (2003), I demean the scaling variable $cgap_t$ in all the empirical investigations of this paper. Fitting an AR(1) for the consumption gap delivers a first-order autocorrelation parameter of 0.904, which implies an half-life of almost two years ($\log(0.5)/\log(0.904) = 6.87$ quarters). The consumption gap correlates with the business cycle: it reaches its highest values around the onset of recessions, drops throughout economic downturns, and rises after recessions, i.e., the *cgap* is procyclical. For instance, the correlation of *cgap* with industrial production annual growth and the output gap is, respectively, 46% and 70%.

The fact that consumption spending levels temporarily deviate from their permanent component, i.e., that the consumption gap is not always zero, naturally leads to a new permanent–transitory decomposition of the levels of consumption. Such a decomposition allows to dig deeper into the question of which component of consumption relates to stock returns. Specifically, in Appendix C.2, I show that assuming that the consumption gap follows a stationary autoregressive process of order 1, log consumption growth Δc_t can be written as:

$$\Delta c_t = \mu + (\rho - 1)cgap_{t-1} + \eta_t + \epsilon_t \ , \tag{4}$$

where η_t represents the shock to the permanent component $\lim_{j\to\infty} \mathrm{E}_t\left[y_{t+j} - j\mu\right]$ in equation (1) and ϵ_t is the shock to the consumption gap.[7] Since $\mid \rho \mid< 1$, $(\rho - 1) < 0$, thus positive consumption gap relates to low future consumption growth.

[7] Pohl, Schmedders, and Wilms (2016) introduce a similar consumption growth decomposition to study a theoretical asset pricing model with non-permanent shocks to consumption. Chernov, Lochstoer, and Song (2021) use a similar specification but with heteroskedastic errors to explain several puzzles about the stock-bond interaction. Also, notice that specification (4) features a time-varying mean in consumption growth, like in long-run risk models (Bansal and Yaron, 2004). However, the consumption gap is significantly less persistent than the predictable consumption component discussed in this literature. For example, Bansal and Yaron (2004) calibrate the AR(1) parameter $\rho = 0.979$, which corresponds to an half-life of more than 8 years. Recent work that study non-i.i.d consumption include Ortu, Tamoni, and Tebaldi (2013), Bidder and Dew-Becker (2016), and Schorfheide, Song, and Yaron (2018).

Equation (4) shows that when consumption levels contain both a permanent and a temporary component (see (C.1)), consumption growth features three elements: news to the permanent and to the temporary component plus the lagged temporary component itself. Thus, consumption growth would be a convolution of this three different elements, and studying the series alone does not allow to distinguish among the different components.[8]

2.2 Origins and Interpretation of Aggregate Consumption Fluctuations

According to the life cycle-permanent income hypothesis (PIH), individuals set consumption at any point in time according to their long-run productivity estimate. Instead, I document that current aggregate consumption temporarily deviates from the trend in real GDP. These deviations are stationary, but persistent and predictable. These findings suggest the following general representation:

$$\begin{aligned} c_t &= \mathrm{E}_t^I \left[y_{t+\infty} \right] \\ &= \mathrm{E}_t \left[y_{t+\infty} \right] + cgap_t \\ \mathrm{E}_t \left[y_{t+\infty} \right] &= \mu + \mathrm{E}_{t-1} \left[y_{t+\infty} \right] + \eta_t \\ cgap_t &= \rho \cdot cgap_{t-1} + \epsilon_t \end{aligned} \tag{5}$$

where y_t is log real GDP, $\mathrm{E}_t \left[y_{t+\infty} \right] \equiv \lim_{j \to \infty} \mathrm{E}_t \left[y_{t+j} - j\mu \right]$, $\rho \in [0, 1]$, and ϵ_t is i.i.d. with zero mean and variance σ_ϵ^2. Aggregate consumption expenditures c_t equals aggregate indi-

[8] An alternative approach is to use some filtering procedure for consumption growth to extract different components; see, e.g., Ortu, Tamoni, and Tebaldi (2013); Schorfheide, Song, and Yaron (2018); Bandi, Perron, Tamoni, and Tebaldi (2019).

viduals' productivity expectations $\mathrm{E}_t^I\left[y_{t+\infty}\right]$. Unlike the standard PIH, these productivity expectations contain both a permanent component, $\mathrm{E}_t\left[y_{t+\infty}\right]$—as predicted by the rational expectations framework—, and a novel temporary component, $cgap_t$.

Which theories can generate such temporary deviations of consumption from its level warranted by productivity? One possibility is that fundamentals are observed with noise (e.g., Lorenzoni, 2009; Blanchard, L'Huillier, and Lorenzoni, 2013; LHuillier and Yoo, 2017). Consumers and firms receive information about the future, which can be news or noise. Ex-post, if the information turns out to be news, the economy would adjust to a new level of activity. If the information turns out to be just noise, the economy would return to its initial state. Thus, this signal extraction problem with learning generates temporary aggregate consumption fluctuations.

Also, several forms of departure from the rational expectations paradigm can lead to temporary consumption fluctuations; recent work include Fuster, Hebert, and Laibson (2012); Bidder and Dew-Becker (2016); Mian, Sufi, and Verner (2017); Bouchaud, Krueger, Landier, and Thesmar (2019); Ilut and Valchev (2020); Angeletos and Lian (2021); Bianchi, Ilut, and Saijo (2021); Chodorow-Reich, Guren, and McQuade (2021). In my sample, I find that the correlation between the shock to the consumption gap ϵ_t and the shock to fundamentals η_t is positive. This fact suggests that good (bad) news about fundamentals are associated with more optimistic (pessimistic) long-term productivity forecasts and higher (lower) consumption spending, consistent with a view of individuals having diagnostic expectations (see Bordalo, Gennaioli, and Shleifer, 2018; Bordalo, Gennaioli, Porta, and Shleifer, 2019).[9]

How to interpret the documented consumption fluctuations? Specification (5) suggests

[9] Several other situations can lead to consumption temporarily deviating from its long-term path, including binding liquidity constraints, hand-to-mouth consumption, and credit supply expansion. Identifying the exact source of consumption fluctuations is beyond the scope of this paper.

that the consumption gap represents deviations of aggregate households' consumption spending from its permanent path. I find that the consumption gap exhibits positive and significant correlation with measures of consumer sentiment. Indeed, periods in which individuals on average over(under)–estimate their productivity expectation path and decide to consume more (less) coincide with periods of high (low) sentiment. Figure 1 overlays the consumption gap (blue line) with two survey measures of consumer sentiment: the University of Michigan Consumer Sentiment Index and the Conference Board Consumer Confidence Index.[10] The three series show an equivalent pattern: They fall during economic contractions and they rise at the onset of economic expansions, reaching their peaks just before recessions.

Regardless of their origins or interpretation, aggregate fluctuations in consumer spending appear to generate business cycles in the data (e.g., Lorenzoni, 2009; Ilut and Schneider, 2014; Mian and Sufi, 2014; Angeletos and Lian, 2021). From an asset pricing perspective, business cycle fluctuations can contain systematic economic information relevant for long-term investors (Cochrane, 2005). The following Section investigates how the consumption gap relates to the time series and the cross-section of returns.

2.2.1 Permanent vs Current Levels to Identify Consumption Fluctuations?

In their seminal paper, Lettau and Ludvigson (2001a) introduce a novel market return predictor *cay*, computed as the cointegrated residual of a regression of consumption c on asset holdings a and labor income y. The predictor *cay* is a proxy for the consumption-to-wealth ratio and follows from the log-linearization of the intertemporal budget constraint. To see

[10] Consumer sentiment data can be found at `http://www.sca.isr.umich.edu/` and `https://conference-board.org/data/consumerconfidence.cfm`.

how the predictor *cgap* is different from *cay*, consider the following specification:

$$c_t = \mathrm{E}_t \left[y_{t+\infty} \right] + cgap_t \tag{6}$$
$$w_t = \mathrm{E}_t \left[w_{t+\infty} \right] + x_t^w$$

where w_t is log aggregate wealth computed from asset holdings and labor income as in Lettau and Ludvigson (2001a), $\mathrm{E}_t \left[w_{t+\infty} \right] \equiv \lim_{j\to\infty} \mathrm{E}_t \left[w_{t+j} - j\mu_w \right]$ is the permanent component in wealth, which follows an I(1) process, and x_t^w is the temporary, stationary component of wealth.[11]

Using w_t instead of its permanent level to compute consumption fluctuations smooths away part of the fluctuations in consumption. Indeed, consumption and wealth are largely exposed to the same permanent shocks: the permanent components of the two series have a correlation of 99%. Hence, $c_t - w_t \approx cgap_t - x_t^w$, which means that the consumption-to-wealth ratio as proxied by *cay* does not contain the same information contained in the *cgap*.[12] Specifically, part of the consumption fluctuations captured by the consumption gap are eliminated by temporary fluctuations in wealth.

Finally, notice that *cay* should predict returns with a positive sign, whereas *cgap* with a negative sign. The classical explanation for the positive relation between *cay* and future returns, motivated by the intertemporal budget constraint, is that rational, forward–looking individuals will consume little today relative to their wealth if they expect future returns to be low. I.e., when *cay* is low, wealth is temporarily high, and future expected returns are low.[13] Differently, when *cgap* is low, consumption is temporarily low, and future expected

[11] Using the estimates and the data provided on Martin Lettau's website, I compute wealth as $w_t = -0.441 + 0.218 \cdot a_t + 0.801 \cdot y_t$.

[12] I use the linear projection method of Hamilton (2018) to calculate the permanent wealth component.

[13] Brennan and Xia (2005) find that replacing consumption with calendar time in the Lettau and Lud-

returns are high.

I document that the consumption gap subsumes the predictive power of other macro-based predictors, including *cay*. This finding suggests that using the permanent level in productivity to extract temporary fluctuations in consumption delivers a variable which is more informative for conditional asset pricing than a variable constructed using current levels.

3 Empirical Investigation

For the main empirical analysis, I focus on U.S. data–NYSE, AMEX, and Nasdaq stocks from the Center for Research in Security Prices (CRSP) and Compustat data required for sorting – for the sample 1967–2020. I use mostly quarterly observations, but I also provide results for annual frequencies. More information on additional data used in the empirical analysis can be found in Appendix A.

3.1 Time Series Analysis

3.1.1 Consumption Growth, its Components, and the Stock Market

I start by investigating the relationship between aggregate stock market returns and different components of consumption. The goal of this exercise is to understand which component(-s) of consumption relates to the time series of returns. To this end, I use the following

vigson (2001a)'s cointegrating regression leads to a predictor—*tay*—with similar predictive ability for the stock market when compared to *cay*. This result suggests that it is the wealth component that is driving predictability rather than temporary consumption fluctuations.

consumption growth decomposition:[14]

$$\Delta c_t = \mu + (\rho - 1)cgap_{t-1} + \eta_t + \epsilon_t \ ,$$

in which η_t and ϵ_t are, respectively, the shock to the stochastic trend in real GDP and the shock to the $cgap$, and $cgap$ is the consumption gap computed as in equation (3).

Table 1 reports the results from regressing log excess market returns on log consumption growth or its components. Column (1) shows that, when using aggregate log consumption growth, the coefficient is not significantly different from zero and the proportion of market returns variation explained by consumption growth is zero. This result confirms the well-known missing empirical link between the stock market and consumption discussed in the literature, i.e., the "consumption disconnect" puzzle.

However, when I substitute aggregate consumption growth for its components, I find that the consumption gap is significantly related the the stock market. Specifically, Column (5) in Table 1 shows that the $cgap$ is the only component of consumption that relates to aggregate stock market returns. Indeed, neither the permanent shock η_t nor the temporary shock ϵ_t significantly relate to aggregate market returns.[15]

Notice that the consumption gap is inversely related to future stock returns. In my sample, a drop in current consumption below its trend of 1% implies an increase in expected market excess return over the next quarter of about 0.9%. The intuition is that a drop in consumption below its trend coincides to periods in which marginal utility is high and investors ask for a larger compensation for bearing market risk, thus risk premia are higher.

[14]This decomposition is derived in Appendix C.2.

[15]Appendix Table D.3 shows a similar exercise for the short-term rate. In this case, I find that the consumption gap and the permanent shock relate to the short-term rate.

Appendix Table D.2 shows similar results in the case of annual returns.

3.1.2 The Consumption Gap and Stock Returns Predictability

I investigate the ability of the consumption gap to predict returns by running the predictive regression:

$$r_{m,t+i} = \gamma_0 + \gamma_1 cgap_t + \epsilon_{t+i} \ , \tag{7}$$

where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded between t and $t+i$ and $cgap$ is the consumption gap.

Table 2 reports results for the predictive regression (7) for different i.[16] Panel A shows the in-sample results. I find that the consumption gap significantly forecasts the aggregate stock market, with an R^2 ranging from 3.9% for quarterly returns up to 12.1% for annual returns. The last row in Table 2 shows that return predictability implied by the consumption gap is economically large: The regression implies that expected returns on the market vary by more than their unconditional level. I confirm the prediction $\gamma_1 < 0$ for every aggregation horizon.

Figure 2 shows the in-sample R^2 for the predictable regression (7) for $i = \{1, 2, \ldots, 40\}$. The R^2 from this predictive regression displays the hump-shaped pattern also documented in Bandi, Perron, Tamoni, and Tebaldi (2019). Predictability is increasing as a function of the aggregation horizon, with the R^2 reaching its maximum and stabilizing after 4 years, and then declining after the 8 years.

Then, I test the ability of the consumption gap to predict the equity risk premium out-of-

[16] For $i > 1$, to account for potential inference issues related to overlapping observations, I follow Ang and Bekaert (2007) and rely on conservative standard errors from reverse regressions as proposed by Hodrick (1992).

sample (OOS). The out-of-sample exercise is particularly important because of two reasons. First, the consumption gap is a stationary but persistent variable which can generate spurious results. Second, if the consumption gap would not perform well OOS, it would be of little use to long-term investors who want to time the market.

I test for out-of-sample predictability using the metric introduced by Campbell and Thompson (2008). Specifically, I compute the R^2_{OOS}

$$R^2_{OOS} = 1 - \frac{\sum_{t=1}^{T} \left(r_{m,t+i} - \widehat{r}_{m,t+i}\right)^2}{\sum_{t=1}^{T} \left(r_{m,t+i} - \bar{r}_{m,t+i}\right)^2}$$

where $\widehat{r}_{m,t+i}$ is the fitted value from my predictive regression estimated through period $t-1$ and $\bar{r}_{m,t+i}$ is the historical average return estimated thorough period $t-1$. If the R^2_{OOS} is positive, then the predictive regression has lower average mean squared prediction error than the historical average return.

Panel B in Table 2 shows the out-of-sample results. I find that the consumption gap significantly predicts expected excess market returns for different aggregation horizons ranging from one quarter to ten years. This result is not associated to any particular sample, as the R^2_{OOS} is positive and significant regardless the OOS period starts in 1990, 2000, or 2010. This finding addresses the concern of Welch and Goyal (2008), who find that most of the traditional equity premium predictors display a negative R^2_{OOS}.

Furthermore, as consumption and potential output data for quarter t are available in the middle of quarter $t+1$, I replicate the results documented in Table 2 using the second lag of *cgap* when predicting stock returns. That is, stock return forecasts at time $t+i$ are based on the consumption gap in time $t-1$. Appendix Table D.1 reports results for these predictive regressions. Both the in-sample and out-of-sample analyses deliver identical conclusions when compared with results reported in Table 2, suggesting that the lag in releasing the macroeconomic data is not driving the documented predictability.

Vintage Data. As consumption data feature measurement errors and are revised and potential output is a latent process that needs to be estimated, a natural question is whether using only data available in real-time for these series would affect the forecasting performance of the consumption gap. To address this issue, I use vintage data for aggregate consumption expenditures on non-durables and services and potential output from the real-time economic data archive at the St. Louis Fed's Economic Research Division.[17]

Vintage data for consumption are available from 1959Q1, while data for potential output are available starting from 1997Q1. I start by computing the consumption gap as in equation (3) using only vintage data in the sample period 1997Q1 to 2020Q4. Then, I run both the in-sample and out-of-sample analysis using the "vintage" version of the consumption gap $cgap^{vintage}$. Also, to ensure that results are immune to look-ahead bias, I use the second lag of the vintage consumption gap when forecasting returns.

Table 3 shows OLS estimates and the R^2 from the predictive regressions. Panel A reports the in-sample results, while Panel B reports results for the out-of-sample forecasting regressions. A few comments are in order. First, the coefficients on the consumption gap are negative for different aggregation horizons, confirming the economic mechanism behind the predictability generated by the consumption gap. Furthermore, all $t-$stats are above 2.7, showing that also using only vintage data the predictability documented using revised data

[17]Data are available at https://alfred.stlouisfed.org.

holds. Second, the in-sample predictive R^2 ranges from 6% at quarterly frequency to almost 20% at annual frequency. These values are higher than the R^2 obtained using revised data both over the entire sample and over the same sample, consistent with Borup and Schütte (2021) who find that unrevised macroeconomic data are more informative for asset pricing than final releases. Finally, Panel B shows the out-of-sample results and documents that the forecasts of market returns based on the consumption gap are more accurate than forecasts based on the historical mean. This result holds for different aggregation horizons and two different samples.

Overall, these findings suggest that using the consumption gap to predict the market is particularly valuable when using only data known to investors at the time the forecasts are made.

Horse Race Against Other Market Predictors. Is the informative predictability content of the consumption gap subsumed by other market return predictors? Table 4 addresses this question. Panel A reports an horse race of *cgap* against four alternative single return-predictors: the (repurchased-adjusted) dividend-price ratio, long-term Treasury bond returns, CPI inflation, and the investment-to-capital ratio (Cochrane, 1991). These four predictors are the only predictors out of the fifteen variables originally considered by Welch and Goyal (2008) plus the price-to-output ratio (Rangvid, 2006) and the repurchase-adjusted dividend-price ratio constructed in Nagel and Xu (2021) that significantly predict the stock market in my sample.[18]

Columns (1) to (4) display the in-sample predictability for the four regressors individually. Market returns are predictable, with an R^2 ranging from 1.4% to 2.6% for quarterly data.

[18]I am very grateful to Zhengyang Xu for sharing his data.

In Column (5) to (8), I include the consumption gap as an additional predictor. The *cgap* is always statistically significant at 5% level of confidence. Also, controlling for the consumption gap, the evidence of predictability for the other predictors is weak.

Table 4 Panel B reports results for several well-known consumption-based predictors: *cay* (Lettau and Ludvigson, 2001a), surplus-to-consumption ratio (s) calculated as in Wachter (2006), long-run consumption (x) calculated as in Bansal, Kiku, and Yaron (2010), and cyclical consumption (cc) calculated as in Atanasov, Møller, and Priestley (2020). For the individual variables, predictability is significant only for the surplus-to-consumption ratio, with an R^2 of 2.2%, and for cyclical consumption, with an R^2 of 2.4%. In Column (5) to (8), I include *cgap*. Controlling for the consumption gap, there is not evidence of predictability for the other consumption-based predictors. Appendix Table F.1 reports results for the same analysis in case of annual returns and shows equivalent conclusions.

Furthermore, Appendix Table F.3 investigates the predictability of the consumer sentiment measures displayed in Figure 1; in Appendix Table F.4, I use as controls for the predictive regressions the first five principal components (PCs) extracted from a large cross-section of macroeconomic variables derived in Ludvigson and Ng (2009). The consumption gap appears as the most relevant single predictor for market returns.

Overall, these results show that the informative content of the consumption gap broadly subsumes the predictability of other macro-based single-return predictors.

Output Gap and Consumption Gap. Cooper and Priestley (2009) find that the output gap—a key production-based macro variable—predicts returns. The consumption gap is one of the components of the output gap. To see this, consider the national income identity $Y = C + I + G + NX$, where Y is total GDP, C is total consumption, I is gross private domestic investment, G is government expenditures, and NX is net exports of goods and services (all in real terms). Assuming a closed economy, ignoring expenditures for durable goods, and dividing both sides of the national income identity by potential output $YPOT$ leads to $ygap \approx cgap + igap + ggap$, where $ygap = \log(Y/YPOT)$ is the output gap, $igap = \log(I/YPOT)$, and $ggap = \log(G/YPOT)$.

Appendix Table F.2 reports results for the market predictive regression in which I consider the various components of the national income identity as predictors. This Table shows that the consumption gap is the only component of the output gap that predicts stock returns.

Predicting Factors Beyond the Market. I ask whether the consumption gap predicts also other sources of systematic risk. Appendix Table F.5 reports results for the predictive regression

$$f_{n,t+1} = \gamma_{n,0} + \gamma_{n,1} cgap_t + \epsilon_{n,t+1} \ ,$$

where $f_{n,t+1}$ is the log factor return for factor n in the period $t : t+1$. Panel A reports results for the four (five minus the market portfolio) factors proposed by Fama and French (2015). Among size, value, investment, and profitability factors, the consumption gap significantly predicts only the value factor HML, delivering an R^2 of 1.3% using quarterly data. Panel B reports results for the four (five minus the market portfolio) q-factors proposed by Hou, Xue, and Zhang (2015); Hou, Mo, Xue, and Zhang (2018). I find that the *cgap* significantly forecasts the investment, ROE, and expected growth factors, with the R^2 ranging from 2% to 6.6%. Panel C and D shows that the consumption gap significantly predicts both the two mispricing factors proposed by Stambaugh and Yuan (2016) and the two behavioral factors proposed by Daniel, Hirshleifer, and Sun (2020).

These results suggest that the consumption gap is not just related to the market portfolio. Rather, it carries economically relevant information, thus predicting several systematic sources of risk.

Dissecting Predictability. To investigate whether the consumption gap carries any information related to asset-specific risk, I consider a predictive regression in which I forecast only the idiosyncratic return component. I use the 10 Portfolios sorted on Book-to-Market as test assets. To compute the idiosyncratic return component, I orthogonalize each asset i returns with respect to the market portfolio. Specifically, I first run the regression $r_{n,t+1} = \alpha_{n,0} + \beta_{n,1} r_{m,t+i} + v_{n,t+1}$ where $r_{n,t+1}$ is the excess return for asset n and $r_{m,t+i}$ is the log market excess return. Then, I compute the idiosyncratic component as $r^I_{n,t+i} = \hat{v}_{n,t+1}$.[19] Appendix Table F.6 reports results for the predictive regression:

$$r^I_{n,t+i} = \gamma_{n,0} + \gamma_{n,1} cgap_t + \epsilon_{n,t+i} \ .$$

Panel A reports results for quarterly returns. I do not find any predictability for the idiosyncratic component of returns using the consumption gap. Panel B reports identical results for annual returns.[20]

3.1.3 The Economic Value of Market Timing

I investigate the economic value of timing the market for an investor using the consumption gap. Specifically, I consider a mean-variance investor who faces the following objective

[19] Appendix Figure F.1 shows actual and fitted values for Growth (decile 1) and Value (decile 10) annual returns in the 10 Portfolios sorted on Book-to-Market using the consumption gap. Across the 10 Portfolios I find an R^2 of 2.6% and 8.5% respectively for quarterly and annual returns.

[20] My analysis is complementary to Favero, Melone, and Tamoni (2021) who find that the cointegrating relationship between portfolio prices and factor prices predicts the idiosyncratic component of returns. Differently, I show that temporary consumption deviations predict the systematic component of returns.

function at the end of quarter t:

$$\arg\max_{w_{t+1|t}} \mathrm{E}_t\left[R_{p,t+1}\right] - 0.5\gamma \mathrm{Var}\left(R_{p,t+1}\right)$$

where $R_{p,t+1} = w_{t+1|t} r_{m,t+1}$, $w_{t+1|t}$ is the allocation to the market portfolio in period $t+1$ given its forecast at time t, and γ represents the coefficient of relative risk aversion. Following, e.g., Rapach, Strauss, and Zhou (2010), I set $\gamma = 3$. Given the optimal portfolio weights, the average utility realized by the investor is given by

$$\overline{U}_j = \overline{R}_{p,t+1} - 0.5\gamma \widehat{\mathrm{Var}}\left(R_{p,t+1}\right)\ \text{, for } j = 0, 1,$$

where a subscript of 0 or 1 indicates the mean and variance for the portfolio return when the investor uses, respectively, the prevailing mean or the consumption gap to predict $r_{m,t+1}$. In the case $j = 1$, optimal portfolio weights are a linear function of the state variable *cgap* (e.g., Campbell and Viceira, 1999; Brandt and Santa-Clara, 2006). Finally, I compute the average utility gain (or increase in certainty equivalent return) when the investor uses the competing forecast in lieu of the prevailing mean benchmark as $\Delta\overline{U} = \overline{U}_1 - \overline{U}_0$.

The average utility gain has the same unit of measure as returns and, multiplied by four, can be interpreted as the annualized portfolio management fee (as a proportion of wealth) that the investor would be willing to pay to have access to the information in the competing forecast relative to that in the prevailing mean benchmark. I find that the annualized utility gain reaches 2.39% for the market timing, which is higher than the 2% threshold for economic significance introduced by Pástor and Stambaugh (2000).[21]

[21] Using the vintage version of the consumption gap $cgap^{vintage}$ leads to an annualized certainty equivalent return of 5.78%.

Such a market timing strategy generates an annualized performance of 7.56% with a $t-$stat of 3.6. Furthermore, this performance cannot be explained by the performance of several systematic factors. Table 5 reports results for the regression

$$MT_t^{cgap} = \alpha + \beta X_t + \varepsilon_t \ ,$$

where MT_t^{cgap} is the performance resulting from the market timing strategy which uses the consumption gap to predict the market and X_t is a matrix containing factor returns for different characteristics-based factor models. Specifically, I consider the CAPM, (Fama and French, 1993) three-factor model (FF3), the Carhart (1997) four-factor model, the (Fama and French, 2015) five-factor model (FF5), the five-factor model proposed by (Hou, Xue, and Zhang, 2015; Hou et al., 2018), the Stambaugh and Yuan (2016) mispricing four-factor model, the Daniel, Hirshleifer, and Sun (2020) behavioral three-factor model, and the CAPM plus "betting against the beta" (BAB) factor proposed by Frazzini and Pedersen (2014). I find that the market timing strategy which uses the *cgap* to predict the market generates a positive and significant alpha which ranges between 0.87 to 1.61% per year across different factor models.

Finally, I also compare the performance of the market timing strategy using the *cgap* with the one using *cay* (Lettau and Ludvigson, 2001a) as a return predictor. A regression of returns resulting from the *cgap*-strategy on returns from the *cay*-strategy delivers an annualized alpha of 3.28% with a $t-$stat of 2.1 and an associated Sharpe ratio which is 58% higher than the Sharpe ratio of the market timing strategy in which one would use *cay* to predict the market.

3.2 Cross-Sectional Analysis

Macro-finance models can be represented via a stochastic discount factor (SDF) which is linear in the factors that span the mean-variance frontier (Hansen and Jagannathan, 1991). A general SDF representation for the CAPM is:

$$m_{t+1} = 1 - b_t'(R_{m,t+1} - \mathrm{E}_t[R_{m,t+1}]) \ , \tag{8}$$

where $R_{m,t+1}$ is the excess market return at time $t+1$. In the absence of arbitrage, the SDF in equation (8) satisfies the fundamental pricing equation $0 = \mathrm{E}_t[m_{t+1}R_{i,t+1}]$ for any excess asset return $R_{i,t+1}$. b_t represents the SDF loadings and can be interpreted as the time-varying price of risk on the market portfolio.[22]

This model implies the familiar expected return-beta representation:

$$\mathrm{E}_t[R_{i,t+1}] = \beta_{i,t}\lambda_{m,t} \tag{9}$$

where $\lambda_{m,t} = \mathrm{E}_t[R_{m,t+1}] = b_t \mathrm{Var}_t[R_{m,t+1}]$ is the time-varying market risk premium. This specification represents a conditional version of the CAPM (Hansen and Richard, 1987). The conditional CAPM in equation (9) is a dynamic single-factor model with the market portfolio as the only fundamental factor.

Next, I estimate a conditional version of the CAPM (8) using the consumption gap as conditioning variable. Then, I study the properties of the implied empirical stochastic discount factor (SDF).

[22] In consumption-based macro-finance models in which consumption growth is the risk factor, b_t represents aggregate risk aversion (e.g., Campbell and Cochrane, 1999). Recent empirical work on time-varying risk aversion include Guiso, Sapienza, and Zingales (2018) and Bekaert, Engstrom, and Xu (2021).

3.2.1 Pricing Tests

Consider the conditional CAPM in equation (9). To test this, following, e.g., Jagannathan and Wang (1996), I parameterize the conditional model assuming that the time-varying coefficients are a linear function of the consumption gap. This characterization allows to specify the conditional model (9) as

$$\mathrm{E}_t[R_{i,t+1}] = \beta_{0,i}\mathrm{E}_t[R_{m,t+1}] + \beta_{1,i}\mathrm{E}_t[cgap_t R_{m,t+1}]$$

and taking unconditional expectations

$$\mathrm{E}[R_{i,t+1}] = \beta_{0,i}\lambda_0 + \beta_{1,i}\lambda_1 \ , \tag{10}$$

where $\lambda_0 = \mathrm{E}[R_{m,t+1}]$ and $\lambda_1 = \mathrm{E}[cgap_t R_{m,t+1}]$. This is a multifactor unconditional model in which the factors are the excess market return and the excess market return scaled by the consumption gap. The market portfolio is the only fundamental factor. The scaled excess market can be interpreted as a managed portfolio which invests more aggressively in the market when the consumption gap is high. Since the consumption gap negatively predicts the stock market, the risk premium associated with this managed portfolio should be negative (Maio and Santa-Clara, 2012; Boons, 2016).

Finally, I can use standard methods to estimate the system of equations

$$\begin{aligned} m_{t+1} &= 1 - \boldsymbol{b}'\boldsymbol{F}_{t+1} \\ 0 &= \mathrm{E}[R_{i,t+1} m_{t+1}] \ , \end{aligned} \tag{11}$$

where $\boldsymbol{F}_{t+1} = \{R_{m,t+1}, cgap_t R_{m,t+1}\}$, $\boldsymbol{b} = \{b_0, b_1\}$, and $R_{i,t}$ is the excess return of asset i at time t. Once I estimate $\boldsymbol{b}$ using the system of equation (11), I can compute the lambdas in equation (10) as $\mathrm{E}[F_t' F_t]\boldsymbol{b}$. Following, e.g., Parker and Julliard (2005) and Lettau and Ludvigson (2001b), I estimate the system of equations (11) via the Generalized Method of Moments (GMM). This estimation methodology allows to correct for autocorrelation and heteroskedasticity as well as for pricing errors correlation across assets. I use a two-step procedure. In the first step, I set the weighting matrix for the moment conditions equal to the identity matrix. In the second step, I use the optimal weighting matrix from the first step. I compute robust standard errors as in Newey and West (1987).

Table 6 reports estimates of lambdas and SDF coefficients for several test assets, together with cross-sectional R^2. Test assets are the 25 double-sorted portfolios formed on book-to-market (B/M), market capitalization (Size), investments (Inv), long-term reversal (LT Rev), operating profitability (Op); the 32 triple-sorted portfolios formed on book-to-market, market capitalization, and investments; the 90 (dec 10 and dec 1) anomaly portfolios constructed in Kelly, Kozak, and Giglio (2020), Haddad, Kozak, and Santosh (2020), and Kozak, Nagel, and Santosh (2020).[23,24] I convert the returns for each test assets to quarterly returns in excess of the risk-free rate for the sample period 1967Q1 to 2020Q4, that is, 216 observations.

The first row shows results for the 25 Portfolio formed on Size and B/M.[25] Both lambdas

[23]Data as well as a detailed description for the construction of the double- and triple-sorted portfolios are available at Kenneth French's website `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`. Appendix A.2 provides details for the larger-cross section of anomaly portfolios.

[24]I consider only decile 1 and decile 10 for each of the 45 anomalies as it allows to retain the most important information for each characteristic (the long and the short portfolios) in a parsimonious way. Indeed, as I use a two-step GMM estimation, I need the time series sample to be larger than the cross-sectional sample size (see, e.g., Ferson and Foerster, 1994; Lettau and Ludvigson, 2001b).

[25]I report results for a general data generating process $R_{i,t+1} = \alpha_{i,t} + \beta_{i,t} R_{m,t+1}$, in which both alphas and betas are time-varying and can be specified as a linear function of the consumption gap. However,

and SDF sensitivities are significant and with the correct sign. The market portfolio is priced; the scaled market portfolio is associated with a negative lambda and a negative SDF coefficient. This is consistent with the fact that a positive consumption gap signals low future consumption growth (see equation (4)) and is associated to high marginal utility and low expected returns. Thus, b_1 should be negative (e.g., Barroso, Boons, and Karehnke, 2021). Results are consistent and robust across the different test assets, including the large cross-section of single-sorted anomaly portfolios.

The last two columns of Table 6 report, for each test assets, the improvement in GLS R^2 (Lewellen, Nagel, and Shanken, 2010) with respect to the CAPM (GLS R^2+ CAPM) or the conditional CAPM of Lettau and Ludvigson (2001a) in which *cay* is used as scaling variable rather than *cgap* (GLS R^2+ LL). For the 25 Portfolio formed on Size and B/M, using the consumption gap as a conditioning variable leads to an improvement in the cross-sectional R^2 of about 14% with respect to the unconditional CAPM and of about 8% with respect to the conditional CAPM of Lettau and Ludvigson (2001b). I get similar results for other test assets.

Overall, these results suggest that the consumption gap significantly relates to time-variation in the SDF exposure to the market portfolio, i.e., it tracks the price of market risk. Furthermore, the conditional version of the CAPM in which I use the consumption gap as a conditioning variable shows a significant ability at explaining cross-sectional variation in betas across different test assets when benchmarked against the CAPM and its conditional version proposed by Lettau and Ludvigson (2001b).

assuming that only the betas are time-varying would deliver similar results.

3.2.2 Properties of the SDF and Economic Conditions

Estimating the system of equations (11) enables to compute the stochastic discount factor (SDF) implied by the empirical equilibrium model (10) and investigate its properties.

Given the estimates reported in Table 6, I calibrate $b_0 = 0.025$ and $b_1 = -0.013$. Thus, I calculate the empirical SDF as: $m_{t+1} = 1 - 0.025 \times R_{m,t+1} + 0.013 \times cgap_t R_{m,t+1}$. Figure 3 shows the conditional variance for the SDF calculated using a standard GARCH(1, 1). The average annualized conditional variance of the SDF is 0.90.[26] Its yearly autocorrelation is 0.93 reflecting slow, long-term changes in economic conditions.

This empirical SDF has similar properties as theoretical counterparts in leading macro-finance models. For example, Bansal and Yaron (2004) report an annualized conditional variance of 0.85. Moreover, the variance of the SDF fluctuates substantially over time in a way consistent with, e.g., habit (e.g., Campbell and Cochrane, 1999) and heterogeneous agents models (e.g., Chan and Kogan, 2002). The SDF variance is related to the state of the business cycle. In Appendix Figure D.1, I plot the cross-correlogram between the variance of the SDF and the credit spread index introduced by Gilchrist and Zakrajšek (2012). The Gilchrist-Zakrajsek spread (GZ) is constructed using prices of individual corporate bonds traded in the secondary market and is a powerful predictor of economic activity. I compute the cross-correlation as $\rho^{m,GZ} = corr(\sigma_t^{SDF}, GZ_{t-i})$, where σ_t^{SDF} is the variance of the SDF at time t and GZ_{t-i} is the GZ spread at time $t - i$. The cross-correlogram shows that, for $i = \{0, \ldots, 4\}$, the correlation between the two variables takes values ranging between 45% and 51%, consistent with the idea that the GZ spread leads the SDF over the business cycle.

[26] These results are not driven by the recent spike in the SDF variance due to the COVID-19 pandemic. Indeed, excluding 2020, the average annualized conditional variance of the SDF is 0.87.

Furthermore, the large spikes in the SDF in Figure 3 coincide with periods of intense economic and financial stress—the 1970s energy crisis, the early 1980s recession, the Dot-com bubble, the Great Financial Crisis 2007–2009, and the COVID-19 pandemic. This fact suggests that the empirical SDF well-captures how the price of risk changes over the business cycle: When marginal utility is more volatile, investors ask higher compensation for bearing risks.

3.3 Robustness

3.3.1 Alternative Permanent Component Measures

The forecasting ability of the consumption gap could be related to the way the CBO potential output is constructed. To address this issue, I run both in-sample and out-of-sample analysis for three alternative measures of the permanent component used to de-trend aggregate consumption expenditures.

First, I use a simple quadratic regression to compute the trend in real GDP. Specifically, I follow, e.g., Clarida, Gali, and Gertler (2000); Cooper and Priestley (2009), and compute the trend in real GDP as the fitted value from the regression $y_t = \pi_0 + \pi_1 t + \pi_2 t^2 + \varepsilon_t$, where y_t is real GDP at time t. Then, I compute the consumption gap subtracting the new estimate for the trend in real GDP from current consumption expenditures. Appendix Table F.7 reports results for the predictive regressions for the market. Panel A reports in-sample results. The coefficients on the consumption gap calculated using the alternative real GDP trend estimate are significant and negative. The R^2 from the regressions ranges between 4.8% (quarterly returns) to about 15% (annual returns). Panel B reports out-of-sample R^2, which is positive for different aggregation horizons and across different sample periods.

Second, I follow Laubach and Williams (2003) and compute the stochastic trend in real GDP using a Kalman filter methodology. Specifically, I compute the consumption gap subtracting the estimate for potential output produced using the Kalman filter from current consumption expenditures. Appendix Table F.8 reports results for predicting the market using the newly obtained consumption gap. Panel A reports in-sample results. The coefficients on the consumption gap computed using the alternative potential output estimate are significant and negative. The R^2 from the regressions ranges between 2.3% (quarterly returns) to about 9% (annual returns). Panel B reports out-of-sample R^2, which is positive for different aggregation horizons up to 40 quarters and across different sample periods.

As a last alternative measure for the permanent component, I follow the original formulation of the PIH of Friedman (1957) and employ real disposable income instead of real GDP in equation (1). I compute the permanent component in income using the linear projection method proposed by Hamilton (2018), which avoids potential spurious dynamics introduced by HP filtering. Specifically, the trend in real disposable income x is the fitted value from the regression: $x_t = a_0 + a_1 x_{t-k} + a_2 x_{t-k-1} + a_3 x_{t-k-2} + a_1 x_{t-k-3} + v_t$; following Atanasov, Møller, and Priestley (2020), I use $k = 24$. Finally, I compute temporary consumption deviations cx as the difference between current consumption c and permanent income x. Appendix Table F.9 reports results for the predictive regression $r_{m,t+i} = \gamma_0 + \gamma_1 cx_t + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t + i$. Panel A reports in-sample predictability results, while Panel B reports results for the out-of-sample analysis. These findings are qualitatively and quantitatively similar to the results reported in Table 2, suggesting that the main findings are robust to different specification for the trend component.

3.3.2 De-Trending Consumption Levels with a Deterministic Time Trend

Appendix Table C.1

3.3.3 A Reduced-Form VAR Approach

I use a Vector Autoregressive Model (VAR) to (a) assess the predictive power of the consumption gap for market returns controlling for the dynamic interactions of market returns and consumption growth; (b) investigate both the short- and the long-run effects of the consumption gap on market returns. Specifically, I fit the following reduced-form VAR:

$$Y_t = C + BY_{t-1} + \zeta \tag{12}$$

where $Y_t = [r_{m,t}, \Delta c_t, cgap_t]$, C is a vector of constants, B is a matrix containing the exposure of each variable to lagged values of all the variables, and ζ is a vector of shocks. Appendix Table E.1 reports OLS estimates and the R^2 for the VAR specified in equation (12). Column (1) reports the central resul the consumption gap negatively predicts aggregate stock returns, even after controlling for past returns and consumption growth. Column (2) shows that the stock market together with consumption growth predict future consumption growth. Column (3) shows that all the variables contribute at predicting the consumption gap.

The VAR allows to assess the response of stock market returns caused by a shock to any of the predictors. To this end, Appendix Figure E.1 shows the impulse response functions (IRF) for the estimated VAR. Panel (a) and Panel (b) shows similar results: one positive standard deviation of, respectively, market returns or consumption growth has no effect on future market returns. Differently, Panel (c) shows that a consumption gap shock is associated with a statistically significant reduction in future market returns, consistent with the evidence reported above.

Finally, I study the long-run effect of a positive consumption gap shock on market returns Appendix Figure E.2 shows the cumulative effect of a one positive standard deviation consumption gap shock. The VAR model produces a significant economic role for consumption gap shocks over the long-run, whit a point estimate of the change in expected five-year ahead returns of more than 10%.

Overall, the VAR approach delivers results comparable with the ones from OLS regressions, supporting the robustness of those empirical findings.

3.3.4 International Evidence

I use international data from the OECD, IMF, and World Bank database, together with stock market data from Compustat/XpressFeed Global database to investigate predictability implied by the consumption gap for Canada and the UK.[27] Appendix Table G.1 reports results for the predictive regression $r_{i,t+1} = \alpha + \beta cgap_t^i + \epsilon_{t+1}$, where $r_{i,t+1}$ is the log market return of country i at time $t+1$ and $cgap^i$ is the consumption gap of country i computed as in equation (3) (but using macro-data for the respective country). For both countries, CAN and GBR, I find that the consumption gap significantly relate to future expected market returns, delivering an R^2 greater than 1% using quarterly data. Furthermore, the coefficient on the consumption gap is negative, supporting the economic mechanism which generates predictability. I also control for the US consumption gap and I find that it has no role at predicting cross-country market returns.

[27]Further details on data are reported in Appendix A.3.

4 Economic Mechanisms

My empirical findings suggest that the market price of risk varies over time and that it is inversely related to the consumption gap.

Several equilibrium theories can generate a time-varying price of risk through various economic mechanisms, including "catching up with the Joneses" utility functions (e.g., Campbell and Cochrane, 1999; Chan and Kogan, 2002), overlapping generation model with heterogeneous agents (Gârleanu and Panageas, 2015), or intermediary constraints (He and Krishnamurthy, 2013).

Given the interpretation of the consumption gap as temporary deviations of current consumption from aggregate productivity expectations, I focus on a simplified version of the model proposed by Bhamra and Uppal (2014), which is flexible enough to allow for heterogeneity in beliefs. This specification is also consistent with the growing empirical literature documenting pervasive heterogeneity in beliefs (e.g., Manski, 2018; Baker, McElroy, and Sheng, 2020; Das, Kuhnen, and Nagel, 2020; Giglio, Maggiori, Stroebel, and Utkus, 2021).

Bhamra and Uppal (2014) consider a continuous-time, pure-exchange economy; financial markets are complete and the time horizon is infinite. There is a unique non-storable consumption good that serves as the numeraire. The stochastic endowment process Y_t is exogenous. There are two types of agents $k \in \{1, 2\}$ who may not have correct beliefs about the aggregate endowment process. Beliefs between the two agents can be different. Agents have power utility over consumption C_k with external habit H_k. Agent k maximizes expected lifetime utility

$$V_{k,t} = \mathrm{E}_t^k \left[\int_t^\infty e^{-\beta_k (u-t)} \frac{1}{1-\gamma} \left(\frac{C_{k,u}}{H_{k,u}} \right)^{1-\gamma} du \right] \ ,$$

where β_k is the constant subjective discount rate, γ is the coefficient of relative risk aversion, and E_t^k is the conditional expectation at time t under belief k, subject to the static budget constraint

$$\mathrm{E}_0^k \left[\int_0^\infty \frac{\pi_{k,t}}{\pi_{k,0}} C_{k,t} dt \right] \leq W_{k,0} \ ,$$

where $W_{k,0}$ is her initial wealth. The equilibrium consumption-sharing rule is given by the first-order condition for optimal consumption for the central planners problem. The true dynamics for the aggregate endowment are

$$\frac{dY_t}{Y_t} = \mu_Y dt + \sigma_Y dZ_t \ .$$

However, agent k believes that the expected growth rate of the endowment process is $\mu_{Y,k}$ Aggregate beliefs about the endowment process, $\tilde{\mu}_{Y,t}$, are given by the consumption-share weighted mean of individual agents' beliefs. Finally, the evolution of the central planner's state-price density π_t is

$$\frac{d\pi_t}{\pi_t} = -r_t dt - b_t dZ_t \ ,$$

where b_t represents the market price of risk and is given by

$$b_t = \gamma \sigma_Y + \frac{\mu_Y - \tilde{\mu}_{Y,t}}{\sigma_Y} \ . \tag{13}$$

Equation (13) predicts that time-variation in the market price of risk depends on whether beliefs are more pessimistic or more optimistic on average. Specifically, when average beliefs are more optimistic, i.e., $\mu_Y - \tilde{\mu}_{Y,t} < 0$, the price of market risk is lower than when average beliefs are more pessimistic, i.e., $\mu_Y - \tilde{\mu}_{Y,t} > 0$. Periods in which aggregate beliefs are pessimistic correspond to periods in which the consumption gap is negative. The intuition for this is that after receiving bad news about the state of the economy, agents have more pessimistic productivity expectations on average. In this situation, current consumption falls below the trend in real GDP, marginal utility is more volatile, and the compensation for bearing risk is higher.

If this intuition is correct, the price of market risk should be inversely related to measures of macroeconomic optimism. In Figure 4, I overlay the market price of risk implied by the empirical model (10) and the macroeconomic optimism index (OPTINDX) constructed in Das, Kuhnen, and Nagel (2020).[28] Two comments are in order. First, the price of market risk is countercyclical. This result is consistent with seminal work by Fama and French (1989) and Ferson and Harvey (1991) who document that the market risk premium increases during economic contractions and peaks near business cycle troughs. Furthermore, Figure 4 confirms the model prediction: more optimistic periods (e.g., during economic expansions) coincide, on average, with periods of lower market price of risk, whereas more pessimistic periods (e.g., recessions) coincide with periods of higher market price of risk.

5 Conclusions

In their survey, Lettau and Ludvigson (2010) discuss the puzzling disconnect between expected returns and real economic variables. This paper addresses this puzzle by documenting

[28] OPTINDX is an aggregate index constructed as the standardized average of several beliefs measures from the Michigan Survey of Consumers. For details on how to construct OPTINDX see Das, Kuhnen, and Nagel (2020).

a novel link between return predictability and consumption expenditures.

I have started from the economic intuition that—like in the life cycle-permanent income hypothesis (PIH)—individuals decide how much to consume according to their long-run productivity expectations. However, differently from the standard PIH, I find that aggregate consumption temporarily deviates from its permanent level warranted by productivity. These transitory consumption deviations—termed the consumption gap—are consistent with a view that individuals overreact to recent information. In this paper, I find that the consumption gap predicts stock returns as it tracks the price of market risk over the business cycle. These results lead to an empirical SDF that varies with the economic conditions, consistent with benchmark macro–finance theories.

These findings challenge the traditional view that economic policies affect consumption only as much as they affect permanent income (Hall, 1978; Cochrane, 1994). Indeed, individuals' expectations, and their volatility over time, play a central role in shaping business cycle consumption fluctuations, which are relevant for understanding countercyclical variation in the market price of risk.

References

Ang, A., and G. Bekaert. 2007. Stock return predictability: Is it there? *The Review of Financial Studies* 20:651–707.

Angeletos, G.-M., and C. Lian. 2021. Confidence and the propagation of demand shocks. *Review of Economic Studies (Forthcoming)* –:–.

Atanasov, V., S. V. Møller, and R. Priestley. 2020. Consumption fluctuations and expected returns. *The Journal of Finance* 75:1677–713.

Baker, S. R., T. S. McElroy, and X. S. Sheng. 2020. Expectation formation following large, unexpected shocks. *Review of Economics and Statistics* 102:287–303.

Bali, T. G. 2008. The intertemporal relation between expected returns and risk. *Journal of Financial Economics* 87:101–31.

Bandi, F., B. Perron, A. Tamoni, and C. Tebaldi. 2019. The scale of predictability. *Journal of Econometrics* 208:120–40.

Bandi, F., and A. Tamoni. 2020. Business-cycle consumption risk and asset prices. Working Paper.

Bansal, R., D. Kiku, and A. Yaron. 2010. Risks for the long run: Estimation and inference. Working Paper.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance* 59:1481–509.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer. 2015. X-capm: An extrapolative capital asset pricing model. *Journal of Financial Economics* 115:1–24.

Barro, R. J. 2009. Rare disasters, asset prices, and welfare costs. *American Economic Review* 99:243–64. doi:10.1257/aer.99.1.243.

Barroso, P., M. Boons, and P. Karehnke. 2021. Time-varying state variable risk premia in the icapm. *Journal of Financial Economics* 139:428–51.

Bekaert, G., E. C. Engstrom, and N. R. Xu. 2021. The time variation in risk appetite and uncertainty. *Management Science* .

Beveridge, S., and C. R. Nelson. 1981. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics* 7:151–74.

Bhamra, H. S., and R. Uppal. 2014. Asset prices with heterogeneity in preferences and beliefs. *The Review of Financial Studies* 27:519–80.

Bianchi, F., C. L. Ilut, and H. Saijo. 2021. Implications of diagnostic expectations: Theory and applications. Working Paper, National Bureau of Economic Research.

Bidder, R., and I. Dew-Becker. 2016. Long-run risk is the worst-case scenario. *American Economic Review* 106:2494–527.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni. 2013. News, noise, and fluctuations: An empirical exploration. *American Economic Review* 103:3045–70.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge. 1988. A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96:116–31.

Boons, M. 2016. State variables, macroeconomic activity, and the cross section of individual stocks. *Journal of Financial Economics* 119:489–511.

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer. 2019. Diagnostic expectations and stock returns. *The Journal of Finance* 74:2839–74.

Bordalo, P., N. Gennaioli, and A. Shleifer. 2018. Diagnostic expectations and credit cycles. *The Journal of Finance* 73:199–227.

Borup, D., and E. C. M. Schütte. 2021. Asset pricing with data revisions. *Journal of Financial Markets* 100620.

Bouchaud, J.-P., P. Krueger, A. Landier, and D. Thesmar. 2019. Sticky expectations and the profitability anomaly. *The Journal of Finance* 74:639–74.

Box, G. E., and G. M. Jenkins. 1970. *Time series analysis: Forecasting and control.*

Brandt, M. W., and P. Santa-Clara. 2006. Dynamic portfolio selection by augmenting the asset space. *The Journal of Finance* 61:2187–217.

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

Bryzgalova, S., and C. Julliard. 2021. Consumption. Working Paper.

Campbell, J. Y., and J. H. Cochrane. 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107:205–51.

Campbell, J. Y., C. Pflueger, and L. M. Viceira. 2020. Macroeconomic drivers of bond and equity risks. *Journal of Political Economy* 128:3148–85.

Campbell, J. Y., and S. B. Thompson. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies* 21:1509–31.

Campbell, J. Y., and L. M. Viceira. 1999. Consumption and Portfolio Decisions when Expected Returns are Time Varying*. *The Quarterly Journal of Economics* 114:433–95.

Carhart, M. M. 1997. On persistence in mutual fund performance. *The Journal of Finance* 52:57–82.

Chan, Y. L., and L. Kogan. 2002. Catching up with the joneses: Heterogeneous preferences and the dynamics of asset prices. *Journal of Political Economy* 110:1255–85.

Chernov, M., L. A. Lochstoer, and D. Song. 2021. The real explanation of nominal bond-stock puzzles. Working Paper.

Chodorow-Reich, G., A. M. Guren, and T. J. McQuade. 2021. The 2000s housing cycle with 2020 hindsight: A neo-kindlebergerian view. Working Paper, National Bureau of Economic Research.

Clarida, R., J. Gali, and M. Gertler. 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. *The Quarterly Journal of Economics* 115:147–80.

Clark, T., and K. West. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138:291–311.

Cochrane, J. H. 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. *The Journal of Finance* 46:209–37.

———. 1994. Permanent and transitory components of gnp and stock prices. *The Quarterly Journal of Economics* 109:241–65.

———. 2005. *Asset pricing: Revised edition.* Princeton university press.

Constantinides, G. M., and D. Duffie. 1996. Asset pricing with heterogeneous consumers. *Journal of Political Economy* 104:219–40.

Cooper, I., and R. Priestley. 2009. Time-varying risk premiums and the output gap. *The Review of Financial Studies* 22:2801–33.

Daniel, K., D. Hirshleifer, and L. Sun. 2020. Short-and long-horizon behavioral factors. *The Review of Financial Studies* 33:1673–736.

Das, S., C. M. Kuhnen, and S. Nagel. 2020. Socioeconomic status and macroeconomic expectations. *The Review of Financial Studies* 33:395–432.

Duffee, G. R. 2005. Time variation in the covariance between stock returns and consumption growth. *The Journal of Finance* 60:1673–712.

Fama, E. F. 1991. Efficient capital markets: Ii. *The Journal of Finance* 46:1575–617.

Fama, E. F., and K. R. French. 1988. Permanent and temporary components of stock prices. *Journal of Political Economy* 96:246–73.

———. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25:23–49.

———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* .

———. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116:1–22.

Favero, C., A. Melone, and A. Tamoni. 2021. Factor models with drifting prices. Working Paper.

Ferson, W. E., and S. R. Foerster. 1994. Finite sample properties of the generalized method of moments in tests of conditional asset pricing models. *Journal of Financial Economics* 36:29–55.

Ferson, W. E., and C. R. Harvey. 1991. The variation of economic risk premiums. *Journal of Political Economy* 99:385–415.

———. 1999. Conditioning variables and the cross section of stock returns. *The Journal of Finance* 54:1325–60.

Ferson, W. E., S. Sarkissian, and T. T. Simin. 2003. Spurious regressions in financial economics? *The Journal of Finance* 58:1393–413.

Ferson, W. E., and R. W. Schadt. 1996. Measuring fund strategy and performance in changing economic conditions. *The Journal of Finance* 51:425–61.

Frazzini, A., and L. H. Pedersen. 2014. Betting against beta. *Journal of Financial Economics* 111:1–25.

Friedman, M. 1957. The permanent income hypothesis. In *A theory of the consumption function*, 20–37. Princeton University Press.

Fuster, A., B. Hebert, and D. Laibson. 2012. Natural expectations, macroeconomic dynamics, and asset pricing. *NBER Macroeconomics Annual* 26:1–48.

Gârleanu, N., and S. Panageas. 2015. Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing. *Journal of Political Economy* 123:670–85.

Gibbons, M. R., and W. Ferson. 1985. Testing asset pricing models with changing expectations and an unobservable market portfolio. *Journal of Financial Economics* 14:217–36.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus. 2021. Five facts about beliefs and portfolios. *American Economic Review* 111:1481–522.

Gilchrist, S., and E. Zakrajšek. 2012. Credit spreads and business cycle fluctuations. *American Economic Review* 102:1692–720.

Gormsen, N. J., and C. S. Jensen. 2021. Conditional risk in global stock returns. Working Paper.

Guiso, L., P. Sapienza, and L. Zingales. 2018. Time varying risk aversion. *Journal of Financial Economics* 128:403–21.

Haddad, V., S. Kozak, and S. Santosh. 2020. Factor timing. *The Review of Financial Studies* 33:1980–2018.

Hall, R. E. 1978. Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *Journal of Political Economy* 86:971–87.

Hamilton, J. D. 2018. Why you should never use the hodrick-prescott filter. *Review of Economics and Statistics* 100:831–43.

Hansen, L. P., and R. Jagannathan. 1991. Implications of security market data for models of dynamic economies. *Journal of Political Economy* 99:225–62.

Hansen, L. P., and S. F. Richard. 1987. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica* 587–613.

Hasler, M., and C. Martineau. 2020. The time series of capm-implied returns. Working Paper.

He, Z., and A. Krishnamurthy. 2013. Intermediary asset pricing. *American Economic Review* 103:732–70.

Hodrick, R. J. 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *The Review of Financial Studies* 5:357–86.

Hou, K., H. Mo, C. Xue, and L. Zhang. 2018. Which factors? *Review of Finance* 23:1–35.

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *The Review of Financial Studies* 28:650–705.

Ilut, C. L., and M. Schneider. 2014. Ambiguous business cycles. *American Economic Review* 104:2368–99.

Ilut, C. L., and R. Valchev. 2020. Economic agents as imperfect problem solvers. Working Paper, National Bureau of Economic Research.

Jagannathan, R., and Y. Wang. 2007. Lazy investors, discretionary consumption, and the cross-section of stock returns. *The Journal of Finance* 62:1623–61.

Jagannathan, R., and Z. Wang. 1996. The conditional capm and the cross-section of expected returns. *The Journal of Finance* 51:3–53.

Johansen, S. 1991. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica* 1551–80.

Kelly, B., S. Kozak, and S. Giglio. 2020. Equity term structures without dividend strips data. Working Paper.

King, R. G., C. I. Plosser, J. H. Stock, and M. W. Watson. 1991. Stochastic trends and economic fluctuations. *American Economic Review* 819–40.

Kozak, S., S. Nagel, and S. Santosh. 2020. Shrinking the cross-section. *Journal of Financial Economics* 135:271 – 292. ISSN 0304-405X. doi:https://doi.org/10.1016/j.jfineco.2019.06.008.

Kroencke, T. A. 2017. Asset pricing without garbage. *The Journal of Finance* 72:47–98.

Laubach, T., and J. C. Williams. 2003. Measuring the natural rate of interest. *Review of Economics and Statistics* 85:1063–70.

Lettau, M., and S. C. Ludvigson. 2001a. Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance* 56:815–49.

———. 2001b. Resurrecting the (c) capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109:1238–87.

———. 2010. Measuring and modeling variation in the risk-return trade-off. *Handbook of financial econometrics: Tools and techniques* 617–90.

Lewellen, J., S. Nagel, and J. Shanken. 2010. A skeptical appraisal of asset pricing tests. *Journal of Financial Economics* 96:175–94.

LHuillier, J.-P., and D. Yoo. 2017. Bad news in the great depression, the great recession, and other us recessions: A comparative study. *Journal of Economic Dynamics and Control* 81:79–98.

Liu, Y., and B. Matthies. 2021. Long run risk: Is it there? *The Journal of Finance (Forthcoming)* –:–.

Lorenzoni, G. 2009. A theory of demand shocks. *American Economic Review* 99:2050–84.

Ludvigson, S. C., and S. Ng. 2009. Macro factors in bond risk premia. *The Review of Financial Studies* 22:5027–67.

MacKinnon, J. G. 2010. Critical values for cointegration tests. Working Paper, Queen's Economics Department.

Maio, P., and P. Santa-Clara. 2012. Multifactor models and their consistency with the icapm. *Journal of Financial Economics* 106:586–613.

Manski, C. F. 2018. Survey measurement of probabilistic macroeconomic expectations: progress and promise. *NBER Macroeconomics Annual* 32:411–71.

Martin, I., and S. Nagel. 2021. Market efficiency in the age of big data. *Journal of Financial Economics (Forthcoming)* –:–.

Mian, A., and A. Sufi. 2014. What explains the 2007–2009 drop in employment? *Econometrica* 82:2197–223.

Mian, A., A. Sufi, and E. Verner. 2017. Household debt and business cycles worldwide. *The Quarterly Journal of Economics* 132:1755–817.

Modigliani, F. 1971. Monetary policy and consumption. *Consumer spending and monetary policy: the linkages* 9–84.

Møller, S. V., and J. Rangvid. 2015. End-of-the-year economic growth and time-varying expected returns. *Journal of Financial Economics* 115:136–54.

Nagel, S., and K. J. Singleton. 2011. Estimation and evaluation of conditional asset pricing models. *The Journal of Finance* 66:873–909.

Nagel, S., and Z. Xu. 2021. Asset Pricing with Fading Memory. *The Review of Financial Studies* ISSN 0893-9454. doi:10.1093/rfs/hhab086. Hhab086.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8.

———. 1994. Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies* 61:631–53.

Ortu, F., A. Tamoni, and C. Tebaldi. 2013. Long-run risk and the persistence of consumption shocks. *The Review of Financial Studies* 26:2876–915.

Parker, J. A., and C. Julliard. 2005. Consumption risk and the cross section of expected returns. *Journal of Political Economy* 113:185–222.

Pástor, L., and R. F. Stambaugh. 2000. Comparing asset pricing models: an investment perspective. *Journal of Financial Economics* 56:335–81.

Pohl, W., K. Schmedders, and O. Wilms. 2016. Asset prices with non-permanent shocks to consumption. *Journal of Economic Dynamics and Control* 69:152–78.

Proietti, T. 2006. Trend–cycle decompositions with correlated components. *Econometric Reviews* 25:61–84.

Rangvid, J. 2006. Output and expected returns. *Journal of Financial Economics* 81:595–624.

Rapach, D. E., J. K. Strauss, and G. Zhou. 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies* 23:821–62.

Roussanov, N. 2014. Composition of wealth, conditioning information, and the cross-section of stock returns. *Journal of Financial Economics* 111:352–80.

Santos, T., and P. Veronesi. 2006. Labor income and predictable stock returns. *The Review of Financial Studies* 19:1–44.

Savov, A. 2011. Asset pricing with garbage. *The Journal of Finance* 66:177–201.

Schorfheide, F., D. Song, and A. Yaron. 2018. Identifying long-run risks: A bayesian mixed-frequency approach. *Econometrica* 86:617–54.

Stambaugh, R. F., and Y. Yuan. 2016. Mispricing factors. *The Review of Financial Studies* 30:1270–315.

Stock, J. H., and M. W. Watson. 1988a. Testing for common trends. *Journal of the American Statistical Association* 83:1097–107.

———. 1988b. Variable trends in economic time series. *Journal of Economic Perspectives* 2:147–74.

———. 1993. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 783–820.

Wachter, J. A. 2006. A consumption-based model of the term structure of interest rates. *Journal of Financial Economics* 79:365–99.

Watson, M. W. 1986. Univariate detrending methods with stochastic trends. *Journal of Monetary Economics* 18:49–75.

Welch, I., and A. Goyal. 2008. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies* 21:1455–508.

Whelan, K. 2008. Consumption and expected asset returns without assumptions about unobservables. *Journal of Monetary Economics* 55:1209–21.

Tables and Figures

Table 1: Stock Market Returns and Consumption

This table reports OLS estimates and the R^2 for the regression: $r_{m,t+1} = \alpha + \beta \Delta c_{t+1} + \epsilon_{t+1}$, where $r_{m,t+1}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+1$ and Δc is log consumption growth. Columns (2) to (5) substitutes log consumption growth for its different components $\Delta c_{t+1} = \mu + (\rho - 1) cgap_t + \eta_{t+1} + \epsilon_{t+1}$, where η and ϵ are respectively the shock to the trend in real GDP and the shock to the *cgap*, and *cgap* is the consumption gap computed as in equation (3). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

	$r_{m,t+1}$				
	(1)	(2)	(3)	(4)	(5)
Δc_{t+1}	0.041				
	(0.647)				
$cgap_t$		-0.861^{***}			-0.990^{***}
		(0.170)			(0.240)
η_{t+1}			-2.850		2.697
			(3.231)		(3.866)
ϵ_{t+1}				-0.305	-0.310
				(0.564)	(0.729)
Constant	1.407^{*}	1.401^{**}	1.320^{**}	1.426^{**}	1.511^{***}
	(0.718)	(0.579)	(0.582)	(0.593)	(0.572)
Observations	216	215	216	216	215
Adjusted R^2	-0.005	0.034	-0.001	-0.004	0.028

Table 2: Predicting Market Returns

This table reports OLS estimates and the R^2 from the predictive regression: $r_{m,t+i} = \gamma_0 + \gamma_1 cgap_t + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t + i$ and *cgap* is the consumption gap computed as in equation (3). Panel A shows the in-sample results. Panel B shows the out-of-sample results. The R^2_{OOS} is computed as in Campbell and Thompson (2008); p-values for R^2_{OOS} are computed as in Clark and West (2007). I use an expanding window for estimating the predictive regressions; the in-sample period starts in 1967Q1 and ends in 1989Q4, 1999Q4, and 2009Q4. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994); for $i = \{2, 3, 4\}$, I use overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

Panel A: In-Sample

	$r_{m,t+1}$ (1)	$r_{m,t+2}$ (2)	$r_{m,t+3}$ (3)	$r_{m,t+4}$ (4)
$cgap_t$	-0.861^{***}	-1.689^{***}	-2.813^{***}	-3.564^{***}
	(0.170)	(0.614)	(0.850)	(1.034)
Constant	1.401^{**}	2.799^{**}	4.319^{**}	5.641^{**}
	(0.579)	(1.205)	(1.807)	(2.408)
Observations	215	214	213	212
R^2	0.039	0.069	0.097	0.121
$\sigma[E_t(r_{m,t+i})]/E(r_{m,t+i})$	1.229	1.216	1.188	1.160

Panel B: Out-Of-Sample

R^2_{OOS}	From 1990	From 2000	From 2010
$r_{m,t+1}$	3.14^{***}	4.31^{***}	8.96^{***}
$r_{m,t+4}$	5.66^{***}	10.43^{***}	30.58^{***}
$r_{m,t+8}$	14.13^{***}	22.86^{***}	42.15^{***}
$r_{m,t+20}$	22.59^{***}	32.22^{***}	65.74^{***}

Table 3: Predicting Market Returns Using Vintage Data

This table reports OLS estimates and the R^2 from the predictive regression: $r_{m,t+i} = \gamma_0 + \gamma_1 cgap_{t-1}^{vintage} + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+i$ and $cgap^{vintage}$ is the consumption gap computed as in equation (3) using only real-time unrevised data on consumption and potential output. Panel A shows the in-sample results. Panel B shows the out-of-sample results. The R^2_{OOS} is computed as in Campbell and Thompson (2008); p-values for R^2_{OOS} are computed as in Clark and West (2007). I use an expanding window for estimating the predictive regressions; the in-sample period starts in 1997Q1 and ends in 2004Q4 and 2009Q4. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994); for $i = \{2, 3, 4\}$, I use overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1997Q1 to 2020Q4.

Panel A: In-Sample

	$r_{m,t+1}$	$r_{m,t+2}$	$r_{m,t+3}$	$r_{m,t+4}$
	(1)	(2)	(3)	(4)
$cgap_{t-1}^{vintage}$	-0.232^{***}	-0.475^{***}	-0.647^{***}	-0.820^{***}
	(0.069)	(0.177)	(0.242)	(0.307)
Constant	1.727^{**}	3.378^{*}	4.937^{*}	6.085^{*}
	(0.805)	(1.872)	(2.780)	(3.698)
Observations	95	94	93	92
R^2	0.060	0.126	0.152	0.195
$\sigma[E_t(r_{m,t+i})]/E(r_{m,t+i})$	1.287	1.429	1.497	1.424

Panel B: Out-Of-Sample

R^2_{OOS}	$r_{m,t+1}$	$r_{m,t+4}$	$r_{m,t+8}$	$r_{m,t+20}$
From 2005	5.57^{**}	13.49^{***}	30.75^{***}	30.20^{***}
From 2010	7.46^{**}	20.19^{***}	34.49^{***}	34.25^{***}

Table 4: The Consumption Gap and Other Market Predictors

This table reports OLS estimates and the R^2 from the predictive regression $r_{m,t+1} = \beta_0 + \beta_1 x_t + \gamma_1 cgap_t + \epsilon_{t+1}$, where $r_{m,t+1}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+1$, x is a market predictor, and *cgap* is the consumption gap computed as in equation (3). In Columns (1) to (4) I study the case $\gamma_1 = 0$. Panel A shows results for the significant predictors among the sample of market predictors analized in Welch and Goyal (2008): long-term Government bond returns (ltr), CPI inflation ($infl$), and the investment-to-capital ratio (ik), plus the repurchase-adjusted dividend-price ratio constructed in Nagel and Xu (2021) ($adj.dp_t$). Panel B shows results for *cay* (Lettau and Ludvigson, 2001a), surplus-consumption ratio (s) calculated as in Wachter (2006), long-run consumption (x) calculated as in Bansal, Kiku, and Yaron (2010), and cyclical consumption (cc) calculated as in Atanasov, Møller, and Priestley (2020). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

Panel A: Goyal-Welch (2006)

	$r_{m,t+1}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$adj.dp_t$	0.018**				0.013			
	(0.009)				(0.009)			
ltr_t		0.232*				0.215*		
		(0.127)				(0.124)		
$infl_t$			−1.276**				−0.834	
			(0.605)				(0.844)	
ik_t				−4.897***				−2.773
				(1.820)				(1.974)
$cgap_t$					−0.736***	−0.825***	−0.754***	−0.636***
					(0.168)	(0.179)	(0.202)	(0.179)
Observations	215	215	215	215	215	215	215	215
Adjusted R^2	0.019	0.019	0.014	0.026	0.041	0.050	0.037	0.037

Panel B: Consumption-Based Variables

	$r_{m,t+1}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
cay_t	0.310				0.443			
	(0.289)				(0.286)			
s_t		-0.720^{***}				-0.047		
		(0.246)				(0.528)		
x_t			-0.269				0.022	
			(0.212)				(0.286)	
cc_t				-0.480^{***}				-0.278
				(0.174)				(0.201)
$cgap_t$					-0.951^{***}	-0.824^{*}	-0.875^{***}	-0.664^{***}
					(0.262)	(0.422)	(0.259)	(0.206)
Observations	215	215	215	215	215	215	215	215
Adjusted R^2	0.001	0.022	0.002	0.024	0.030	0.030	0.030	0.037

Table 5: Disconnect Between Market Timing Performance and Other Factors

This table reports the results from regressing returns resulting from the market timing strategy which uses the consumption gap to predict the market (MT^{cgap}) on several characteristics-based risk factors. Column (1) is the CAPM. Column (2) is the (Fama and French, 1993) three-factor model (FF3). Column (3) is the Carhart (1997) four-factor model. Column (4) is the (Fama and French, 2015) five-factor model (FF5). Column (5) is the five-factor model proposed by (Hou, Xue, and Zhang, 2015; Hou et al., 2018). Column (6) is the Stambaugh and Yuan (2016) mispricing four-factor model. Column (7) is the Daniel, Hirshleifer, and Sun (2020) behavioral three-factor model. Column (8) is CAPM plus "betting against the beta" (BAB) factor proposed by Frazzini and Pedersen (2014). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4; in Column (7), the sample starts in 1972Q3.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\alpha_{MT^{cgap}}$	1.14**	1.01**	1.11**	0.87*	1.55*	1.18**	1.61***	1.03**
	(0.48)	(0.49)	(0.46)	(0.49)	(0.85)	(0.51)	(0.52)	(0.46)
MKT	0.14***	0.15***	0.14***	0.15***	0.14***	0.14***	0.15***	0.13***
	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
SMB		−0.01	−0.02	−0.01		−0.01		
		(0.02)	(0.02)	(0.02)		(0.02)		
HML		0.04**	0.04	0.01				
		(0.02)	(0.02)	(0.03)				
Mom			−0.01					
			(0.02)					
RMW				−0.01				
				(0.03)				
CMA				0.07*				
				(0.03)				
ME					−0.03			
					(0.02)			
IA					0.06			
					(0.05)			
ROE					−0.07**			
					(0.03)			
EG					−0.02			
					(0.04)			
Mgmt						0.04*		
						(0.02)		
Perf						−0.04***		
						(0.02)		
PEAD							−0.08**	
							(0.03)	
FIN							0.02	
							(0.02)	
BAB								−0.00
								(0.02)
Observations	215	215	215	215	215	215	194	215
Adjusted R^2	0.40	0.40	0.40	0.41	0.43	0.42	0.43	0.41

Table 6: Prices of Risk and SDF Loadings

This table reports factor risk premia and SDF coefficients for the unconditional two-factor model in equation (11). I use the generalized method of moments (GMM) to estimate the system of equations (11). Test assets are the 25 double-sorted portfolios formed on book-to-market (B/M), market capitalization (Size), investments (Inv), long-term reversal (LT Rev), operating profitability (Op); the 32 triple-sorted portfolios formed on book-to-market, market capitalization, and investments; the 90 (dec 10 and dec 1) anomaly portfolios constructed in Kelly, Kozak, and Giglio (2020), Haddad, Kozak, and Santosh (2020), and Kozak, Nagel, and Santosh (2020). GLS R^2+ CAPM and LL denote the improvement in GLS R^2 with respect to the CAPM or using *cay* (Lettau and Ludvigson, 2001a) rather than *cgap* as scaling variable. The consumption gap *cgap* is computed as in equation (3). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors. ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

Test Assets	λ_0	λ_1	b_0	b_1	GLS R^2+ CAPM	GLS R^2+ LL
25 Portfolios Size and B/M	1.839*** (0.635)	−2.592*** (0.978)	0.025*** (0.008)	−0.013*** (0.005)	13.73	7.55
25 Portfolios Size and Inv	1.878*** (0.633)	−2.602*** (0.947)	0.025*** (0.008)	−0.013*** (0.005)	12.30	4.27
25 Portfolios Size and LT Rev	1.982*** (0.637)	−2.383** (0.948)	0.026*** (0.009)	−0.012** (0.005)	15.49	7.69
25 Portfolios B/M and Inv	1.877*** (0.644)	−2.382** (0.966)	0.025*** (0.009)	−0.012** (0.005)	13.72	9.03
25 Portfolios B/M and Op	1.714*** (0.632)	−2.753*** (1.000)	0.023*** (0.008)	−0.014*** (0.005)	12.70	5.15
32 Portfolios Size, B/M, and Inv	1.936*** (0.641)	−2.509** (0.981)	0.026*** (0.009)	−0.013** (0.005)	13.49	6.25
90 Anomaly Portfolios	1.398** (0.610)	−2.616*** (0.985)	0.024*** (0.009)	−0.021*** (0.007)	11.20	8.84

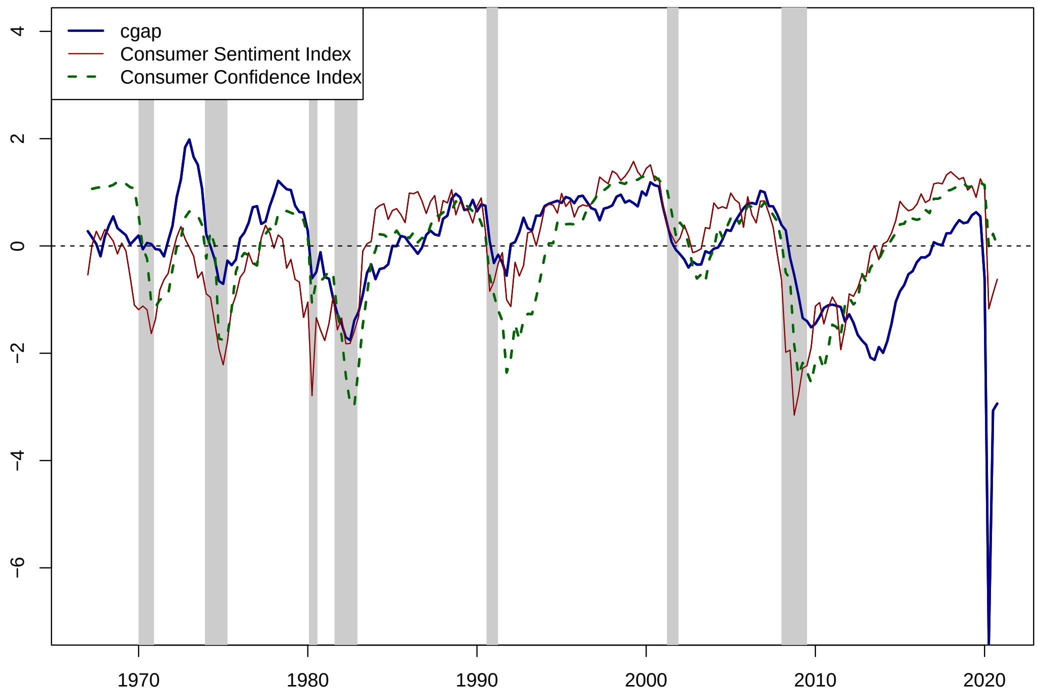


Figure 1: Consumption Gap and Consumer Sentiment. This figure shows the consumption gap with two survey-based measures of consumer sentiment. The consumption gap *cgap* is computed as in equation (3). Consumer Sentiment Index is from the University of Michigan Surveys of Consumers. Consumer Confidence Index is from the Conference Board. All series are standardized to have zero mean and unit standard deviation. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

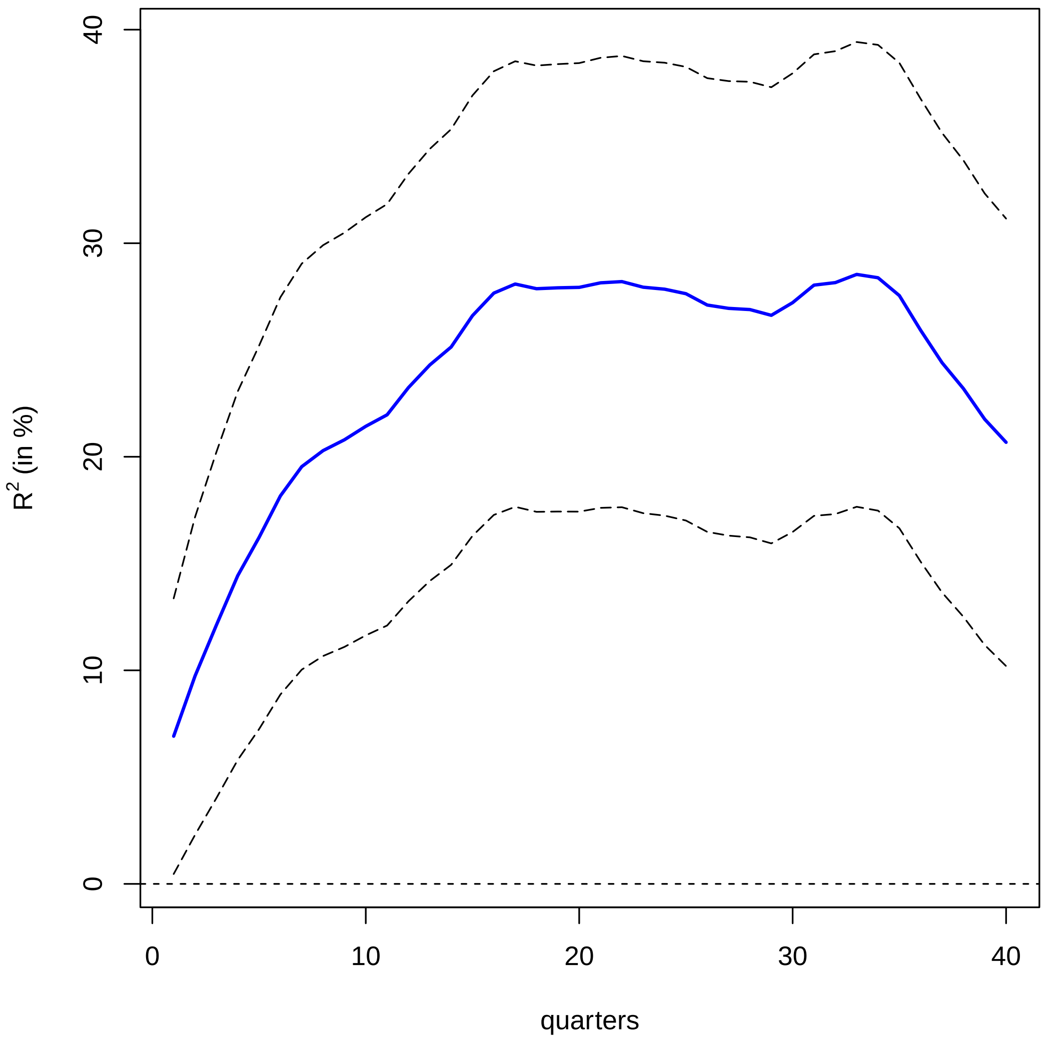


Figure 2: Long-Horizon Predictive Regressions R^2. This figure shows the in-sample R^2 for the predictive regression $r_{m,t+i} = \gamma_0 + \gamma_1 cgap_t + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+i$ for $i = \{1, 2, \ldots, 40\}$ and $cgap$ is the consumption gap computed as in equation (3). The dashed black lines represent 5% and 95% confidence intervals for the R^2 computed by bootstrap resampling. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

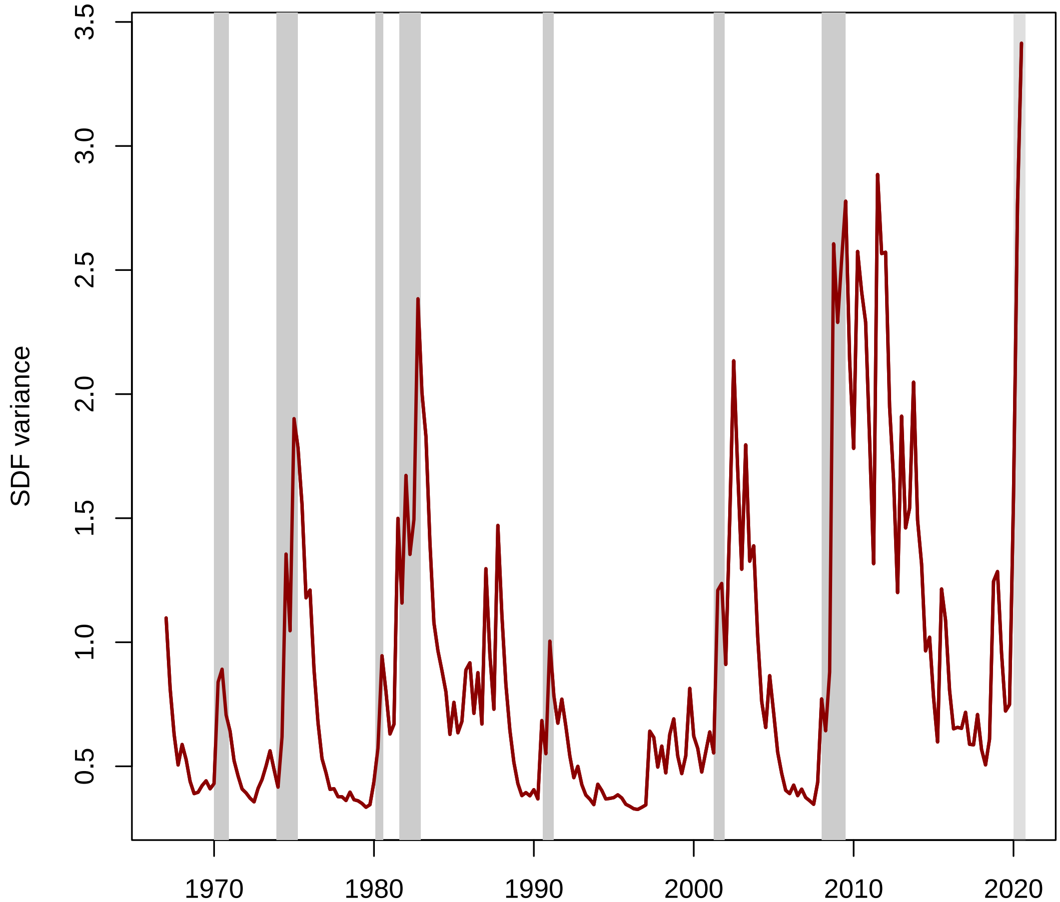


Figure 3: SDF Conditional Variance. This figure shows the annualized conditional variance of the CAPM-SDF derived using the consumption gap as conditioning variable. I calculate the conditional variance using a GARCH(1, 1) for the SDF computed as $m_{t+1} = 1 - 0.025 \times R_{m,t+1} + 0.013 \times cgap_t R_{m,t+1}$. Shaded areas are NBER recessions. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

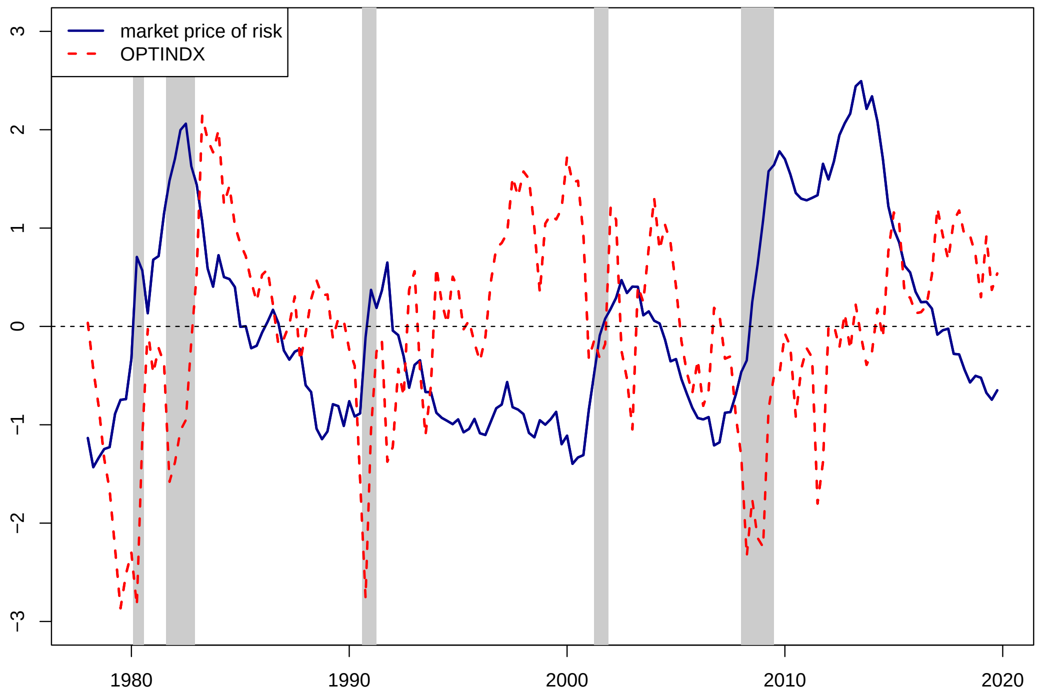


Figure 4: The Price of Market Risk and Macro Beliefs. This figure shows the consumption gap with the macroeconomic optimism index (OPTINDX) constructed as in Das, Kuhnen, and Nagel (2020). The consumption gap *cgap* is computed as in equation (3). Both series are standardized to have zero mean and unit standard deviation. Quarterly observations. The sample period is 1978Q1 to 2019Q4.

Appendix

A Data

A.1 CBO Potential Output

The US Congressional Budget Office (CBO) defines potential output as the trend growth in the productive capacity of the economy. It is a measure of maximum sustainable real GDP—the level of real GDP that is consistent with a stable rate of inflation. Current real GDP exceeds potential output when the rate of unemployment is below the natural rate of unemployment. Conversely, when the unemployment rate exceeds its natural rate, real GDP falls below potential output.

CBO uses the Solow growth model in which the real GDP growth is the product of three input factors: capital, labor, and technology. They attribute GDP into five sectors: nonfarm business, government, farm, households and nonprofits, and housing. For every sector, the CBO estimates a standard production function based on labor, capital, and total factor productivity (TFP). Specifically, sector-specific real GDP Y_i is assumed to be generated by sector-specific TFP A_i, number of hours worked L_i, and level of capital stock K_i combined into a Cobb-Douglas production function:

$$Y_{i,t} = A_{i,t} L_{i,t}^{(1-\alpha)} K_{i,t}^{(\alpha)} ,$$

where α is the capital share—the capital contribution to output growth; it is set equal to 0.3, based on historical growth accounting data that measure payments to owners of capital as 30 percent of total income. Aggregate real GDP is the sum of real GDP across the five sectors.

To estimate potential output, CBO estimates for each sector the potential level of labor force $N_{i,t}^*$ which is a function of the unemployment gap and business cycle dummies. With $N_{i,t}^*$ in hands, they compute the potential level of hours worked $L_{i,t}^*$. Then, they cyclically-adjust TFP to remove business cycle fluctuations. Finally, they compute for each sector potential output as $Y_{i,t}^* = A_{i,t} L_{i,t}^{*(0.7)} K_{i,t}^{(0.3)}$. For a complete description of the methodology used by the CBO see `CBO METHOD FOR ESTIMATING POTENTIAL OUTPUT`.

A.2 A Large Cross-Section of Test Assets

I use the large cross-section of anomaly portfolios constructed in Kelly, Kozak, and Giglio (2020), Haddad, Kozak, and Santosh (2020), and Kozak, Nagel, and Santosh (2020).

The original sample includes 51 anomaly characteristics. Then, I consider only anomaly characteristics whose observations start not later than Jan 1967 and end no earlier than Dec 2019. After this filtering process, 46 anomaly characteristics are left.

In total, my final sample consists of 92 decile 1 and decile 10 anomaly portfolios. Specifically, I use the following anomaly portfolios: accruals, age, aturnover, betaarb, cfp, ciss, divg, divp, dur, ep, exchsw, fscore, gltnoa, gmargins, growth, igrowth, indmom, indmomrev, indrrev, indrrevlv, inv, invcap, ivol, lev, lrrev, mom, mom12, momrev, nissa, nissm, noa, price, prof, roaa, roea, season, sgrowth, shvol, size, sp, strev, valmom, valmomprof, valprof, value.

Detailed anomaly definitions are from the aforementioned papers and can be found also at Serhiy Kozak's website https://sites.google.com/site/serhiykozak/data.

A.3 International Data

Aggregate consumption and potential output data are from the OECD database; consumption is quarterly, potential output is annual and linearly interpolated to get quarterly observations. The IMF database provides the CPI series, while population data are from the World Bank database.

Stock market data include all available common stocks on the Compustat/XpressFeed Global database for Canada and UK.

B Cointegration Analysis: Consumption Expenditures and Real GDP

I test whether log aggregate NIPA non-durable and service expenditures and log real GDP are cointegrated. To do this, I use testing procedure suggested by Johansen (1991). This procedure presumes a p-dimensional vector autoregressive (VAR) model with k lags, where p corresponds to the number of stochastic variables among which the investigator wishes to test for cointegration. For my application, $p = 2$ and I choose the number of lags in the VAR according to the Akaike information criterion (AIC). The Johansen procedure provides two tests for cointegration: Under the null hypothesis, H_0, that there are exactly r cointegrating relations, the "Trace" statistic supplies a likelihood ratio test of H_0 against the alternative, H_1, that there are p cointegrating relations. A second approach uses the "L-max" statistic to test the null hypothesis of r cointegrating relations against the alternative of $r + 1$ cointegrating relations.

Table B.1 presents the test results (along with the 90, 95, and 99 percent critical values for these statistics). Both tests reject at 99% level of confidence the null of no cointegration between consumption expenditures and real GDP. The estimated cointegrating vector is $(1, -1)$.

Table B.1: Johansen Cointegration Test

This table reports results for the Johansen (1991) procedure to test for cointegration between the I(1) series consumption expenditures and real GDP. Panel A reports results for the "Trace" statistics. Panel B reports results for the "L-Max" statistics. The null hypothesis is that the number of cointegration vectors is at most r. Testing is sequential for $r^* = 0, 1, ..., k-1$ and the first non-rejection of the null represents an estimate of r. The number of lags for the Johansen test is chosen according to the Akaike information criterion (AIC) and is equal to 3; I include a linear trend in the long-run regression. Quarterly observations. The sample period is 1967:Q1 to 2020:Q4.

Panel A: Trace

	Test statistics	90% CV	95% CV	99% CV
r <= 1 |	6.004	10.490	12.250	16.260
r = 0 |	40.694	22.760	25.320	30.450

Panel B: L-Max

	Test statistics	90% CV	95% CV	99% CV
r <= 1 |	6.004	10.490	12.250	16.260
r = 0 |	34.690	16.850	18.960	23.650

C Consumption: Levels vs Growth

C.1 Are Consumption Levels Trend-Stationary?

Aggregate household consumption expenditures are non-stationary. Non-stationarity in time series processes usually generates from a trend in the mean, which can be due both to the presence of a unit root or of a deterministic time trend. In the case of a deterministic trend, the time series process is trend-stationary, and shocks have only transitory effects, after which the process tends toward the time trend. In the case of a unit root, shocks have permanent effects on the process, and the deterministically de-trended process is not mean-reverting.

Appendix Figure C.1 shows aggregate log consumption with its time trend in the top panel. If the time trend would describe consumption levels, then the residuals from regressing log consumption levels on their time trend would be stationary. The bottom panel of Figure C.1 shows the deterministic trend regression residuals. The non-stationarity of the residuals is evident; furthermore, the ADF test statistics cannot reject the null of non-stationarity with a p-value of 0.99. This finding suggests that aggregate log consumption levels are not trend-stationary.

C.2 Permanent and Transitory Components of Consumption Levels

Aggregate consumption levels are integrated of order 1: they contain a stochastic trend (Beveridge and Nelson, 1981). Largely influenced by the Box and Jenkins (1970) approach, the problem of non-stationarity is commonly addressed by taking first differences of the series, which implies that consumption *growth* is specified as a stationary i.i.d. process. However, it is well-known that differencing rules out the possibility to study the relation between the levels of the series (see, e.g., Stock and Watson, 1988b).

Consider the situation in which consumption levels are described by a macroeconomic stochastic trend. In this case, some linear combination of these two variables is stationary. It means that although different economic scenarios can cause permanent changes in aggregate consumption or its macroeconomic stochastic trend, there is a long-run relationship binding

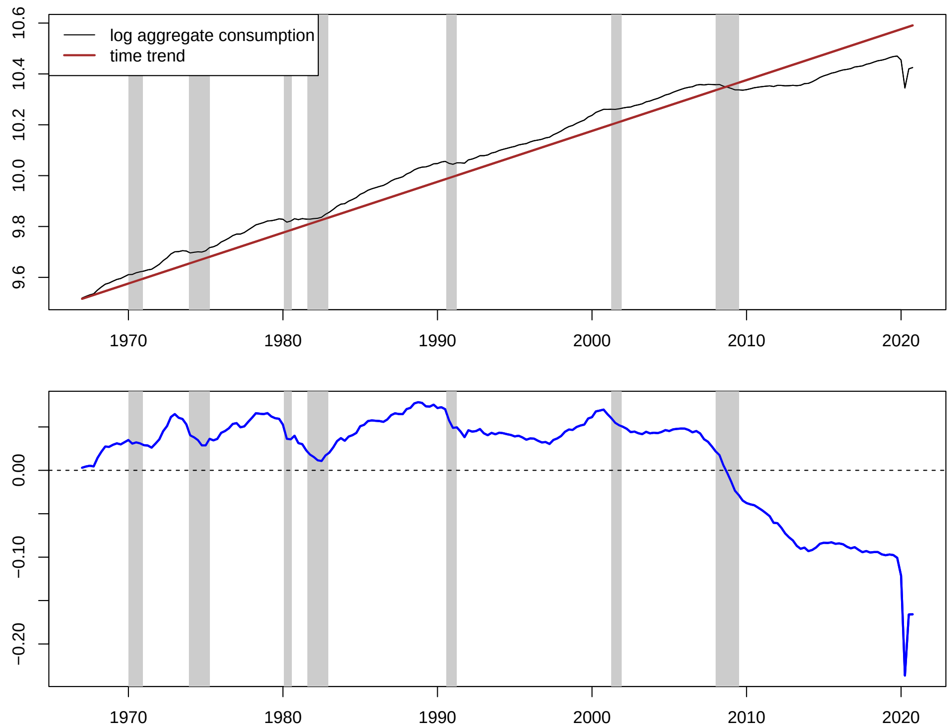


Figure C.1: Aggregate Consumption and its Time Trend. This figure shows log aggregate consumption level (expenditures on non-durables and services from NIPA, seasonally adjusted, in real per capita terms, 2012 chain-weighted dollars) and its time trend (top panel), and residuals from regressing consumption levels on the time trend plus a constant (bottom panel). The ADF test statistic for the residuals is 0.63 (p-value = 0.99). Shaded areas are NBER recessions. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

Table C.1: Predicting Market Returns: De-Trending Consumption with a Time Trend

This table reports OLS estimates and the R^2 from the predictive regression: $r_{m,t+i} = \gamma_0 + \gamma_1 cons\_time_t + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+i$ and $cons\_time$ is the residual from regressing aggregate consumption on a deterministic time trend. Panel A shows the in-sample results. Panel B shows the out-of-sample results. The R^2_{OOS} is computed as in Campbell and Thompson (2008); p-values for R^2_{OOS} are computed as in Clark and West (2007). I use an expanding window for estimating the predictive regressions; the in-sample period starts in 1967Q1 and ends in 1989Q4, 1999Q4, and 2009Q4. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994); for $i = \{2, 3, 4\}$, I use overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

	$r_{m,t+1}$	$r_{m,t+2}$	$r_{m,t+3}$	$r_{m,t+4}$
	(1)	(2)	(3)	(4)
$cons\_time_t$	-0.218^{*}	-0.397	-0.504	-0.574
	(0.122)	(0.285)	(0.416)	(0.536)
Constant	1.385^{**}	2.730^{**}	4.038^{**}	5.237^{**}
	(0.603)	(1.205)	(1.807)	(2.408)
Observations	215	214	213	212
R^2	0.012	0.018	0.018	0.019

the two series together.

Thus, to retain all the economic information present in consumption expenditures levels, instead of taking first-differences to address the issue of non-stationarity I propose to decompose consumption levels in a permanent (i.e., non-stationary) and a temporary (i.e., stationary) component. This idea borrows from the well-known permanent-transitory decomposition for stock prices proposed by Fama and French (1988). Specifically, log consumption levels are the sum of two components:

$$c_t = \tau_t + x_t \ ,$$

where τ_t represents the stochastic trend in aggregate consumption and x_t is the transitory, stationary component. The level of consumption is affected by changes in its stochastic trend, τ_t, i.e., the stochastic trend captures shocks that have a permanent effect on consumption. The stationary component, x_t, captures shocks that only have a temporary effect on aggregate consumption.

To provide a full representation of consumption level dynamics, one needs to assume some processes for the permanent and the temporary components. I specify τ_t as a random walk with drift and x_t as a stationary AR(1) process with homoskedastic errors. Thus, the system of equations specifying the dynamics of aggregate consumption levels is:

$$\begin{aligned} c_t &= \tau_t + x_t \\ \tau_t &= \mu + \tau_{t-1} + \eta_t \\ x_t &= \rho x_{t-1} + \epsilon_t \ , \end{aligned} \tag{C.1}$$

with $\mid \rho \mid < 1$, $\eta_t, \epsilon_t \sim i.i.d.\ \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, and $\mathbf{\Sigma} = diag(\sigma_\eta^2, \sigma_\epsilon^2)$.[29]

The stochastic trend τ_t can be treated as a latent variable, and various econometric frameworks have been proposed to filter trend and cycle from the series of interest (see, e.g., Beveridge and Nelson, 1981; Watson, 1986; Hamilton, 2018). Alternatively, in the spirit of, e.g., Stock and Watson (1988a), Stock and Watson (1988b), and King, Plosser, Stock, and

[29] I assume that the permanent and non-permanent shocks are uncorrelated. However, it is possible to model the correlation between the two shocks (see, e.g., Proietti, 2006). Future work could investigate the asset pricing implications of permanent shocks also affecting transitory innovations.

Watson (1991), I consider an empirical counterpart for τ_t which is observable. Specifically, inspired by the Friedman-Modigliani permanent income hypothesis (PIH), I assume that the stochastic trend in real GDP represents the permanent consumption component, i.e., $\tau_t \equiv ypot_t$, where $ypot$ is the log of potential output. Using this assumption together with the representation (5), allows to interpret the cyclical component of consumption as the consumption gap, i.e., $x_t \equiv cgap_t$.

Finally, the permanent-transitory decomposition for log consumption *levels* also implies a specification for log consumption *growth*. Rewriting (C.1) in terms of the consumption gap and taking the first difference, I obtain:

$$\begin{aligned} \Delta c_t &= \mu + \eta_t + \Delta cgap_t \\ &= \mu + (\rho - 1)cgap_{t-1} + \eta_t + \epsilon_t \end{aligned} \tag{C.2}$$

where Δc_t and $\Delta cgap_t$ are respectively log consumption growth and log growth in the cyclical component of consumption. The analytical formulation of $cgap_t$ as an AR(1) allows to write $\Delta cgap_t = (\rho - 1)cgap_{t-1} + \epsilon_t$. Thus, log consumption growth features two shocks: a permanent shock, η_t, and a transitory shock, ϵ_t. μ and $(\rho - 1)cgap_{t-1}$ represent respectively the unconditional and the conditional mean of log consumption growth. In my sample, $\mu = 0.48$ and $\rho = 0.85$. For the $\rho = 1$ and $\sigma_\epsilon = 0$, specification (C.2) collapses to the standard log consumption growth i.i.d. process.

D Additional Results

Table D.1: Predicting Market Returns Using the Second Lag

This table reports OLS estimates and the R^2 from the predictive regression: $r_{m,t+i} = \gamma_0 + \gamma_1 cgap_{t-1} + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+i$ and $cgap$ is the consumption gap computed as in equation (3). Panel A shows the in-sample results. Panel B shows the out-of-sample results. The R^2_{OOS} is computed as in Campbell and Thompson (2008); p-values for R^2_{OOS} are computed as in Clark and West (2007). I use an expanding window for estimating the predictive regressions; the in-sample period starts in 1967Q1 and ends in 1989Q4, 1999Q4, and 2009Q4. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994); for $i = \{2, 3, 4\}$, I use overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

Panel A: In-Sample

	$r_{m,t+1}$	$r_{m,t+2}$	$r_{m,t+3}$	$r_{m,t+4}$
	(1)	(2)	(3)	(4)
$cgap_{t-1}$	-0.886^{***}	-1.914^{***}	-2.759^{***}	-3.439^{***}
	(0.201)	(0.614)	(0.850)	(1.034)
Constant	1.431^{***}	2.920^{**}	4.333^{**}	5.665^{**}
	(0.547)	(1.205)	(1.807)	(2.408)
Observations	214	213	212	211
R^2	0.038	0.065	0.092	0.112
$\sigma[E_t(r_{m,t+i})]/E(r_{m,t+i})$	1.225	1.184	1.156	1.120

Panel B: Out-Of-Sample

R^2_{OOS}	From 1990	From 2000	From 2010
$r_{m,t+1}$	2.99^{**}	5.11^{***}	7.89^{**}
$r_{m,t+4}$	5.50^{***}	11.46^{***}	25.42^{***}
$r_{m,t+8}$	14.29^{***}	22.68^{***}	41.10^{***}
$r_{m,t+20}$	20.45^{***}	27.98^{***}	59.54^{***}
$r_{m,t+40}$	22.60^{***}	23.98^{***}	42.36^{***}

Table D.2: Annual Market Returns and Consumption

This table reports OLS estimates and the R^2 for the regression: $r_{m,t+4} = \alpha + \beta \Delta c_{t+4} + \epsilon_{t+4}$, where $r_{m,t+4}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+4$ and Δc is the log consumption growth. Columns (2) to (5) substitutes log consumption growth for its different components $\Delta c_{t+4} = \mu + (\rho - 1) cgap_t + \eta_{t+4} + \epsilon_{t+4}$, where η and ϵ are respectively the shock to the stochastic trend in real GDP and the shock to the *cgap*, and *cgap* is the consumption gap computed as in equation (3). Values in parenthesis are overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Annual returns of quarterly observations. The sample period is 1967Q1 to 2020Q4.

	$r_{m,t+4}$				
	(1)	(2)	(3)	(4)	(5)
Δc_{t+4}	1.685				
	(1.400)				
$cgap_t$		-4.200^{***}			-4.018^{***}
		(1.579)			(1.548)
η_{t+4}			-3.109		-1.183
			(2.808)		(2.586)
ϵ_{t+4}				1.844	2.096
				(1.722)	(1.931)
Constant	2.355	6.326^{***}	4.706^{**}	5.200^{**}	6.133^{***}
	(3.745)	(1.669)	(2.091)	(2.081)	(1.735)
Observations	213	213	213	213	213
Adjusted R^2	0.023	0.105	0.012	0.018	0.124

Table D.3: Short-Term Rate and Consumption

This table reports OLS estimates and the R^2 for the regression: $i_{t+1} = \alpha + \beta \Delta c_{t+1} + \epsilon_{t+1}$, where i_{t+1} is the log risk-free rate at $t+1$ and Δc is log consumption growth. Columns (2) to (5) substitutes log consumption growth for its different components $\Delta c_{t+1} = \mu + (\rho - 1)cgap_t + \eta_{t+1} + \epsilon_{t+1}$, where η and ϵ are respectively the shock to the stochastic trend in real GDP and the shock to the *cgap*, and *cgap* is the consumption gap computed as in equation (3). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

			i_{t+1}		
	(1)	(2)	(3)	(4)	(5)
Δc_{t+1}	0.052				
	(0.085)				
$cgap_t$		0.212***			0.123*
		(0.057)			(0.073)
η_{t+1}			2.312**		1.649*
			(1.043)		(0.945)
ϵ_{t+1}				0.045	0.003
				(0.126)	(0.105)
Constant	1.102*	1.118***	1.209***	1.124**	1.184***
	(0.566)	(0.323)	(0.309)	(0.509)	(0.233)
Observations	216	215	216	216	215
Adjusted R^2	0.000	0.242	0.276	0.000	0.333

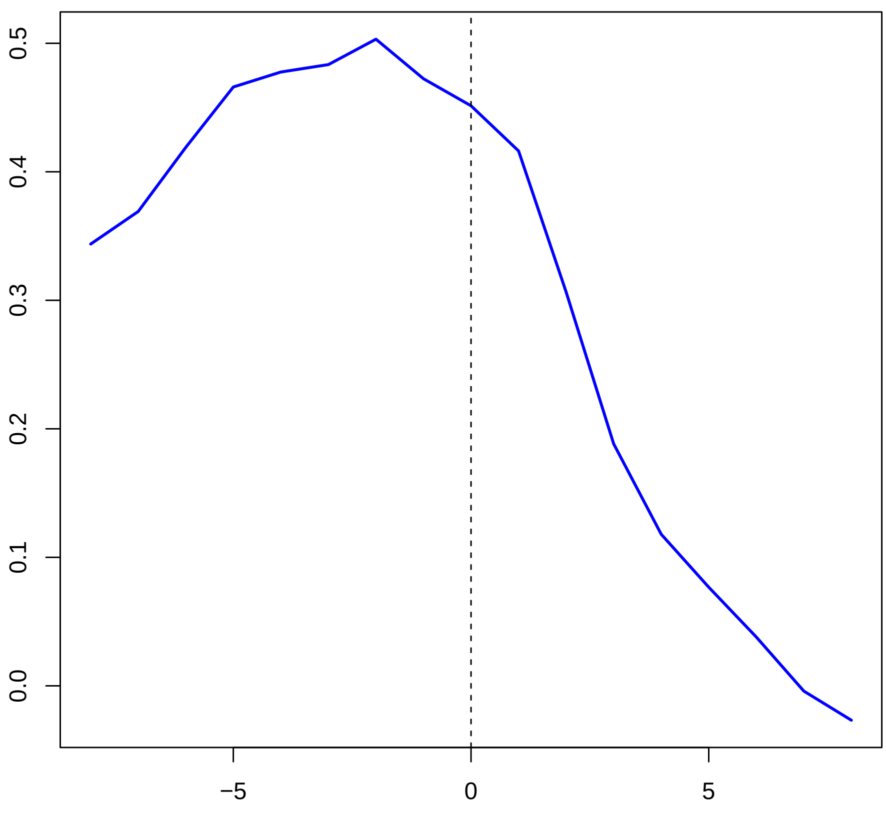


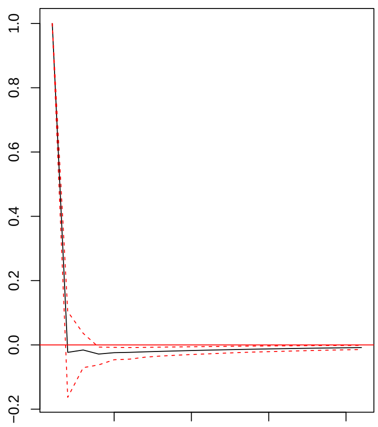
Figure D.1: Cross-Correlogram betweeen the SDF Variance and the Gilchrist-Zakrajsek (2012) spread. This figure shows the cross-correlation $\rho^{m,GZ} = corr(\sigma_t^{SDF}, GZ_{t-i})$, where σ_t^{SDF} is the variance of the SDF at time t and GZ_{t-i} is the credit spread index introduced by Gilchrist and Zakrajšek (2012) at time $t-i$. . Quarterly observations. The sample period is 1967Q1 to 2020Q4.

E VAR

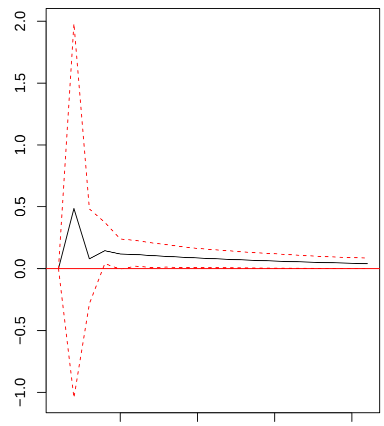
Table E.1: Stock Market Returns and Consumption: A VAR Approach

This table reports OLS estimates and the R^2 for the VAR specified in equation (12). The consumption gap *cgap* is computed as in equation (3). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

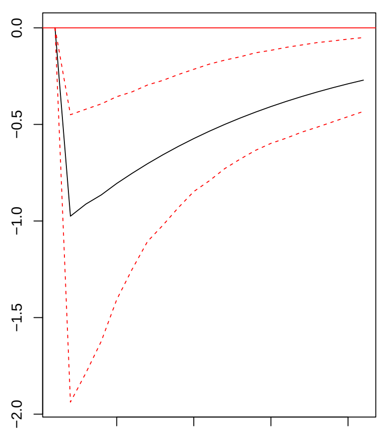
	$r_{m,t+1}$	Δc_{t+1}	$cgap_{t+1}$
	(1)	(2)	(3)
$r_{m,t}$	-0.023	0.036^{***}	0.035^{***}
	(0.069)	(0.008)	(0.007)
Δc_t	0.459	-0.186^{***}	-0.184^{***}
	(0.639)	(0.070)	(0.068)
$cgap_t$	-0.967^{***}	0.010	0.965^{***}
	(0.325)	(0.036)	(0.035)
const	1.242^{*}	0.451^{***}	-0.010
	(0.659)	(0.072)	(0.070)
Observations	215	215	215
R^2	0.041	0.126	0.803



(a) Market returns



(b) Consumption growth



(c) Consumption gap

Figure E.1: VAR Impulse Response Function. This figure shows the impulse response functions (IRF) for the VAR specified in equation (12). The plots display the estimated response of market return to one standard deviation impulses in market return (Panel (a)), log consumption growth (Panel (b)), and the consumption gap (Panel (c)). The consumption gap is computed as in equation (3). Quarterly observations. The sample period is 1967Q1 to 2020Q4.

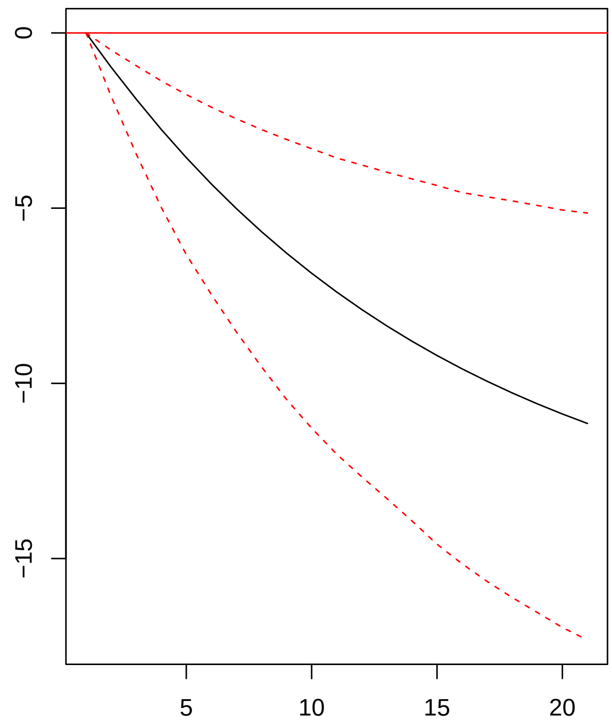


Figure E.2: Long-Run Effect of a Consumption Gap Shock on Market Returns. This figure shows the cumulative effect of a one standard deviation positive consumption gap shock on market returns implied by the VAR specified in equation (12). The consumption gap *cgap* is computed as in equation (3). Quarterly observations. The sample period is 1967Q1 to 2020Q4.

F Returns Predictability: Further Results

F.1 Other Market Predictors: Annual

Table F.1: The Consumption Gap and Other Market Predictors

This table reports OLS estimates and the R^2 from the predictive regression $r_{m,t+4} = \beta_0 + \beta_1 x_t + \gamma_1 cgap_t + \epsilon_{t+4}$, where $r_{m,t+4}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+4$, x is a market predictor, and $cgap$ is the consumption gap computed as in equation (3). In Columns (1) to (4) I study the case $\gamma_1 = 0$. Panel A shows results for the significant predictors among the sample of market predictors analized in Welch and Goyal (2008): long-term rate of returns (ltr), CPI inflation ($infl$), and the invetment-to-capital ratio (ik). Panel B shows results for cay (Lettau and Ludvigson, 2001a), surplus-consumption ratio (s) calculated as in Wachter (2006), long-run consumption (x) calculated as in Bansal, Kiku, and Yaron (2010), and cyclical consumption (cc) calculated as in Atanasov, Møller, and Priestley (2020). Values in parenthesis are overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

Panel A: Goyal-Welch (2006)

	$r_{m,t+4}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ltr_t	0.360			0.332	0.295			0.293
	(0.212)			(0.226)	(0.206)			(0.223)
$infl_t$		−2.210		−1.289		−0.491		−0.485
		(2.235)		(2.330)		(2.227)		(2.321)
ik_t			-16.237^{**}	-14.970^{*}			−8.572	−8.202
			(7.897)	(8.015)			(8.701)	(8.803)
$cgap_t$					-3.500^{***}	-3.487^{***}	-2.721^{**}	-2.616^{**}
					(1.026)	(1.031)	(1.112)	(1.101)
Observations	212	212	212	212	212	212	212	212
Adjusted R^2	0.010	0.010	0.083	0.090	0.123	0.113	0.131	0.132

Panel B: Consumption-Based Variables

	$r_{m,t+4}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
cay_t	1.392				1.906			
	(1.112)				(1.160)			
s_t		-2.537^{**}				0.104		
		(1.169)				(2.024)		
x_t			-0.834				0.258	
			(0.858)				(0.954)	
cc_t				-1.685^{**}				-1.001
				(0.747)				(0.912)
$cgap_t$					-3.899^{***}	-3.666^{**}	-3.761^{***}	-2.740^{**}
					(1.268)	(1.800)	(1.147)	(1.276)
Observations	212	212	212	212	212	212	212	212
Adjusted R^2	0.023	0.078	0.013	0.085	0.162	0.113	0.114	0.138

F.2 Alternative Market Predictors

Table F.2: Output Gap and Consumption Gap

This table reports OLS estimates and the R^2 from the predictive regression $r_{m,t+1} = \beta_0 + \beta_1 X_t + \epsilon_{t+1}$, where $r_{m,t+1}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+1$ and X is a matrix containing the output gap $ygap = \log(Y/YPOT)$, the consumption gap $cgap$ computed as in equation (3), $igap = \log(I/YPOT)$, and $ggap = \log(G/YPOT)$, where Y is real GDP, I is gross private domestic investment, G is government expenditures, and $YPOT$ is potential output. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

	$r_{m,t+1}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ygap	-0.748^{***}				-0.576^{**}	-0.767^{***}	-0.741^{***}
	(0.208)				(0.246)	(0.227)	(0.237)
cgap		-0.861^{***}			-0.393^{**}		
		(0.170)			(0.179)		
igap			-0.024			0.010	
			(0.036)			(0.033)	
ggap				-0.035			-0.003
				(0.028)			(0.029)
Constant	0.599	1.401^{**}	-3.197	-3.682	0.736	2.524	0.235
	(0.631)	(0.579)	(6.933)	(4.214)	(0.666)	(6.430)	(4.268)
Observations	215	215	215	215	215	215	215
Adjusted R^2	0.035	0.034	-0.003	0.000	0.041	0.031	0.030

Table F.3: Consumption Gap and Consumer Sentiment

This table reports OLS estimates and the R^2 from the predictive regression $r_{m,t+1} = \beta_0 + \beta_1 x_t + \gamma_1 cgap_t + \epsilon_{t+1}$, where $r_{m,t+1}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+1$, x is a consumer sentiment index, and $cgap$ is the consumption gap computed as in equation (3). CSI is the Consumer Sentiment Index from the University of Michigan Surveys of Consumers. CCI is the Consumer Confidence Index from the Conference Board. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

	$r_{m,t+1}$			
	(1)	(2)	(3)	(4)
CSI_t	-0.008		0.008	
	(0.022)		(0.022)	
CCI_t		-0.026^{**}		-0.012
		(0.012)		(0.012)
$cgap_t$			-0.949^{***}	-0.781^{***}
			(0.210)	(0.196)
Observations	216	214	215	214
Adjusted R^2	0.000	0.013	0.029	0.031

Table F.4: Consumption Gap and Macro-Principal Components

This table reports OLS estimates and the R^2 from the predictive regression $r_{m,t+1} = \beta_0 + \beta_1 PC_t + \gamma_1 cgap_t + \epsilon_{t+1}$, where $r_{m,t+1}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+1$ and $cgap$ is the consumption gap. PC represents one (or a combination) of the principal components extracted from a large cross-section of macroeconomic variables constructed in Ludvigson and Ng (2009). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

	$r_{m,t+1}$											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PC_t^1	0.939					0.714	0.414					0.414
	(0.717)					(0.955)	(0.784)					(0.927)
PC_t^2		-2.154^{***}				-1.835^{**}		-1.356				-0.936
		(0.762)				(0.920)		(0.861)				(0.968)
PC_t^3			-2.617			-1.814			-1.819			-1.646
			(2.046)			(2.228)			(2.042)			(2.321)
PC_t^4				-1.027		-1.412				-1.940^{**}		-1.817^{*}
				(0.759)		(1.015)				(0.813)		(0.935)
PC_t^5					-0.479	0.167					-1.354	-0.412
					(1.539)	(1.697)					(1.525)	(1.663)
$cgap_t$							-0.814^{***}	-0.626^{***}	-0.821^{***}	-1.070^{***}	-0.978^{***}	-0.738^{***}
							(0.201)	(0.233)	(0.214)	(0.214)	(0.240)	(0.224)
Observations	215	215	215	215	215	215	215	215	215	215	215	215
Adjusted R^2	0.008	0.029	0.009	0.000	-0.004	0.032	0.030	0.037	0.034	0.044	0.033	0.041

F.3 Consumption Gap and Common Risk Factors

Table F.5: Predicting Other Risk Factors

This table reports OLS estimates and the R^2 from the predictive regression: $f_{n,t+1} = \gamma_{n,0} + \gamma_{n,1} cgap_t + \epsilon_{n,t+1}$, where $f_{n,t+1}$ is the log factor return for factor n in the period $t : t+1$ and $cgap$ is the consumption gap computed as in equation (3). Panel A shows the results for the Fama and French (2015) five-factor model. Panel B shows the results for the Hou, Xue, and Zhang (2015); Hou, Mo, Xue, and Zhang (2018) five-factor model. Panel C shows the results for the Stambaugh and Yuan (2016) mispricing factor model. Panel D shows the results for the Daniel, Hirshleifer, and Sun (2020) behavioral factor model. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4; for Panel D, the sample starts in 1972Q3.

Panel A: Fama-French (2015)

	SMB_{t+1}	HML_{t+1}	RMW_{t+1}	CMA_{t+1}
$cgap_t$	-0.253	0.368^{**}	-0.007	0.169
	(0.203)	(0.184)	(0.134)	(0.112)
Observations	215	215	215	215
R^2	0.007	0.013	0.000	0.006

Panel B: Hou-Mo-Xue-Zhang (2018)

	ME_{t+1}	IA_{t+1}	ROE_{t+1}	EG_{t+1}
$cgap_t$	-0.155	0.264^{**}	0.402^{***}	0.517^{***}
	(0.185)	(0.121)	(0.134)	(0.105)
Observations	215	215	215	215
R^2	0.003	0.020	0.024	0.066

Panel C: Stambaugh-Yuan (2016)

	Mgmt_{t+1}	Perf_{t+1}
$cgap_t$	0.394^{**}	0.382^{**}
	(0.158)	(0.195)
Observations	215	215
R^2	0.016	0.013

Panel D: Daniel-Hirshleifer-Sun (2020)

	PEAD_{t+1}	FIN_{t+1}
$cgap_t$	0.294^{**}	0.592^{***}
	(0.123)	(0.170)
Observations	194	194
R^2	0.030	0.025

F.4 Value Portfolios

Table F.6: Idiosyncratic Component Predictability

This table reports OLS estimates and the R^2 from the predictive regression: $r^I_{n,t+i} = \gamma_{n,0} + \gamma_{n,1} cgap_t + \epsilon_{n,t+i}$, where $r^I_{n,t+i}$ is the idiosyncratic component of the log excess return for asset n at time $t+i$ and $cgap$ is the consumption gap computed as in equation (3). First, I run the regression $r_{n,t+i} = \alpha_{n,0} + \beta_{n,1} r_{m,t+i} + v_{n,t+i}$ where $r_{n,t+i}$ is the excess return for asset n and $r_{m,t+i}$ is the log market excess return. Then, I compute the idiosyncratic component as $r^I_{n,t+i} = \hat{v}_{n,t+i}$. Test assets are the 10 Portfolios sorted on Book-to-Market. Panel A shows the results for $r^I_{n,t+1}$. Panel B shows the results for $r^I_{n,t+4}$. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

Panel A: Quarterly Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$cgap_t$	−0.043	0.044	0.049	0.062	−0.028	0.138	0.282	0.185	0.178	0.204
	(0.133)	(0.095)	(0.096)	(0.135)	(0.140)	(0.173)	(0.231)	(0.205)	(0.205)	(0.270)
Constant	0.008	−0.008	−0.009	−0.012	0.005	−0.026	−0.053	−0.035	−0.034	−0.038
	(0.290)	(0.162)	(0.161)	(0.244)	(0.254)	(0.294)	(0.409)	(0.354)	(0.394)	(0.507)
Observations	216	216	216	216	216	216	216	216	216	216
R^2	0.000	0.001	0.001	0.001	0.000	0.004	0.011	0.004	0.003	0.002

Panel B: Annual Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$cgap_t$	0.664	0.295	0.036	−0.108	−0.320	−0.264	−0.407	−0.062	−0.133	−0.243
	(1.179)	(0.561)	(0.616)	(1.109)	(1.334)	(1.448)	(2.239)	(1.552)	(1.573)	(2.856)
Constant	−0.183	−0.081	−0.010	0.030	0.088	0.073	0.112	0.017	0.037	0.067
	(2.012)	(0.828)	(0.826)	(1.522)	(1.617)	(1.752)	(3.497)	(2.236)	(2.218)	(3.449)
Observations	213	213	213	213	213	213	213	213	213	213
R^2	0.012	0.006	0.000	0.000	0.003	0.002	0.002	0.000	0.000	0.000

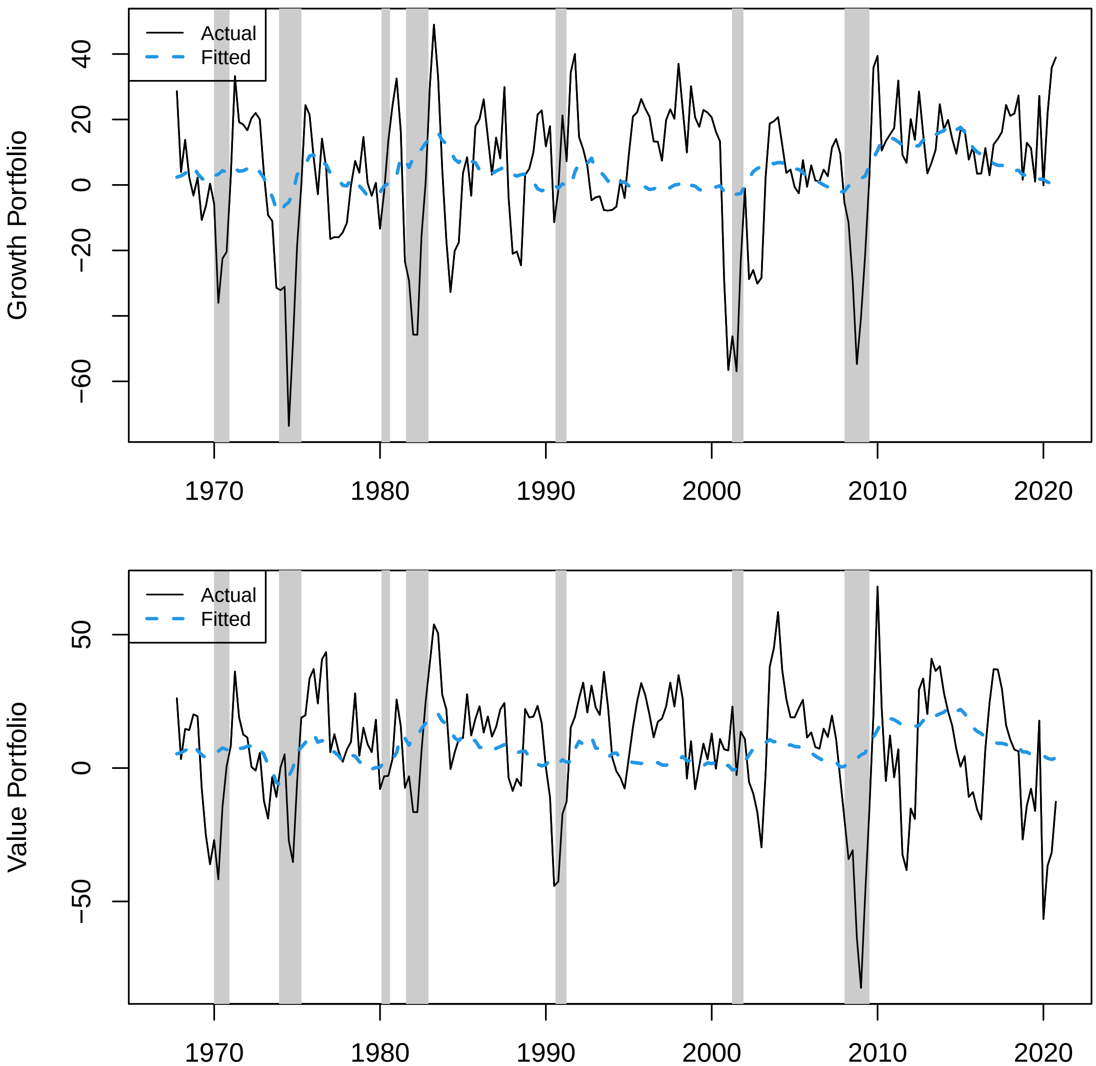


Figure F.1: Value Portfolios and the Consumption Gap. This figure shows log excess returns and fitted values for the predictive regression: $r_{i,t+4} = \gamma_{i,0} + \gamma_{i,1} cgap_t + \epsilon_{i,t+4}$, where $r_{i,t+4}$ is the log excess return for asset i at time $t+4$ and $cgap$ is the consumption gap computed as in equation (3). Test assets are the 10 Portfolios sorted on Book-to-Market. The average R^2 across the 10 Value Portfolios is 8.5%. Shaded areas are NBER recessions. Annual returns of quarterly observations. The sample period is 1967Q1 to 2020Q4.

F.5 Alternative Trends in Real GDP

Table F.7: Predicting Market Returns: Quadratic Trend

This table reports OLS estimates and the R^2 from the predictive regression: $r_{m,t+i} = \gamma_0 + \gamma_1 cgap_t^{quadr} + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+i$ and $cgap_t^{quadr}$ is the difference between aggregate consumption and the trend in real GDP computed as the fitted value of a regression of real GDP on a deterministic trend plus the quadratic trend component. Panel A shows the in-sample results. Panel B shows the out-of-sample results. The R^2_{OOS} is computed as in Campbell and Thompson (2008); p-values for R^2_{OOS} are computed as in Clark and West (2007). I use an expanding window for estimating both the predictive regressions and the trend in real GDP; the in-sample period starts in 1967Q1 and ends in 1989Q4, 1999Q4, and 2009Q4. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994); for $i = \{2, 3, 4\}$, I use overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q2.

Panel A: In-Sample

	$r_{m,t+1}$	$r_{m,t+2}$	$r_{m,t+3}$	$r_{m,t+4}$
	(1)	(2)	(3)	(4)
$cgap_t^{quadr}$	-0.780^{***}	-1.519^{***}	-2.350^{***}	-2.985^{***}
	(0.203)	(0.483)	(0.691)	(0.868)
Constant	1.398^{***}	2.787^{**}	4.256^{**}	5.554^{**}
	(0.524)	(1.205)	(1.807)	(2.408)
Observations	215	214	213	212
R^2	0.048	0.087	0.119	0.151

Panel B: Out-Of-Sample

R^2_{OOS}	From 1990	From 2000	From 2010
$r_{m,t+1}$	1.40^{*}	1.34^{*}	4.79^{**}
$r_{m,t+4}$	3.64^{***}	3.35^{***}	13.87^{***}
$r_{m,t+8}$	10.73^{***}	15.79^{***}	30.49^{***}
$r_{m,t+20}$	20.86^{***}	38.92^{***}	51.04^{***}
$r_{m,t+40}$	23.44^{***}	39.21^{***}	5.61^{**}

Table F.8: Predicting Market Returns: Laubach and Williams (2003)

This table reports OLS estimates and the R^2 from the predictive regression: $r_{m,t+i} = \gamma_0 + \gamma_1 cgap_t^{LW2003} + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+i$ and $cgap_t^{LW2003}$ is the difference between aggregate consumption and the potential output computed using a Kalman filter as in Laubach and Williams (2003). Panel A shows the in-sample results. Panel B shows the out-of-sample results. The R^2_{OOS} is computed as in Campbell and Thompson (2008); p-values for R^2_{OOS} are computed as in Clark and West (2007). I use an expanding window for estimating both the predictive regressions and the potential output; the in-sample period starts in 1967Q1 and ends in 1989Q4, 1999Q4, and 2009Q4. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994); for $i = \{2, 3, 4\}$, I use overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q2.

Panel A: In-Sample

	$r_{m,t+1}$	$r_{m,t+2}$	$r_{m,t+3}$	$r_{m,t+4}$
	(1)	(2)	(3)	(4)
$cgap_t^{LW2003}$	-0.673^{***}	-1.415^{**}	-2.623^{***}	-3.349^{***}
	(0.194)	(0.720)	(0.919)	(1.120)
Constant	1.313^{**}	2.684^{**}	4.150^{**}	5.463^{**}
	(0.595)	(1.208)	(1.812)	(2.414)
Observations	215	214	213	212
R^2	0.023	0.049	0.074	0.089

Panel B: Out-Of-Sample

R^2_{OOS}	From 1990	From 2000	From 2010
$r_{m,t+1}$	2.19^{*}	1.24	4.15
$r_{m,t+4}$	8.43^{***}	4.33^{***}	9.14^{***}
$r_{m,t+8}$	14.50^{***}	8.94^{***}	13.01^{***}
$r_{m,t+20}$	9.37^{***}	-1.46	17.64^{***}
$r_{m,t+40}$	-1.28	-8.78	-6.37

F.6 Alternative Permanent Component Measures

Table F.9: Predicting Market Returns: Real Disposable Income

This table reports OLS estimates and the R^2 from the predictive regression: $r_{m,t+i} = \gamma_0 + \gamma_1 cx_t + \epsilon_{t+i}$, where $r_{m,t+i}$ is the log market return in excess of the risk-free rate compounded in the period $t : t+i$ and cx is the difference between aggregate consumption and the trend in real disposable income computed utilizing the methodology proposed by Hamilton (2018). Panel A shows the in-sample results. Panel B shows the out-of-sample results. The R^2_{OOS} is computed as in Campbell and Thompson (2008); p-values for R^2_{OOS} are computed as in Clark and West (2007). I use an expanding window for estimating the predictive regressions; the in-sample period starts in 1967Q1 and ends in 1989Q4, 1999Q4, and 2009Q4. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994); for $i = \{2, 3, 4\}$, I use overlapping observations-corrected standard errors as in Hodrick (1992). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1967Q1 to 2020Q4.

Panel A: In-Sample

	$r_{m,t+1}$ (1)	$r_{m,t+2}$ (2)	$r_{m,t+3}$ (3)	$r_{m,t+4}$ (4)
cx_t	-0.377^{***}	-0.752^{***}	-1.096^{**}	-1.450^{***}
	(0.127)	(0.290)	(0.427)	(0.558)
Constant	1.386^{**}	2.737^{**}	4.083^{**}	5.309^{**}
	(0.563)	(1.205)	(1.807)	(2.408)
Observations	215	214	213	212
R^2	0.030	0.059	0.082	0.113

Panel B: Out-Of-Sample

R^2_{OOS}	From 1990	From 2000	From 2010
$r_{m,t+1}$	4.32^{**}	5.78^{**}	9.57^{***}
$r_{m,t+4}$	14.47^{***}	16.2^{***}	38.62^{***}
$r_{m,t+8}$	20.85^{***}	25.98^{***}	48.23^{***}
$r_{m,t+20}$	26.39^{***}	48.40^{***}	68.32^{***}
$r_{m,t+40}$	6.52^{***}	30.08^{***}	41.74^{***}

G International Evidence

Table G.1: Stock Market Predictability: Canada and UK

This table reports OLS estimates and the R^2 for the predictive regression: $r_{i,t+1} = \alpha + \beta cgap_t^i + \epsilon_{t+1}$, where $r_{i,t+1}$ is the log market return of country i compounded in the period $t : t+1$ and $cgap^i$ is the consumption gap of country i computed as in equation (3) using consumption and potential output data for country i. In Columns (2) and (4), I also control for the US consumption gap. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. Quarterly observations. The sample period is 1985Q1 to 2020Q4.

	$r_{CAN,t+1}$		$r_{GBR,t+1}$	
	(1)	(2)	(3)	(4)
$cgap_t^{CAN}$	-0.432^{**}	-0.414^{**}		
	(0.208)	(0.205)		
$cgap_t^{GBR}$			-0.436^{**}	-0.517^{**}
			(0.182)	(0.215)
$cgap_t$		-0.246		0.301
		(0.295)		(0.370)
Observations	140	140	140	140
R^2	0.011	0.014	0.029	0.031