The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXII, NO. 1 • FEBRUARY 2017

# Asset Pricing without Garbage

TIM A. KROENCKE*

**ABSTRACT**

This paper provides an explanation for why garbage implies a much lower relative risk aversion in the consumption-based asset pricing model than National Income and Product Accounts (NIPA) consumption expenditure: Unlike garbage, NIPA consumption is filtered to mitigate measurement error. I apply a simple model of the filtering process that allows one to undo the filtering inherent in NIPA consumption. "Unfiltered NIPA consumption" well explains the equity premium and is priced in the cross-section of stock returns. I discuss the likely properties of true consumption (i.e., without measurement error and filtering) and quantify implications for habit and long-run risk models.

In RESPONSE TO THE FAILURE of the classic consumption-based capital asset pricing model to explain the equity premium, the risk-free rate, and return predictability, recent literature has largely focused on developing alternative models of investor behavior. However, the deficiencies of the classic model might be attributed, at least in part, to a failure to measure consumption correctly. In a seminal contribution, Savov (2011) finds that using garbage to capture consumption, the classic model matches the equity premium and risk-free rate with a coefficient of relative risk aversion that is several times lower compared to using any other consumption measure based on National Income and Product Accounts (NIPA). A possible explanation for the relative success of garbage is that NIPA consumption fails to measure consumption properly (Savov (2011, p. 200)). Yet dozens of statisticians have tried to estimate NIPA consumption

*Tim A. Kroencke is at University of Basel. This paper is a revised version of the first chapter of my PhD thesis at the University of Mannheim. I am grateful for invaluable discussions and advice to my dissertation committee, Erik Theissen and Stefan Ruenzi. I would like to give special thanks to Kenneth Singleton (the Editor) and two anonymous referees for comments that substantially improved the paper. I have also benefited from suggestions by Yakov Amihud, Nicole Branger, Michael Burda, John Cochrane, Patrick Gruening, Lena Jaroszek, Paolo Maio, Frieder Mokinski, Stig Møller, Jonathan Parker, Marco Poltera, Felix Schindler, Andreas Schrimpf, Christian Walkshäusl, Paul Whelan, participants at conferences of the European Finance Association Lugano 2014, Financial Management Association Maastricht 2014, Swiss Finance Association (SGF) Zurich 2014, the German Finance Association (DGF) Wuppertal 2013, and seminar participants at Aarhus University, Humboldt University Berlin, University of Mannheim, University of Muenster, and the ZEW Mannheim Brownbag. I also thank Alexi Savov for sharing his data. I have no relevant or material financial interests that relate to the research described in this paper. An Internet Appendix and the unfiltered NIPA consumption series can be found on the *Journal of Finance* website.

DOI: 10.1111/jofi.12438

as precisely as possible. It is therefore highly unlikely that NIPA consumption is simply a poor proxy for consumption and that the story ends here.

This paper offers an alternative explanation for the dismal performance of NIPA consumption. Observable consumption is subject to measurement error, which is uncorrelated with stock market returns. From an asset pricing perspective, observable consumption growth would be eligible to measure the consumption risk of stock returns, that is, should produce unbiased estimates of consumption covariances. However, NIPA statisticians do not attempt to provide a consumption series to measure stock market consumption risk. Instead, they try to estimate the level of consumption as precisely as possible. As a result, they optimally filter observable consumption to generate their series of reported NIPA consumption. Concerning asset pricing, however, filtered consumption leads to disastrous results when the consumption risk of stocks is estimated. The simpler measure of garbage, in contrast, is not subject to filtering; garbage covariances are accordingly not "distorted." On top, filtering is intensified by the well-known bias stemming from time-aggregation: While reported consumption is an estimate of consumption flow during a specific period, the consumption-based asset pricing model relates asset returns to consumption at a specific point in time (Breeden, Gibbons, and Litzenberger (1989)).

I find that a simple model of the filter process involved in the estimation of reported NIPA consumption is surprisingly effective at explaining the characteristics of alternative consumption measures sampled at the annual frequency. Importantly, the filter model allows the effects of the filtering that are inherent in NIPA consumption to be stripped out using a simple backward recursion. I refer to this backward recursion as the "unfilter method" and call the measure resulting from this process "unfiltered NIPA consumption." I can also substantially mitigate time-aggregation bias by adding a correction factor to the unfilter recursion and applying a time-aggregation timing adjustment to stock returns, similar to Cochrane (1996).

More specifically, I first find that unfiltered NIPA consumption is able to explain the equity premium together with constant relative risk aversion (CRRA) preferences with a coefficient of relative risk aversion between 19 and 23 in the postwar period (1960–2014), which is close to the coefficient of 16 for garbage. A coefficient below 10 is well within two standard errors. A feature of unfiltered NIPA consumption is a sample covering a longer period. When including the prewar observations in the sample (1928–2014), I find a point estimate for the coefficient of relative risk aversion as low as 10 for unfiltered NIPA consumption.

Second, I find that the proposed filter model of NIPA consumption provides a unifying explanation for the success of three-year consumption (Parker and Julliard (2005)) and fourth-quarter to fourth-quarter consumption (Jagannathan and Wang (2007)). These alternative NIPA-based consumption measures have been shown in previous research to be priced in the cross-section of stock returns. In support of my hypothesis, this implies that there is valuable latent information in NIPA consumption. A simulation experiment using the filter model shows that fourth-quarter to fourth-quarter consumption is effective in removing time-aggregation bias, while the three-year consumption measure is effective in removing filtering. In this sense, both can be viewed as simple ad hoc unfilter rules.

Third, I find that unfiltered NIPA consumption can explain a substantial fraction of the average returns of decile portfolios sorted by size, book-to-market, and investment growth. Pricing errors are low and $R^2$s are above 60%. At the same time, cross-sectional slope coefficients imply a coefficient of relative risk aversion of around 30. Three-year and fourth-quarter consumption produce similar low pricing errors, but only with a substantially larger coefficient of relative risk aversion. However, these pricing results do not hold for profitability portfolios. On the contrary, highly profitable stocks have low consumption risk despite earning high average returns. Put differently, (unfiltered) consumption risk fails to explain the profitability premium, just as the market return factor fails to explain the value premium (Fama and French (1993)).

Fourth, I find that unfiltered NIPA consumption together with CRRA preferences cannot account for stock return predictability. In view of this result, I investigate the implications of unfiltered NIPA consumption for asset pricing models that are able to generate time-varying risk premia, particularly the external habit model developed by Campbell and Cochrane (1999) and the long-run risk model of Bansal and Yaron (2004).[1] Reported NIPA consumption has a substantially lower one-year standard deviation than true consumption and has a persistent component due to the filtering. In contrast, observed consumption growth has a higher standard deviation than true consumption and is mean-reverting due to measurement error. Hence, a researcher calibrating consumption growth to reported NIPA consumption, and as a result not accounting for the evidence provided by observed consumption (unfiltered NIPA consumption or garbage), may (i) underestimate the one-year consumption standard deviation and the contemporaneous stock market covariance, or (ii) overstate the predictability of future consumption growth and future stock market covariances.

I show that the filter model is in line with the postwar data when true consumption growth has no persistent component and its one-year standard deviation is equal to its long-run standard deviation of 2.5%.[2] Regarding the habit model, this implies that relative risk aversion is reduced and the fit of the model to the data can be considerably improved. In the long-run risk model, the filter model strengthens the role of short-run consumption risk at the expense of long-run consumption risk. At the same time, there remains the possibility of consumption volatility risk having a significant role. These results

[1] In principle, the proposed filter model of reported consumption is complementary to virtually any consumption-based asset pricing model, for example, models that include reduced-form variants of time-varying price of consumption risk, income from labor, complementary goods, housing, time-varying disasters, or small departures from rationality (Lettau and Ludvigson (2001), Santos and Veronesi (2006), Yogo (2006), Piazzesi, Schneider, and Tuzel (2007), Wachter (2013), Adam, Marcet, and Nicolini (2016)). Ludvigson (2013) provides an overview.

[2] I obtain a long-run standard deviation of 2.5% from calibrating the filter model to postwar consumption data. Interestingly, Dew-Becker (2016) reports a benchmark estimate of 2.5% as well.

are quite different from when the models are calibrated to reported NIPA consumption and illustrate the importance of a better consumption measure for assessing advanced asset pricing models.

*Background & Literature*: The unfiltering approach applied in this paper was first introduced by Quan and Quigley (1989) and Geltner (1989, 1993) in a different context. Research in real estate finds that appraisal-based real estate indices have a lower standard deviation, substantially larger autocorrelation, and lower stock market covariance than market value-based real estate indices. Before the work of Quan, Quigley, and Geltner, it was widely accepted that appraisal-based values of real estate are prone to measurement errors and thus are less reliable than market values. Quan, Quigley, and Geltner shifted focus to the appraisal process and documented possible reasons for—and the effects of—filtering on the time series of real estate indices. In this paper, I argue that "appraised" consumption is subject to similar issues. Indeed, related research supports this view.

Virtually every paper on consumption-based asset pricing mentions possible problems related to imperfections in the measurement of consumption (among the very earliest, Hansen and Singleton (1983)). However, surprisingly little work is devoted to analyzing the influence of the measurement methodology on the properties of reported NIPA consumption. In a rare exception, Wilcox (1992) and Bell and Wilcox (1993) investigate the properties of the *monthly* retail trade survey, which is the most important ingredient of NIPA nondurable consumption sampled quarterly (or monthly). They focus on the sampling error component of the measurement error that occurs because consumption estimates are derived from a selected sample and not from the total population. Bell and Wilcox (1993) study the implications for empirical studies and warn that the model parameters are not necessarily identifiable in empirical work due to the "imperfections in the data." I study annual consumption data, which are based on the considerably more comprehensive *annual* retail trade survey, which means that sampling error should be much less of a problem (Wilcox (1992) and Bell and Wilcox (1993)). However, nonsampling error (e.g., reporting error, misclassification, and measurement imperfections) is still present in annual consumption data and is central in this paper.

I argue that the problem for consumption-based asset pricing is not the measurement error per se, but rather what happens when the measurement error is recognized and treated in an optimal way by statisticians of the Bureau of Economic Analysis (BEA). Of course, garbage is far from free of measurement error as well. The internet appendix to Savov (2011) provides a detailed description of how the garbage time series is constructed. Several potential sources of measurement error are described. However, the garbage data are considerably more puristic and do not underlie such complex procedures, as is the case for NIPA consumption.

I deduce, in a similar vein as Daniel and Marshall (1997), that if BEA statisticians recognize measurement error in observed consumption data, they smooth their series of reported consumption and as a result stock market covariances appear to be too small. As a possible solution, Daniel and Marshall (1997) study consumption covariances measured over long horizons. Earlier work exploring long-horizon moments, for similar reasons, includes Singleton (1990) and Cochrane and Hansen (1992). I provide evidence that one needs to look at horizons of multiple years to make this approach work. However, the available observations in the postwar period are not sufficient to identify the coefficient of relative risk aversion, which is in line with the unsatisfying results in the earlier literature.

A complementary, measurement-related, explanation for the smoothness of reported NIPA consumption data is that households make infrequent and nonsynchronized consumption decisions (Lynch (1996), Gabaix and Laibson (2001)). However, frictions at the level of households are difficult to reconcile with the properties of observed consumption as measured by the garbage series (at least at the annual frequency). In related work, Kolev (2013) studies Gallup consumption data, a household survey that is available weekly since 2008. He finds large stock market covariances and low estimates of relative risk aversion, which supports the view that, from an asset pricing perspective, there are issues with official consumption data.

As mentioned before, two recent contributions propose the direct use of alternative NIPA-based consumption measures. Parker and Julliard (2005) argue that ultimate consumption risk can provide a "correct measure of risk under several extant explanations of slow consumption adjustment." As their first example for slow consumption adjustment, Parker and Julliard (2005, p. 186) mention "measurement error in consumption." Jagannathan and Wang (2007) argue that investors are more likely to simultaneously make consumption and investment decisions in the fourth-quarter of each year, when investors' tax year ends, and that end-of-year consumption is thus a better measure for asset pricing. However, using fourth-quarter to fourth-quarter consumption is a straightforward way to mitigate time-aggregation and to bring the data closer to point consumption growth as well (Breeden, Gibbons, and Litzenberger (1989), Savov (2011)).

More work addresses the time-aggregation bias. For example, Hall (1988), Grossman, Melino, and Shiller (1987), and Hansen and Singleton (1996) propose procedures to account for time-aggregation in the estimation method. Aït-Sahalia, Parker, and Yogo (2004) report estimates of the coefficient of relative risk aversion that are corrected for time-aggregation by dividing estimates by two. Calibrations of advanced asset pricing models, for example, Campbell and Cochrane (1999) and Bansal and Yaron (2004), carefully account for the time-aggregation bias. Cochrane (1996) relies on time-aggregated stock returns to mitigate the time-aggregation bias in empirical tests. I adopt Cochrane's treatment and show that it is highly effective in combination with unfiltered NIPA consumption.

Breeden, Gibbons, and Litzenberger (1989) discuss the possible influence of both interpolation and time-aggregation in consumption data. They are particularly concerned that interpolation is "exacerbated" by time-aggregation, and conclude that "it is difficult to disentangle the two effects" (Breeden, Gibbons, and Litzenberger (1989, p. 243)). Indeed, to my knowledge, no prior research tries to disentangle the two effects. This paper provides a first attempt.

Finally, I find evidence that there is valuable information in macroeconomic data for the purpose of asset pricing, but this information is obscured due to the filter process. My findings provide an explanation for why the growing literature identifies macroeconomic factors as spurious factors in cross-sectional asset pricing tests (e.g., Gospodinov, Kan, and Robotti (2014), Bryzgalova (2014), Burnside (2016)), which might be useful to further improve tests.

Section I presents the filter model and shows how I recover unfiltered NIPA consumption from reported NIPA consumption. Section II discusses the empirical properties of the alternative consumption measures and reports estimates of the coefficient of relative risk aversion. I provide evidence that the unfiltering approach is robust to several considerations. Section III presents results for linear cross-sectional asset pricing tests. I discuss the implications for advanced asset pricing models in Section IV. Finally, Section V concludes.

## I. Filtering and Time-Aggregation in Consumption Data

This section introduces the filter model of NIPA consumption data. The true process undertaken by the BEA to estimate NIPA consumption is extremely comprehensive and includes collecting a huge amount of data from different sources, from differently designed surveys, sampled at different frequencies, dealing with seasonality issues, and more.[3] Many of the steps involved are not fully known, and the raw data themselves are not available. For that reason, I reduce the unavoidably comprehensive estimation process undertaken by the BEA to a more tractable problem and assume that an optimal filter is applied to the data. After introducing the filter model, I show how the filter can be reversed to generate unfiltered NIPA consumption.

Furthermore, I discuss how unfiltered NIPA consumption can be modified to account for the time-aggregation bias inherent in consumption data. A simulation experiment investigates how well the unfilter method works in a controlled environment. The simulation also serves as a benchmark to calibrate the unfilter method to empirical data.

### *A. Setting*

I assume a standard endowment economy in which the state of "true" log consumption, $c_t = \log(C_t)$, follows the process

$$c_t = c_{t-1} + \mu_{c,t} + \sigma_{\eta,t}\eta_t, \tag{1}$$

where $\mu_{c,t}$ is a time-varying drift, and the mean-zero economic shocks, $\eta_t \sim N(0,\ 1)$, are levered by the time-varying volatility parameter $\sigma_{\eta,t}$. This consumption process nests the random walk model of consumption (e.g., Hall

[3] A discussion of how the BEA actually filters its raw data is given in the Internet Appendix. The Internet Appendix may be found in the online version of this article.

(1978)), the long-run risk models studied by Bansal and Yaron (2004), Bollerslev, Tauchen, and Zhou (2009), and Bansal and Shaliastovich (2012), as well as habit-based models such as those in Campbell and Cochrane (1999), Wachter (2006), and Verdelhan (2010).

I explicitly account for the fact that the true state of consumption $c_t$ is not observable to NIPA statisticians. That is, the observed time series of consumption $y_t$ satisfies

$$y_t = c_t + \sigma_\xi \xi_t, \tag{2}$$

where $\xi_t \sim N(0,\ 1)$ represents measurement error that is uncorrelated with stock market log returns $r_{m,t}$.[4] Observable consumption could be, for example, a consumption indicator based on a survey of sales of goods or the amount of garbage that is produced by consuming goods. Mismeasurement of consumption does not bias estimates of stock market risk, as measured by the covariance of consumption growth ($\triangle c_t$) and stock returns ($r_{m,t}$):

$$\mathrm{cov}(\triangle c_t, r_{m,t}) = \mathrm{cov}(\triangle y_t, r_{m,t}). \tag{3}$$

Thus, for the purpose of asset pricing, $\triangle y_t$ is well behaved to measure (contemporaneous) consumption risk of stock returns. Unfortunately, NIPA statisticians are not researchers in asset pricing and this has consequences for what they perceive as an optimal estimate of consumed goods. The objective of NIPA statisticians is to measure the level of consumption $c_t$ as precisely as possible and not $\mathrm{cov}(\triangle c_t, r_{m,t})$.

### *B. A Model of Filtered Consumption Data*

I define NIPA consumption $\hat{c}_t = E_t(c_t)$ as a conditional time $t$ estimate of true but unobservable consumption $c_t$ when a new data point of the observed variable $y_t$ becomes available. The optimal solution to this problem depends on the assumptions that the NIPA statisticians impose on $\mu_{c,t}$ and $\sigma_{\eta,t}$.

Throughout the paper, I assume that NIPA statisticians treat the drift term $\mu_{c,t}$ as a constant, that is, there is no persistent or predictable component in true consumption growth. This assumption does not require that true consumption necessarily has a constant drift. It only says that NIPA statisticians choose a constant drift to filter observed consumption. Interestingly, the filter model implies that, in a world with a constant drift in true consumption, a researcher will find a persistent component due to filtering in reported NIPA consumption and mean-reversion due to measurement error in observed consumption. This pattern is exactly what I find in the empirical data and is the motivation behind

[4] This modeling of measurement error in consumption is used, for example, by Daniel and Marshall (1997) and Parker (2001). In the Internet Appendix, I argue that this choice is reasonable for consumption data sampled at the annual frequency.

the constant drift assumption.[5] For ease of notation, I set the constant drift to zero without changing the implications of the model.

With respect to $\sigma_{\eta,t}$, there is strong evidence in favor of heteroskedasticity in empirical consumption data.[6] For expositional purposes, for the moment I proceed as if NIPA statisticians assume that consumption volatility is constant. Later, I generalize the filter model and allow NIPA statisticians to take time-varying consumption volatility into account.

### *B.1. Constant Consumption Volatility*

The optimal estimate of $\hat{c}_t$ can be found using the Kalman filter model:[7]

$$\hat{c}_t = \hat{c}_{t-1} + K_t(y_t - \hat{c}_{t-1}), \tag{4}$$

$$K_t = \frac{P_t}{P_t + \sigma_\xi^2}, \tag{5}$$

$$P_t = P_{t-1}(1 - K_{t-1}) + \sigma_{\eta,t}^2, \tag{6}$$

where $K_t$ is the filter parameter, or Kalman gain, and $P_t$ is the conditional variance of $c_t$. Because $\sigma_{\eta,t}^2 = \bar{\sigma}_\eta^2$ does not change over time, the conditional variance converges to $\bar{P} = \mathit{constant} + \bar{\sigma}_\eta^2$ after a few recursions.[8] As a result, one can verify that the optimal filter parameter converges to a constant steady state $\bar{K} = \bar{P}/(\bar{P} + \sigma_\xi^2)$ as well.

The filter model provides a simple economic rationale. The best estimate of NIPA consumption is the NIPA consumption estimate of the past period plus a weighted surprise $(y_t - \hat{c}_{t-1})$. If the measurement error is large relative to economic shocks, $\bar{K}$ will be small and the expectation of the state of consumption should only be slightly updated. In contrast, if there is no measurement error, it is optimal to adjust the expectation on the state of consumption one-for-one according to observable consumption.

To see how $\bar{K}$ influences asset pricing results, equation (4) might be rearranged and expressed in terms of log growth rates:

$$\Delta\hat{c}_t \cong \bar{K}\Delta y_t + (1 - \bar{K})\Delta\hat{c}_{t-1}. \tag{7}$$

[5] In Section IV, I discuss the implications of this finding for advanced asset pricing models. There is an ongoing debate about whether there is a persistent component in the consumption process or not. Some argue that the process is (close) to a random walk, while others are in favor of a persistent component (or even a mean-reverting component in the very long run). See, for example, Hansen, Heaton, and Li (2008), Marakani (2009), Constantinides and Gosh (2011), Beeler and Campbell (2012), Bansal, Kiku, and Yaron (2012), and Dew-Becker (2016).

[6] See, for example, Kandel and Stambaugh (1990), Bansal, Khatchatrian, and Yaron (2005), Lettau, Ludvigson, and Wachter (2008), Boguth and Kuehn (2013), and Bansal et al. (2014).

[7] See, for example, Durbin and Koopman (2012).

[8] More precisely, the equation is $\bar{P} = \bar{P}(1 - \bar{P}/(\bar{P} + \sigma_\xi^2)) + \sigma_\eta^2$, which has only one sensible solution, $\bar{P} = \sigma_\xi^2(x + (x^2 + 4x)^{1/2})/2$, with $x = \sigma_\eta^2/\sigma_\xi^2$.

The stock market covariance based on $\Delta\hat{c}_t$ will only be a fraction $\bar{K}$ of the true stock market covariance. High measurement error implies a low optimal $\bar{K}$ and a low stock market covariance for $\Delta\hat{c}_t$.

### *B.2. Time-Varying Consumption Volatility*

If consumption volatility $\sigma_{\eta,t}$ changes over time, the filter no longer converges to a constant value $\bar{K}$. For example, observed consumption becomes more informative in high consumption volatility states, which means that, in such states, consumption expectations should be updated more than otherwise ($K_t > \bar{K}$).[9]

I assume that consumption volatility can be characterized by a generalized autoregressive conditional heteroskedasticity (GARCH) process.[10] Harvey, Ruiz, and Sentana (1992) examine in detail how GARCH shocks can be incorporated into the Kalman filter model. Following Harvey, Ruiz, and Sentana, I model the economic disturbances according to

$$\sigma^2_{\eta,t} = a_0 + a_1\eta^{*2}_{t-1} + a_2\sigma^2_{\eta,t-1}, \tag{8}$$

$$\eta^*_t = \eta_t\sigma_{\eta,t}, \ \ \eta_t \sim N(0,1). \tag{9}$$

In these equations, the past values of economic disturbances $\eta^{*2}_{t-1}$ are not observable. As proposed by Harvey, Ruiz, and Sentana (1992), I approximate $\eta^{*2}_{t-1}$ by its conditional expectation $E_{t-1}(\eta^{*2}_{t-1})$ to get operable filter equations. Details on how $E_{t-1}(\eta^{*2}_{t-1})$ is computed are provided in the Appendix.

Practicability is the main motivation behind the GARCH specification of consumption volatility. An observation-driven volatility model makes it convenient and transparent to unfilter reported NIPA consumption. In fact, the proposed specification allows unfiltered NIPA consumption to be computed from reported NIPA consumption in a simple way using any spreadsheet application. As before, the choice of volatility model does not imply that true consumption necessarily follows such a process. It only says that NIPA statisticians choose this representation to characterize consumption volatility when filtering observed consumption.[11]

In the robustness section, I consider a more data-driven approach to model $\sigma_{\eta,t}$ that is similar to recent attempts to model consumption volatility (e.g.,

[9] Accordingly, reported NIPA consumption should be more informative in times with very high consumption volatility. This is in line with Grossman and Shiller (1980, 1981), Campbell (2003), Engsted and Møller (2011, 2015), among others, who find that the stock market covariance of reported NIPA consumption is substantially higher when the more volatile prewar observations are included compared to sample periods that are restricted to postwar data.

[10] Previous work by, for example, Bansal, Khatchatrian, and Yaron (2005), Duffee (2005), and Tédongap (2014) apply a GARCH model to consumption volatility.

[11] Broto and Ruiz (2006) provide extensions for GARCH models with asymmetric conditional variances. In general, the unobserved component models are not limited to the use of a GARCH specification. See, for example, Durbin and Koopman (2012) or Kim and Nelson (1999) for an overview of alternative specifications.

Bansal et al. (2014)). I find that unfiltered NIPA consumption *growth* is not very sensitive to alternative consumption volatility specifications. Even a specification that pretends constant volatility delivers fairly similar results, at least in the postwar sample period.[12]

### C. Unfiltered NIPA Consumption

Unfiltered NIPA consumption, $\hat{y}_t$, is obtained by solving equation (4) for $y_t$:[13]

$$\hat{y}_t = \frac{\hat{c}_t - (1 - K_t)\hat{c}_{t-1}}{K_t}, \tag{10}$$

where $\hat{c}_t$ corresponds to reported NIPA consumption. To initialize the unfiltering procedure, I use the steady-state filter $\bar{K}$ setting $\sigma^2_{\eta,t=1} = a_0/(1 - a_1 - a_2)$.[14] In terms of the model, garbage can be thought of as observable consumption $y_t$ without filtering. Unfiltered NIPA consumption growth may be viewed as observed consumption growth plus an error term,

$$\triangle \hat{y}_t = \triangle y_t + e_t, \tag{11}$$

where $e_t$ reflects the fact that the filter model is a simplification of the true and more comprehensive estimation process behind reported NIPA consumption. If my filter model provides a good approximation of reality, $e_t$ should be small and pure noise. It follows that unfiltered NIPA consumption $(\triangle\hat{y}_t)$ is as informative as garbage $(\triangle y_t)$ when measuring the consumption risk of stock market returns. Whether this is the case is ultimately an empirical question, which I investigate in Section II.

### D. Accounting for the Time-Aggregation Bias

NIPA consumption is measured as the flow of consumption over a specific period, not from point-to-point in time. Working (1960) shows that the variance of the growth rate of a time-aggregated random walk is approximately two-thirds that of the true point-to-point growth rate. Furthermore, the first-order autocorrelation is shifted from zero to one-quarter, and, Taio (1972) shows that the covariance with a second variable (e.g., the stock market) is reduced by half. The latter implies that asset pricing tests using time-aggregated consumption and stock returns measured from point-to-point in time will suffer from time-aggregation bias (e.g., Breeden, Gibbons, and Litzenberger (1989)).

I tackle time-aggregation bias from two sides. First, I add a simple time-aggregation bias adjustment to unfilter NIPA consumption. This adjustment

[12] The observed degree of time-variation in postwar consumption volatility simply does not suffice for large variation in $K_t$.

[13] This backward induction was proposed by Geltner (1993) to estimate real estate market indices from appraisal-based real estate indices.

[14] To apply the unfilter to the data, I de-mean simple computed consumption growth rates, accumulate to levels, take logs, then unfilter, and convert back in reversed order.

corrects the standard deviation of consumption growth. In addition, I follow Cochrane (1996) and time-aggregate stock returns so that their timing is naturally more aligned with time-aggregated consumption data.

### *D.1. Adjusted Unfilter Method*

The most obvious way to address time-aggregation bias is to use December-to-December (or fourth-quarter to fourth-quarter) consumption growth to reduce time-aggregation in the first place (Breeden, Gibbons, and Litzenberger (1989), Savov (2011)). Unfortunately, this option is not available for consumption data in the prewar sample, where only annual consumption data are available. Alternatively, Hall (1988) suggests accounting for time-aggregation using a simple autoregressive representation as an approximation.[15] In this spirit, equation (10) can be applied to adjust for time-aggregation as follows:

$$\Delta c_t^{adTA} = \left[\Delta c_t^{TA} - (1-\alpha)\Delta c_{t-1}^{TA}\right]/\alpha, \tag{12}$$

where $\Delta c_t^{adTA}$ is the time-aggregation bias-adjusted estimate of consumption and $\Delta c_t^{TA}$ is the time-aggregated measure. The parameter $\alpha = 0.80$ ensures that $\Delta c_t^{adTA}$ has the same standard deviation as consumption measured from point-to-point in time.[16] Following this reasoning, an unfilter rule that accounts for both time-aggregation and filtering can be set up as

$$\hat{y}_t = \left[\hat{c}_t - (1-\Omega_t)\Delta\hat{c}_{t-1}\right]/\Omega_t, \tag{13}$$

where $\Omega_t = K_t \times \alpha$. Put simply, this adjustment brings the standard deviation of time-aggregated consumption back to its point-to-point in time value. The higher standard deviation helps improve the covariance. Still, the timing of consumption data and stock returns is misaligned, which leaves room for further improvement.

### *D.2. Adjusted Timing of Stock Returns*

In addition to leveraging the filter parameter, I apply a modified version of the ad hoc correction of stock returns as in Cochrane (1996). In a first step, I

[15] More specifically, to apply the instrumental variables estimator of Hayashi and Sims (1983), Hall (1988) uses a second-order forward autoregressive (un)filter to remove autocorrelation in the time-aggregated variables. Grossman, Melino, and Shiller (1987) and Hansen and Singleton (1996) propose estimation strategies to take time-aggregation directly into account.

[16] The value $\alpha = 0.80$ solves:

$$\mathrm{Var}\left(\Delta c_t^{TA}\right) = \frac{\alpha}{2-\alpha}\mathrm{Var}\left(\Delta c_t^{adTA}\right) = \frac{2}{3}\mathrm{Var}\left(\Delta c_t^{adTA}\right),$$

where the first equality is implied by equation (12) and the second equality uses the results in Working (1960) and Breeden, Gibbons, and Litzenberger (1989).

compute time-aggregated stock returns by summing end-of-month levels $\Pi_{m,t+1}$:

$$\triangle R_{t+1}^{TA} = \sum_{m=1}^{12} \Pi_{m,t+1} / \sum_{m=1}^{12} \Pi_{m,t} - 1, \tag{14}$$

where $\triangle R_{t+1}^{TA}$ is the time-aggregated stock return sampled at the annual frequency.[17] Naturally, this timing of stock returns is now more aligned with the timing of consumption data. However, the standard deviation and other return properties are now subject to time-aggregation bias. To match time-aggregated stock returns with unfiltered NIPA consumption adjusted for time-aggregation, I apply the adjustment

$$\triangle R_{t+1}^{adTA} = \frac{\triangle R_{t+1}^{TA} - E\left(\triangle R_{t+1}^{TA}\right)}{\sigma\left(\triangle R_{t+1}^{TA}\right)} \times \sigma\left(\triangle R_{t+1}^{D-D}\right) + E\left(\triangle R_{t+1}^{D-D}\right), \tag{15}$$

where $\triangle R_{t+1}^{D-D} = \Pi_{12,t+1}/\Pi_{12,t} - 1$ are the usual point-to-point stock returns measured from December-to-December. For simplicity, I call $\triangle R_{t+1}^{adTA}$ the "time-aggregated stock return." However, the first and second moments of stock returns remain unchanged while the timing of the returns is now comparable to the timing of consumption data. The following simulation experiment shows that the proposed treatment of the time-aggregation bias works reasonably well.

## *E. Simulation Experiment*

Intuitively, filtering and time-aggregation destroy the covariance between consumption growth and stock returns and drive a wedge between observed consumption and reported NIPA consumption. To quantify these effects, I simulate a model economy in which consumption-based asset pricing works.

### *E.1. The Simulated Asset Pricing Economy*

The unconditional annual standard deviation of true consumption is set to $\bar{\sigma}_\eta = 2.5\%$. This value is obtained by calibrating the simulated long-run standard deviations of alternative consumption measures to their postwar sample counterparts. Consumption growth has an unobservable stochastic volatility component but no predictable component.[18]

In the data, aggregate NIPA consumption volatility is low, which indicates low aggregate measurement error. However, individual components of

[17] Cochrane (1996) uses time-aggregated quarterly stock returns to match them with quarterly investment returns (which are based on time-aggregated macroeconomic variables). He gives credit for the use of time-aggregated stock returns to Cam Harvey.

[18] I set the monthly persistence in consumption volatility to 0.99 and the volatility of consumption volatility at 0.0003%. These values are in the range of those reported in Bansal and Yaron (2004), and Bansal, Kiku, and Yaron (2012).

consumption (e.g., food, clothing, health care) are considerably more volatile, which indicates high measurement error at this level of the data.[19] NIPA consumption is constructed bottom-up, that is, each individual component of consumption is estimated separately and is then summed up to aggregate consumption. As a result, the optimal filter for the individual components will be carried over to aggregate consumption. To account for the fact that the measurement error is not fully diversified in observed consumption, I split annual true consumption into 10 components and add an individual measurement error term with a standard deviation of $\sigma_{\xi} = 2.8\%$ to each series.[20]

Observable consumption (i.e., garbage) is just the simple sum of all 10 components before any filtering. NIPA statisticians then apply the filter model to obtain the aggregate "reported" NIPA consumption series.[21] As a reference point on the degree of filtering, the optimal filter would be $\bar{K} = 0.58$ in the absence of volatility shocks. Finally, unfiltered NIPA consumption is obtained by applying the unfilter method as in equation (13) to aggregate reported NIPA consumption.

I get returns of the market portfolio in excess of the risk-free rate (equity premium) by imposing that the pricing equation of the consumption-based model with CRRA preferences holds.[22] I set the true coefficient of relative risk aversion to 15. The correlation with the stock market is 65% and matches the empirical properties of garbage (i.e., observable consumption). That is, I treat garbage as neither filtered nor time-aggregated.[23] The resulting annual equity premium of this economy is 4.5%, with a standard deviation of roughly 20%. I generate 10,000 time series of the model with monthly observations of consumption growth and the equity premium. I convert these monthly observations so that I have 55 annual time-aggregated observations. Further details on the simulations can be found in the Internet Appendix.

[19] A similar pattern can be observed for aggregate garbage and the components of garbage (see Savov (2011)).

[20] It is straightforward to incorporate individual consumption components into the filter model without changing the main equations.

[21] For simplicity, I impose that all individual components of consumption are filtered by the same parameter at any given point in time. I set the GARCH parameters to 0.00004375, 0.08, and 0.85, such that the unconditional consumption volatility is 2.5%. I find that these parameter values do a reasonably good job in approximating the assumed time-varying volatility when a very long time series of 1,000 simulated years is available.

[22] I use CRRA preferences here to focus on contemporaneous consumption covariance risk. I investigate external habits and recursive preferences in Section IV.

[23] The assumption that garbage is not filtered is in line with the empirical properties of garbage and is based on the estimation methodology of garbage discussed in the internet appendix to Savov (2011). The assumption that garbage is not subject to time-aggregation is more difficult to make. However, the empirical properties of garbage hardly point to a time-aggregation effect (this observation is also mentioned in the internet appendix to Savov (2011)). Moreover, Savov (2011) uses the "beginning-of-period" timing convention for garbage. This is the same as giving garbage a "lag" of one period—effectively an adjustment for the lag between consumption and disposal—and therefore is not included in the simulation.

### E.2. Results

Table I summarizes the properties of the alternative consumption measures in the simulated asset pricing economy. By construction, true consumption has a one-period standard deviation of 2.5%. The long-run standard deviation measured by six-year growth rates (scaled by $1/\sqrt{6}$) is essentially unchanged.[24] Observed consumption is more volatile when measured over one year (2.8%) but not when measured over six years (2.5%). There is mild, short-lived mean reversion in observed consumption due to measurement error. Both true and observed consumption have a stock market covariance of 0.29%. Measurement error does not affect covariances after averaging over many simulations. Estimating relative risk aversion using true or observed consumption gives results close to the true coefficient of risk aversion of 15 (bottom row).

Reported NIPA consumption is filtered and time-aggregated. As a result, its one-period standard deviation is just 1.5%. A high first- and second-order autocorrelation indicate a sizable persistent component. Over the six-year horizon, the standard deviation is 2.1%. In contrast to measurement error, filtering and time-aggregation are fatal for asset pricing. The contemporaneous stock return covariance is extremely low at 0.08%. As a result, the estimate of relative risk aversion is as large as 55. Moreover, the distribution of the relative risk aversion estimates gets extremely wide, as shown in Figure 1. A relative risk aversion far above 100 is not unusual for reported NIPA consumption. Measuring covariances with true consumption is already difficult, and it becomes a mess when filtering is added to the data.

The simulation also reproduces other well-known properties of the empirical data. For example, leading consumption by one period delivers larger covariances than for contemporaneous consumption. Thus, using the "beginning-of-period" timing convention put forth by Campbell (2003) is helpful. There are also positive covariances when giving reported NIPA consumption a lead of two or three years, as Parker and Julliard (2005; P-J) find. P-J and Q4-Q4 consumption provide lower estimates of relative risk aversion than reported NIPA consumption.

Unfiltered NIPA consumption in combination with time-aggregated stock returns perform fairly well in recovering the properties of observed consumption. The one-year standard deviation is 2.8%, hand in hand with the stock market covariance of 0.28%. This is only marginally smaller than the true covariance. The estimate of relative risk aversion is 16.6. The distribution of estimates across simulations is tight, similar to the distribution obtained from observed consumption (Figure 1). Admittedly, the unfilter approach does not recover observed consumption perfectly. For example, unfiltered NIPA consumption spreads the (small) negative first-order autocorrelation of observed consumption over a horizon of up to four years.

The Internet Appendix provides simulation results when the data are only filtered or only time-aggregated, that is, when the two effects are isolated.

[24] It is slightly lower because of the well-known downward bias for long-horizon standard deviations in finite samples (e.g., Campbell, Lo, and MacKinlay (1997)).

**Table I**

## Alternative Consumption Measures in a Simulated Asset Pricing Economy

The statistics displayed are medians of 10,000 simulations of alternative consumption measures. Data are simulated monthly and then converted to an annual frequency with 55 observations. The reported sample moments ($\times$100) are the mean, standard deviation, long-run standard deviation (computed over a six-year horizon), autocorrelation (one to five lags), correlation with the equity premium, cross-covariances ($\times 100 \times 100$) with the equity premium (contemporaneous and up to five years' future consumption), and the GMM estimate of the coefficient of relative risk aversion. The parameters for observed consumption are such that the series matches the empirical characteristics of garbage. Stocks are priced by $E[\beta(C_{t+1}/C_t)^{-\gamma} R^e_{m,t+1}] = 0$ with a true coefficient of relative risk aversion ($\gamma$) of 15. The equity premium is 4.5%. Reported NIPA consumption is time-aggregated and filtered from observable consumption. The columns to the right compare three alternative NIPA-based measures: (i) three-year consumption growth (P-J) as in Parker and Julliard (2005), (ii) fourth-quarter to fourth-quarter consumption growth (Q4-Q4) as in Jagannathan and Wang (2007), and (iii) unfiltered NIPA consumption. Results are reported when the unfilter is correctly calibrated ($\sigma_\xi$ is exactly as in the data) and stock returns are measured from December-to-December (Dec-Dec) or are time-aggregated (Time-Ag.) to align with the timing of NIPA consumption.

| | | | Simulation with Time-Aggregation and Filtering ($\sigma_\xi = 2.8\%$; $\bar{K} = 0.58$) | | | | |
|---|---|---|---|---|---|---|---|
| | | | NIPA Consumption | | | | |
| | True Consumption | Observed Consumption Garbage | Reported | Unfiltered | | | |
| Unfilter | $\sigma_\xi = 0\%$ | $\sigma_\xi = 2.8\%$ | Filtered | P-J | Q4-Q4 | $\bar{K} = 0.58$ | $\bar{K} = 0.58$ |
| Timing of $r^e_m$ | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Time-Ag. |
| E[$\triangle c$ ] | 1.90 | 1.90 | 1.90 | 5.61 | 1.90 | 1.90 | 1.90 |
| $\sigma$ [$\triangle c$ ] | 2.48 | 2.78 | 1.46 | 3.29 | 1.94 | 2.78 | 2.78 |
| $\sigma$ [$\sum_{q=1}^{6} \triangle c_q$ ]/$\sqrt{6}$ | 2.39 | 2.46 | 2.10 | 5.96 | 2.17 | 2.52 | 2.52 |
| AC(1)[$\triangle c$ ] | −1.54 | −11.48 | 43.46 | 78.52 | 1.14 | −5.02 | −5.02 |
| AC(2)[$\triangle c$ ] | −1.90 | −1.49 | 15.56 | 44.33 | 14.09 | −4.92 | −4.92 |
| AC(3)[$\triangle c$ ] | −1.82 | −1.43 | 3.41 | 13.45 | 1.02 | −3.00 | −3.00 |
| AC(4)[$\triangle c$ ] | −1.72 | −1.53 | −1.23 | 0.36 | −0.57 | −2.10 | −2.10 |
| AC(5)[$\triangle c$ ] | −1.73 | −1.26 | −3.03 | −4.41 | −1.78 | −1.52 | −1.52 |
| Corr(0,0)[$\triangle c, r^e_m$ ] | 65.54 | 58.45 | 31.41 | 39.06 | 55.03 | 36.72 | 54.45 |
| Cov(0,0)[$\triangle c, r^e_m$ ] | 29.28 | 29.28 | 8.18 | 22.83 | 19.18 | 18.28 | 27.60 |
| Cov(1,0)[$\triangle c, r^e_m$ ] | −0.44 | −0.32 | 10.59 | 16.14 | 3.26 | 13.41 | 3.66 |
| Cov(2,0)[$\triangle c, r^e_m$ ] | −0.57 | −0.50 | 4.21 | 5.80 | 4.47 | −2.31 | −2.28 |
| Cov(3,0)[$\triangle c, r^e_m$ ] | −0.38 | −0.37 | 1.56 | 1.59 | 0.65 | −1.18 | −1.24 |
| Cov(4,0)[$\triangle c, r^e_m$ ] | −0.49 | −0.52 | 0.31 | −0.10 | 0.23 | −0.98 | −0.89 |
| Cov(5,0)[$\triangle c, r^e_m$ ] | −0.55 | −0.44 | −0.09 | −0.81 | −0.19 | −0.63 | −0.74 |
| RRA($\hat{\gamma}$) | 15.46 | 15.54 | 54.79 | 20.37 | 23.73 | 24.13 | 16.64 |

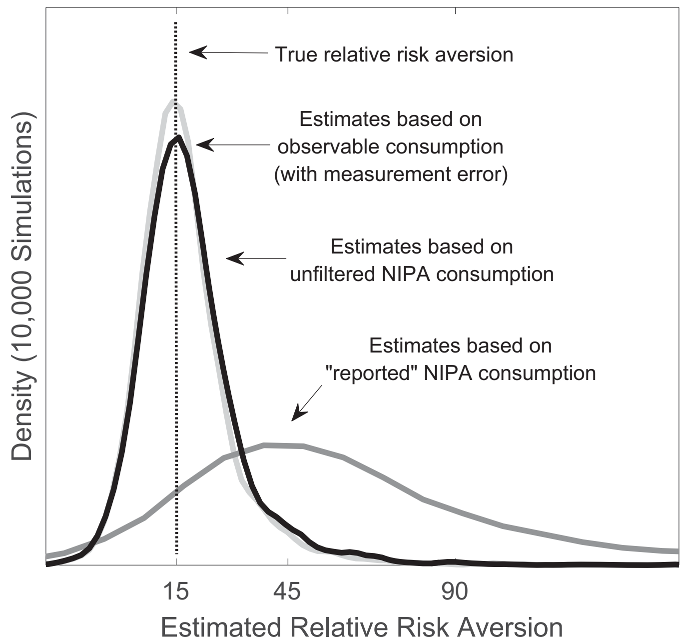


**Figure 1. Relative risk aversion estimates in a simulated asset pricing economy.** The figure provides the simulated distribution of GMM-based relative risk aversion estimates of observable consumption, time-aggregated and filtered ("reported") NIPA consumption, and unfiltered NIPA consumption (using time-aggregated stock returns).

These results show that three-year consumption growth (Parker and Julliard (2005)) works well when the data are only filtered. Fourth-quarter consumption growth (Jagannathan and Wang (2007)), as well as time-aggregated stock returns, recover the true consumption covariance when the data are only time-aggregated.

### *F. Calibration Strategy*

Setting the unfilter parameter too low results in over-unfiltering of the data. The standard deviation of unfiltered NIPA consumption will increase while the stock market correlation will only marginally decrease. As a result, arbitrary large stock market covariances can be generated in the absence of imposing discipline on the unfilter procedure.[25] I use long-run consumption standard deviations of reported and unfiltered NIPA consumption to restrict the degree of unfiltering. Measured over a long horizon, the effects of measurement error, filtering, and time-aggregation wash out, such that the long-run standard

[25] The effects of over-unfiltering and under-unfiltering are discussed in detail in the Internet Appendix.

deviation of true, observed, reported, and unfiltered NIPA consumption are identical.[26]

Which horizon should one pick as the long horizon for the calibration? Table I shows strong convergence when moving from one-year to six-year standard deviations. The standard deviations of unfiltered and reported NIPA consumption are 2.8% and 1.5% at the one-year horizon. Below, the standard deviations are 2.5% and 2.1% at the six-year horizon. Figure 2 (left side) compares the simulated long-run standard deviations at horizons from one to nine years. One can see that, after six years, further convergence is quite slow. Because at longer horizons the number of nonoverlapping observations becomes small quickly, which gives rise to small-sample issues, I impose the condition that the six-year standard deviation of unfiltered NIPA consumption is no larger than 1.2 (=2.52%/2.10%) times the value of reported NIPA consumption.

As an additional plausibility check, I also compare one-year standard deviations between unfiltered NIPA consumption and garbage. The idea is that the measurement error of aggregate NIPA consumption should not be (substantially) larger than the measurement error of garbage. Accordingly, the one-year standard deviation of unfiltered NIPA consumption should not be larger than the one-year standard deviation for garbage (2.9%).

## II. Asset Pricing without Garbage

I begin this section by calibrating unfiltered NIPA consumption and comparing the new series with the properties of alternative consumption measures. Next, I test whether unfiltered NIPA consumption can explain the equity premium and the risk-free rate with a reasonable coefficient of relative risk aversion. Moreover, I analyze further predictions of the filter model and provide tests on the robustness of the baseline results. In the data, the properties of reported NIPA services are considerably different from reported NIPA nondurables. I find that the filter model can also explain this heterogeneity between the two main subcategories. All these results are robust to using alternative unfilter specifications based on fourth-quarter consumption and other consumption volatility specifications. Finally, I test long-horizon Euler equations and conditional moment restrictions.

### A. Comparing Alternative Consumption Measures

Table II summarizes the empirical properties of the alternative consumption measures. All consumption data are measured over an annual frequency and on a real per-capita basis. The table reports results for the 1928–2014 period as well as the 1960–2014 postwar period. Garbage (1960–2007) and the Q4-Q4 measure are only available over the postwar sample (or a slightly shorter period). For unfiltered NIPA consumption, I focus the discussion on the results when stock returns are time-aggregated ("Time-Ag."). A column showing

[26] I thank an anonymous referee for this suggestion.

**Table II**

**Empirical Properties of Alternative Consumption Measures**

The table reports empirical moments of alternative consumption growth measures for the full sample (1928–2014) and the postwar period (1960–2014). The garbage measure is from Savov (2011) and is available for a slightly shorter postwar era (1960–2007). The reported consumption measure is real per-capita nondurables and services obtained from the NIPA statistics. The other columns report characteristics for three alternative NIPA-based measures: P-J is three-year NIPA consumption growth (Parker and Julliard (2005)). Q4-Q4 is fourth-quarter to fourth-quarter NIPA consumption growth (Jagannathan and Wang (2007)). The third is unfiltered NIPA consumption. The parameter $\sigma_\xi$ is calibrated to 2.8%, which implies an "average" unfilter of $\bar{K} = 0.5792$. Stock returns are measured from December-to-December (Dec-Dec) or are time-aggregated (Time-Ag.) to align with the timing of NIPA consumption.

| | Postwar Sample (1960–2014) | | | | | | Full Sample (1928–2014) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Garbage | NIPA Consumption | | | | | NIPA Consumption | | |
| | | Reported | Unfiltered | | | | Reported | Unfiltered | |
| Unfilter | – | Filtered | P-J | Q4-Q4 | $\bar{K} = 0.58$ | | Filtered | $\bar{K} = 0.58$ | |
| Timing of $r_m^e$ | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Time-Ag. | Dec-Dec | Dec-Dec | Time-Ag. |
| $E[\triangle c]$ | 1.42 | 1.88 | 5.60 | 1.88 | 1.85 | 1.85 | 1.79 | 1.79 | 1.79 |
| $\sigma\ [\triangle c]$ | 2.86 | 1.30 | 3.06 | 1.45 | 2.57 | 2.57 | 2.18 | 4.07 | 4.07 |
| $\sigma\ [\sum_{q=1}^{6} \triangle c_q\ ]/\sqrt{6}$ | 2.44 | 2.02 | 5.95 | 2.04 | 2.44 | 2.44 | 2.69 | 3.08 | 3.08 |
| $AC(1)[\triangle c]$ | −14.26 | 49.40 | 81.06 | 37.98 | 0.56 | 0.56 | 39.33 | −10.89 | −10.89 |
| $AC(2)[\triangle c]$ | −11.23 | 16.26 | 49.46 | −1.75 | −11.43 | −11.43 | 13.44 | 5.28 | 5.28 |
| $AC(3)[\triangle c]$ | −0.48 | 3.98 | 22.42 | 7.46 | −6.84 | −6.84 | −13.40 | −14.21 | −14.21 |
| $AC(4)[\triangle c]$ | 0.43 | 2.90 | 13.90 | 7.02 | 0.09 | 0.09 | −22.24 | −22.95 | −22.95 |
| $AC(5)[\triangle c]$ | 8.69 | 0.41 | 15.48 | −8.32 | −6.25 | −6.25 | 2.67 | 14.03 | 14.03 |
| $Corr(0,0)[\triangle c, r_m^e]$ | 60.70 | 6.69 | 16.84 | 33.15 | 25.42 | 47.95 | 15.76 | 25.96 | 62.97 |
| $Cov(0,0)[\triangle c, r_m^e]$ | 27.99 | 1.50 | 8.90 | 8.27 | 11.25 | 21.18 | 6.93 | 21.32 | 53.74 |
| $Cov(1,0)[\triangle c, r_m^e]$ | −2.74 | 9.56 | 5.80 | 7.85 | 19.56 | 7.01 | 28.79 | 50.13 | 15.77 |
| $Cov(2,0)[\triangle c, r_m^e]$ | −10.11 | −0.80 | −3.50 | −4.46 | −14.22 | −12.47 | 6.35 | −18.01 | −19.77 |
| $Cov(3,0)[\triangle c, r_m^e]$ | −1.21 | −0.79 | −1.27 | −0.21 | −0.99 | −1.33 | 2.00 | −1.75 | −11.87 |
| $Cov(4,0)[\triangle c, r_m^e]$ | 6.57 | −1.28 | −3.00 | 0.60 | −1.95 | −0.16 | −8.88 | −18.69 | −23.12 |
| $Cov(5,0)[\triangle c, r_m^e]$ | −2.87 | 0.97 | 4.36 | −0.55 | 3.48 | 3.07 | −9.13 | −8.00 | 2.12 |
| $E[R_m^e]$ | | | 5.77 | | | | | 7.27 | |
| $\sigma\ [R_m^e]$ | | | 16.89 | | | | | 20.10 | |

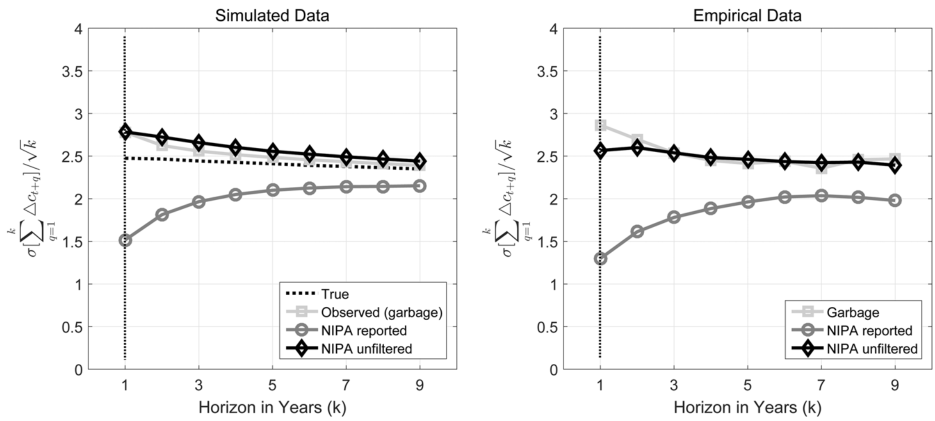


**Figure 2. NIPA consumption measures.** The figure shows annual real per-capita NIPA consumption growth from 1928 to 2014. The consumption measures are reported NIPA nondurables and services, three-year NIPA consumption (P-J) as in Parker and Julliard (2005), fourth-quarter to fourth-quarter NIPA consumption (Q4-Q4) as in Jagannathan and Wang (2007), and unfiltered NIPA consumption. Table II provides further details.

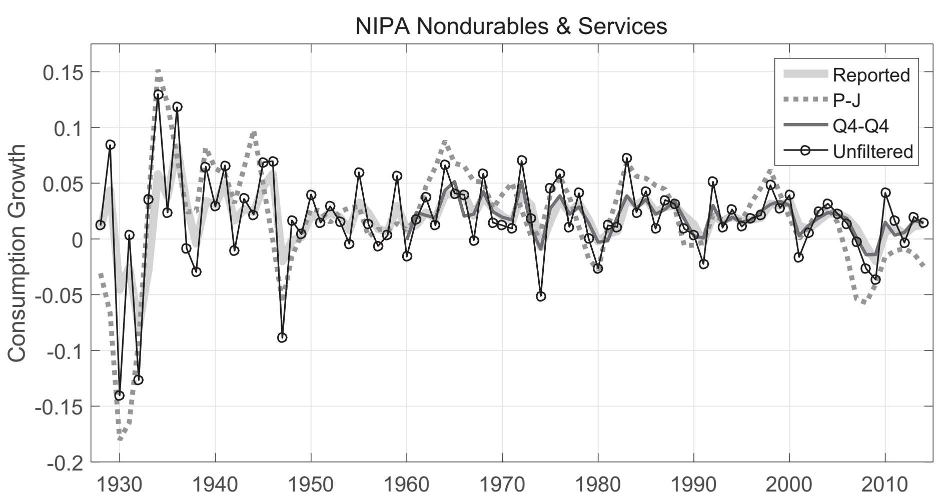


**Figure 3. Simulated and empirical long-run consumption standard deviation.** The figure on the left shows simulated long-run consumption standard deviations for alternative consumption measures. The figure on the right provides empirical counterparts for the period 1960 to 2014. Horizons $k = 1$ and $k = 6$ correspond to the values that can be found in Tables I and II.

results with stock returns based on the usual timing from December to December ("Dec–Dec") is provided for completeness. The time series of the alternative consumption measures can be inspected in Figure 3. Further details on data sources and construction are provided in the Internet Appendix.

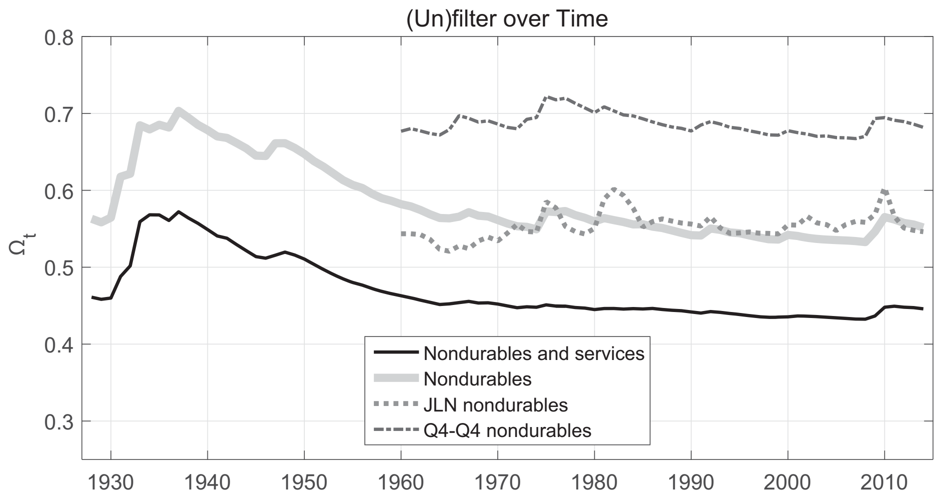


**Figure 4. Unfilter over time.** The figure shows the time series of alternative unfilter $\Omega_t$. "Nondurables and services" corresponds to the unfilter in the baseline results of Table II. "Nondurables" is the unfilter used in Table IV. "Q4-Q4 nondurables" and "JLN nondurables" provide the unfilter for the alternative specifications of Table V.

### *A.1. Calibration to Data*

I set the (un)filter parameters so that the unconditional consumption volatility is $\bar{\sigma}_\eta = 2.5\%$.[27] This value matches simulated and empirical long-run standard deviations of the growth rates of observed, reported NIPA, and unfiltered NIPA consumption in the postwar data. Moreover, this calibration matches recent empirical evidence provided by Dew-Becker (2016). Based on a nonparametric method, he pins down the long-run standard deviation of reported NIPA consumption at 2.5% as well. I set the degree of measurement error to $\sigma_\xi = 2.8\%$, such that the six-year standard deviation of unfiltered NIPA consumption is 1.2 times the value for reported NIPA consumption. The resulting (un)filter parameter has a steady-state value of $\bar{\Omega} = 0.46 = 0.58 \times 0.80$; its time series is provided in Figure 4.

### *A.2. Results*

Measured over one year, garbage is more volatile than reported NIPA consumption (2.9% versus 1.3%). This difference is considerably lower at the six-year horizon as the standard deviations of both series move toward each other (2.4% versus 2.0%). Indeed, garbage has negative short-term autocorrelation,

[27] The free parameters controlling consumption volatility dynamics are $a_1 = 0.08$, and $a_2 = 0.85$; these values are motivated by the findings of a simulation experiment as in the previous section. I find that unfiltered NIPA consumption is not very sensitive to a wide range of values for $a_1$ and $a_2$. Results in the robustness section show that unfiltered NIPA consumption is even robust to more fundamental changes in the specification of the consumption volatility dynamics (Table V).

which is in line with measurement error, pushing long-run standard deviations down. On the other hand, reported NIPA consumption has high positive short-term autocorrelation, which is in line with filtering pulling long-run standard deviations up.

By construction, the six-year standard deviation of unfiltered NIPA consumption (2.44%) is 1.2 times the one-year volatility of reported NIPA consumption (2.02%). Unfiltered NIPA consumption is robust to the choice of horizon for the calibration. Figure 3 shows that the empirical consumption standard deviations (right side) match the predictions of the model (left side) for *all horizons* from one to nine years. In addition, unfiltered NIPA consumption also tightly follows garbage in the figure. This does not need to be the case as garbage is not used for the calibration.[28]

The stock market covariance of reported NIPA consumption is virtually zero, in contrast to 0.28% for garbage. The covariance for unfiltered NIPA consumption is 0.21%—much larger than that for the other two NIPA-based measures, P-J (0.09%) and Q4-Q4 (0.08%). One advantage of unfiltered NIPA consumption compared to garbage is that it can be computed for a longer time period. In the full sample, unfiltered NIPA consumption has a stock market covariance as large as 0.54%. Average excess returns are also somewhat larger in the full sample (7.3%) compared to the postwar sample (5.8%).

Reported NIPA consumption growth has a low stock market covariance and shows a persistent component. Garbage has a high stock market covariance and is mean-reverting. These seemingly incompatible empirical properties of alternative consumption measures are hard to explain using standard models of the consumption process that do not try to account for the estimation process of reported NIPA consumption. Although the filter model proposed in this paper is a clear simplification of reality, it is able to match important aspects of the data surprisingly well and brings the properties of NIPA consumption close to garbage.

### *B. Baseline Results*

I test the Euler equation implied by the consumption-based asset pricing model (Lucas (1978), Breeden (1979)),

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^e_{m,t+1}\right] = 0. \tag{16}$$

Following Savov (2011), among others, I fix $\beta = 0.95$ to focus on relative risk aversion. It is well known that estimates of the relative risk aversion obtained from this Euler equation are typically too high from a theoretical

[28] The largest difference between the two series may be observed for the one-period standard deviation, which is somewhat lower for unfiltered NIPA consumption (2.6%) than for garbage (2.9%). A difference of this magnitude at the one-year horizon is well in line with the conjecture that measurement error in *aggregate* NIPA consumption should be small and not larger as in garbage.

perspective (Grossman and Shiller (1980, 1981), Hansen and Singleton (1982, 1983), Mehra and Prescott (1985)). The risk-free rate puzzle is related to the equity premium puzzle (Weil (1989)). High coefficients of relative risk aversion imply a high desire for consumption smoothing, resulting in implausibly large real risk-free interest rates. Taking a log approximation of the Euler equation (e.g., Campbell, Lo, and MacKinlay (1997)), the risk-free rate is given as

$$r_f = -\log(\beta) + \gamma E[log(C_{t+1}/C_t)] - \frac{1}{2}\gamma^2 \mathrm{Var}[log(C_{t+1}/C_t)]. \tag{17}$$

I report this implied risk-free rate when testing equation (16), again following Savov (2011). Furthermore, I provide results for when the empirically observed risk-free rate is included as a test asset in addition to the equity premium,

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{f,t+1} - 1\right] = 0, \tag{18}$$

which means that the estimation procedure is forced to match the equity premium jointly with the empirical risk-free rate. Estimation is carried out by generalized method of moments (GMM) with Newey and West (1987) standard errors (five lags, removed moment means). When the risk-free rate is included as a test asset, the two moments are weighted by the identity matrix. I report the mean absolute error (MAE) of the moment condition(s) and the $p$-value to the $J$-test of overidentified restrictions when multiple moments are considered. Further estimation details can be found in the Internet Appendix.

Table III provides the GMM estimates of relative risk aversion for all consumption measures considered. In the postwar period, the traditional NIPA consumption measure requires a coefficient of relative risk aversion of 67 when the equity premium is the only test asset. The implied risk-free rate is 94%. Strikingly, the high relative risk aversion is not enough to price the equity premium, as indicated by an MAE of 2.5%. In contrast, unfiltered NIPA consumption in combination with time-aggregated stock returns requires a relatively low estimate of 23, an MAE of zero, and an implied risk-free rate of 30%. A relative risk aversion of 10 is well within two standard errors, and the estimate for garbage (16) is within one. Three-year consumption growth (P-J) produces a coefficient of relative risk aversion of 42, and the corresponding estimate for fourth-quarter consumption (Q4-Q4) is 64.[29] For the full sample period, I find that reported NIPA consumption requires a relative risk aversion of 37, compared to 10 for unfiltered NIPA consumption.

In the postwar sample, adding the risk-free rate as a moment condition increases the estimate of relative risk aversion for reported NIPA consumption well above 100. P-J and Q4-Q4 consumption run into similar trouble as reported

[29] The discount factor in Parker and Julliard (2005) includes a three-year risk-free rate term. I ignore this term to focus on three-year consumption as an ad hoc unfilter. Including the risk-free rate term has only a small effect on the results (as pointed out by Parker and Julliard (2005)); for example, the estimate of 42 increases to 43.

**Table III**

## GMM Estimates of Relative Risk Aversion

The table reports GMM estimates of the relative risk aversion coefficient, RRA ($\gamma$). The moment restrictions are

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^e_{m,t+1} - 0\right] \text{ and } E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{f,t+1} - 1\right],$$

where $R^e_m$ is the excess return of the market portfolio and $R_f$ is the gross risk-free rate. The discount parameter is fixed, $\beta = 0.95$. The first panel uses the equity premium of the aggregate stock market as the only moment restriction. In the second panel, the risk-free rate is added as a second moment restriction to the equity premium. The consumption measures are reported NIPA consumption, three-year NIPA consumption (P-J, Parker and Julliard (2005)), fourth-quarter to fourth-quarter NIPA consumption (Q4-Q4, Jagannathan and Wang (2007)), and unfiltered NIPA consumption. Asset returns are measured from December-to-December ("Dec-Dec") or are time-aggregated ("Time-Ag."). Estimates on garbage are based on a shorter sample period from 1960 to 2007. MAE is the mean absolute error of the moment restrictions. The $p$-values correspond to the J-test of overidentified restrictions.

| | Postwar Sample (1960–2014) | | | | | | Full Sample (1928–2014) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Garbage | NIPA Consumption | | | | | NIPA Consumption | | |
| | | Reported | Unfiltered | | | | Reported | Unfiltered | |
| Unfilter | – | Filtered | P-J | Q4-Q4 | $\bar{K} = 0.58$ | | Filtered | $\bar{K} = 0.58$ | |
| Timing of $R^e_i$ | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Time-Ag. | Dec-Dec | Dec-Dec | Time-Ag. |
| **Equity Premium (Exactly Identified)** | | | | | | | | | |
| RRA ($\gamma$) | 15.63 | 67.32 | 42.35 | 64.05 | 37.02 | 22.53 | 36.86 | 15.76 | 10.32 |
| (s.e.) | (8.38) | – | (22.87) | (39.61) | (24.83) | (11.98) | (13.13) | (6.02) | (4.53) |
| Implied $r_f$, % | 17.25 | 93.77 | 158.19 | 82.60 | 28.47 | 30.09 | 38.81 | 12.84 | 14.83 |
| MAE % | 0.00 | 2.45 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Equity Premium & Risk-Free Rate (One-Stage GMM)** | | | | | | | | | |
| RRA ($\gamma$) | 28.72 | 186.91 | 171.84 | 215.16 | 52.48 | 50.84 | 41.14 | 15.90 | 15.11 |
| (s.e.) | (9.39) | (48.86) | (48.06) | (58.14) | (12.92) | (12.99) | (14.26) | (5.98) | (5.96) |
| MAE % | 4.57 | 11.07 | 12.81 | 4.44 | 2.51 | 7.27 | 0.93 | 0.07 | 4.09 |
| $p$-value ($J$) | 0.63 | 0.18 | 0.12 | 0.55 | 0.75 | 0.39 | 0.88 | 0.99 | 0.57 |

NIPA consumption. In contrast, garbage gives an estimate of 29 and unfiltered NIPA consumption of 51. The risk-free rate is much less of an issue when looking at the full sample period. Unfiltered NIPA consumption can price the equity premium and the risk-free rate with a relative risk aversion of 15.

I conclude that accounting for the filtering of reported consumption data, which is magnified by time-aggregation, is crucial for asset pricing. Unfiltered NIPA consumption performs well in explaining the equity premium and helps reduce the risk-free rate puzzle. In the longer sample period, I get point estimates very close to theoretically reasonable values.

### *C. Separating NIPA Nondurables and Services*

I now test a further prediction implied by the filter model. So far I have considered NIPA consumption measured as NIPA nondurables and services. This is arguably the most widely used measure of consumption in the literature. However, there is evidence of heterogeneity in the degree of measurement error in the two main components of NIPA consumption. Indeed, the official NIPA handbook emphasizes that services are considerably more difficult to measure than nondurables (BEA (2014)). The literature has also been skeptical of the services components of NIPA consumption. For example, Hansen and Singleton (1982) and Hall (1988) test nondurables excluding services. Wilcox (1992) suggests distinguishing between nondurables and services in empirical work, and Piazzesi, Schneider, and Tuzel (2007) provide a detailed discussion of the problems surrounding the services component related to housing. Savov (2011) further argues that the way services are estimated is particularly problematic for the NIPA measure and may help explain differences with respect to garbage.

The filter model is compatible with this reasoning. It predicts that the components of reported NIPA consumption that are more difficult to measure are more heavily filtered, that is, $\bar{K}$ is closer to unity for NIPA nondurables than for NIPA services. As a result, NIPA consumption—including services—should have a lower standard deviation, a higher autocorrelation, and, importantly, a lower covariance with stock returns than NIPA consumption excluding services.

Table IV summarizes the empirical properties of NIPA consumption separated into NIPA nondurables and NIPA services. As documented elsewhere, compared to NIPA services, postwar NIPA nondurables are more volatile (1.6% versus 1.3%) and less autocorrelated (35% versus 57%). Moreover, excluding services more than doubles the stock market correlation, from 7% to 19%. These results are in line with less measurement error and therefore less filtering of NIPA nondurables.

Accordingly, I recalibrate the unfilter recursion for NIPA nondurables by setting $\sigma_\xi$ to 1.9% and keep the other parameters unchanged. The implied degree of unfiltering in the steady state is now $\bar{\Omega} = \bar{K} \times a = 0.71 \times 0.80 = 0.57$, closer to unity, and in line with less filtering than for the sum of NIPA nondurables

**Table IV**

**Separating NIPA Nondurables and Services**

The table provides results when the two main components of NIPA consumption, NIPA nondurables and NIPA services, are separated. Panel A provides empirical moments (as in Table II) and Panel B provides GMM estimates of relative risk aversion (as in Table III).

| | Postwar Sample (1960–2014) | | | | Full Sample (1928–2014) | | | |
|---|---|---|---|---|---|---|---|---|
| | NIPA Consumption | | | | NIPA Consumption | | | |
| | Reported | | | Unfiltered | Reported | | | Unfiltered |
| Unfilter | Nond. & Serv. Filtered | Nondur. Filtered | Services Filtered | Nondur. $\bar{K} = 0.71$ | Nond. & Serv. Filtered | Nondur. Filtered | Services Filtered | Nondur. $\bar{K} = 0.71$ |
| Timing of $r_m^e$ | Dec-Dec | Dec-Dec | Dec-Dec | Time-Ag. | Dec-Dec | Dec-Dec | Dec-Dec | Time-Ag. |
| **Panel A: Empirical Moments** | | | | | | | | |
| $\sigma$ $[\triangle c]$ | 1.30 | 1.56 | 1.33 | 2.68 | 2.18 | 2.60 | 2.10 | 4.15 |
| $\sigma$ $[\sum_{q=1}^{6} \triangle c_q\ ]/\sqrt{6}$ | 2.02 | 1.95 | 2.38 | 2.30 | 2.69 | 2.77 | 3.10 | 3.12 |
| AC(1)$[\triangle c]$ | 49.40 | 34.76 | 57.04 | 0.77 | 39.33 | 27.22 | 52.52 | −11.31 |
| AC(2)$[\triangle c]$ | 16.26 | −6.30 | 36.57 | −23.71 | 13.44 | 2.94 | 28.00 | 1.87 |
| AC(3)$[\triangle c]$ | 3.98 | −5.55 | 19.35 | 0.53 | −13.40 | −20.00 | 1.70 | −13.93 |
| Corr(0,0)$[\triangle c, r_m^e]$ | 6.69 | 18.87 | −1.05 | 45.49 | 15.76 | 19.40 | 12.01 | 57.83 |
| Cov(0,0)$[\triangle c, r_m^e]$ | 1.50 | 5.09 | −0.24 | 21.01 | 6.93 | 10.20 | 5.09 | 50.32 |
| Cov(1,0)$[\triangle c, r_m^e]$ | 9.56 | 13.54 | 7.02 | 6.90 | 28.79 | 33.27 | 25.41 | 15.45 |
| **Panel B: GMM Estimates of Relative Risk Aversion** | | | | | | | | |
| Equity Premium (Exactly Identified) | | | | | | | | |
| RRA ($\gamma$) | 67.32 | 60.31 | 64.46 | 20.08 | 36.86 | 30.38 | 52.92 | 10.76 |
| (s.e.) | – | (37.67) | – | (10.38) | (13.13) | (10.93) | (25.02) | (4.82) |
| Implied $r_f$, % | 93.77 | 39.30 | 112.01 | 16.38 | 38.81 | 13.94 | 54.98 | 9.40 |
| MAE % | 2.45 | 0.00 | 2.66 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Equity Premium and Risk-Free Rate (One-Stage GMM) | | | | | | | | |
| RRA ($\gamma$) | 186.91 | 84.59 | 220.27 | 30.71 | 41.14 | 29.17 | 50.21 | 13.18 |
| (s.e.) | (48.86) | (21.68) | (56.07) | (9.76) | (14.26) | (11.15) | (16.75) | (5.32) |
| MAE % | 11.07 | 2.59 | 15.36 | 3.44 | 0.93 | 0.30 | 0.32 | 1.95 |
| $p$-value ($J$) | 0.18 | 0.74 | 0.07 | 0.69 | 0.88 | 0.96 | 0.95 | 0.79 |

and services ($\bar{\Omega} = 0.58 \times 0.80 = 0.46$).[30] By construction, I get a six-year consumption standard deviation of unfiltered NIPA nondurables of 2.3%, which is 1.2 times the one-year standard deviation of 1.95%. The resulting one-period standard deviation of unfiltered NIPA nondurables is 2.7%—slightly larger compared to the unfiltered NIPA nondurables and services (2.5%). Unfiltered NIPA nondurables has a stock market covariance of 0.21% and allows the equity premium to be priced with a relative risk aversion of 20 in the postwar period. Moreover, unfiltered NIPA nondurables can better price the equity premium together with the risk-free rate. The estimated risk aversion is 31—almost the same as the estimate of 29 for garbage. In the full sample period, I find a relative risk aversion of about 11, which is similar to unfiltered NIPA nondurables and services.

In the Internet Appendix, I report additional results for unfiltered NIPA services and a finer decomposition of the components of NIPA consumption. Unfiltering each component separately and aggregating does not lead to better (or worse) results. Overall, the results for unfiltered NIPA nondurables are similar to those for unfiltered NIPA nondurables and services. In line with the predictions of the filter model, excluding services is helpful for asset pricing and allows NIPA consumption to be unfiltered with a filter parameter closer to unity.

### *D. Unfiltered Fourth-Quarter Consumption*

Fourth-quarter consumption is a natural alternative to mitigate the time-aggregation bias.[31] Replacing the proposed time-aggregation bias adjustments with the use of fourth-quarter consumption should lead to similar results and serves as a robustness check. Of course, a limitation of fourth-quarter consumption is that consumption data are only available for the postwar period.

Table V (left-hand side) reports results for unfiltered fourth-quarter NIPA consumption. Since the use of fourth-quarter NIPA consumption already accounts for time-aggregation, I calibrate the unfilter recursion as in the previous calibrations and set the time-aggregation adjustment to one. This gives a steady-state unfilter of $\bar{\Omega} = \bar{K} \times 1 = 0.58$ for NIPA nondurables and services and $\bar{\Omega} = \bar{K} \times 1 = 0.71$ for NIPA nondurables excluding services.[32] For the same reason, stock returns are always measured from December to December.

The one-period standard deviation of unfiltered *fourth-quarter* NIPA nondurables (NIPA nondurables and services) is 2.9% (2.5%), with a six-year standard deviation of 2.3% (2.3%) and a stock market covariance of 0.26% (0.20%). Interestingly, the stochastic properties of unfiltered fourth-quarter NIPA nondurables are essentially the same as for garbage (see Table II). I find that the estimates of relative risk aversion based on unfiltered fourth-quarter NIPA

[30] The time series of the unfilter $\Omega_t$ is shown in Figure 4.

[31] Simulation results provided in the Internet Appendix suggest that fourth-quarter consumption is almost sufficient to mitigate time-aggregation bias in annual data.

[32] The actual time series of $\Omega_t$ is shown in Figure 4.

**Table V**

## Alternative Unfilter Specifications

"Fourth-Quarter Consumption" results correspond to reported and unfiltered fourth-quarter NIPA consumption. For "JLN Volatility," I use the (rescaled) macroeconomic uncertainty index estimated by Jurado, Ludvigson, and Ng (2015) as a measure of consumption volatility in the unfilter $K_t$. "Constant Volatility" results obtain when I assume constant consumption volatility to unfilter the data. Panel A provides empirical moments (as in Table II) and Panel B provides GMM estimates of relative risk aversion (as in Table III) for the alternative consumption measures.

| | Postwar Sample (1960–2014) NIPA Consumption | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Fourth-Quarter Consumption | | | | JLN Volatility | | Constant Volatility | |
| | Reported | | Unfiltered | | Unfiltered | | Unfiltered | |
| NIPA series | N & S | Nond. | N & S | Nond. | N & S | Nond. | N & S | Nond. |
| Unfilter | Filtered | Filtered | $\bar{K}=0.58$ | $\bar{K}=0.71$ | $\bar{K}=0.58$ | $\bar{K}=0.71$ | $\bar{K}=0.58$ | $\bar{K}=0.71$ |
| Timing of $r_m^e$ | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Time-Ag. | Time-Ag. | Time-Ag. | Time-Ag. |
| | Panel A: Empirical Moments | | | | | | | |
| $\sigma\ [\triangle c]$ | 1.45 | 1.99 | 2.46 | 2.90 | 2.55 | 2.69 | 2.46 | 2.62 |
| $\sigma\ [\sum_{q=1}^{6} \triangle c_q\ ]/\sqrt{6}$ | 2.04 | 2.03 | 2.32 | 2.27 | 2.38 | 2.26 | 2.40 | 2.27 |
| AC(1)[$\triangle c$] | 37.98 | 16.85 | 1.47 | −4.68 | −0.42 | −1.06 | 1.78 | 0.96 |
| AC(2)[$\triangle c$] | −1.75 | −25.56 | −26.00 | −33.67 | −9.46 | −22.32 | −10.59 | −22.57 |
| AC(3)[$\triangle c$] | 7.46 | −0.07 | 7.33 | 7.87 | −6.21 | −0.05 | −5.92 | 0.94 |
| Corr(0,0)[$\triangle c, r_m^e$] | 33.15 | 45.16 | 46.77 | 52.26 | 46.53 | 44.84 | 47.39 | 45.31 |
| Cov(0,0)[$\triangle c, r_m^e$] | 8.27 | 15.53 | 19.83 | 26.12 | 20.49 | 20.82 | 20.10 | 20.45 |
| Cov(0,1)[$\triangle c, r_m^e$] | 7.85 | 10.86 | 7.47 | 8.54 | 7.04 | 7.25 | 7.18 | 7.31 |
| | Panel B: GMM Estimates of Relative Risk Aversion | | | | | | | |
| | Equity Premium (Exactly Identified) | | | | | | | |
| RRA ($\gamma$) | 64.05 | 28.80 | 27.86 | 19.02 | 22.35 | 19.67 | 23.54 | 20.71 |
| (s.e.) | (39.61) | (14.30) | (15.12) | (9.88) | (12.44) | (10.64) | (12.45) | (10.71) |
| Implied $r_f$, % | 82.60 | 26.35 | 33.64 | 14.85 | 31.78 | 18.04 | 31.85 | 16.94 |
| MAE % | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | Equity Premium and Risk-free Rate (One-Stage GMM) | | | | | | | |
| RRA ($\gamma$) | 215.16 | 55.83 | 76.49 | 31.32 | 51.56 | 31.46 | 54.64 | 32.20 |
| (s.e.) | (58.14) | (15.81) | (17.94) | (9.78) | (13.43) | (10.07) | (13.82) | (10.14) |
| MAE % | 4.44 | 6.11 | 6.12 | 3.24 | 7.90 | 3.96 | 7.74 | 3.58 |
| $p$-value ($J$) | 0.55 | 0.48 | 0.45 | 0.70 | 0.37 | 0.65 | 0.36 | 0.68 |

consumption are very similar to their time-aggregation bias-adjusted counterparts reported earlier. I conclude that the baseline results do not crucially depend on the use of time-aggregated stock returns when quarterly consumption data are available.

### *E. Alternative Consumption Volatility Measures*

The baseline results rely on a GARCH specification to capture consumption volatility in the filter model. This specification allows unfiltered NIPA consumption to be computed solely from reported NIPA consumption. In this section, I address potential concerns with respect to the GARCH specification. First, the GARCH model may fail to capture high-frequency time-variation because it infers volatility from lagged and annual consumption data (Bansal and Yaron (2004)). Second, a GARCH specification may not capture low-frequency shifts in consumption volatility because it is not designed to account for such a component (Lettau, Ludvigson, and Wachter (2008), Campbell et al. (2015)).

A more data-driven approach to measuring consumption volatility should mitigate both of these concerns. Bansal et al. (2014) measure annual consumption volatility from a VAR in which (rescaled) monthly observations of industrial production serve as a state variable. In a similar spirit, I use the macroeconomic uncertainty index recently developed by Jurado, Ludvigson, and Ng (2015) to construct a measure of consumption volatility that I then directly feed into the unfilter $\Omega_t$. The Jurado, Ludvigson, and Ng (2015; JLN) macroeconomic uncertainty index is as free from specific theoretical models as possible and exploits a rich cross-section of economic indicators sampled at the monthly frequency.[33] To ensure consistency, I rescale the JLN macroeconomic uncertainty index such that the mean of consumption volatility matches the baseline results. I fix the volatility of consumption volatility to 0.0003%. Due to data availability, I can only test this specification for the postwar sample period. Further details are delegated to the Internet Appendix, which also provides a discussion of the benefits and limitations of this approach.

Figure 4 compares the resulting time series of the unfilter based on the JLN macroeconomic uncertainty index ("JLN-unfilter"), with the baseline specification. Overall, the JLN-unfilter and the baseline unfilter are closely related over longer periods. The JLN-unfilter catches recent innovations in consumption volatility more quickly. However, these movements are small from a historical perspective. As can be inferred from the figure, the unfilter using the standard GARCH specification of volatility picks up the volatile period around 1930–1940. For example, the unfilter corresponding to NIPA nondurables is as large as 0.70 in 1940 and then lingers within a narrow band around its steady-state value of 0.57 ($= 0.71 \times 0.80$) in the postwar period. Given that the JLN-unfilter and the baseline unfilter track each other closely, it is not surprising that I find very similar results for JLN-unfiltered NIPA consumption compared to the earlier results (see Tables II–V). For example, JLN-unfiltered NIPA non-

[33] Industrial production is one of over 100 series used to construct this index.

durables price the equity premium with a $\gamma$ coefficient of 20, as in the baseline results.

To better see the importance of the consumption volatility model for the unfilter, the right-hand side of Table V reports results for the unfilter assuming constant consumption volatility. Here, each observation is unfiltered by the same steady-state value of the unfilter. I again find that the effect is arguably not too large for the postwar period. For example, constantly unfiltered NIPA nondurables and services (NIPA nondurables) price the equity premium with a $\gamma$ coefficient of 24 (21).

For unfiltered NIPA consumption in the postwar period, the steady-state level of the unfilter is more important than short-lived movements induced by changing volatility. However, the early observations in the 1930s are much more volatile, and it is more important to account for changing consumption volatility between the early observations around the 1930s and the postwar period. In line with this observation, I find that the unfilter accounting for time-varying volatility is much closer to unity in the prewar period than in the postwar data. For further robustness, I test a consumption volatility specification with a regime shift in 1950, that is, I recalibrate the unfilter for the observations before 1950. The results, reported in the Internet Appendix, are similar to the findings presented so far.

### *F. Long-Horizon Returns*

Filtering and time-aggregation bias destroy the covariance between consumption growth and stock returns, particularly over short horizons. As pointed out by Singleton (1990), Cochrane and Hansen (1992), and Daniel and Marshall (1997), both kinds of frictions diminish over long horizons and should be overcome by testing the multiperiod Euler equation, where consumption growth and returns are measured over multiple years $h$:

$$E\left[\beta^h\left(\frac{C_{t+h}}{C_t}\right)^{-\gamma} R^e_{m,t\to t+h}\right] = 0. \tag{19}$$

When increasing $h$, estimates of the coefficient of relative risk aversion based on reported NIPA consumption should converge to estimates based on observed consumption (unfiltered NIPA consumption and garbage). How quick is convergence according to the filter model? In the Internet Appendix, simulation evidence shows that convergence is relatively slow and thus one needs to look at rather long horizons. At the three-year (six-year) horizon, the estimate of relative risk aversion based on reported consumption is still about 2 (1.2) times larger than the true value.

Table VI provides empirical results for the three-year ($h = 3$) and six-year ($h = 6$) horizons. The previous one-year horizon results ($h = 1$) are reported to facilitate comparison. In the postwar sample, reported NIPA consumption fits the three-year equity premium with a similar coefficient of relative risk aversion as the one-year equity premium. However, the MAE drops from 2.5%

**Table VI**
## Long-Horizon Returns

The table reports GMM estimates of the relative risk aversion coefficient, RRA ($\gamma$). The moment restriction is

$$E\left[\beta^h \left(\frac{C_{t+h}}{C_t}\right)^{-\gamma} R^e_{m,t\to t+h} - 0\right],$$

where $R^e_m$ is the $h$-year excess return of the market portfolio. The discount parameter is fixed, $\beta = 0.95$. The consumption measures are garbage as in Savov (2011), reported NIPA nondurables and services ("N&S"), unfiltered NIPA nondurables and services, unfiltered NIPA nondurables ("Ndr."), and unfiltered fourth-quarter NIPA nondurables ("Q4-N"). Stock returns are measured from December-to-December ("D-D") or are time-aggregated ("T-A"). Estimates on garbage are based on a shorter sample period from 1960 to 2007. MAE is the mean absolute error of the moment restriction. $\#\triangle C_{t+h} < 0$ is the number of observations with negative consumption growth.

| | Postwar Sample (1960–2014) | | | | | Full Sample (1928–2014) | | |
|---|---|---|---|---|---|---|---|---|
| | Garbage | NIPA Consumption | | | | NIPA Consumption | | |
| | | Reported | Unfiltered | | | Reported | Unfiltered | |
| Unfilter | – | Filtered | $\bar{K} = 0.58$ | $\bar{K} = 0.71$ | $\bar{K} = 0.71$ | Filtered | $\bar{K} = 0.58$ | $\bar{K} = 0.71$ |
| Timing of $R^e_m$ | D-D | D-D | T-A | T-A | D-D | D-D | T-A | T-A |
| NIPA eries | | N&S | N&S | Ndr. | Q4-N | N&S | N&S | Ndr. |
| One-Year Equity Premium ($h = 1$) | | | | | | | | |
| RRA ($\gamma$) | 15.63 | 67.32 | 22.53 | 20.08 | 19.02 | 36.86 | 10.32 | 10.76 |
| (s.e.) | (8.38) | – | (11.98) | (10.38) | (9.88) | (13.13) | (4.53) | (4.82) |
| MAE % | 0.00 | 2.45 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $\#\triangle C_{t+h} < 0$ | 13 | 5 | 10 | 14 | 15 | 12 | 18 | 25 |
| Three-Year Equity Premium ($h = 3$) | | | | | | | | |
| RRA ($\gamma$) | 27.75 | 72.94 | 22.06 | 20.48 | 23.39 | 14.45 | 8.05 | 8.95 |
| (s.e.) | (18.73) | (27.69) | (8.89) | (8.53) | (10.68) | (5.21) | (3.16) | (3.54) |
| MAE % | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $\#\triangle C_{t+h} < 0$ | 8 | 3 | 6 | 9 | 10 | 8 | 12 | 16 |
| Six-Year Equity Premium ($h = 6$) | | | | | | | | |
| RRA ($\gamma$) | 71.80 | >1,000 | 77.00 | 59.20 | 55.07 | 16.80 | 13.61 | 17.25 |
| (s.e.) | (74.84) | – | – | – | – | (5.33) | (4.98) | (6.65) |
| MAE % | 0.00 | 0.00 | 0.30 | 0.33 | 0.76 | 0.00 | 0.00 | 0.00 |
| $\#\triangle C_{t+h} < 0$ | 2 | 0 | 1 | 5 | 4 | 3 | 4 | 9 |

to zero and reflects an improved model fit. In the full sample period, I find an estimate of relative risk aversion for reported NIPA consumption as low as 15. As expected, results for unfiltered NIPA consumption remain unchanged in both sample periods.

At the six-year horizon, I find that results deteriorate for all consumption measures in the postwar sample. As reported in the table, there are no negative six-year consumption growth rates for reported NIPA consumption and only a few for garbage and unfiltered NIPA consumption. Cochrane and Hansen (1992) make a similar observation for reported NIPA consumption at the five-year horizon. Estimates increase at very long horizons, because there are fewer and fewer instances of negative consumption growth rates. In the extreme, if there are no negative consumption growth rates at all, the sample counterpart of the moment condition in equation (19) can be set to zero for any large number $\gamma$ (after computational rounding).

I conclude that the postwar sample is not informative for estimation with long-horizon moments.[34] Likewise, the full sample depends on only a few negative observations. Using one-year unfiltered NIPA consumption is a much more revealing way to estimate the model.

### *G. Conditioning Information*

This section considers conditional moment restrictions and explores whether unfiltered NIPA consumption can account for the predictability of stock returns. If conditional expected returns are high simply because conditional consumption covariances are high, then the consumption-based model with CRRA preferences is able to account for return predictability (e.g., Parker (2003), Duffee (2005)). However, this is not what I find in the data.

Table VII reports the results. As instruments, I employ lagged consumption growth, the lagged stock market excess return, and the lagged log price-dividend ratio. The log price-dividend ratio is demeaned and scaled by $-5$ ($\approx -1/\sigma(pd)$).[35] The point estimates of relative risk aversion are close to the results based on unconditional moments, but the MAEs increase. A closer look at the individual moment conditions reveals that the average pricing error of the return scaled by the price-dividend ratio is particularly large. Giving this moment restriction a larger weight in the GMM estimation leads to rejection of the model. The reason is that consumption covariances of the "managed" portfolio (Cochrane (2005)) that goes long when the price-dividend ratio predicts high future returns are close to zero for unfiltered NIPA consumption, and even

[34] For simulated data with a sample size of 55 years, I find in the Internet Appendix that in 68% of the draws no negative six-year growth rates for reported consumption are evident. Thus, even when the null is true, it is quite likely that estimation using six-year moments fails. Calculations of the rate of convergence mentioned above are conditional on whether one is drawing a sample with at least one negative growth rate.

[35] I find that the price-dividend ratio significantly predicts next year's excess market return with a slope of $-0.30$ in the postwar sample ($-0.27$ in the full sample).

**Table VII**
## Conditioning Information

The table reports GMM estimates of the relative risk aversion coefficient, RRA ($\gamma$). The moment restrictions are

$$E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R^e_{m,t+1} \otimes z_t - 0\right],$$

where $\beta = 0.95$, $R^e_m$ is the excess return of the market portfolio, and $z_t$ is a vector of instruments including the constant one, lagged (gross) consumption growth $(C_t/C_{t-1})$, one plus the lagged excess market return $(1+R^e_{m,t})$, and the demeaned lagged log price-dividend ratio $(-5 \times (pd_t - E[pd]))$, where $-5 \approx -1/\sigma(pd)$. The log price-dividend ratio is structural-break adjusted (Lettau and Nieuwerburgh (2008)) and mid-year dividends are not reinvested in the stock market. When returns are measured from December-to-December ("D-D"), the instruments are lagged by one period. When returns are time-aggregated ("T-A"), the instruments are lagged by two periods to avoid an overlap with the measurement interval of returns. MAE is the mean absolute error of the moment restriction. $b_{pd}$ is the slope coefficient when predicting discounted returns with the lagged log price-dividend, $t$-statistics are reported in square brackets ratio; estimation is based on robust least squares using the Andrews weighting function to down-weight extreme observations.

| | Postwar Sample (1960–2014) | | | | | Full Sample (1928–2014) | | |
|---|---|---|---|---|---|---|---|---|
| | Garbage | NIPA Consumption | | | | NIPA Consumption | | |
| | | Reported | Unfiltered | | | Reported | Unfiltered | |
| Unfilter | – | Filtered | $\bar{K} = 0.58$ | $\bar{K} = 0.71$ | $\bar{K} = 0.71$ | Filtered | $\bar{K} = 0.58$ | $\bar{K} = 0.71$ |
| Timing of $R^e_m$ | D-D | D-D | T-A | T-A | D-D | D-D | T-A | T-A |
| NIPA series | | N&S | N&S | Ndr. | Q4-N | N&S | N&S | Ndr. |
| Equity Premium & Instruments (One-Stage GMM) | | | | | | | | |
| RRA ($\gamma$) | 16.96 | 45.29 | 19.41 | 15.38 | 18.23 | 38.31 | 9.28 | 9.92 |
| (s.e.) | (8.15) | (37.45) | (13.70) | (11.04) | (9.91) | (11.89) | (4.79) | (5.10) |
| MAE % | 1.25 | 3.65 | 2.65 | 2.79 | 2.06 | 0.80 | 2.56 | 2.55 |
| $p$-value ($J$) | 0.99 | 0.95 | 0.96 | 0.96 | 0.96 | 1.00 | 0.89 | 0.88 |
| Conditional Pricing Errors: $\beta(C_{t+1}/C_t)^{-\gamma} R^e_{m,t+1} = a + b_{pd}\, pd_t + e_{t+1}$ | | | | | | | | |
| $b_{pd}$ | −0.20 | −0.17 | −0.21 | −0.21 | −0.30 | −0.23 | −0.33 | −0.31 |
| $t(b_{pd})$ | [−1.99] | [−3.07] | [−2.28] | [−2.20] | [−3.33] | [−3.73] | [−4.28] | [−3.96] |

slightly negative for reported NIPA consumption (in line with Parker (2003) and Duffee (2005)).

Another way to see the model's failure to account for time-variation in expected returns is to look at conditional pricing errors instead of average pricing errors (Nagel and Singleton (2011)). As shown by Hansen and Singleton (1982), discounted returns ($\beta(C_{t+1}/C_t)^{-\gamma} R^e_{m,t+1}$) should not be predicted by the instruments. In the lower part of Table VII, I find that discounted returns are virtually as predictable as returns.[36]

Unfiltered NIPA consumption in combination with CRRA preferences helps explain the level of the equity premium but not its predictability. In Section IV, I discuss the implications of unfiltered NIPA consumption for advanced asset pricing models that can generate time-varying risk premia.

## III. Cross-Section of Stock Returns

In cross-sectional pricing tests, garbage allows for low estimates of risk aversion but the pricing errors are large (Savov (2011)). The P-J and fourth-quarter consumption measures have low pricing errors in cross-sectional pricing tests but produce high estimates of risk aversion (Parker and Julliard (2005), Jagannathan and Wang (2007)). This is a highly unsatisfying state of affairs. As pointed out by Lettau and Ludvigson (2009), among others, a reasonable measure of consumption should allow for a low estimate of relative risk aversion *and* low pricing errors.

In this section, I investigate how unfiltered NIPA consumption compares in cross-sectional asset pricing tests. A linearized version of the consumption-based asset pricing model implies the following (cross-sectional) one-factor return-beta representation (Breeden, Gibbons, and Litzenberger (1989)):

$$E\left[R^e_{i,t+1}\right] = \lambda_0 + \lambda_C \beta_i, \tag{20}$$

where $E[R^e_{i,t+1}]$ is the expected excess return of asset $i$, $\lambda_0$ is a common pricing error equal to zero in theory, and $\lambda_C$ is the price of consumption risk. In the benchmark model with CRRA preferences, the price of consumption risk is linked to the coefficient of relative risk aversion by $\gamma \approx \lambda_C/\mathrm{Var}(\triangle C_t)$.

### A. Portfolio Consumption Covariances

Consumption-based asset pricing models predict that assets with high average returns have high consumption covariances while assets with low average returns have low consumption covariances. To investigate this prediction, I sort firms according to their size (market equity), value (book-to-market ratio), investment (growth of total assets), and operating profitability (revenues minus cost of goods sold, minus selling, general, and administrative

[36] In these regressions, I use robust least squares to down-weight extreme observations of pricing errors. Simple OLS results are similar except for reported NIPA consumption in the full sample period.

expenses, minus interest expenses, divided by book equity) into decile portfolios. The definition and timing of firm characteristics and portfolio construction closely follow Fama and French (1993, 1996, 2015, 2016). Portfolio data are restricted to the period from 1960 to 2014 because information on investment and operating profitability is only available in the postwar period. Further portfolio construction details can be found in the Internet Appendix.

Size and value portfolios are very much the standard test assets in the literature. Furthermore, investment and operating profitability portfolios produce large spreads in average returns that are a challenge to return-based factor models like the three-factor model of Fama and French (1993). This observation motivated Fama and French (2015) to propose a new five-factor model that accounts for the investment and profitability premium. Moreover, Fama and French (2016) find that their five-factor model also well captures the average returns of several other anomalies not targeted by the model (see also Hou, Xue, and Zhang (2015)).

Table VIII reports the mean excess return and consumption covariances for each of the 40 portfolios. Results are given for unfiltered NIPA nondurables and services, as in the baseline analysis, as well as for unfiltered NIPA nondurables and fourth-quarter unfiltered NIPA nondurables for robustness. To conserve space, I do not report individual $t$-statistics below each consumption covariance but I provide in the right-hand column the number of portfolios with GMM-based $t$-statistics larger than two.

All portfolios earn positive excess returns, ranging from 4% (high investment growth) to 12% (high value and small size). However, all portfolio consumption covariances computed with reported NIPA consumption are small and insignificant. The largest consumption covariance is 0.03% (high value). As many as 21 of 40 covariances are negative. There is no compelling systematic relation between consumption covariances and average returns.

In contrast, all three unfiltered NIPA-based consumption measures provide large covariances that are in an economically wide range between 0.14% and 0.40%. These covariances are mainly significant. In line with average excess returns, consumption covariances decrease from small to large size portfolios and increase from low to high value portfolios. Similarly, investment portfolios have consumption covariances that mirror average returns when moving from low to high investment deciles, except for the highest investment decile portfolio. It has the lowest return despite a relatively large unfiltered consumption covariance between 0.17% and 0.23% (depending on the measure).

Interestingly, consumption covariances monotonically decrease from low to high operating profitability portfolios. This is the reverse pattern of average profitability returns. Put differently, buying profitable stocks and selling unprofitable stocks insures against consumption risk despite earning positive average returns. From this observation, it is clear that the one-factor consumption-based model cannot price the profitability premium. In the following, I only consider size, value, and investment portfolios in cross-sectional asset pricing tests. When adding operating profitability portfolios to the test

**Table VIII**
**Portfolio Consumption Covariances**

This table reports real average excess returns (×100) and consumption covariances (×100 × 100) of 40 portfolios independently sorted by size, book-to-market, investment growth, and operating profitability. The consumption measures are reported NIPA nondurables and services ("N&S"), unfiltered NIPA nondurables and services, unfiltered NIPA nondurables ("Ndr."), and unfiltered fourth-quarter NIPA nondurables ("Q4-N"). Asset returns are measured from December-to-December ("D-D") or are time-aggregated ("T-A"). $\#t > 2$ indicates the number of GMM-HAC-robust $t$-statistics for consumption covariances that are above two in a given row.

| Consumption Measure | | Postwar Sample (1960–2014) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Portfolio Decile | | | | | | | | | | |
| and Timing of $R_i^e$ | | Low | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | High | $\#t > 2$ |
| Firms Sorted by Size | | | | | | | | | | | | |
| Mean Return, % | | 12.09 | 10.14 | 9.12 | 8.61 | 9.20 | 8.31 | 8.15 | 7.67 | 6.11 | 5.27 | |
| NIPA Consumption Covariances: | | | | | | | | | | | | |
| Reported D-D | N&S | −2.2 | −0.4 | 0.2 | −0.9 | −0.4 | −1.4 | −1.5 | −1.0 | −0.2 | 0.6 | 0 |
| Unfiltered T-A | N&S | 35.9 | 31.3 | 25.3 | 22.2 | 22.2 | 20.6 | 19.6 | 19.0 | 16.5 | 17.3 | 10 |
| Unfiltered T-A | Ndr. | 35.5 | 29.5 | 25.3 | 21.6 | 21.8 | 19.5 | 19.5 | 18.3 | 16.2 | 15.8 | 10 |
| Unfiltered D-D | Q4-N | 39.6 | 35.1 | 30.5 | 28.1 | 29.2 | 25.5 | 24.9 | 23.6 | 21.6 | 20.2 | 10 |
| Firms Sorted by Book-to-Market | | | | | | | | | | | | |
| Mean Return, % | | 4.81 | 6.28 | 5.79 | 5.80 | 6.28 | 7.34 | 7.93 | 8.25 | 9.22 | 11.50 | |
| NIPA Consumption Covariances: | | | | | | | | | | | | |
| Reported D-D | N&S | −1.6 | −0.8 | −1.1 | −1.0 | 1.5 | 1.2 | 1.6 | 2.1 | 2.1 | 3.0 | 0 |
| Unfiltered T-A | N&S | 19.0 | 17.3 | 14.8 | 13.6 | 15.0 | 17.5 | 18.7 | 22.2 | 19.7 | 28.0 | 9 |
| Unfiltered T-A | Ndr. | 16.8 | 16.1 | 14.8 | 14.3 | 16.7 | 17.1 | 18.8 | 21.4 | 20.2 | 26.7 | 7 |
| Unfiltered D-D | Q4-N | 19.3 | 19.8 | 19.9 | 20.2 | 20.1 | 21.3 | 25.0 | 27.0 | 27.0 | 33.0 | 10 |

(*Continued*)

**Table VIII**—*Continued*

| Consumption Measure and Timing of $R_i^e$ | | Low | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | High | $\#t > 2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Postwar Sample (1960–2014) Portfolio Decile | | | | | | | | | | |
| Firms Sorted by Investment Growth | | | | | | | | | | | | |
| Mean Return, % | | 8.77 | 8.19 | 6.82 | 6.16 | 6.09 | 6.13 | 6.70 | 6.39 | 6.33 | 4.05 | |
| | | NIPA Consumption Covariances: | | | | | | | | | | |
| Reported D-D | N&S | −0.9 | −0.2 | 0.6 | 0.8 | −0.8 | 1.8 | 1.9 | 0.0 | −1.5 | −2.4 | 0 |
| Unfiltered T-A | N&S | 25.9 | 21.0 | 15.1 | 15.0 | 14.8 | 16.9 | 18.2 | 20.4 | 18.0 | 19.5 | 10 |
| Unfiltered T-A | Ndr. | 25.6 | 20.9 | 14.7 | 14.4 | 15.1 | 17.0 | 16.7 | 18.6 | 17.3 | 17.0 | 8 |
| Unfiltered D-D | Q4-N | 29.8 | 24.6 | 18.8 | 20.3 | 18.2 | 22.8 | 22.5 | 22.0 | 22.1 | 22.6 | 9 |
| Firms Sorted by Operating Profitability | | | | | | | | | | | | |
| Mean Return, % | | 5.00 | 5.10 | 4.29 | 6.24 | 5.63 | 5.18 | 5.40 | 6.01 | 6.72 | 7.07 | |
| | | NIPA Consumption Covariances: | | | | | | | | | | |
| Reported D-D | N&S | −0.7 | 1.0 | 0.2 | −0.7 | 0.6 | 0.3 | 0.8 | −0.9 | 1.9 | −1.7 | 0 |
| Unfiltered T-A | N&S | 29.2 | 25.9 | 20.0 | 19.4 | 17.2 | 17.0 | 19.4 | 14.8 | 16.1 | 18.4 | 10 |
| Unfiltered T-A | Ndr. | 30.1 | 24.3 | 18.3 | 19.3 | 14.7 | 16.8 | 18.9 | 14.3 | 13.6 | 17.5 | 8 |
| Unfiltered D-D | Q4-N | 33.9 | 30.8 | 23.5 | 24.5 | 20.7 | 19.4 | 24.6 | 20.7 | 21.2 | 18.8 | 9 |

asset space, the results reported in the Internet Appendix corroborate my conjecture.

### *B. Cross-Sectional Asset Pricing Tests*

To estimate equation (20), I apply a GMM variant of the two-stage cross-sectional regression method of Fama and MacBeth (1973). The weighting matrix is set up such that the point estimates are identical to the traditional Fama-MacBeth regressions. Following Cochrane (2005) and Burnside (2011), the first-step consumption growth betas and the second-step risk factor price ($\lambda_C$) are estimated simultaneously via GMM to obtain robust standard errors. More specifically, I compute (i) Fama-MacBeth standard errors that include the Shanken (1992) correction term and (ii) GMM standard errors that also account for heteroskedasticity and autocorrelation (HAC).[37] To check for cross-sectional variation in the first-step betas, I report a $p$-value ($p,\ \beta_i = \beta$) for the test that all assets have the same exposure to consumption growth.[38] The cross-sectional $R^2$, the MAE, and a $\chi^2$-test on the pricing errors of the model (Cochrane (2005)) are reported as measures of model fit. Equation (20) can be estimated with and without including the constant to the estimation problem. I report results for both specifications and provide the formulas used for estimation in the Internet Appendix.

When including the constant ($\lambda_0$), the tested returns are essentially $E[R^e_{i,t+1}] - \lambda_0$. A positive (negative) constant allows the estimation method to reduce (increase) the common *level* of returns among all test assets in order to better fit differences in average returns. However, this means that even if betas perfectly explain the spread in average returns, such that the cross-sectional $R^2$ is one, a large constant would still indicate that the model does not explain the equity premium of the aggregate market with a reasonable coefficient of risk aversion. Estimation without the constant is by construction immune against this issue.

#### *B.1. Results*

Table IX provides estimates of equation (20) based on 31 portfolios sorted by size, value, and investment, plus the equity premium as the test assets. All models are rejected by the $\chi^2$-test on pricing errors when relying on GMM-based inference. However, there are interesting differences in the relative performance of the alternative consumption measures.

Not surprisingly, reported NIPA consumption fails in pricing the test assets. The constant, when included, is 7.4% and thus simply absorbs the average

[37] For the latter, I apply the parametric VARHAC method described by den Haan and Levin (2000). Yogo (2006) and Burnside (2011) argue that if some factors are persistent, as is particularly the case for the P-J measure, the GMM-VARHAC should be preferred to the Newey-West (1987) approach to calculate standard errors. I find that the difference is small for my data.

[38] Large consumption covariances (betas) alone do not rule out the case of weak identification in the cross-sectional pricing regression.

**Table IX**

**Cross-Sectional Tests with 31 Portfolio Returns:Fama-MacBeth Estimates**

The table displays estimates on the consumption risk premium ($\lambda_C$) from the cross-sectional regression of 31 portfolios:

$$\bar{R}_i^e = \lambda_0 + \lambda_C \beta_i + u_i,$$

where $\bar{R}_i^e$ is mean excess return of portfolio $i$, $\lambda_0$ is a constant (common pricing error), and $\beta_i$ is the time-series beta from a first-pass regression of the portfolio return on consumption growth. $t$-statistics are based on FMB-Shanken [or GMM-HAC-robust] standard errors. The lower panel is estimated imposing a zero constant, $\lambda_0 = 0$. $\gamma$ is the implied coefficient of relative risk aversion approximated as $\lambda_C/\text{var}(\triangle c)$. The 31 test assets are independent decile portfolios sorted by size, book-to-market, and investment growth, plus the market portfolio. The consumption measures are garbage as in Savov (2011), reported NIPA consumption, three-year NIPA consumption (P-J, Parker and Julliard (2005)), fourth-quarter to fourth-quarter NIPA consumption (Q4-Q4, Jagannathan and Wang (2007)), unfiltered NIPA nondurables and services, unfiltered NIPA nondurables, and unfiltered fourth-quarter NIPA nondurables. Estimates on garbage are based on a shorter sample period from 1960 to 2007.

| | Postwar Sample (1960–2014) Equity Premium, 10 Size, 10 Value, and 10 Investment Portfolios | | | | | | |
|---|---|---|---|---|---|---|---|
| | Garbage | NIPA Consumption | | | | | |
| | | Reported | Unfiltered | | | | |
| NIPA series | | N&S | N&S | Q4-N&S | N&S | Ndr. | Q4-Ndr. |
| Unfilter | – | Filtered | P-J | Q4-Q4 | $\bar{K} = 0.58$ | $\bar{K} = 0.71$ | $\bar{K} = 0.71$ |
| Timing of $R_i^e$ | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Time-Ag. | Time-Ag. | Dec-Dec |
| | Estimation with Intercept | | | | | | |
| $\lambda_0$ | 1.24 | 7.40 | 4.57 | 0.94 | 1.27 | 0.86 | −0.63 |
| *t-shanken* | (0.28) | (3.14) | (1.19) | (0.21) | (0.36) | (0.22) | (−0.15) |
| *t-varhac* | [0.28] | [2.74] | [0.89] | [0.16] | [0.36] | [0.21] | [−0.14] |
| $\lambda_C$ | 1.71 | 0.37 | 4.42 | 2.00 | 2.04 | 2.40 | 2.76 |
| *t-shanken* | (1.05) | (0.70) | (1.83) | (1.81) | (1.75) | (1.70) | (1.85) |
| *t-varhac* | [1.13] | [0.61] | [2.07] | [2.06] | [2.25] | [2.03] | [2.19] |
| $\gamma$ | 20.69 | 21.48 | 42.56 | 92.16 | 30.06 | 33.05 | 32.20 |
| $R^2$ | 0.15 | 0.03 | 0.59 | 0.64 | 0.67 | 0.75 | 0.81 |

(*Continued*)

**Table IX**—*Continued*

| | Postwar Sample (1960–2014) Equity Premium, 10 Size, 10 Value, & 10 Investment Portfolios | | | | | | |
|---|---|---|---|---|---|---|---|
| | Garbage | NIPA Consumption | | | | | |
| | | Reported | Unfiltered | | | | |
| NIPA series | | N&S | N&S | Q4-N&S | N&S | Ndr. | Q4-Ndr. |
| Unfilter | – | filtered | P-J | Q4-Q4 | $\bar{K} = 0.58$ | $\bar{K} = 0.71$ | $\bar{K} = 0.71$ |
| Timing of $R_i^e$ | Dec-Dec | Dec-Dec | Dec-Dec | Dec-Dec | Time-Ag. | Time-Ag. | Dec-Dec |
| MAE | 1.36 | 1.52 | 0.99 | 0.87 | 0.78 | 0.71 | 0.60 |
| $p,\ \beta_i = \beta$ | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| $\chi^2$ | (0.183) | (0.004) | (0.904) | (0.888) | (0.023) | (0.067) | (0.415) |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| | Estimation without Intercept ($\lambda_0 = 0$) | | | | | | |
| $\lambda_C$ | 2.09 | 0.31 | 10.14 | 2.28 | 2.44 | 2.69 | 2.55 |
| *t-shanken* | (2.51) | (0.57) | (1.05) | (1.87) | (2.59) | (2.50) | (2.64) |
| *t-varhac* | [2.36] | [0.56] | [0.87] | [1.36] | [2.40] | [2.36] | [2.24] |
| $\gamma$ | 25.25 | 17.64 | 97.75 | 104.84 | 35.95 | 37.15 | 29.79 |
| $R^2$ | 0.14 | −16.19 | −0.65 | 0.63 | 0.64 | 0.74 | 0.80 |

of the test asset returns (about 7%). The $R^2$ is zero, and the MAE is large at 1.5%. Reported NIPA consumption does not explain the level of average returns or their cross-sectional dispersion. For garbage, the constant is considerably smaller (1.2%) and insignificant. However, the $R^2$ is small, and the MAE of 1.4% is not much lower than that of reported NIPA consumption. Dropping the constant leads to a consumption factor price of 2.1%, which implies a coefficient of relative risk aversion of 25. Garbage can explain the level of average returns with a (relatively) low coefficient of relative risk aversion but not the spread in portfolio returns.

For the P-J measure, I find an $R^2$ close to 60% and a lower MAE of 1%. The cross-sectional slope $\lambda_C$ is significant but implies a relative risk aversion of 43 when I allow for a common pricing error. However, the large constant of 4.6% already indicates that these results are not robust to estimation without a constant. Indeed, in the lower panel with a zero constant imposed, the high $R^2$ vanishes and the implied relative risk aversion is as high as 98. Q4-Q4 consumption produces high $R^2$ above 60% and a low MAE of 0.9%. These results are robust to whether or not a constant is included, as pointed out by Jagannathan and Wang (2007). However, the slope $\lambda_C$ is quite steep in both specifications and implies that relative risk aversion is 92 (105, excluding the constant). Both the P-J and Q4-Q4 measures can explain much of the cross-sectional dispersion in average returns but not the level of average returns.

Against this backdrop, I find for all three unfiltered NIPA consumption measures that the $R^2$ is above 60% and the MAE is 0.8% or even lower. The price of consumption risk $\lambda_C$ is between 2.0% and 2.8%, and these slopes imply a relative risk aversion between 30 and 33. The constant is tiny, and thus these results are fairly robust to estimation with or without a constant.

In summary, I find that unfiltered NIPA consumption can explain a large fraction of average returns of portfolios sorted by size, value, and investment. Compared to other alternative consumption measures, I find that unfiltered NIPA consumption allows for a relatively low estimate of relative risk aversion, going hand in hand with relatively low pricing errors.

### *B.2. Additional Results*

The Internet Appendix provides further results. The conclusions mentioned above are fairly robust to adding profitability portfolios to the test assets. Estimates of $\lambda_C$ remain largely unchanged, and as expected, $R^2$s decrease and MAEs increase for all alternative consumption measures. Further, excluding investment portfolios from the test assets gives estimates of $\lambda_C$ and $\gamma$ that are largely unchanged in terms of model fit. For value and size portfolios, I can examine results for the full sample period. Results for reported NIPA consumption do not change much but further improve for unfiltered NIPA consumption. In the longer sample, $t$-statistics for the price of consumption risk $\lambda_C$ are well above two (three) when estimation is with (without) a constant, and the implied coefficient of relative risk aversion is below 20.

I also apply the novel method proposed by Dittmar and Lundblad (2015) to compute the consumption risk exposure of individual firms. These results do not depend on how portfolios are constructed and exploit the full cross-section of individual firms. For unfiltered NIPA consumption, the spread in average returns between high and low consumption risk decile portfolios is about 7% per annum with a $t$-statistic of 2.6. Stocks with high consumption risk tend to be small value. Stocks with low consumption risk are big growth. The relationship between consumption risk and investment is negative but somewhat flatter. These findings are in line with the portfolio-level results reported in this section.

## IV. Implications for Advanced Asset Pricing Models

CRRA preferences are an informative benchmark to focus on consumption risk measured by contemporaneous stock market covariances. The literature has long moved on to more complex preference functions to explain the equity premium, return predictability by the price-dividend ratio, and the risk-free rate puzzle. However, there is an ongoing debate on the true properties of consumption growth and, at this point, it is still pretty unclear how a reasonable consumption process should be calibrated in asset pricing models (see, e.g., Hansen, Heaton, and Li (2008), Marakani (2009), Constantinides and Gosh (2011), Beeler and Campbell (2012), Bansal, Kiku, and Yaron (2012), and Dew-Becker (2016)).

The filter model generates a process for reported NIPA consumption that matches consumption data fairly well (recall Figure 2). For that reason, it is interesting to investigate the implications of filtered consumption data for asset pricing models that go beyond CRRA preferences. In particular, I review implications for preferences with external habit formation (Campbell and Cochrane (1999)) and for recursive preferences in a long-run risk model (Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2012)). To this end, I simply feed my "true" consumption growth process, that is, consumption before time-aggregation, filtering, or measurement error—as inferred from the filter model—into the habit and long-run risk model and keep (most of) the other model parameters unchanged. The discussion is therefore intended to be illustrative rather than to provide a comprehensive re-calibration of the models.[39] Further details on the models and the computational steps can be found in the Internet Appendix.

### *A. External Habit Model*

The external habit model of Campbell and Cochrane (1999; CC) relies on a standard i.i.d. consumption growth process and a slow-moving external habit

[39] See Ludvigson (2013) for a review of the empirical performance of both models from a broader perspective.

that is added to standard power utility preferences.[40] The filter model mitigates three issues when the external habit model is confronted with empirical data.

First, CC calibrate the consumption standard deviation to 1.5% to match reported NIPA consumption in the postwar data. Because consumption growth is i.i.d., the long-run consumption standard deviation equals its short-run counterpart. However, in the data, the long-run standard deviation is 2.5% (e.g., implied by unfiltered NIPA consumption or the estimate of Dew-Becker (2016)). The filter model resolves this issue as it drives a wedge between true consumption (which enters preferences) and reported consumption (which is traditionally used by the researcher for calibration). Second, the relative risk aversion ($\gamma/S$) in the habit model is high—about 35 when measured in the steady state.[41] According to the filter model, the economy is riskier by a factor of 1.67 (=2.5%/1.5%) and therefore the relative risk aversion is (approximately) 40% (=1−1/1.67) lower. Third, the habit model implies an unconditional correlation between stock returns and consumption growth that is larger than what can be measured with contemporaneous reported NIPA consumption. However, replacing reported NIPA consumption with unfiltered NIPA consumption resolves this issue. Model-implied correlations are now even slightly lower, as in the data.

Table X (left-hand side) quantifies the possible improvement of the external habit model when accounting for filtering inherent in reported NIPA consumption. The first column provides simulation results based on the well-known original CC calibration.[42] In the second column, I increase the consumption standard deviation to 2.5%. The equity premium is virtually unaffected by this change. However, the relative risk aversion in the steady state ($\gamma/\bar{S}$) drops from 35 to 21. Increasing the consumption standard deviation increases the covariance of observed consumption (e.g., unfiltered NIPA consumption or garbage) from 0.14% to 0.24%. This covariance is substantially larger than what can be measured when using reported NIPA consumption but is close to the values as measured by unfiltered NIPA consumption (or garbage) in the postwar data. The third column provides evidence on how much the correlation between consumption and stock returns must be increased to get an equity premium of ~5%, while fixing the relative risk aversion at 15. To this end, I increase the correlation between consumption and dividends from 20% (the CC value) to 27%. Furthermore, I decrease the curvature parameter $\gamma$ from 2 to 1.015, such that, in the steady state, I get $\gamma/\bar{S} = 15$. The model-implied consumption

[40] Earlier theoretical studies on preferences with habits are Ryder and Heal (1973), Sundaresan (1989), and Abel (1990). Wachter (2006) uses habit preferences to explain bond term premia and Verdelhan (2010) to explain FX premia.

[41] However, Campbell and Cochrane (1999) point out that the habit model avoids the risk-free rate puzzle even with a high relative risk aversion.

[42] The results differ from those reported in Campbell and Cochrane (1999) for two reasons. First, following Wachter (2005), I use an improved, very fine grid when solving the model. Second, I simulate the model for a sample period of 55 years to get finite-sample moments (in particular, autocorrelations show the finite-sample downward bias). In contrast, Campbell and Cochrane (1999) provide results for one draw of a very large sample, that is, population moments.

**Table X**

**Asset Pricing without Garbage in the External Habit or Long-Run Risk Model**

The statistics displayed are medians of 10,000 simulations of the habit and long-run risk models. Data are simulated monthly and then converted to an annual frequency with 55 observations. Key model parameters are reported at the top. CC results correspond to the parameter calibration as in Campbell and Cochrane (1999). H-I increases the consumption volatility to 2.5%. H-II increases the stock return correlation and decreases risk aversion to 15. BKY results are for the parameter calibration as in Bansal, Kiku, and Yaron (2012). LRR-I switches long-run risks off (the persistent component in consumption growth) and increases risk aversion to 12.5. LRR-II increases the stock return correlation via the cash flow channel. The consumption measure to compute correlations and covariances is observed consumption (i.e., true consumption measured with error).

| | External Habit Model | | | Long-Run Risk Model | | |
|---|---|---|---|---|---|---|
| | CC | H-I | H-II | BKY | LRR-I | LRR-II |
| $\gamma/\bar{S}$, or $\gamma$ | 35 | 21 | 15 | 10 | 12.5 | 12.5 |
| $\sigma(\triangle c)$ | 1.5% | 2.5% | 2.5% | 2.5% | 2.5% | 2.5% |
| LR Risk ($\rho$) | 0 | 0 | 0 | .975 | 0 | 0 |
| Vola Risk ($\nu$) | 0 | 0 | 0 | .999 | .999 | .999 |
| corr($\triangle c$,$\triangle d$) | 20% | 20% | 27% | 40% | 40% | 65% |
| | | | Equity Premium | | | |
| E$[R^e_m]$ | 5.28 | 5.19 | 4.99 | 6.67 | 2.31 | 5.59 |
| $\sigma$ $[R^e_m]$ | 15.71 | 15.58 | 18.39 | 19.90 | 17.20 | 20.95 |
| Corr(0,0)$[\triangle c, r^e_m]$ | 50.16 | 57.84 | 64.40 | 33.70 | 35.90 | 56.57 |
| Cov(0,0)$[\triangle c, r^e_m]$ | 14.37 | 23.47 | 30.96 | 18.78 | 16.27 | 31.13 |
| Cov(0,1)$[\triangle c, r^e_m]$ | −0.05 | 0.02 | −0.22 | 3.99 | −0.32 | −0.54 |
| | | | log P/D Ratios | | | |
| E$[pd]$ | 3.52 | 3.54 | 3.79 | 3.15 | 4.68 | 3.12 |
| $\sigma$ $[pd]$ | 0.13 | 0.13 | 0.16 | 0.17 | 0.09 | 0.15 |
| AC(1)$[pd]$ | 80.24 | 76.58 | 76.73 | 56.43 | 1.71 | 44.32 |
| | | | Return Predictability: $r^e_{m;t+1} = a + b_{pd}\, pd_t + e_{t+1}$ | | | |
| $b_{pd}$ | −0.20 | −0.20 | −0.21 | −0.09 | −0.03 | −0.10 |
| $R^2_{pd}$ | 0.03 | 0.03 | 0.04 | 0.01 | 0.01 | 0.01 |

covariance is now 0.31%. This is somewhat larger than what I find for unfiltered NIPA consumption in the postwar data, but this value is arguably still close enough to be plausible (and still smaller compared to the full sample period).

The table also reports the properties of the log price-dividend ratio and results from return predictability regressions. The results remain mainly unchanged as the properties of the price-dividend ratio are dominated by the parameter that controls the persistence of the consumption-surplus ratio, which is the same across all three calibrations. Notice that the properties of the price-dividend ratio are quite close to recent empirical evidence when the price-dividend ratio is structural break-adjusted and mid-year dividends are not reinvested in the stock market (Lettau and Nieuwerburgh (2008), Binsbergen and Koijen (2010)).[43]

To summarize, adding the filter model to the habit model results in a consumption process that matches the postwar data well. Consumption correlations and covariances are in line with unfiltered NIPA consumption. The external habit model can explain the equity premium with a lower coefficient of relative risk aversion than before. And as in the original calibration, there is no risk-free rate puzzle and the equity premium can be predicted by the price-dividend ratio.

### *B. Long-Run Risk Model*

In the long-run risk model of Bansal and Yaron (2004), consumption growth has a persistent component and its volatility is time-varying.[44] Investors have recursive preferences (Epstein and Zin (1989, 1991)). The persistent component in consumption growth together with recursive preferences imply that the covariance of contemporaneous and future consumption growth determine the equity premium. Time-varying consumption volatility adds time-variation to the conditional expected equity premium. As a result, the long-run risk model generates a large equity premium, return predictability by the price-dividend ratio, and a low risk-free rate. The persistent component in consumption growth plays a dominant role in explaining the equity premium in the original calibration of the long-run risk model of Bansal and Yaron (2004). Movements in consumption volatility are more important for stock prices in the more recent calibration of Bansal, Kiku, and Yaron (2012).[45]

The degree of persistence in consumption growth has received substantial attention in recent literature. An interesting feature of the filter model is that it generates a persistent component in reported NIPA consumption even when it is not there in true consumption. Thus, a researcher may falsely conclude that consumption growth is persistent when inference is based solely on reported NIPA consumption and ignores the properties of observed consumption.

Figure 5 compares short-run and long-run consumption standard deviations of observed, time-aggregated, and both filtered and time-aggregated consumption according to (i) the Bansal, Kiku, and Yaron (2012 BKY) calibration of the consumption process, (ii) a version that switches the persistent component off (which is almost the same process as in Table I), and (iii) the empirical postwar data. The consumption standard deviations according to the BKY calibration

[43] The structural break adjustment reduces the autocorrelation and the standard deviation of the dividend-price ratio from 0.91 to 0.61 and from 0.42 to 0.20 (Lettau and Nieuwerburgh (2008)). Assuming that dividends are reinvested at the risk-free rate and taking structural breaks into account, Koijen and Nieuwerburgh (2011) get a return predictability coefficient of −0.21 (−0.32) in the 1926–2009 (1945–2009) sample period.

[44] Bansal and Shaliastovich (2012) use the long-run risk model to explain bond and FX premia; Bollerslev, Tauchen, and Zhou (2009) extend the model to study the variance risk premium.

[45] I focus on the Bansal, Kiku, and Yaron (2012) calibration of the long-run risk model. See Beeler and Campbell (2012) for a detailed comparison of both calibrations.

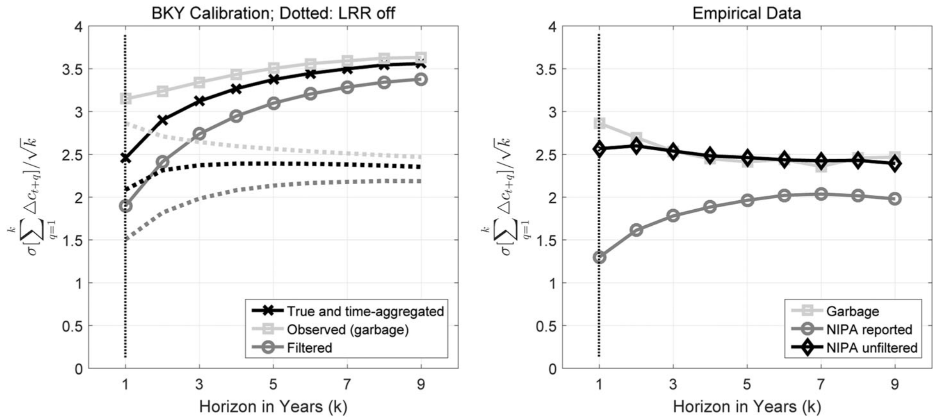


**Figure 5. Long-run consumption standard deviation in the long-run risk model.** The figure on the left shows simulated long-run consumption standard deviations of alternative consumption measures according to the long-run risk model. Bold lines are based on the calibration as in Bansal, Kiku, and Yaron (2012). Dotted lines show results when the long-run risk channel is switched off (LRR-II/III in Table X). The figure on the right provides empirical counterparts for the period 1960 to 2014.

are above their empirical counterparts at every horizon. In contrast to empirical evidence, consumption standard deviations of observed consumption increase with the horizon. However, switching the persistent component off and adding the filter to reported NIPA consumption matches the data reasonably well.[46]

Table X provides simulation results for the equity premium in the long-run risk model. The first column on the right side of the table reports results for the well-known BKY calibration with a persistent component in true consumption growth. The next column shows how the properties of the equity premium change when the persistent component is switched off and the coefficient of risk aversion is increased slightly from 10 to 12.5. The equity premium is only 2.3%, and the price-dividend ratio is no longer persistent or shows strong return predictability. However, the consumption covariance is now lower than what can be measured when relying on unfiltered NIPA consumption. Therefore, the third calibration increases the consumption sensitivity of stock returns to bring the contemporaneous consumption covariance close to the empirical data. I get a large equity premium of 5.6%, and a volatile and persistent price-dividend ratio with similar return predictability properties as in the BKY calibration. What drives this result? Contemporaneous consumption is now as important for the

[46] BKY calibrate the long-run risk model to the full sample period and not to postwar data. I consider whether there is a better way to fit a consumption process with a persistent component to the postwar data. This turns out to be difficult. First, because of the high degree of consumption volatility persistence, it is not sufficient to lower the short-run consumption standard deviation to match the postwar data. Second, a lower short-run consumption standard deviation also implies lower consumption covariances, and the equity premium is driven almost completely by consumption volatility risk. But then neither garbage nor unfiltered NIPA consumption growth should work.

equity premium as contemporaneous and future consumption in the BKY calibration. It turns out that this is sufficient for time-varying consumption volatility to matter and to induce return predictability by the price-dividend ratio.[47]

Of course, the filter model does not provide definitive proof that a persistent component in consumption does not exist. However, the filter model without a persistent component in true consumption matches the postwar data better. It is able to explain the observed differences between the properties of reported NIPA nondurables and NIPA services. The high consumption covariances of observed consumption (unfiltered NIPA consumption or garbage) remove the necessity of explaining a large equity premium with covariances with future consumption growth. What this exercise shows is that a persistent component is not needed at all costs to get a large equity premium and return predictability.

## V. Conclusion

This paper provides an explanation for why garbage-based asset pricing works better than NIPA consumption-based asset pricing. NIPA statisticians filter observable consumption to mitigate measurement error and obtain "optimal" estimates of the consumption level. Garbage is not subject to such filtering and shares the properties of observable consumption.

From an asset pricing perspective, observable consumption is eligible for calculating consumption covariances and there is no rationale for applying a filter to it. On the contrary, I show that filtering in combination with time-aggregation is disastrous for gauging the consumption risk of stock market returns. I apply a model to abstract the comprehensive and complex NIPA consumption estimation procedure. The filter model allows the filtering of the data to be reversed and the time-aggregation bias to be mitigated. I find that unfiltered NIPA consumption works well in explaining the average return of the market portfolio and the cross-section of stock returns with a relatively low coefficient of relative risk aversion.

Unfiltered NIPA consumption together with CRRA preferences cannot account for the predictability of stock returns. However, the new measure of consumption is complementary to any model that relies on a more elaborate specification of preferences and is able to generate time-varying expected returns. Specifically, I find that the fit of the habit model to the data can be improved. In the long-run risk model, short-run consumption risk becomes more important at the cost of long-run risk, while there remains an important role for consumption volatility risk.

In this paper, I use a simplified model of the true "appraisal" process applied by BEA statisticians to estimate NIPA consumption. My measure of unfiltered

[47] In the long-run risk model, the equity premium can be decomposed into short-run consumption risk, long-run consumption risk, and consumption volatility risk. In the third calibration, consumption volatility risk accounts for 1.70% of the equity premium in contrast to 2.75% in the BKY calibration. In this sense, consumption volatility risk is somewhat less important than in the original BKY calibration. The complete return decomposition is provided in the Internet Appendix.

NIPA consumption is a first step to account for the fact that reported NIPA consumption is estimated. Future research could evaluate how to modify the true appraisal process such that macroeconomic data optimized for the needs of economic research can be made available, thus making asset pricing without garbage ultimately viable.

Initial submission: November 21, 2014; Accepted: January 10, 2016
Editors: Bruno Biais, Michael R. Roberts, and Kenneth J. Singleton

## Appendix: Kalman Filter with GARCH Disturbances

To run the (un)filter requires an empirical counterpart for $\eta^{*2}_{t-1}$ in equation (8). Following Harvey, Ruiz, and Sentana (1992) and Broto and Ruiz (2006), I extend the model representation such that the Kalman equations allow me to derive such an expression.

The measurement equation is extended to capture true consumption, lagged true consumption, and the economic shocks,

$$\boldsymbol{X}_t = \begin{bmatrix} c_t \\ c_{t-1} \\ \eta_t \end{bmatrix}, \tag{A1}$$

and satisfies

$$\boldsymbol{X}_t = \mathbf{T}\boldsymbol{X}_{t-1} + \boldsymbol{R}\eta^*_t = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \boldsymbol{X}_{t-1} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \eta^*_t, \tag{A2}$$

where the economic shocks follow the GARCH process

$$\eta^*_t = \eta_t \sigma_{\eta,t}, \;\; \eta_t \sim N(0,1), \tag{A3}$$

$$\sigma^2_{\eta,t} = a_0 + a_1 \eta^{*2}_{t-1} + a_2 \sigma^2_{\eta,t-1}. \tag{A4}$$

The measurement equation (observed consumption) is

$$y_t = \boldsymbol{Z}\boldsymbol{X}_t + \sigma_\xi \xi_t = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \boldsymbol{X}_t + \sigma_\xi \xi_t, \tag{A5}$$

where the measurement error $\xi_t$ is standard normally distributed; $\xi_t$ and $\eta_t$ are mutually independent. The usual Kalman filter equations are

$$\boldsymbol{X}_{t|t-1} = \boldsymbol{T}\boldsymbol{X}_{t-1|t-1}, \tag{A6}$$

$$\boldsymbol{P}_{t|t-1} = \boldsymbol{T}\boldsymbol{P}_{t-1|t-1}\boldsymbol{T}' + \boldsymbol{R}\sigma^2_{\eta,t}\boldsymbol{R}', \tag{A7}$$

$$v_t = y_t - \boldsymbol{Z}\boldsymbol{X}_{t|t-1}, \tag{A8}$$

$$f_t = \boldsymbol{Z}\boldsymbol{P}_{t|t-1}\boldsymbol{Z}' + \sigma^2_\xi, \tag{A9}$$

$$\boldsymbol{X}_{t|t} = \boldsymbol{X}_{t|t-1} + \boldsymbol{P}_{t|t-1}\boldsymbol{Z}' f_t^{-1} v_t, \tag{A10}$$

$$\boldsymbol{P}_{t|t} = \boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\boldsymbol{Z}' f_t^{-1} \boldsymbol{Z}\boldsymbol{P}_{t|t-1}, \tag{A11}$$

where $t \mid t-1$ refers to an estimate conditional on information up to time $t-1$. The Kalman gain is defined as $\boldsymbol{K}_t = \boldsymbol{T}\,\boldsymbol{P}_{t|t-1}\boldsymbol{Z}' f_t^{-1}$. Running the recursion requires the calculation of $\sigma^2_{\eta,t}$ in $\boldsymbol{P}_{t|t-1}$. The variance $\sigma^2_{\eta,t}$ is a function of unobservable shocks $\eta^{*2}_{t-1}$. Harvey, Ruiz, and Sentana (1992) propose that $\eta^{*2}_{t-1}$ be replaced by its conditional expectation $E_{t-1}[\eta^{*2}_{t-1}] = \eta^{*2}_{t-1|t-1} + P^{\eta}_{t-1|t-1}$. In the extended state-space representation, $\eta^{*2}_{t-1|t-1}$ is directly identified as the third element of $\boldsymbol{X}_{t-1|t-1}$. Similarly, the conditional variance $P^{\eta}_{t-1|t-1}$ is provided in the lower right corner of $\boldsymbol{P}_{t-1|t-1}$,

$$\boldsymbol{P}_{t-1|t-1} = \begin{bmatrix} P^{c_t}_{t-1|t-1} & \bullet & \bullet \\ \bullet & P^{c_{t-1}}_{t-1|t-1} & \bullet \\ \bullet & \bullet & P^{\eta}_{t-1|t-1} \end{bmatrix}. \tag{A12}$$

A closer look at the Kalman filter equations reveals that $E_t[\eta^{*2}_t]$ (and therefore $E_{t-1}[\eta^{*2}_{t-1}]$) can be also directly computed by

$$\eta^{*}_{t|t} = v_t \sigma^2_{\eta,t-1}\left(P^{c_t}_{t|t-1} + \sigma^2_{\xi}\right)^{-1} \tag{A13}$$

and

$$P^{\eta}_{t|t} = \sigma^2_{\eta,t-1}\left(1 - \sigma^2_{\eta,t-1}\left(P^{c_t}_{t|t-1} + \sigma^2_{\xi}\right)^{-1}\right), \tag{A14}$$

where $P^{c_t}_{t|t-1}$ is identical to $P_t$ in equation (6) of the main paper. Equations (A13) and (A14) are a short cut. They allow me to process the Kalman filter with GARCH disturbances using the reduced filter equations in the main paper.

## REFERENCES

Abel, Andrew B., 1990, Asset prices under habit formation and catching up with the Joneses, *American Economic Review: Papers & Proceedings* 80, 38–42.

Adam, Klaus, Albert Marcet, and Juan P. Nicolini, 2016, Stock market volatility and learning, *Journal of Finance* 71, 33–82.

Aït-Sahalia, Yacine, Jonathan A. Parker, and Motohiro Yogo, 2004, Luxury goods and the equity premium, *Journal of Finance* 59, 2959–3004.

Bansal, Ravi, Varoujan Khatchatrian, and Amir Yaron, 2005, Interpretable asset markets? *European Economic Review* 49, 531–560.

Bansal, Ravi, Dana Kiku, Ivan Shaliastovich, and Amir Yaron, 2014, Volatility, the macroeconomy, and asset prices, *Journal of Finance* 69, 2471–2511.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An empirical evaluation of the long-run risks model for asset prices, *Critical Finance Review* 1, 183–221.

Bansal, Ravi, and Ivan Shaliastovich, 2012, A long-run risks explanation of predictability puzzles in bond and currency markets, *Review of Financial Studies* 26, 1–33.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

BEA, 2014, *Concepts and Methods of the U.S. National Income and Product Accounts* (The NIPA Handbook 2014, U.S. Bureau of Economic Analysis (BEA), Washington DC).

Beeler, Jason, and John Y. Campbell, 2012, The long-run risks model and aggregate asset prices: An empirical assessment, *Critical Finance Review* 1, 141–182.

Bell, William R., and David W. Wilcox, 1993, The effect of sampling error on the time series behavior of consumption data, *Journal of Econometrics* 55, 235–265.

Binsbergen, Jules Van, and Ralph S.J. Koijen, 2010, Predictive regressions: A present-value approach, *Journal of Finance* 65, 1439–1471.

Boguth, Oliver, and Lars-Alexander Kuehn, 2013, Consumption volatility risk, *Journal of Finance* 68, 2589–2615.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265–296.

Breeden, Douglas T., Michael Gibbons, and Robert Litzenberger, 1989, Empirical tests of the consumption-oriented CAPM, *Journal of Finance* 44, 231–262.

Broto, Carmen, and Esther Ruiz, 2006, Unobserved component models with asymmetric conditional variances, *Computational Statistics & Data Analysis* 50, 2146–2166.

Bryzgalova, Svetlana, 2014, Spurious factors in linear asset pricing models, Working paper, Stanford University.

Burnside, Craig, 2011, The forward rate premium is still a puzzle, a comment on "The cross-section of foreign currency risk premia and consumption growth risk," *American Economic Review* 101, 3456–3476.

Burnside, Craig, 2016, Identification and inference in linear stochastic discount factor models, *Journal of Financial Econometrics* 14, 295–330.

Campbell, John Y., 2003, Consumption-based asset pricing, in G. Constantinides, M. Harris, and R. Stulz, eds.: *Handbook of the Economics of Finance* (North-Holland, Amsterdam).

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Campbell, John Y., Stefano Giglio, Christopher Polk, and Robert Turley, 2015, An intertemporal CAPM with stochastic volatility, Working paper, Harvard University.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton, NJ).

Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572–621.

Cochrane, John H., 2005, *Asset Pricing* (Princeton University Press, Princeton, NJ).

Cochrane, John H., and Lars Peter Hansen, 1992, Asset pricing explorations for macroeconomics, *NBER Macroeconomics Annual* 7, 115–165.

Constantinides, George M., and Anisha Gosh, 2011, Asset pricing tests with long-run risk consumption growth, *Review of Asset Pricing Studies* 1, 96–136.

Daniel, Kent, and David Marshall, 1997, Equity-premium and risk-free rate puzzle at long horizons, *Macroeconomic Dynamics* 1, 452–484.

den Haan, Wouter J., and Andrew T. Levin, 2000, Robust covariance matrix estimation with data-dependent VAR prewhitening order, NBER Technical Working paper 255.

Dew-Becker, Ian, 2016, How risky is consumption in the long-run? Benchmark estimates from a robust estimator, *Review of Financial Studies* doi: 10.1093/rfs/hhw015

Dittmar, Robert F., and Christian Lundblad, 2015, Firm characteristics, consumption risk, and firm-level risk exposures, Working paper, University of Michigan.

Duffee, Gregory R., 2005, Time variation in the covariance between stock returns and consumption growth, *Journal of Finance* 60, 1673–1712.

Durbin, James, and Siem Jan Koopman, 2012, *Time Series Analysis by State Space Methods* (Oxford Statistical Science Series, Oxford University Press, Oxford, UK).

Engsted, Tom, and Stig V. Møller, 2011, Cross-sectional consumption-based asset pricing: The importance of consumption timing and the inclusion of severe crises, CREATES Research Paper.

Engsted, Tom, and Stig V. Møller, 2015, Cross-sectional consumption-based asset pricing: A reappraisal, *Economics Letters* 132, 101–104.

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Epstein, Larry G., and Stanley E. Zin, 1991, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis, *Journal of Political Economy* 99, 263–286.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and Kenneth R. French, 2016, Dissecting anomalies with a five-factor model, *Review of Financial Studies* 29, 69–103.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Gabaix, Xavier, and David Laibson, 2001, The 6D bias and the equity premium puzzle, *NBER Macroeconomics Annual* 16, 257–312.

Geltner, David, 1989, Estimating real estate's systematic risk from aggregate level appraisal-based returns, *AREUEA Journal* 17, 463–481.

Geltner, David, 1993, Estimating market values from appraised values without assuming an efficient market, *Journal of Real Estate Research* 8, 325–345.

Gospodinov, Nikolay, Raymond Kan, and Cesare Robotti, 2014, Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors, *Review of Financial Studies* 27, 2139–2170.

Grossman, Sanford J., Angelo Melino, and Robert J. Shiller, 1987, Estimating the continuous-time consumption-based asset pricing model, *Journal of Business and Economic Statistics* 5, 315–327.

Grossman, Sanford J., and Robert J. Shiller, 1980, Preliminary results on the determinants of the variability of stock market prices, Working paper, University of Pennsylvania.

Grossman, Sanford J., and Robert J. Shiller, 1981, The determinants of the variability of stock market prices, *American Economic Review: Papers & Proceedings* 71, 222–227.

Hall, Robert E., 1978, Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence, *Journal of Political Economy* 86, 971–987.

Hall, Robert E., 1988, Intertemporal substitution in consumption, *Journal of Political Economy* 96, 339–357.

Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long-run risk, *Journal of Political Economy* 116, 260–302.

Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.

Hansen, Lars Peter, and Kenneth J. Singleton, 1983, Stochastic consumption, risk aversion, and the temporal behavior of asset returns, *Journal of Political Economy* 91, 249–268.

Hansen, Lars Peter, and Kenneth J. Singleton, 1996, Efficient estimation of linear asset-pricing models with moving average errors, *Journal of Business & Economic Statistics* 14, 53–68.

Harvey, Andrew, Esther Ruiz, and Enrique Sentana, 1992, Unobserved component time series models with ARCH disturbances, *Journal of Econometrics* 52, 129–157.

Hayashi, Fumio, and Christopher Sims, 1983, Nearly efficient estimation of time series models with predetermined, but not exogenous instruments, *Econometrica* 51, 783–798.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Jagannathan, Ravi, and Yong Wang, 2007, Lazy investors, discretionary consumption, and the cross-section of stock returns, *Journal of Finance* 62, 1623–1661.

Jurado, Kylie, Sydney C. Ludvigson, and Serena Ng, 2015, Measuring uncertainty, *American Economic Review* 105, 1177–1216.

Kandel, Shmuel, and Robert F. Stambaugh, 1990, Expectations and volatility of consumption and asset returns, *Review of Financial Studies* 3, 207–232.

Kim, Chang-Jin, and Charles R. Nelson, 1999, *State-Space Models with Regime Switching* (MIT Press, Cambridge, MA).

Koijen, Ralph S.J., and Stijn Van Nieuwerburgh, 2011, Predictability of returns and cash flows, *Annual Review of Financial Economics* 3, 467–491.

Kolev, Gueorgui I., 2013, What generates the equity premium puzzle: Bad model or bad data? Working paper, EDHEC Business School.

Lettau, Martin, and Sydney C. Ludvigson, 2001, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238–1287.

Lettau, Martin, and Sydney C. Ludvigson, 2009, Euler equation errors, *Review of Economic Dynamics* 12, 255–283.

Lettau, Martin, Sydney C. Ludvigson, and Jessica A. Wachter, 2008, The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies* 21, 1653–1687.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the return predictability evidence, *Review of Financial Studies* 21, 1607–1652.

Lucas, Robert E., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429–1445.

Ludvigson, Sydney C., 2013, Advances in consumption-based asset pricing: Empirical tests, in George M. Constantinides, Milton Harris, and René M. Stulz, eds.: *Handbook of the Economics of Finance* (Elsevier Science B.V., North Holland, Amsterdam).

Lynch, Anthony W., 1996, Decision frequency and synchronization across agents: Implications for aggregate consumption and equity return, *Journal of Finance* 51, 1479–1497.

Marakani, Srikant, 2009, Long-run consumption risks: Are they there? Working paper, City University of Hong Kong.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Nagel, Stefan, and Kenneth J. Singleton, 2011, Estimation and evaluation of conditional asset pricing models, *Journal of Finance* 66, 873–909.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Parker, Jonathan A., 2001, The consumption risk of the stock market, *Brookings Papers on Economic Activity* 2, 279–348.

Parker, Jonathan A., 2003, Consumption risk and expected stock returns, *American Economic Review: Papers & Proceedings* 93, 376–382.

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

Piazzesi, Monika, Martin Schneider, and Selale Tuzel, 2007, Housing, consumption and asset pricing, *Journal of Financial Economics* 83, 531–569.

Quan, Daniel C., and John M. Quigley, 1989, Inferring an investment return series for real estate from observations on sales, *AREUEA Journal* 17, 218–234.

Ryder, Harl E., and Geoffrey M. Heal, 1973, Optimum growth with intertemporal dependent preferences, *Review of Economic Dynamics* 40, 1–33.

Santos, Tano, and Pietro Veronesi, 2006, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1–44.

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

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–33.

Singleton, Kenneth J., 1990, Specification and estimation of intertemporal asset pricing models, in B.M. Friedman, and F.H. Hahn, eds.: *Handbook of Monetary Economics* (North Holland, Amsterdam).

Sundaresan, Suresh M., 1989, Intertemorally dependent preferences and the volatility of consumption and wealth, *Review of Financial Studies* 2, 73–99.

Taio, George C., 1972, Asymptotic behaviour of temporal aggregates of time series, *Biometrika* 59, 525–531.

Tédongap, Roméo, 2014, Consumption volatility and the cross-section of stock returns, *Review of Finance* 19, 1–39.

Verdelhan, Adrien, 2010, A habit-based explanation of the exchange rate risk premium, *Journal of Finance* 65, 123–146.

Wachter, Jessica A., 2005, Solving models with external habit, *Finance Research Letters* 2, 210–226.

Wachter, Jessica A., 2006, A consumption-based model of the term structure of interest rates, *Journal of Financial Economics* 79, 365–399.

Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance* 68, 987–1035.

Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 401–422.

Wilcox, David W., 1992, The construction of U.S. consumption data: Some facts and their implications for empirical work, *American Economic Review* 82, 922–941.

Working, Holbrook, 1960, Note on the correlation of first differences of averages in a random chain, *Econometrica* 28, 916–918.

Yogo, Motohiro, 2006, A consumption-based explanation of expected stock returns, *Journal of Finance* 61, 539–580.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Appendix S1**: Internet Appendix.