Asset pricing in the frequency domain: theory and empirics

Ian Dew-Becker and Stefano Giglio*

November 13, 2015

Abstract

We quantify investors' preferences over the dynamics of shocks by deriving frequency-specific risk prices that capture the price of risk of consumption fluctuations at each frequency. The frequency-specific risk prices are derived analytically for leading models. The decomposition helps measure the importance of economic fluctuations at different frequencies. We precisely quantify the meaning of "long-run" in the context of Epstein–Zin preferences – centuries – and measure the exact relevance of business-cycle fluctuations. Last, we estimate frequency-specific risk prices and show that cycles longer than the business cycle – long-run risks – are significantly priced in the equity market.

*Dew-Becker: Northwestern University. Giglio: University of Chicago and NBER. We appreciate helpful comments and discussions from Fernando Alvarez, Francisco Barillas, Rhys Bidder, Jarda Borovicka, John Campbell, Hui Chen, Mike Chernov, John Cochrane, George Constantinides, Darrell Duffie, Lars Hansen, John Heaton, Urban Jermann, Bryan Kelly, Aytek Malkhozov, Nikola Mirkov, Marius Rodriguez, Eric Swanson, Jessica Wachter, and seminar participants at the San Francisco Fed, University of Bergen, University of Wisconsin, Chicago Booth, UC Santa Cruz, Bank of Canada, Kellogg Junior Finance Conference, SED, ITAM, the Macro Finance Society, Kellogg, Fuqua, UNC Kenan-Flagler, UCLA Anderson, the SFS Cavalcade, the Tinbergen Institute, the Asset Pricing Retreat (Tilburg), the SoFiE annual meeting, the Red Rock Finance Conference, and the Western Finance Association.

1 Introduction

This paper develops a novel frequency-domain decomposition of innovations to the pricing kernel. The decomposition quantifies exactly how economic fluctuations at different frequencies are priced and reveals previously overlooked constraints imposed by widely used preference specifications. The frequency-domain tools also lead to a novel estimation method that provides new evidence on the pricing of economic risks that is both stronger and more statistically powerful than previous methods. Low-frequency fluctuations in the economy are significantly priced across a wide range of specifications, while business-cycle and higher-frequency fluctuations are not, which highlights the importance of long-run risks in determining risk premia.

Our frequency-domain decomposition applies to affine asset pricing models, including the CAPM, the consumption CAPM, the standard log-linearized version of Epstein–Zin (1991) preferences, and the ICAPM (Merton, 1973; Campbell, 1993). The dynamic effects of shocks have become central in the recent asset pricing literature, and we argue that the frequency domain is the natural setting in which to analyze dynamics.

The dynamic response of the economy to a shock is usually summarized in the time domain by an impulse response function (IRF). Long-run shocks to consumption growth that have large risk prices under Epstein–Zin preferences – for example, those studied in Bansal and Yaron (2004) – have IRFs that decay slowly. We map the IRF of a shock into the frequency domain. A shock that has strong long-run effects has high power at low frequencies, whereas shocks that dissipate rapidly have relatively more power at high frequencies. We refer to the frequency-domain version of the IRF as the impulse transfer function.

Our theoretical result is that the price of risk for a shock depends on the integral of the impulse transfer function over the set of all frequencies ω, weighted by a function $Z(\omega)$. $Z(\omega)$ measures the frequency-specific price of risk and is determined purely by investor preferences, not the dynamics of consumption. We derive $Z(\omega)$ in closed form for various theoretical models and estimate it empirically in equity markets.

The spectral representation we describe is useful for two main reasons. First, it yields *quantitative* insights about the importance of the dynamics of shocks for asset prices in different models. We show that standard calibrations of Epstein–Zin preferences imply that more than half of the mass of the spectral weighting function lies on cycles lasting a century or longer. While it is certainly understood that Epstein–Zin preferences place weight on low-frequency shocks, this is the first paper to quantify exactly what "long-run" means and show how large the weight on those frequencies is. Similarly, we show that models with internal habit formation place the majority of their mass on high frequencies.[1]

The analysis also reveals that standard preference specifications are very tightly constrained in certain regards, which leads to sharp and testable predictions. Epstein–Zin preferences isolate their weight almost exclusively on very low frequencies, internal habit formation isolates its weight on high frequencies, and both have monotone weighting functions. So if we can even just measure the average slope of the spectral weighting function, we can empirically distinguish the two models. Moreover, due to the monotonicity of the weighting functions, neither model allows investors to express a particular aversion to fluctuations at mid-range frequencies, e.g. business cycles, a constraint that has not been highlighted previously.

The second contribution of the paper is to provide estimates of the spectral weighting function in US equity markets. We begin by showing that when we apply the standard Euler equation estimation methodology of Hansen and Singleton (1982) using the specific functional forms implied by utility-based models, no coefficients are consistently significant across various groups of test assets (and the implied risk aversion parameters at the point estimates are implausibly high, though the confidence bands also cover more plausible val-

[1] See also Epstein et al. (2014), who quantitatively analyze the preference for the timing of the resolution of uncertainty under Epstein–Zin preferences.

ues). That result would usually be taken to imply that consumption is unpriced and that the models are not a good description of risk premia in the equity market. We argue, though, that such a conclusion is premature, and depends on the extremely tight constraints imposed by the models. When we generalize the models to allow investors to price fluctuations in broader ranges of frequencies, we find that low-frequency shocks to consumption growth are in fact consistently priced in equity markets. The key is simply that we must allow for the possibility that the concept of "long-run" that is priced is a shock that lasts longer than the business cycle, rather than a shock that lasts hundreds of years as implied by Epstein–Zin preferences.

In addition to allowing us to estimate more general specifications for the pricing kernel, the frequency-domain analysis also suggests a novel way to test asset pricing models. Standard tests, e.g. the Gibbons, Ross, and Shanken (1989) and GMM overidentifying tests, are often difficult to interpret, because their rejections are in some sense statistical and do not have a clear economic interpretation. The tests tell us that some portfolios are unpriced, but they do not tell us anything about the economic source of the failure.

We suggest instead that models can be tested against parametric alternatives (as suggested by Andrews and Ploberger (1996)). As an example, consider the problem of testing whether habit formation gives a good description of investor preferences. Habit formation implies that the covariance of an asset's return with high-frequency shocks to consumption growth should determine its average return. We test the model by asking whether long-run risks are also priced. When we find that those long-run shocks are significantly priced, not only do we reject habit formation, but we give economic meaning to its failure – it is inconsistent with the fact that investors are averse to long-run risks. Similarly, we do not simply fail to reject that the price of risk for high-frequency fluctuations is zero; we will rather show that low-frequency fluctuations are priced.

To summarize, then, the analysis of asset pricing models in the frequency domain gives two novel results: it quantifies precisely how models place weight on different frequencies (and clarifies exactly what "long-run risk" actually means for standard Epstein–Zin investors), and it delivers novel and consistent evidence showing that investors are averse to low-frequency economic fluctuations.

There is very little extant analysis of preference-based asset pricing in the frequency domain.[2] Otrok, Ravikumar, and Whiteman (2002) and Yu (2012) are two recent examples. While those papers also present spectral decompositions of prices and consumption fluctuations, the object of the decomposition is different from ours. Instead of studying how shocks at different frequencies are priced by an investor, they ask how the price of a consumption claim depends on the spectral density of consumption and its relation with returns. Since the price of the asset reflects a combination of preferences and dynamics, it is impossible to evaluate the relative importance of the two beyond very specific cases.[3,4]

Our paper is closely related to a vast empirical literature studying the importance of dynamics for asset pricing in the time domain. A number of papers study the relationship between asset returns and consumption growth at long horizons as methods of testing the implications of Epstein–Zin or power utility.[5] We complement that work by estimating how fluctuations in consumption growth at different frequencies are priced, and in a way that imposes weaker restrictions.

Finally, our work is related to other important decompositions of the stochastic discount factor (SDF), most notably Alvarez and Jermann (2005), Hansen and Scheinkman (2009)

[2] Frequency-domain tools have been applied in finance for other purposes, for example for estimation or valuation of derivatives, as in Carr and Madan (1999), Duffie, Pan and Singleton (2000), and Singleton (2001).

[3] Calvet and Fisher (2007), Ortu, Tamoni and Tebaldi (2013) and Bandi and Tamoni (2014) study a different decomposition of the consumption and returns processes into components operating at different time scales, exploring their covariance and relation with expected returns at different time scales, in reduced form and within the framework of Epstein–Zin utility. The focus of these papers is in disentangling the different components of the consumption and returns processes, while we provide a decomposition of both the consumption processes and, most importantly, the agent's risk preferences, for any utility function.

[4] See also Alvarez and Jermann (2004), who measure the cost of business-cycle fluctuations by computing the price of a claim to the business-cycle component of consumption.

[5] For example, Parker and Julliard (2005); Malloy, Moskowitz, and Vissing-Jorgensen (2009); Bansal, Dittmar, and Lundblad (2005); Yu (2012); Daniel and Marshall (1997); van Binsbergen, Brandt and Koijen (2012); Hansen, Heaton, and Li (2008).

and Borovicka et al. (2011). Those decompositions study the dynamic effects of shocks for the evolution of the stochastic discount factor over time and are closely related to work on the term structure of risk premia.[6] Rather than studying how a single shock today affects the SDF in the *future*, we study how the innovation to the SDF *today* depends on news about consumption in the future. In other words, those papers analyze the impulse response function of the SDF, while we study the impulse response function of consumption and how it affects the one-period innovation in the SDF. The two approaches are complementary. Our decomposition explains risk premia (since the risk premium of any asset depends only on the one-period innovation in the SDF), rather than the term structure of prices of claim to future consumption.

1.1 Notation

We use the following conventions in our notation:

c_t: lower-case italic type represents a scalar variable or function

Z: upper-case italic type represents a scalar-valued function in the frequency domain

$\mathbf{x}_t$: lower-case bold type represents a vector

$\mathbf{\Phi}$: upper-case bold type represents a matrix or matrix-valued function

We also follow standard conventions for denoting widely used operators such as expectations, lags, and first differences.

2 Frequency-specific risk prices

We derive our spectral decomposition of the pricing kernel under two main assumptions. First, the log pricing kernel, m_t, depends on the current and future values of a scalar *priced variable*, x_t (often consumption growth or market returns). Second, the dynamics of the

[6] See, for example, Hansen, Heaton, and Li (2008) and Lettau and Wachter (2007). Backus, Chernov, and Zin (2014) study how the dispersion of the pricing kernel varies by horizon.

economy are described by a vector moving average process $\mathbf{x}_t$ which includes x_t as an element.

Assumption 1: Structure of the SDF.

Denote the log pricing kernel (or stochastic discount factor, SDF) m_{t+1}.[7] We assume that m_t depends on current and future values of x_t:

$$m_{t+1} = f(\mathcal{I}_t) - \Delta\mathbb{E}_{t+1} \sum_{k=0}^{\infty} z_k x_{t+1+k} \tag{1}$$

where $f(\mathcal{I}_t)$ is some unspecified function of the time-t information set $\mathcal{I}_t$, $\mathbb{E}_t$ is the expectation operator conditional on information available on date t, and $\Delta\mathbb{E}_{t+1} \equiv \mathbb{E}_{t+1} - \mathbb{E}_t$ denotes the innovation in expectations. This specification is sufficiently flexible to match standard empirical applications of power utility, habit formation, Epstein–Zin preferences, the CAPM and the ICAPM (in some cases under log-linearization). Intuitively, equation (1) simply says that the innovation to the SDF depends on news about the priced variable in the future, with weights z_k at each future horizon k. It implies that risk prices are constant, but we discuss how to relax that assumption below.

Assumption 2: Dynamics of the economy.

$\mathbf{x}_t$ is driven by an n-dimensional vector moving average process

$$x_t = \mathbf{b}_1 \mathbf{x}_t \tag{2}$$

$$\mathbf{x}_t = \mathbf{\Gamma}(L)\, \boldsymbol{\varepsilon}_t \tag{3}$$

where $\mathbf{x}_t$ has dimension $n \times 1$, L is the lag operator, $\mathbf{\Gamma}(L)$ is an $n \times n$ matrix lag polynomial,

$$\mathbf{\Gamma}(L) = \sum_{k=0}^{\infty} \mathbf{\Gamma}_k L^k \tag{4}$$

[7] In our derivation we assume that the log SDF, m_{t+1}, is linear in the news about future values of the priced variable x_t, because the most widely used models specify an affine form for the log SDF. The same decomposition holds if we assume that the level of the SDF, $\exp(m_{t+1})$, is linear in the news terms. We also do not take a position on whether m_t is the pricing kernel for all markets or whether there is some sort of market segmentation, nor do we assume at this point that there is a representative investor.

and $\boldsymbol{\varepsilon}_t$ is an $n \times 1$ vector of (potentially correlated) martingale difference sequences. Note that we make no assumptions about the conditional distribution of the innovations $\boldsymbol{\varepsilon}_t$ except that it has a mean of zero. $\boldsymbol{\varepsilon}_t$ therefore could include disasters, it could be heteroskedastic, and it could have fat tails. Furthermore, note that the elements of $\boldsymbol{\varepsilon}_t$ need not be orthogonal or in any sense represent identified "structural" shocks as in the structural vector autoregression literature. For example, they could all have nonzero correlations. Finally, we do not at this point make any specific assumptions about the function $\boldsymbol{\Gamma}(L)$. Different asset pricing models will place different constraints on admissible forms of $\boldsymbol{\Gamma}(L)$. We instead make the high-level assumption that the variance of m_{t+1} is finite, which will imply constraints on $\boldsymbol{\Gamma}(L)$ depending on $\{z_k\}$.

Throughout the paper $\mathbf{b}_j$ denotes a conformable (here, $1 \times n$) vector equal to 1 in element j and zero elsewhere. We assume without loss of generality that x_t is the first element of $\mathbf{x}_t$. Furthermore, we require $\boldsymbol{\Gamma}(L)$ to have properties sufficient to ensure that x_t is covariance stationary.

Combining assumptions 1 and 2, we can write the innovations to the pricing kernel as a function of the impulse-response functions (IRFs) of x_t to each of the shocks. In particular, for the jth element of $\boldsymbol{\varepsilon}_t$, $\varepsilon_{j,t}$, the IRF of x_t is the set of $g_{j,k}$ for all horizons k defined as:

$$g_{j,k} \equiv \begin{cases} \mathbf{b}_1 \boldsymbol{\Gamma}_k \mathbf{b}_j' \text{ for } k \geq 0 \\ 0 \text{ otherwise} \end{cases} \tag{5}$$

We can then rewrite the innovation to the log SDF as:

$$\Delta \mathbb{E}_{t+1} m_{t+1} = -\sum_j \left(\sum_{k=0}^{\infty} z_k \, g_{j,k} \right) \varepsilon_{j,t+1} \tag{6}$$

and we refer to $\sum_{k=0}^{\infty} z_k g_{j,k}$ as the *price of risk* for shock j. In this representation, the effect of a shock $\varepsilon_{j,t+1}$ on the pricing kernel is decomposed by horizon: for every horizon k, the effect of the shock depends on the response of x at that horizon (captured by $g_{j,k}$) and on the horizon-specific price of risk z_k.

Our main result is a spectral decomposition in which the price of risk of a shock j depends on the response of x to that shock at each frequency ω ($G_j(\omega)$) and on a frequency-specific price of risk, $Z(\omega)$ (see the appendix for all proofs).

Result 1. *Under Assumptions 1 and 2, the innovations to the log SDF are*

$$\Delta\mathbb{E}_{t+1}m_{t+1} = -\sum_j \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} Z(\omega)\, G_j(\omega)\, d\omega\right)\varepsilon_{j,t+1} \tag{7}$$

where $Z(\omega)$ is a weighting function depending only on the risk prices $\{z_k\}$, and $G_j(\omega)$ measures the dynamic effects of $\varepsilon_{j,t}$ on x in the frequency domain,

$$Z(\omega) \equiv z_0 + 2\sum_{k=1}^{\infty} z_k \cos(\omega k) \tag{8}$$

$$G_j(\omega) \equiv \sum_{k=0}^{\infty} \cos(\omega k)\, g_{j,k} \tag{9}$$

The price of risk for a shock ε_j is then

$$\sum_{k=0}^{\infty} z_k g_{j,k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} Z(\omega)\, G_j(\omega)\, d\omega \tag{10}$$

Equation (10) allows us to represent the information contained in the infinitely long IRF $\{g_{j,k}\}$ and the infinite set of weights $\{z_k\}$ in a compact and interpretable pair of functions on a bounded interval. The function $G_j(\omega)$ decomposes the effects of a shock by frequency. If $\varepsilon_{j,t}$ has very long-lasting effects on x, it induces low-frequency cycles in consumption, and most of the mass of $G_j(\omega)$ will lie at low frequencies. If $\varepsilon_{j,t}$ induces mainly transitory dynamics in x, then $G_j(\omega)$ will isolate high frequencies. We refer to G_j as the *impulse transfer function* (ITF) of shock j since it is the real part of the transfer function associated with the filter $\sum_{k=0}^{\infty} g_{j,k}L^k$.[8]

[8] There is an alternative way to see how $G_j(\omega)$ maps into the IRF which is clearest in continuous time. Denote the IRF at horizon k as $g_j(k)$. The ITF is then $G_j(\omega) = \int_0^{\infty} g_j(k)\cos(\omega k)\, dk$. The inverse

The price of risk for shock ε_j thus depends on an integral over the function $G_j(\omega)$, with weights $Z(\omega)$. Since $G_j(\omega)$ tells us the effect of ε_j on x at frequency ω, we interpret $Z(\omega)$ as the price of risk for any shock to the variable x at frequency ω. $Z(\omega)$ is also akin to a density that G is integrated over (though this density may be negative). We will thus often discuss shifts in the mass of Z as changes that increase Z for certain ranges of frequencies and reduce it elsewhere.

The weighting function $Z(\omega)$ does not tell us anything about the dynamics of the pricing kernel, m_{t+1}. Rather, $Z(\omega)$ tells us how the dynamic features of any given shock $\varepsilon_{j,t+1}$ map into the *innovation* in the pricing kernel, $\Delta\mathbb{E}_{t+1}m_{t+1}$, which is what is relevant for calculating risk premia and expected returns. So whereas, for example, Borovicka et al. (2011) study generalizations of the IRF of the SDF itself, we study how the IRF of consumption growth affects the SDF purely on date $t+1$.

It is possibly surprising that the distribution of $\boldsymbol{\varepsilon}_t$ is irrelevant for our analysis. The irrelevance is due to the fact that we separate the *price* of risk from the *quantity* of risk. The volatility of the pricing kernel, and hence the size of risk premia in the economy, depends on both the price of risk of each shock j, $\sum_{k=0}^{\infty} z_k\, g_{j,k}$, and the quantity of risk, determined by the distribution of $\boldsymbol{\varepsilon}_{j,t}$. Since we decompose only the price of risk, our result holds independently of any assumptions about the distribution of the shocks. We thus have a three-way separation between frequency-specific risk prices, $Z(\omega)$, frequency-specific power, $G_j(\omega)$, and the quantity of risk, determined by the distribution of ε_j.[9]

Finally, note that, unlike analyses of the term structure of risk premia, including Borovicka et al. (2011), equation (10) gives a complete separation between the dynamics of the economy and risk prices. In particular, the frequency-specific risk prices $Z(\omega)$ depend purely

transformation is $g_j(k) = \pi^{-1}\int_{-\pi}^{\pi} G_j(\omega)\cos(\omega k)\,d\omega$. G_j can be viewed as a decomposition of the IRF into cosines, and $G_j(\omega)$ is the contribution to the IRF from the cosine with frequency ω.

[9] Note also that we are not unique in using frequency-domain analysis in the presence of potential heteroskedasticity. For example, the Newey–West (1987) estimator of the spectral density at frequency zero is specifically intended to be used in the presence of heteroskedasticity. Standard frequency-domain results in the econometrics literature rely on second-order stationarity rather than homoskedasticity or serial independence.

on preferences, so our results about Z in different models are general characterizations of preferences, rather than statements that depend on specific calibrations of a consumption process.

2.1 Examples of impulse transfer functions $G_j(\omega)$

Figure 1 plots the impulse response and impulse transfer functions for four different hypothetical shocks. For the sake of concreteness, think of the priced variable x_t as log consumption growth, Δc_t. While we are ultimately interested in the effects of the shocks on log consumption growth, Δc_t, since that is what enters the log SDF, for ease of interpretation we plot the IRF in terms of the *level* of consumption, c_t.

The first shock is a simple one-time increase in consumption. This shock has a flat impulse transfer function, indicating it has power at all frequencies. The second shock is a long-run-risk type shock, inducing persistently positive growth, with the level of consumption ultimately reaching the same level as that induced by the first shock. In this case, there is much less power at high frequencies, but the power at frequency zero is identical, since $G(0)$ depends only on the long-run effect of the shock on the level of consumption, $G_j(0) = \sum_{k=0}^{\infty} g_{j,k}$.

The next two shocks have purely transitory effects. The third shock raises consumption for just a single period, and we now see zero power at frequency zero and positive power at high frequencies. The fourth shock is more interesting. Consumption rises initially, turns negative in the second period, and returns to its initial level in the third period. The transfer function is again equal to zero at $\omega = 0$, but it now actually has negative power at low and middle frequencies. This is a result of the fact that the impulse response of consumption is actually negative in some periods. The sign of G reflects the direction in which the shock drives consumption. If we had reversed the signs of the impulse responses for the first three shocks, their transfer functions would all have been negative.

3 Weighting functions in consumption-based models

This section applies the spectral decomposition to power utility, internal and external habit formation, and Epstein–Zin preferences.[10] Similar results can be obtained for other models, such as affine term structure models and Campbell's (1993) specification where the market return is the priced variable.

3.1 Power utility

Under power utility, the log pricing kernel is

$$m_{t+1} = \log \beta - \alpha \Delta c_{t+1} \tag{11}$$

where c_t is log consumption, α is the coefficient of relative risk aversion, and $-\log \beta$ is the rate of pure time preference. The associated weighting function (equation (8)) is

$$Z^{power}(\omega) = \alpha \tag{12}$$

Z^{power} is flat and exactly equal to the coefficient of relative risk aversion. Z^{power} is constant because the only determinant of the innovation to the SDF is the innovation to consumption on date $t+1$. A shock to consumption growth has the same effect on the pricing kernel regardless of how long the innovation is expected to last, so future dynamics do not matter.

[10] While these models of preferences are often applied under the assumption of the existence of a representative agent, that assumption is not strictly necessary for our results. The pricing kernel generated by an agent's Euler equation will hold for any market in which she participates.

3.2 Habit formation

Adding an internal habit to the preferences, in a simplified version of Constantinides (1990), yields the lifetime utility function

$$\sum_{j=0}^{\infty} \beta^j \frac{\left(\exp\left(c_{t+j}\right) - b\exp\left(c_{t+j-1}\right)\right)^{1-\alpha}}{1-\alpha} \tag{13}$$

where b is a parameter determining the importance of the habit. The pricing kernel is

$$\exp\left(m_{t+1}\right) = \beta \frac{\left(\exp\left(c_{t+1}\right) - b\exp\left(c_t\right)\right)^{-\alpha} - \mathbb{E}_{t+1} b\left(\exp\left(c_{t+2}\right) - b\exp\left(c_{t+1}\right)\right)^{-\alpha}}{\left(\exp\left(c_t\right) - b\exp\left(c_{t-1}\right)\right)^{-\alpha} - \mathbb{E}_t b\left(\exp\left(c_{t+1}\right) - b\exp\left(c_t\right)\right)^{-\alpha}} \tag{14}$$

Linearizing in terms of Δc_{t+1} and Δc_{t+2} around a zero-growth steady-state yields

$$\Delta\mathbb{E}_{t+1} m_{t+1} \approx -\alpha\left(b\left(1-b\right)^{-2} + 1\right)\Delta\mathbb{E}_{t+1}\Delta c_{t+1} + \alpha b\left(1-b\right)^{-2}\Delta\mathbb{E}_{t+1}\Delta c_{t+2} \tag{15}$$

With internal habits the pricing kernel depends on both the innovation to current consumption growth and also the change in expected consumption growth between dates $t+1$ and $t+2$. The spectral weighting function is then

$$Z^{internal}\left(\omega\right) = \alpha\left(1 + b\left(1-b\right)^{-2}\right) - \alpha b\left(1-b\right)^{-2} 2\cos\left(\omega\right) \tag{16}$$

$Z^{internal}(\omega)$ is equal to a constant plus a negative multiple of $\cos(\omega)$. As we would expect, $Z^{internal}(\omega) = Z^{power}(\omega)$ when $b = 0$. The left panel of figure 2 plots $Z^{internal}$ for various values of b (here and in all cases below we only plot Z between 0 and π since Z is even and periodic). The x-axis lists the wavelength of the cycles.[11]

An increase in b has two effects on $Z^{internal}$: its total mass (its integral) rises, and the mass shifts to higher frequencies. The shift in mass is consistent with the usual intuition

[11] Given a frequency of ω, the corresponding cycle has length $2\pi/\omega$ periods (the smallest cycle we can discern from discretely sampled data lasts two periods).

about internal habit formation that households prefer to smooth consumption growth and avoid high-frequency fluctuations to a greater extent than they would under power utility.

It is also useful to consider the case where $b < 0$, which corresponds to durable consumption – people get utility both from current and also past consumption expenditures. In that case, the effects all reverse – $Z^{internal}$ is equal to a constant plus a positive multiple of $\cos(\omega)$. So with durable consumption, investors place relatively more weight on low- than high-frequency fluctuations.

One lesson from the equation for $Z^{internal}$ is that as long as b is the only parameter we can vary, there is little flexibility in controlling preferences over different frequencies. $Z^{internal}$ is monotone, regardless of the value of b, so habit formation does not ever allow business-cycle frequencies to carry more weight than any other frequency; that is, habit formation cannot induce an investor to be particularly averse to business-cycle frequency fluctuations compared to those at other frequencies.[12]

In contrast to internal habit formation, under external habit formation (e.g. Campbell and Cochrane (1999)) the SDF is

$$\exp(m_{t+1}) = \beta \frac{(\exp(c_{t+1}) - b\exp(\bar{c}_t))^{-\alpha}}{(\exp(c_t) - b\exp(\bar{c}_{t-1}))^{-\alpha}} \tag{17}$$

where $\bar{c}$ denotes some external measure of log consumption (e.g. aggregate consumption or that of an agent's neighbors). In this case, the innovation to the SDF depends only on the innovation to c_{t+1} – news about the future is irrelevant. The weighting function with an external habit is therefore completely flat. But since the local sensitivity of the pricing kernel to a shock depends on the distance between consumption and the habit, the level of the weighting function shifts over time. When consumption is close to the habit and risk aversion is high, the level of the weighting function is also high, while when consumption is

[12]The log-linearization of the SDF eliminates time-variation in the price of risk. A simple extension of the analysis is to model the SDF as being conditionally log-linear in consumption growth, with the slope coefficients, and thus the shape of the weighting function, varying over time. We tackle this case in section 5.

farther above the habit, risk aversion and the level of the weighting function are low.[13]

3.3 Epstein–Zin preferences

An alternative way of incorporating non-separabilities in utility over time is Epstein and Zin's (1991) generalized recursive preferences. In general, under recursive preferences, anything that affects an agent's welfare affects the pricing kernel. So not only shocks to current and future consumption growth, but also innovations to higher moments will be priced.

Lifetime utility follows the recursion

$$v_t = \left\{ (1-\beta) \exp(c_t)^{1-\rho} + \beta \left(E_t \left[v_{t+1}^{1-\alpha} \right] \right)^{1-\rho} \right\}^{\frac{1}{1-\rho}} \tag{18}$$

where ρ is the inverse elasticity of intertemporal substitution (EIS), and α is the coefficient of relative risk aversion. Campbell (1993) and Restoy and Weil (1998) show that if the expected excess return on aggregate wealth is constant (i.e. if the quantity of risk in the economy is constant), the stochastic discount factor for these preferences can be log-linearized as

$$\Delta \mathbb{E}_{t+1} m_{t+1} \approx - \left(\alpha \Delta \mathbb{E}_{t+1} \Delta c_{t+1} + (\alpha - \rho) \Delta \mathbb{E}_{t+1} \sum_{j=1}^{\infty} \theta^j \Delta c_{t+1+j} \right) \tag{19}$$

θ is a parameter (generally close to 1) that comes from the log-linearization of the return on the agent's wealth portfolio (Campbell and Shiller, 1988).[14,15]

[13] Otrok, Ravikumar, and Whiteman (2002) show that the external habit has a strong effect on what weight *utility* places on consumption cycles of different frequencies, but what we show here is that the *SDF* is driven entirely by one-period innovations, so all cycles receive the same weight in pricing assets.

[14] $\theta = \left(1 + \overline{DP}\right)^{-1}$, where $\overline{DP}$ is the dividend-price ratio for the wealth portfolio around which we approximate. θ generalizes the rate of pure time preference and depends somewhat on discounting due to uncertainty about future consumption. The separation between preferences and dynamics is thus not totally complete in this case. Hansen, Heaton, and Li (2008), however, derive an alternative approximation in which $\theta = \beta$, in which case the separation is again complete. Their result suggests (as do numerical results) that θ is only minimally affected by consumption dynamics.

[15] In the case where $\rho = 1$, equation (19) is exact and $\theta = \beta$. The approximation used to derive (19) is a linearization of the definition of the return on a consumption claim around a constant consumption/wealth ratio. Since the consumption-wealth ratio is constant when $\rho = 1$, equation (19) holds exactly in that case. We do not assume here that consumption growth is distributed log-normally. The assumption that the

The weighting function associated with equation (19) is

$$Z^{EZ}(\omega) \equiv \alpha + (\alpha - \rho) \sum_{j=1}^{\infty} \theta^j 2\cos(\omega j) \tag{20}$$

Under power utility, $\alpha = \rho$ and $Z^{EZ}(\omega) = \alpha$ is flat, so all frequencies receive equal weight, as discussed above. On the other hand, if $\alpha \neq \rho$, then weights can vary across frequencies.[16]

The right-hand panel of figure 2 plots Z^{EZ} for a variety of parameterizations. The parameterizations are meant to correspond to annual data, so we take $\theta = 0.975$ as our benchmark, which corresponds to a 2.5 percent annual dividend yield. For $\alpha = 5$ and $\rho = 0.5$ (an EIS of 2), we see a large peak near frequency zero, with little weight elsewhere. In fact, half the mass of Z^{EZ} in this case lies on cycles with length of 210 *years* or more, and 75 percent lies on cycles with length 70 years or more.

In this parameterization, it is effectively only the very longest cycles in consumption (up to permanent shocks) that carry any substantial weight in the pricing kernel. Purely temporary shocks to the level of consumption (which is what are induced by shocks to monetary policy in standard models, for example) are essentially unpriced.

The line that is highly negative near $\omega = 0$ is for $\alpha = 0.5$ and $\rho = 5$, where households prefer a late resolution of uncertainty. In that case, the mass of Z^{EZ} is still isolated near zero, but because households now prefer a late resolution of uncertainty, Z^{EZ} is negative at that point (since marginal utility is increasing in good news about long-run consumption growth). The integral of Z^{EZ} is still equal to α, though, so it turns positive at higher frequencies.[17]

Z^{EZ} is much richer than what we obtain in the case of power utility and it has a number

---

expected return on wealth is constant in deriving (19) is what allows us to write the pricing kernel only in terms of expected future consumption growth. The appendix examines the accuracy of the approximation used here in simulations of Bansal and Yaron's (2004) long-run risk model with various calibrations.

[16] As with external habit formation, it is natural here to also imagine variation in the weighting function Z over time. For example, movements in the coefficient of relative risk aversion, α_t, as in Melino and Yang (2003) or Dew-Becker (2013) would induce a time-varying weighting function, $Z_t(\omega) = \alpha_t + (\alpha_t - \rho) \sum_{j=1}^{\infty} \theta^j 2\cos(\omega j)$. In periods when α_t is higher, the weighting function would then have more mass, and the mass would be shifted relatively more towards low frequencies.

[17] Note, though, that the case where $\rho > \alpha$ is not taken as a benchmark and is not widely viewed as empirically relevant (see, e.g., Bansal and Yaron, 2004).

of important properties. First, as with power utility, its average value is exactly equal to the coefficient of relative risk aversion,

$$\frac{1}{\pi}\int_{0}^{\pi} Z^{EZ}(\omega)\, d\omega = \alpha \tag{21}$$

So the total mass of Z^{EZ} depends only on risk aversion. The effect of Epstein–Zin preferences is therefore not to raise overall risk aversion compared to power utility, but to shift that mass to low frequencies, changing the features of the consumption process that an investor is averse to.

In the limit as $\theta \to 1$, i.e. where the effective rate of time preference approaches zero, $Z^{EZ}(\omega)$ approaches

$$Z^{EZ}(\omega) \to (\alpha - \rho)\, \delta_p(\omega) + \rho \tag{22}$$

for ω in the interval $(-\pi, \pi)$, where $\delta_p(\omega)$ is a periodic extension to the Dirac delta function with $\frac{1}{2\pi}\int_{-\pi}^{\pi} \delta_p(\omega) = 1$.[18] So Z^{EZ} can be thought of as approximately a unit point mass weighted by $(\alpha - \rho)$ plus a constant ρ (figure 2 shows that this is a reasonable approximation). In the limit, only two features of the consumption process matter: the permanent innovations at $\omega = 0$ ($\lim_{j\to\infty} \Delta\mathbb{E}_{t+1}c_{t+j}$), which are weighted by $\alpha - \rho$, and all transitory innovations, which have no effect on $\lim_{j\to\infty} \Delta\mathbb{E}_{t+1}c_{t+j}$, and are weighted by ρ. The fraction of the total mass on frequency zero is $\frac{\alpha-\rho}{\alpha}$. The larger is α relative to ρ, the larger is the fraction of the mass of the weighting function that lies at frequency zero. For example, in our benchmark calibration with $\alpha = 5$ and $\rho = 1/2$, $\frac{\alpha-\rho}{\alpha} = 0.9$, so 90 percent of the mass of the weighting function is local to frequency zero.

So in terms of consumption, Epstein–Zin preferences differ from power utility because they add a point mass at zero with weight $(\alpha - \rho)$. They are otherwise nearly identical for cycles of all frequencies away from zero. The large amount of weight placed on very low

[18] Technically, we should use the limit of the Dirichlet kernel, which is a periodic extension of the delta function. On the interval $(-\pi, \pi)$, though, they deliver the same result.

frequencies obviously also makes the estimation problem underlying Epstein–Zin preferences, both for investors and economists, potentially much more difficult than that for power utility.

3.4 Weights on frequency ranges and the cost of business cycles

The spectral weighting functions allow us to directly quantify what fraction of risk prices are driven by any set of frequencies, such as business cycle frequencies or lower frequencies. As an example, while it is known from simple calibrations that under Epstein–Zin preferences fluctuations at business cycle frequencies are relatively unimportant for asset prices (Bansal, Kiku, and Yaron (2010)), we are able to precisely quantify that statement in a more general way, for any possible consumption process.

Given any pricing kernel that satisfies Assumption 1, we can easily compute the total weight that investors give to cycles of a certain length by integrating the associated weighting function $Z(\omega)$. Specifically, the fraction of the mass in the range of frequencies between ω_1 and ω_2 is

$$\frac{\int_{\omega_1}^{\omega_2} Z(\omega)\, d\omega}{\int_0^{\pi} Z(\omega)\, d\omega} \tag{23}$$

Table 1 reports the fraction of the mass of the weighting function Z^{EZ} for Epstein–Zin preferences and $Z^{internal}$ for internal habits in various frequency ranges, under various calibrations of the models.[19] The left-hand columns list parameters for the calibrations. The remaining columns then report results in different frequency ranges. The last column of the table reports the median cycle length for each calibration – the cycle such that exactly half the mass is on either side.

The top panel reports four sets of calibrations of Epstein–Zin preferences. The first three sets consider various combinations of values for α, ρ and θ that all satisfy $\alpha > \rho$, which implies a preference for an early resolution of uncertainty that is assumed in the majority of

[19] For Epstein–Zin, the weight in any frequency range can be computed in closed form as $(\pi\alpha)^{-1}\left(Q^{EZ}(\omega_2) - Q^{EZ}(\omega_1)\right)$ where $Q^{EZ}(\omega) \equiv \rho\omega - 2(\alpha-\rho)\tan^{-1}\left[\frac{(\theta+1)\tan(\frac{\omega}{2})}{\theta-1}\right]$. All the results are obtained from quarterly calibrations; the cycle length and parameters are reported in years for convenience.

the literature (most notably Bansal and Yaron (2004)). The last set of rows has $\alpha < \rho$ to help understand the case when investors prefer a late resolution of uncertainty.

Three notable results emerge from table 1. First, the median cycle is greater than 100 years long in all calibrations for which $\alpha > \rho$ (the only exception being the low-risk-aversion case of $\alpha = 2.5$ and $\rho = 1$, where the median cycle is still 68 years). These results show that when we say that investors with Epstein–Zin preferences are averse to long-run risk, "long-run" should be thought as cycles in consumption lasting *centuries*.

Second, the table shows that Epstein–Zin preferences give an extremely small role to business-cycle fluctuations. Across all calibrations with $\alpha > \rho$ we examine, business-cycle frequencies carry at most 12 percent of the weight of the pricing kernel. Table 1 thus provides a clear and robust result: under Epstein–Zin preferences, business cycles are quantitatively irrelevant, while cycles lasting centuries are priced most strongly.

Third, results change dramatically when $\alpha \leq \rho$. When $\alpha = \rho$, Epstein–Zin collapses to standard power utility, whose median cycle length is 1 year (for quarterly data), and that places 94 percent of the weight on fluctuations of 8 years or shorter. When α is strictly less than ρ, the pricing kernel places negative weight on low frequencies because low-frequency increases in consumption raise marginal utility, so there is relatively more weight on higher frequencies (which is also clear in figure 2).

The bottom panel of table 1 reports the calibration results for internal habits. The only parameter that affects the relative weight across frequency ranges is the habit parameter b. Across all calibrations, internal habit investors place essentially all weight on very high frequencies with cycles shorter than 1.5 years.

The quantitative results we obtain here are independent of the particular consumption process chosen: they are obtained solely from the utility function. So we are able to extend standard results from macroeconomics about the cost of business cycle frequencies in a more general way.

Table 1 is one of the central results of the paper because it *quantifies* how various specifications of utility functions place weight on cycles of different frequencies. While some of the results were perhaps understood qualitatively as part of a folk wisdom, this paper pins down exactly what frequency fluctuations investors are averse to. In the previous literature, for example, Bansal, Kiku, and Yaron (2010) also show that a business-cycle type shock to the economy carries a small risk price, but the analysis here quantifies the importance of business cycle frequencies and derives the result in a much more general setting, not requiring a specific calibration of consumption dynamics.

3.5 Implications of the theoretical result

Our theoretical results on the weight that Epstein–Zin preferences place on cycles of hundreds of years suggest two limitations of models based on this utility specification. First, they endow investors with seemingly implausible amounts of information (an argument related to that in Chen, Dou, and Kogan (2013)). Second, even if investors do have such information, for an econometrician to test Epstein–Zin preferences may require either very strong assumptions about the consumption process or centuries of data. Both of those arguments can be made formally, based on the asymptotic distribution of the sample spectrum of consumption growth. In particular, we show that obtaining *direct* information on the behavior of consumption growth at the frequencies that carry the majority of the weight under Epstein–Zin preferences requires 210 or more years of data.

As an example, suppose consumption growth is driven by a linear univariate Gaussian process (but note that everything here easily extends to a multivariate setting). It is well known that the frequency-domain features of a particular process can be estimated by taking the discrete Fourier transform of an observed sample, which is known as the periodogram. In a sample of length T, the periodogram is defined at a set of equally spaced frequencies, $\omega_k = 2\pi k/T$ for $k \in \{1, 2, ..., T\}$. The periodogram ordinates are asymptotically independent and provide information about the dynamics of consumption growth at frequency ω_k.[20]

[20] Specifically, the periodogram is equal to the true spectrum multiplied by an $Exp\,(1)$ random variable.

In a given sample, the lowest frequency that we have information about is $2\pi/T$ and the associated wavelength is exactly T periods. That is, in a sample of length T, the longest fluctuation that we directly observe lasts T periods. For a post-war quarterly sample, $T \approx 70$ years. At our benchmark calibration in that case, more than 75 percent of the mass of the Epstein–Zin weighting function lies below $\omega_1 = 2\pi/70$. That is, 75 percent of the weight that determines risk premia in the model lies on frequencies about which we have no direct information. To have even a single observation at the median frequency would require having a sample 210 years long.

So in a non-parametric sense we have almost no direct evidence about the frequencies that we must estimate in order to know what Epstein–Zin preferences imply for risk premia. Based purely on the periodogram, models involving Epstein–Zin preferences are essentially untestable – the frequencies that determine risk premia under the model are not directly observable, so it is impossible to test the prediction that power at low frequencies determines the risk premium without adding external information to the estimation method.[21]

But the difficulty in estimation is obviously not just a problem for econometricians; it also affects investors. It seems rather implausible to assume that investors are sure of the dynamics of consumption growth at frequencies that cannot be observed without centuries of data. A number of recent papers build on precisely that point, including Chen, Dou, and Kogan (2013), Bidder and Dew-Becker (2015), and Collin-Dufresne, Johannes, and Lochstoer (2015).

So the fact that most of the mass of the weighting function under Epstein–Zin preferences is located on cycles of 210 years or more is problematic for two reasons – it makes the central asset pricing predictions of the model extremely difficult to test, and it relies on investors

This type of analysis is the basis of both the Whittle (1962) likelihood and also of non-parametric time series estimation (see Brillinger (1981) and Priestley (1981)).

[21] One option, for example, is to impose parametric restrictions on the consumption process, which then allows one to estimate low-frequency dynamics based on higher-frequency data. But at that point, one faces the joint hypothesis problem that the test of the model is only valid if the restrictions on consumption growth are true.

having firm beliefs about features of the consumption process that they have never directly observed.

3.6 Model contamination

One of the advantages of looking at the frequency domain decomposition is that it gives a compact and quantitative representation of the *entire* dynamic process driving the economy. To give a closer comparison to the time-domain methods used in the previous literature to study the role of dynamics in models (e.g. Bansal, Kiku, and Yaron (2010, 2012), or Beeler and Campbell (2012)) we now consider a small perturbation of Bansal and Yaron's (2004) long-run risk model. We show that typical time-domain analysis of the dynamics is not robust to small changes in model specification; frequency-domain analysis is. The results in this section thus give a concrete example of the advantage of the frequency domain analysis that we have thus far provided, beyond the novel quantitative metrics presented above.

In the benchmark long-run risk model, consumption growth follows

$$\Delta c_t = x_{t-1} + \varepsilon_{\Delta c,t} \tag{24}$$

$$x_t = \rho x_{t-1} + \varepsilon_{x,t} = \sum_{j=0}^{\infty} \rho^j \varepsilon_{x,t-j} \tag{25}$$

As a closely related alternative, we examine the contaminated model

$$\Delta c_t^{con} = x_{t-1}^{con} + \varepsilon_{\Delta c,t} \tag{26}$$

$$x_t^{con} = \left(1.8 \sum_{j=0}^{\infty} \rho^j - 0.24 \sum_{j=0}^{\infty} \theta^j\right) \varepsilon_{x,t-j} \tag{27}$$

Instead of following an AR(1) like x_t, x_t^{con} is the difference between two AR(1) processes (with perfectly correlated innovations). While the difference between Δc_t and Δc_t^{con} is visible in these equations, we will see that autocorrelations and IRFs, which are typically examined in the literature (especially for production-based models where the full dynamic process for consumption growth is not known analytically, e.g. Kaltenbrunner and Lochstoer (2010)), suggest that in fact their dynamics are nearly identical.

Using a quarterly calibration following Bansal and Yaron (2004), we set $\rho = 0.938$, $std\left(\varepsilon_{\Delta c,t}\right) = 0.0135$, and $std\left(\varepsilon_{x,t}\right) = 0.000584$. θ is set to take the same value as in the calibration of Epstein–Zin preferences above, $0.975^{1/4}$.[22] The models thus only differ in the dynamics of the persistent component of consumption growth.

The top-left panel of figure 3 plots the first 20 quarterly autocorrelations of consumption growth in the original and contaminated long-run risk model. The choice to examine correlations out to 5 years follows the empirical evaluations of the long-run risk model in Bansal, Kiku, and Yaron (2012) and Beeler and Campbell (2012). The autocorrelations of Δc^{con} are in fact *higher* than in the original long-run risk model, suggesting that the contaminated model should be more risky. The unconditional standard deviation of consumption growth is also slightly higher, at 1.37 instead of 1.36 percent.

The top-right panel of figure 3 shows that the first ten years of the impulse-response function of consumption to $\varepsilon_{x,t}$ is higher at every horizon in the contaminated model. At ten years, the IRF for the contaminated model is 23 percent higher than that for the original. So by standard measures, the autocorrelation and the IRF, the contaminated model seems far more risky than the original calibration.

But appearances deceive us: rather than being more risky than the original, ε_x is in fact far *less* risky in the contaminated model – its risk price is smaller by exactly half. Looking at the impulse transfer functions would have made this immediately clear. The bottom panels of figure 3 plot the impulse transfer functions G for ε_x in the two models along with the weighting function Z under the benchmark Epstein–Zin calibration from above (the right-hand panel zooms in on cycles longer than 5 years). While the ITF is higher in the contaminated model at most frequencies, it rapidly falls at the lowest frequencies, exactly

[22] The power 1/4 is to account for the change from an annual to a quarterly calibration. The choice to align the persistence of the perturbation with the effective discount factor is not a coincidence. See Bidder and Dew-Becker (2014) for an explanation of why this is the most powerful perturbation.

where the frequency-specific risk prices are highest. When we look at the ITFs, the much smaller risk price for the contaminated shock is not at all surprising.

The results in this section help emphasize our motivation for studying the frequency domain. Standard time-domain tools, IRFs and autocorrelations, in this case clearly do not adequately measure risk, while the impulse transfer function, particularly near frequency zero, does.

4 Multiple priced variables

Our frequency decomposition also holds when there is more than one variable that drives utility. A benchmark example is in Bansal and Yaron's (2004) long-run risk model where volatility varies over time and is a priced factor. But there are also many studies in which other higher-order moments of the consumption process vary, e.g. Drechsler and Yaron (2011); Gourio (2012); Wachter (2013); and Constantinides and Ghosh (2013). We show that our analysis easily extends to such cases, with the only difference being that there will be a frequency-domain weighting function for each priced variable (all proofs are reported in the Appendix).

4.1 General pricing result

Assumption 1a: Structure of the SDF

Instead of there being a single priced variable x_t, suppose there is an $m \times 1$ vector of priced variables, $\mathbf{x}_t$, with

$$m_{t+1} = f\left(\mathcal{I}_t\right) - \Delta \mathbb{E}_{t+1} \sum_{k=0}^{\infty} \mathbf{z}_k \mathbf{x}_{t+1+k} \tag{28}$$

where $\mathbf{z}_k$ is a $1 \times m$ vector of weights and $f\left(\mathcal{I}_t\right)$ is an unspecified scalar valued function as before.

Assumption 2a: Dynamics of the economy

We assume that $\mathbf{x}_t$ is driven by a vector moving average process as before,

$$\mathbf{x}_t = \mathbf{J}\bar{\mathbf{x}}_t \tag{29}$$

$$\bar{\mathbf{x}}_t = \mathbf{\Gamma}(L)\boldsymbol{\varepsilon}_t \tag{30}$$

for some matrix $\mathbf{J}$ of dimension $m \times n$, and where Γ is an $n \times n$ matrix-valued power series in the lag operator with coefficients Γ_k.

We have the following extension of Result 1,

Result 2. *Under Assumptions 1a and 2a, we can write the innovations to the SDF as,*

$$\Delta \mathbb{E}_{t+1} m_{t+1} = -\sum_j \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{Z}(\omega)\mathbf{G}(\omega)\, d\omega \right) \varepsilon_{j,t+1} \tag{31}$$

where $\mathbf{Z}(\omega)$ is a $1 \times m$ vector-valued weighting function and $\mathbf{G}(\omega)$ is an $m \times n$ transfer function that measures the dynamic effects of $\boldsymbol{\varepsilon}_t$ on $\mathbf{x}$ in the frequency domain,

$$\mathbf{Z}(\omega) \equiv \mathbf{Z}_0 + 2\sum_{k=1}^{\infty} \mathbf{Z}_k \cos(\omega k) \tag{32}$$

$$\mathbf{G}(\omega) \equiv \sum_{k=0}^{\infty} \cos(\omega k)\, \mathbf{g}_k \tag{33}$$

and $\mathbf{g}_k$ is the matrix of impulse response functions,

$$\mathbf{g}_k \equiv \mathbf{J}\mathbf{\Gamma}_k \tag{34}$$

In this case, then, we have multiple variables whose impulse responses we track in $\mathbf{G}$, and each of the priced variables has its own weighting function, represented as one of the elements of $\mathbf{Z}(\omega)$.

4.2 Epstein–Zin with time-varying higher moments

In our main analysis of Epstein–Zin preferences, we examined the case where the expected excess return on a consumption claim is constant. That case requires that the conditional moments of consumption growth above order 1 be constant. We now show how to extend the result to a case where any of the higher moments of consumption may vary.

We assume that consumption growth follows a vector moving average process as before,

$$\Delta c_t = \sum_j g_j(L)\,\varepsilon_{j,t} \tag{35}$$

We now assume, though, that rather than the $\varepsilon_{j,t}$ having fixed distributions over time, their distributions are driven by a factor $\tilde{x}_t$, which, without loss of generality, we restrict to have zero mean (here we assume $\tilde{x}_t$ is a scalar for simplicity, but the analysis trivially extends to the case where $\tilde{x}_t$ is a vector). We assume that $\tilde{x}_t$ affects the distribution of $\varepsilon_{j,t}$ linearly in its cumulant-generating function,

$$\log E_t \exp\left(\tau\varepsilon_{j,t+1}\right) = f_{j,0}(\tau) + f_{j,1}(\tau)\,\tilde{x}_t \tag{36}$$

A special case of the above is where $\varepsilon_{j,t+1}$ is conditionally normally distributed, in which case $f_{j,0}(\tau) + f_{j,1}(\tau)\,\tilde{x}_t = \frac{1}{2}var_t(\varepsilon_{j,t+1})$, so the conditional variance of $\varepsilon_{j,t+1}$ would be linear in $\tilde{x}_t$ (as in Bansal and Yaron (2004)).

$\tilde{x}_t$ is also assumed to follow a VMA process that depends on the same innovations that drive consumption growth (though note that $\tilde{x}$ and Δc may be made independent with particular choices of the lag polynomials).

$$\tilde{x}_t = \sum_k \tilde{g}_k(L)\,\varepsilon_{k,t} \tag{37}$$

(as is common in the literature, for example in Bansal et al. (2014) or Campbell et al.

(2015), we ignore here the fact that this specification implies that volatilities can become negative).

The appendix then shows that the price of risk for $\varepsilon_{j,t}$ is

$$\int Z_{\Delta c}(\omega) G_j(\omega) d\omega + \int Z_{\tilde{x}}(\omega) \tilde{G}_j(\omega) d\omega \tag{38}$$

where

$$G_j(\omega) \equiv \sum_{j=0}^{\infty} g_j \cos(\omega j) \tag{39}$$

$$\tilde{G}_j(\omega) \equiv \sum_{j=0}^{\infty} \tilde{g}_j \cos(\omega j) \tag{40}$$

and

$$Z_{\Delta c}(\omega) \equiv \rho + (\alpha - \rho) \sum_{j=1}^{\infty} \theta^j 2 \cos(\omega j) \tag{41}$$

$$Z_{\tilde{x}}(\omega) \equiv k_1 \frac{\rho - \alpha}{1 - \rho} \left(1 + \sum_{j=1}^{\infty} \theta^j 2 \cos(\omega j) \right) \tag{42}$$

where k_1 is an equilibrium coefficient that determines the effect of a unit increase in $\tilde{x}_t$ on the expected excess return on wealth.

Each shock is now associated with a pair of transfer functions that measure the effects of the shock on consumption and on the distribution of the innovations, and those transfer functions are interacted with a pair of spectral weighting functions, $Z_{\Delta c}$ and $Z_{\tilde{x}}$. The weighting functions have a number of notable features. First, $Z_{\Delta c}$ is identical to the weighting function obtained for consumption in the case where the innovations were identically distributed over time. That is, adding variation in higher moments (linearly dependent on a variable $\tilde{x}$) does not affect the pricing of innovations to expected consumption growth.

Second, note that the weighting functions for Δc and $\tilde{x}$ are almost identical up to a scaling factor. When $\rho = 0$, they are in fact proportional to each other, and for $\alpha \gg \rho$, they are nearly proportional. So Epstein–Zin preferences imply nearly identical treatment of variation in higher moments of consumption growth to variation in the first moment.

Third, the *shape* of $Z_{\tilde{x}}$ does not depend on how $\tilde{x}$ affects the distribution of consumption growth. Regardless of how $\tilde{x}$ affects the cumulant-generating function of consumption growth (36), its risk price depends purely on its effects on the expected excess return on wealth, through the coefficient k_1.

Finally, note that if $\tilde{G}_j(\omega) = 0$ for some shock ε_j – i.e., if it does not affect $\tilde{x}$ – then the shock is priced in this setting exactly the same way that it is in a homoskedastic model.

In the empirical analysis below, in addition to trying to estimate weighting functions for consumption growth, we also examine the pricing of shocks to expected future volatility. We leave the empirical analysis of variation in other higher moments (such as disaster risk) to future work.

5 Estimation

We now proceed to estimate the weighting function Z in US equity markets. In addition to providing novel evidence on what model of preferences best describes the pricing of risks, the estimation also demonstrates how difficult the estimation problem is that investors face.

We begin by describing the parametric specifications of the weighting function that we estimate. We then carry out the full estimation, focusing in particular on the ability of our frequency domain analysis to help increase the precision of estimates of consumption dynamics.

5.1 Parameterized weighting functions

5.1.1 The utility specification

The analysis of the utility functions in Section 3 suggests modeling Z as:

$$Z^U(\omega; \mathbf{q}) = q_1 \sum_{j=1}^{\infty} \theta^j \cos(\omega j) + q_2 + q_3 \cos(\omega) \tag{43}$$

where q_1, q_2, and q_3 are unknown coefficients and $\mathbf{q} \equiv [q_1, q_2, q_3]$. We call (43) the *utility specification* because it *exactly* nests the weighting functions derived from utility-based models. If $q_3 = 0$, (43) precisely matches the weighting function for Epstein–Zin preferences in (20). If $q_1 = 0$, the long-run component that is crucial in the Epstein–Zin case is shut off, and we obtain the specific weighting function for internal habit formation in (16). Finally, if both $q_1 = 0$ and $q_3 = 0$, then we have the weighting function for power utility.

So tests of hypotheses that those coefficients are equal to zero represent tests of the different specifications of the utility functions. Specifically, if we were to find values for q_1 and q_2 that are significantly different from zero while q_3 is not, that would imply that the data is consistent with Epstein–Zin preferences. On the other hand, if q_2 and q_3 are significant but q_1 is not, the data would support habit formation.

Moreover, the parameters map directly to preferences. For example, if investors have Epstein–Zin preferences, then $q_1 = 2(\alpha - \rho)$, $q_2 = \alpha$, and $q_3 = 0$. Estimation of Z^U is *identical* to estimating the pricing kernel as though the three utility functions hold. That is, the fact that we do this estimation in the frequency domain has no implications for the results.

For the long-run component, we choose $\theta = 0.975^{1/4}$ for quarterly data, corresponding to a 2.5 percent annual consumption/wealth ratio as above.[23]

Because the utility specification is composed of the weighting functions we derived under

[23] In principle, we could estimate θ. However, we find that it is poorly identified in the data, so we proceed to calibrate it to a value widely used in the literature.

various preference specifications, the constituent functions are already plotted in figure 2. In particular, the lines in the right-hand panel represent the first function, $\sum_{j=1}^{\infty} \theta^j \cos(\omega j)$, shifted upward by a constant. This function clearly isolates very low frequencies, and the extent to which the lowest frequencies are isolated depends on the parameter θ. Estimates of the utility specification allow us to empirically measure the ability of the three different utility functions to explain the cross-section of risk premia.

5.1.2 The bandpass specification

As an alternative to the strict utility specification, we also model Z more flexibly. We break the interval $[0, \pi]$ into three economically motivated intervals, corresponding to business-cycle length fluctuations, with wavelength between 6 and 32 quarters (as is standard in the macro literature, e.g. Christiano and Fitzgerald (2003)), and frequencies above and below that window. Z then takes the form of a step function on those three frequency windows, and the levels of the steps are the free parameters to estimate.

We refer to the set of three step functions as the *bandpass specification*, $Z^{BP}(\omega; \mathbf{q})$, since it is composed of the sum of three bandpass filters,

$$Z^{BP}(\omega; \mathbf{q}) = q_1 Z^{(0,2\pi/32)}(\omega) + q_2 Z^{(2\pi/32,2\pi/6)}(\omega) + q_3 Z^{(2\pi/6,\pi)}(\omega) \tag{44}$$

$$\text{where } Z^{(a,b)}(\omega) \equiv \begin{cases} 1 \text{ if } a < |\omega| \leq b \\ 0 \text{ otherwise} \end{cases} \tag{45}$$

If investors are averse to long-run risks, we would expect risk prices to be highest below business cycle frequencies, while habit formation type preferences imply that the risk prices should be highest at higher frequencies. That is, Epstein–Zin preferences imply a high value for q_1 compared to q_2 and q_3, while habit formation implies a high value for q_3 compared to q_1 and q_2. Note that what is considered "long-run risk" here is not the literal interpretation of Epstein–Zin preferences (centuries); rather, the long-run is anything longer than the business cycle.

The bandpass specification demonstrates one of the key features of our approach: we are able to estimate the sources of risk premia in a way that is clearly linked to underlying economic risks – fluctuations in consumption at meaningfully chosen frequencies – but we do not necessarily need to use the highly constrained specifications required by structural models. That said, if the number of steps in the bandpass specification were allowed to increase, it could eventually represent arbitrary preferences accurately. We also expect that the bandpass specification will increase estimation power as it does not isolate its mass nearly as close to frequency zero as the utility specification.

One potential drawback of the bandpass specification is that since it is somewhat reduced-form, it is difficult to see the precise link to microeconomic behavior. The appendix therefore describes a setting in which the bandpass specification would arise endogenously if investors estimate consumption dynamics using a restricted model (e.g. due to information processing constraints).

5.2 Estimates

We now estimate the weighting function Z in the US equity market. The overall estimation method has three basic steps:

1. Estimate a model of consumption dynamics with news at various horizons, based on a factor-augmented vector autoregression (FAVAR) model.

2. Estimate transfer functions and a rotation from the time to the frequency domain.

3. Estimate risk prices on the innovations to consumption growth in the frequency domain.

5.2.1 Step 1: Estimation of the dynamics

We estimate the dynamics of consumption by specifying and estimating a factor-augmented vector autoregression (FAVAR).[24] The FAVAR specification combines the advantages of the VAR methodology with the dimension reduction properties of factor models. Vector autoregressions (VARs) have the advantage that they are easy to estimate (in that they require no numerical optimization), they have been widely used in the previous literature,[25] and because when the number of lags in the VAR increases, the VAR structure can asymptotically capture arbitrarily rich dynamics (Lewis and Reinsel (1985); Mitchell and Brockwell (1997); Schorfheide (2005)).

Our estimation of the dynamics follows Jurado, Ludvigson and Ng (2015), not only because we apply the same methodology to set up and estimate the FAVAR system, but also because we use the same 131 macroeconomic series.

The FAVAR we estimate has three variables: log real consumption growth (observable factor), and two latent factors, F_{1t} and F_{2t}. The VAR is therefore specified as:

$$\bar{\mathbf{x}}_t = \bar{\mathbf{\Phi}}\left(L\right)\bar{\mathbf{x}}_{t-1} + \boldsymbol{\varepsilon}_t \tag{46}$$

where $\bar{\mathbf{x}}_t$ contains Δc_t, F_{1t}, F_{2t}.

Note that if the lag polynomial $\bar{\mathbf{\Phi}}\left(L\right)$ has order k, then we can stack k consecutive observations of $\bar{\mathbf{x}}_t$ so that $\mathbf{x}_t \equiv \left[\bar{\mathbf{x}}_t', \bar{\mathbf{x}}_{t-1}', ..., \bar{\mathbf{x}}_{t-k+1}'\right]'$ follows a VAR(1)

$$\mathbf{x}_t = \mathbf{\Phi}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t \tag{47}$$

and $\Delta c_t = \mathbf{b}_1 \mathbf{x}_t$.[26] F_{1t} and F_{2t} are estimated from the 131 macroeconomic series of Jurado,

[24] See Bernanke, Boivin and Eliasz (2005), Ludvigson and Ng (2007), and Jurado, Ludvigson and Ng (2015).

[25] E.g. Vuolteenaho (2002); Campbell and Vuolteenaho (2004); Bansal, Dittmar, and Lundblad (2005); Larrain and Yogo (2008); Lustig and van Nieuwerburgh (2008); Campbell, Polk, and Vuolteenaho (2010); and Campbell et al. (2015).

[26] Recall that $\mathbf{b}_j$ represents a conformable selection vector equal to 1 in element j and 0 elsewhere.

Ludvigson and Ng (2015) using principal component analysis. The VAR is then estimated through OLS, using the estimated factors $\hat{F}_{1t}$ and $\hat{F}_{2t}$ yielding estimates of $\mathbf{\Phi}$ and the innovations $\boldsymbol{\varepsilon}_t$. For readability, the two principal components are scaled to have the same variance as consumption growth.

We use quarterly data over the longest sample for which all the variables are available, 1961–2011.[27] We select three lags for the VAR, as recommended by cross-validation.[28] Table A1 in the appendix reports the estimated VAR matrix $\mathbf{\Phi}$.

5.2.2 Step 2: Estimate transfer functions and a rotation

Given the estimated FAVAR, the transfer function for shock j is

$$G_j(\omega) = \sum_{k=0}^{\infty} \cos(\omega k)\, \mathbf{b}_1 \mathbf{\Phi}^k \mathbf{b}_j' \tag{48}$$

The two finite-order specifications for Z, the utility and the bandpass specification, both take the form,

$$Z(\omega; \mathbf{q}) = q_1 Z_1(\omega) + q_2 Z_2(\omega) + q_3 Z_3(\omega) \tag{49}$$

$$= \mathbf{q}\left[Z_1(\omega), Z_2(\omega), Z_3(\omega)\right]' \tag{50}$$

(for different sets of functions Z_1, Z_2, Z_3). Denoting the risk price for shock j as p_j, we have

[27] Our primary sample uses quarterly data because that is the highest frequency at which consumption data is available for the full post-war sample. Parker and Julliard (2005) and Malloy, Moskowitz, and Vissing-Jorgenson (2009) find that using lower frequency or time-aggregated data can producer stronger evidence in favor of the consumption CAPM. Note that in our setting, the shortest wavelength cycle that we can price is two periods long. When the unit of observation is quarterly, we can potentially price fluctuations as short as two quarters. The effect of aggregating consumption to a lower (e.g. annual) frequency is thus to eliminate our ability to price higher frequency fluctuations. We examine results using annual data below and find results consistent with the main quarterly analysis.

[28] We have explored other criteria for lag selection. The cross-validation criterion is the most natural in this context, as it is based on the forecasting ability of the model, which plays a central role in our analysis (since the VAR is used to construct news about future consumption at different horizon). Among the other criteria we analyzed, the AIC and the FPE criteria favored using three lags, consistent with the cross-validation approach. BIC suggests two lags, while HQIC favors two lags only slightly against three lags. Appendix G reports robustness tests using two lags in the VAR.

(from Result 1)

$$p_j = \frac{1}{\pi}\mathbf{q}\int_0^{\pi}\left[Z_1(\omega), Z_2(\omega), Z_3(\omega)\right]' G_j(\omega)\, d\omega \tag{51}$$

It is straightforward then to show that the vector of risk prices is,

$$\begin{aligned}
\mathbf{p} &\equiv [p_1, p_2, p_3] &(52)\\
&= \mathbf{qW} &(53)
\end{aligned}$$

where the (i,j)th element of $\mathbf{W}$ is

$$W_{i,j} = \frac{1}{\pi}\int_0^{\pi} Z_i(\omega) G_j(\omega) d\omega \tag{54}$$

This result tells us that once we estimate the rotation matrix $\mathbf{W}$ from the consumption dynamics, we can express the coefficients of the Z function, $\mathbf{q}$, as a function of the risk prices of the VAR innovations, $\mathbf{p}$, as:[29]

$$\mathbf{q} = \mathbf{pW^{-1}} \tag{55}$$

The matrix $\mathbf{W}$ summarizes the interactions of the transfer functions with the components of the weighting function, Z_i (which, given a choice of a set of functions Z_1, Z_2, Z_3, are fully known and need not be estimated). $\mathbf{W}$ allows us to rotate between the risk prices on the reduced-form shocks, $\mathbf{p}$, and the frequency-domain risk prices, $\mathbf{q}$. The entire point of estimating the VAR for consumption growth is to develop estimates of the consumption dynamics, and all the relevant information for asset prices originating from the model dynamics is contained in $\mathbf{W}$.

It is important to note here that there is no need to make any assumptions to identify "structural" shocks in the VAR. Nowhere in the derivations above did we make any assump-

[29] For this last step, $\mathbf{W}$ needs to be invertible.

tions about the shocks ε_t being somehow structural; for example, their covariance matrix is entirely unrestricted. Our results are therefore analytically identical regardless of how the estimated shocks are rotated. That is a major advantage of our approach – the frequency-domain risk prices may be estimated without having to make assumptions to identify a structural VAR. The appendix provides a full derivation of that result.

Estimates of the transfer functions The VAR has three innovations: one to consumption growth and two to the two factors, so we have three impulse transfer functions. Figure 4 plots the estimated impulse transfer functions for each shock. The shaded regions in each figure are pointwise 95-percent confidence intervals. The vertical bar in each plot corresponds to cycles of 8 years – that determines the barrier between business-cycle and below-business-cycle fluctuations. Note that there are meaningful qualitative and quantitative differences across the functions in how power is distributed, which will help identify the underlying risk prices of different frequencies. If the transfer functions were all highly similar, then we would not expect to be able to distinguish risk prices across frequencies very well.

Rotation matrix The ultimate reason that we estimate the VAR for consumption growth is to generate a rotation matrix $\mathbf{W}$. Table 2 reports the rotation matrix $\mathbf{W}$ for the utility and bandpass specifications.

Since the frequency-domain risk prices, $\mathbf{q}$, are rotated from the time-domain prices, $\mathbf{p}$, using $\mathbf{W}$, estimation error in $\mathbf{W}$ is a key factor in determining the standard errors of $\mathbf{q}$. Heuristically, we can think of the moment conditions for the estimate of $\mathbf{q}$ as being the vector $(\mathbf{p} - \mathbf{qW})$ (we discuss the precise moment conditions used for $\mathbf{q}$ below). Using the formula for the optimal-GMM standard errors for $\mathbf{q}$ (and ignoring uncertainty in $\mathbf{p}$ for the moment), we have[30]

$$cov\left(\mathbf{q}\right) \approx \mathbf{W}'^{-1}\mathbf{q}cov\left(\mathbf{W}\right)\mathbf{q}'\mathbf{W}^{-1} \tag{56}$$

This equation approximates the uncertainty in estimates of $\mathbf{q}$ coming from uncertainty in $\mathbf{W}$. The covariance matrix of the estimates $\mathbf{q}$ depends crucially on the covariance matrix of the estimates of $\mathbf{W}$. Doubling the variance of $\mathbf{W}$ doubles the variance of the estimates of $\mathbf{q}$.[31]

5.2.3 Step 3: Estimation of frequency-domain risk prices q

Moment conditions We now proceed to estimate the full model, using the sequential GMM estimation described in Hansen (2008). We account for heteroskedasticity and serial correlation in the errors using Newey-West standard errors with 12 quarterly lags.

Under the assumption that returns are log-normally distributed, the risk prices can be estimated from the asset pricing condition (see the appendix for the derivation)

$$\begin{aligned} \mathbb{E}[\exp\left(r_{it+1}\right) - \exp\left(r_{t+1}^{f}\right)] &= -cov(m_{t+1}, r_{it+1}) & (57)\\ &= \mathbb{E}\left[\mathbf{q}\mathbf{W}\boldsymbol{\varepsilon}_{t+1}r_{it+1}\right] & (58)\end{aligned}$$

where r_{it} are log test asset returns, $\mathbf{r}_t$ is the corresponding vector, and r_t^f is the log risk-free

[30] Specifically, given a moment condition $m\left(\mathbf{q}\right) \equiv \mathbf{p} - \mathbf{q}\mathbf{W}$, the covariance matrix of the estimates of $\mathbf{q}$ is $\left(\Gamma\Delta^{-1}\Gamma'\right)^{-1} = \Gamma'^{-1}\Delta\Gamma^{-1}$, where $\Gamma = \frac{dm(\mathbf{q})}{d\mathbf{q}} = -\mathbf{W}$ and $\Delta = cov\left(m\left(\mathbf{q}\right)\right) = \mathbf{q}cov\left(\mathbf{W}\right)\mathbf{q}'$. This all assumes that the model is correctly specified and is simply meant for illustrative purposes.

[31] In the utility specification, the constant term appears without error because we have normalized the component functions by the variance of consumption growth. Since the constant is simply a one-standard-deviation shock to consumption growth, there is no uncertainty left in it.

rate.[32] Our full set of moment conditions identifying the parameters of the model is

$$H_{t+1}(\mathbf{\Phi}, \mathbf{x}) = \left[\underbrace{(\mathbf{x}_{t+1} - \mathbf{\Phi}\mathbf{x}_t) \otimes \mathbf{x}_t}_{\text{VAR moments}}, \underbrace{\exp \mathbf{r}_{t+1} - \exp r^f_{t+1} - \left(\overbrace{\widehat{\mathbf{q}\mathbf{W}}}^{\text{Mapping into frequency domain}} (\mathbf{x}_{t+1} - \mathbf{\Phi}\mathbf{x}_t) \right) \mathbf{r}_{t+1}}_{\text{Asset pricing moments}} \right] \tag{59}$$

The first set of moment conditions identifies the dynamics and therefore $\mathbf{W}$. The second set of moment conditions are the cross-sectional asset pricing moments that identify $\mathbf{q}$.

While we could in principle minimize the GMM objective function for all the parameters simultaneously, that method has the drawbacks that the optimization is difficult to perform (due to the large number of parameters) and that it allows errors in the asset pricing model to affect the VAR estimates. We therefore construct estimates of $\mathbf{\Phi}$ and $\mathbf{q}$ by minimizing the two sets of moment conditions separately. That is, $\mathbf{\Phi}$ is simply estimated through OLS and then $\mathbf{q}$ is estimated taking $\mathbf{\Phi}$ as given, using standard two-step GMM.[33] This is precisely the sequential GMM procedure described by Hansen (2008), and we calculate standard errors following that paper. The appendix describes the details.

Estimates of $Z(\omega)$ Table 3 reports the estimated risk prices. We repeat the estimation using different test assets, sequentially adding groups of test assets (all but the last one obtained from Ken French's website). The first column uses the set of 25 size- and book/market-sorted portfolios; the second column adds a set of 49 industry portfolios (we drop six industry portfolios that have missing data in the period considered). The third column adds a set of 25 portfolios sorted by investment and operating profitability. The last column reports our most comprehensive test, which also adds 9 risk-sorted portfolios (double sorted based on the exposure to the estimated low- and business-cycle frequency shocks, as

[32] The log-normality assumption for the empirics is standard in the literature. See, e.g., Campbell and Vuolteenaho, 2004; Campbell et al., 2015; Bansal et al., 2014. While not necessary for our theoretical result, assuming log-normality allows us to avoid making assumptions about the mean of the conditional log SDF when estimating the model and focus purely on its innovations; in addition, it yields a linear factor model, which is easy to estimate and interpret.

[33] The same sequential method is used in Campbell and Vuolteenaho (2004) and Campbell et al. (2015).

described in the appendix). For each portfolio set we estimate both the bandpass and the utility specification.

For the utility specification in the top set of rows, no coefficients are significant at the five-percent level and only two out of four are significant at the 10 percent level. That is, none of the three structural models nested in our specification is robustly significant. This result would normally be taken as showing that consumption is not meaningfully priced in the cross-section of returns.

That conclusion would be premature, though. The second set of rows shows that when we use our three-window bandpass specification, low-frequency shocks are in fact priced significantly for all sets of test assets, at the 10 percent level in two specifications and five percent in the other two. Business-cycle and higher-frequency shocks, on the other hand, are not priced.[34] Tests for equality of the coefficients (p-values of the chi2 test are reported in the table) strongly reject the null of equality in all but the last case.[35]

The third set of rows reports the results of the bandpass estimation when we constrain $q_2 = q_3$, or, on other words, we use only two bandpass windows: cycles longer than 8 years (low frequencies) and shorter than 8 years. The coefficients on the low frequency shocks change only minimally and are significant in all cases; cycles shorter than 8 years in fact have average risk prices with the wrong sign. A t-test for the difference is statistically significant for all sets of test assets.

We conclude that, when we use the bandpass specification, we find clear evidence that low-frequency shocks to consumption growth are actually priced, and the price of risk is significantly different than for higher-frequency fluctuations.

Figure 5 plots the estimated spectral weighting functions obtained using the 25 Fama–French portfolios for the utility specification (top row) and the bandpass specification (bot-

[34] The appendix reports results using bootstrapped t-statistics instead of the asymptotic approximation and we obtain similar results.

[35] In addition, the hypothesis that the weighting function is monotonically downward sloping cannot be rejected statistically, consistent with the view that investors are relatively more averse to low-frequency fluctuations.

tom row).[36] The left panels plot all frequencies, while the right panels zoom in on the cycles longer than 5 years. The lighter shaded area corresponds to the 95-percent confidence intervals, and the darker shaded area reports 1 standard deviation intervals.

Consistent with the results in table 3, the figure shows significant weight at low frequencies for both the bandpass and the utility specification. The price of low-frequency shocks is estimated quite precisely using the bandpass specification (and is significantly different from zero at the 95 percent level), while the standard errors of the utility specification estimates diverge quickly as we look at frequencies closer to zero, confirming the large amount of statistical uncertainty exactly in the frequency range most important for Epstein–Zin preferences.

Table 3 and figure 5 together show that when we use the bandpass specification to estimate *average* risk prices in the three frequency ranges, we find that low-frequency shocks are significantly priced, consistent with the economic intuition underlying Epstein–Zin preferences. Using the frequency-domain decomposition leads us to very different conclusions about the underlying theories than standard time-domain techniques would have. The results that employ the utility specification show little support for Epstein–Zin preferences. Looking at the problem using the bandpass filter and targeting the economically relevant set of frequencies instead yields strong and robust support for the idea that low-frequency shocks to the economy are priced in equity markets.

To further assess the role of estimation uncertainty, we also report in figure A1 confidence intervals for the risk prices that ignore the estimation uncertainty of the VAR, and therefore treat the VAR and the associated $G_j(\omega)$ functions as certain. These (tighter) confidence intervals are captured by the darker shaded area. For the utility specification, the difference is dramatic: ignoring the fact that the very lowest frequencies are hard to measure, one would conclude that Epstein–Zin preferences are strongly supported statistically, since the

[36] The appendix reports the factor loadings of the size and book/market sorted portfolios on low-frequency and business-cycle-frequency fluctuations.

dark confidence intervals are tight even around frequency zero. The light shaded area, which reports the confidence intervals including the estimation uncertainty at those frequencies, reveals that the power of a test of the Epstein–Zin model is much smaller, since all the weight of the model is on frequencies that are extremely difficult to measure.

A comparison of the light and shaded confidence intervals for the bandpass specification reveals that treating long-run shocks as all shocks with cycles longer than the business cycle makes the estimation results much more robust. Since the first section of $Z^{BP}(\omega)$ covers a range of frequencies that not only includes those isolated by Epstein–Zin preferences (cycles lasting 210 years or more) but also shorter frequencies which are much better identified in the data (as low as 8 years), adding the estimation uncertainty from the VAR increases the width of the confidence bands in a less dramatic way.

5.3 Alternative specifications of the model: stochastic volatility and external habit formation

In our baseline specification (reported in table 3 and figure 5) we have estimated the pricing of different fluctuations in consumption by estimating two specifications of the Z function.

In this section we report the results of two additional model specifications. The first one estimates the price of risk for fluctuations in volatility in addition to consumption growth. The second one estimates the prices of risk allowing these prices to vary with the surplus consumption ratio (calculated as in Campbell and Cochrane (1999)): it therefore considers a conditional version of the model, where each element that enters the function $Z_t(\omega)$ is allowed to depend on a time-t conditioning variable (the surplus consumption ratio).

In a model where both consumption growth and volatility are priced (as described in Section 4), the stochastic discount factor will depend on *two* weighting functions, one for each priced variable: $Z_c(\omega)$ for consumption and $Z_v(\omega)$ for volatility. In our implementation, we use realized variance as a proxy for consumption volatility, since it can be estimated much more precisely than the variance of consumption growth using high-frequency data (in addition, the two are closely related in those models).[37] For each weighting function, we employ either the utility or the bandpass basis. It is important to note that the estimation of the weighting functions requires observing as many shocks as the parameters of the Z functions to estimate. For example, estimating a 3-parameter utility specification for Z_c and Z_v requires estimating a VAR with at least 6 variables. We do so by adding 3 extra principal components from the Ludvigson and Ng (2007) data; we report the results in Appendix table A3. Given the large dimension of the VAR, it is not surprising that the estimates of the risk prices are not statistically significant.[38] Overall, we do not have enough power in our data to be able to statistically discern the pricing of volatility in addition to consumption.

Table A4 reports instead the specification in which Z_t is allowed to vary with the surplus-consumption ratio, in the spirit of Campbell and Cochrane (1999). The left side of each panel reports the results when each element of Z_t is allowed an unrestricted interaction with the surplus consumption ratio, constructed exactly as in Campbell and Cochrane (1999) and then normalized to have zero mean and unit variance. The "level" estimates indicate the average price of risk for each rotated shock; the "interaction" coefficients are negative when the risk prices are higher in times of low surplus consumption (as predicted by the habit model). For both the utility and the bandpass specifications, the estimation provides evidence that prices of risk for the low-frequency shocks are indeed higher when the surplus consumption ratio is higher; in addition, the unconditional level of the price of risk is positive as in the baseline estimate. This analysis reinforces our main finding that low-frequency fluctuations are significantly priced, and adds evidence that risk prices for these fluctuations increase in bad times (times of low consumption relative to the habit).

The appendix describes a wide range of robustness tests and extensions to the results in

[37] Realized variance has been often used in studies of long-run risks and intertemporal CAPM, for example by Bansal et al. (2014) and Campbell et al. (2015).

[38] The table also reports a restricted specification, with only the low-frequency component of consumption and the three components of volatility. Again, the results are not statistically significant.

addition to those mentioned above. We show that the results are qualitatively unchanged when we use annual rather than quarterly data, when we use alternative data to form the factors for the FAVAR, when we use alternative methods to construct confidence intervals, and when we use two lags in the VAR instead of three (though in that case the results are no longer statistically significant). Overall, the results reported in table A5 confirm that low-frequency fluctuations are significantly priced across a number of different data sources, specifications, and estimation methods.

6 Conclusion

This paper studies risk prices in the frequency domain. The impulse response of consumption growth to a given shock to the economy can be decomposed into components of varying frequencies. We show that in any log-linear asset pricing model, we can analytically derive the price of risk that investors assign to fluctuations in consumption at different frequencies. In addition, this frequency-specific price of risk depends only on the investor's preferences, not on the underlying consumption dynamics in the economy.

First, we show *quantitatively* how important consumption fluctuations at different frequencies are for investors in different models. In standard calibrations of Epstein–Zin preferences investors place most weight on consumption cycles that last a century or more when pricing assets. Conversely, very little weight is placed on fluctuations shorter than 100 years, and there is essentially no weight on business-cycle fluctuations.

Second, we provide estimates of the spectral weighting function in US equity markets. While the highly constrained preferences for dynamics implied by standard models (e.g., Epstein–Zin, habit formation, etc.) fail to explain asset prices empirically in a robust way, we show that a generalization of preferences for dynamics in the frequency domain yields strong support for aversion to low-frequency fluctuations by investors in the equity market.

References

Alvarez, Fernando and Urban J. Jermann, "Using Asset Prices to Measure the Cost of Business Cycles," *Journal of Political Economy*, 2004, *112(6)*, 1223–1256.

____ **and** ____ , "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica*, 2005, *73(6)*, 1977–2016.

Andrews, Donald W.K. and Werner Ploberger, "Testing for Serial Correlation Against an ARMA(1,1) Process," *Journal of the American Statistical Association*, 1996, *91(435)*, 1331–1342.

Backus, David, Mikhail Chernov, and Stanley Zin, "Sources of Entropy in Representative Agent Models," *Journal of Finance*, 2014, *69(1)*, 51–99.

Bai, Jushan and Serena Ng, "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions," *Econometrica*, 2006, *74* (4), 1133–1150.

Bandi, Federico and Andrea Tamoni, "Business-cycle consumption risk and asset prices," *Working Paper*, 2014.

Bansal, Ravi and Amir Yaron, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 2004, *59* (4), 1481–1509.

____ **, Dana Kiku, and Amir Yaron**, "Long-Run Risks, the Macroeconomy, and Asset Prices," *American Economic Review*, 2010, *100(2)*, 542–546.

____ **,** ____ **, and** ____ , "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 2012, *1(1)*, 183–221.

____ **,** ____ **, Ivan Shaliastovich, and Amir Yaron**, "Volatility, the Macroeconomy and Asset Prices," *Journal of Finance*, 2014, *69* (6), 2471–2511.

———, **Robert F. Dittmar, and Christian T. Lundblad**, "Consumption, Dividends, and the Cross-Section of Equity Returns," *Journal of Finance*, 2005, *60(4)*, 1639–1672.

Beeler, Jason and John Y. Campbell, "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 2012, *1(1)*, 141–182.

Bernanke, Ben S., Jean Boivin, and Piotr Eliasz, "Measuring the Effects of Monetary Policy: A Factor-augmented Vector Autoregressive (FAVAR) Approach.," *Quarterly Journal of Economics*, 2005, *120* (1), 387–422.

Bidder, Rhys and Ian Dew-Becker, "Long-run risk is the Worst-Case Scenario," 2015. Working paper.

Borovicka, Jaroslav, Lars P. Hansen, Mark Hendricks, and Jose A. Scheinkman, "Risk-price dynamics," *Journal of Financial Econometrics*, 2011, *9*, 3–65.

Breusch, Trevor S and Adrian R Pagan, "A simple test for heteroscedasticity and random coefficient variation," *Econometrica: Journal of the Econometric Society*, 1979, pp. 1287–1294.

Brillinger, David R., *Time Series: Data Analysis and Theory*, McGraw Hill, 1981.

Calvet, Laurent and Adlai Fisher, "Multifrequency news and stock returns," *Journal of Financial Economics*, 2007, *86*, 178–212.

Campbell, John Y., "Intertemporal Asset Pricing Without Consumption Data," *American Economic Review*, 1993, *83(3)*, 487–512.

——— **and John H. Cochrane**, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 1999, *107* (2), 205–251.

____ **and N. Gregory Mankiw**, "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," in Olivier Jean Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual*, 1989.

____ **and Robert J. Shiller**, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1988, *1(3)* (3), 195–228.

____ **and Tuomo Vuolteenaho**, "Bad Beta, Good Beta," *American Economic Review*, 2004, *94* (5), 1249–1275.

____ **, Christopher Polk, and Tuomo Vuolteenaho**, "Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns," *Review of Financial Studies*, 2010, *23(1)*, 305–344.

____ **, Stefano Giglio, Christopher Polk, and Robert Turley**, "An Intertemporal CAPM with Stochastic Volatility," 2015. Working paper.

Carr, Peter and Dilip Madan, "Option valuation using the fast Fourier transform," *Journal of Computational Finance*, 1999, *2*, 61–73.

Chen, Hui, Winston Dou, and Leonid Kogan, "Measuring the "Dark Matter" in Asset Pricing Models," 2013. Working paper.

Christiano, Lawrence J. and Terry J. Fitzgerald, "The Band Pass Filter," *International Economic Review*, 2003, *44(2)*, 435–465.

Collin-Dufresne, Pierre, Michael Johannes, and Lars Lochstoer, "Parameter Learning in General Equilibrium: The Asset Pricing Implications," *American Economic Review*, Forthcoming.

Constantinides, George M., "Habit Formation: A Resolution of the Equity Premium Puzzle," *The Journal of Political Economy*, 1990, *98(3)*, 519–543.

——— **and Anisha Ghosh**, "Asset Pricing with Countercyclical Household Consumption Risk," 2013. Working paper.

Daniel, Kent D. and David A. Marshall, "The Equity Premium Puzzle and the Risk-Free Rate Puzzle at Long Horizons," *Macroeconomic Dynamics*, 1997, *1(2)*, 452–484.

Dew-Becker, Ian, "A Model of Time-Varying Risk Premia with Habits and Production," 2011. Working paper.

Drechsler, Itamar and Amir Yaron, "What's Vol Got to Do with it?," *The Review of Financial Studies*, 2011, *24(1)*, 1–45.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, "Transform analysis and asset pricing for affine jump-diffusions," *Econometrica*, 2000, *68*, 1343–1376.

Efron, Bradley and R.J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, 1994.

Epstein, Larry G. and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *The Journal of Political Economy*, 1991, *99(2)*, 263–286.

———**, Emmanuel Farhi, and Tomasz Strzalecki**, "How Much Would You Pay to Resolve Long-Run Risk?," *American Economic Review*, 2014, *104*, 2680–2697.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 1989, *57(5)*, 1121–1152.

Gourio, Francois, "Disaster Risk and Business Cycles," *American Economic Review*, 2012, *102(6)*, 2734–2766.

Hansen, Lars P, "Generalized Method of Moments Estimation," in Stephen N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave: A Dictionary of Economics*, Macmillan, 2008.

Hansen, Lars P. and Jose A. Scheinkman, "Long-term risk: an operator approach," *Econometrica*, 2009, *77*, 177–234.

—— **and Kenneth J. Singleton**, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 1982, *50(5)*, 1269–1286.

——**, John C. Heaton, and Nan Li**, "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 2008, *116(2)*, 260–302.

Hansen, Lars Peter and Ravi Jagannathan, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 1991, *99(2)*, 225–262.

Judd, Kenneth L., *Numerical Methods for Economists*, MIT Press, Cambridge, MA, 1999.

Jurado, Kyle, Sydney Ludvigson, and Serena Ng, "Measuring Uncertainty," *American Economic Review*, 2015, *105* (3), 1177–1216.

Kaltenbrunner, Georg and Lars A. Lochstoer, "Long-Run Risk Through Consumption Smoothing," *The Review of Financial Studies*, 2010, *23(8)*, 3141–3189.

Larrain, Borja and Motohiro Yogo, "Does firm value move too much to be justified by subsequent changes in cash flow?," *Journal of Financial Economics*, 2008, *87(1)*, 200–226.

Lettau, Martin and Jessica A. Wachter, "Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium," *Journal of Finance*, 2007, *62* (1), 55–92.

—— **and Sydney Ludvigson**, "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 2001, *56* (3), 815–849.

Lewis, Richard and Gregory C. Reinsel, "Prediction of Multivariate Time Series by Autoregressive Model Fitting," *Journal of Multivariate Analysis*, 1985, *16(3)*, 393–411.

Ludvigson, Sydney C. and Serena Ng, "The Empirical Risk-Return Relation: A Factor Analysis Approach," *Journal of Financial Economics*, 2007, *83(1)*, 171–222.

Lustig, Hanno and Stijn van Nieuwerburgh, "The Returns on Human Capital: Good News on Wall Street is Bad News on Main Street," *Review of Financial Studies*, 2008, *21(5)*, 2097–2137.

Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-Jørgensen, "Long-Run Stockholder Consumption Risk and Asset Returns," *Journal of Finance*, 2009, *64(6)*, 2427–2479.

Melino, Angelo and Alan X. Yang, "State-Dependent Preferences can Explain the Equity Premium Puzzle," *Review of Economic Dynamics*, 2003, *6(4)*, 806–830.

Merton, Robert C., "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 1973, *41* (5), 867–887.

Mitchell, Heather and Peter Brockwell, "Estimation of the Coefficients of a Multivariate Linear Filter Using the Innovations Algorithm," *Journal of Time Series Analysis*, 1997, *18(2)*, 157–179.

Newey, Whitney K. and Kenneth D. West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 1987, *55(3)*, 703–708.

Ortu, Fulvio, Andrea Tamoni, and Claudio Tebaldi, "Long-Run Risk and the Persistence of Consumption Shocks," *Review of Financial Studies*, 2013, *26*, 2876–2915.

Otrok, Christopher, B. Ravikumar, and Charles H. Whiteman, "Habit Formation: A Resolution of the Equity Premium Puzzle?," *Journal of Monetary Economics*, 2002, *49(6)*, 1261–1288.

Parker, Jonathan A. and Christian Julliard, "Consumption Risk and the Cross Section of Expected Returns," *Journal of Political Economy*, 2005, *113(1)*, 185–222.

Priestley, M.B., *Spectral Analysis and Time Series*, Elsevier, 1981.

Restoy, Fernando and Philippe Weil, "Approximate Equilibrium Asset Prices," 1998. NBER Working Paper.

Schorfheide, Frank, "VAR Forecasting Under Misspecification," *Journal of Econometrics*, 2005, *128(1)*, 99–136.

Singleton, Kenneth, "Estimation of affine asset pricing models using the empirical characteristic function," *Journal of Econometrics*, 2001, pp. 111–141.

Szroeter, Jerzy, "A class of parametric tests for heteroscedasticity in linear econometric models," *Econometrica: Journal of the Econometric Society*, 1978, pp. 1311–1327.

van Binsbergen, Jules, Michael Brandt, and Ralph Koijen, "On the Timing and Pricing of Dividends," *The American Economic Review*, 2012, *102(4)*, 1596–1618.

Vuolteenaho, Tuomo, "What Drives Firm-Level Stock Returns?," *Journal of Finance*, 2002, *57(1)*, 233–264.

Wachter, Jessica A., "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?," *Journal of Finance*, 2013, *68(3)*, 987–1035.

Whittle, Peter, "Gaussian estimation in stationary time series," *Bulletin of the International Statistical Institute*, 1962, *39*, 105–129.

Yu, Jianfeng, "Using Long-Run Consumption-Return Correlations to Test Asset Pricing Models," 2012. Working paper.

A Derivation of result 1

For any $g_{j,k}$, we have

$$g_{j,k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{G}_j(\omega)(\cos(\omega k) + i\sin(\omega k))\, d\omega \tag{60}$$

Now since $g_{j,k} = 0$ for $k < 0$, for any $k > 0$ we have

$$\begin{aligned} g_{j,k} &= g_{j,k} + g_{j,-k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{G}_j(\omega) \begin{pmatrix} \cos(\omega k) + i\sin(\omega k) \\ + \cos(-\omega k) + i\sin(-\omega k) \end{pmatrix} d\omega && (61) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{G}_j(\omega)\, 2\cos(\omega k)\, d\omega && (62) \end{aligned}$$

Furthermore, note that the complex part of $\tilde{G}(\omega)$ multiplied by any $\cos(\omega k)$ for integer k integrates to zero, which is why we can just study $G \equiv re\left(\tilde{G}\right)$. We thus have

$$\sum_{k=0}^{\infty} z_k g_{j,k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} G_j(\omega)\left(z_0 + 2\sum_{k=1}^{\infty} z_k \cos(\omega k)\right) d\omega \tag{63}$$

The result is related to Parseval's theorem, but it has the advantage of yielding a decomposition that is entirely real-valued, which is achieved by exploiting the fact that $g_{j,k} = 0$ for $k < 0$.

B Quality of the linear approximation for Epstein–Zin preferences

This section examines the quality of the linear approximation used in analyzing Epstein–Zin preferences. The linear approximation is compared to the solution from a high-order projection of Bansal and Yaron's (2004) long-run risk model which is useful for having highly volatile and persistent state variables.

B.1 Model

The model from Bansal and Yaron (2004) is

$$\Delta c_t = x_{t-1} + \sigma_{t-1}\varepsilon_{c,t} \tag{64}$$

$$x_t = \phi_x x_{t-1} + \varphi_e \sigma_{t-1}\varepsilon_{x,t} \tag{65}$$

$$\sigma_t^2 = v_1\left(\sigma_{t-1}^2 - \bar{\sigma}^2\right) + \bar{\sigma}^2 + \sigma_w \varepsilon_{\sigma,t} \tag{66}$$

Time in this model is monthly. Investors are assumed to have Epstein–Zin preferences with the time discount factor of β, an elasticity of intertemporal substitution of ρ^{-1} and risk aversion of α.

B.2 Solution and simulation of the model

We solve the model using projection onto Chebyshev polynomials, which are solved through collocation (see Judd (1999)). Both lifetime utility and also all asset prices are solved for as 9th-order polynomials in the two state variables, x_t and σ_t^2. Expectations are calculated using Gaussian quadrature with 15 points. All results involving simulations are calculated based on 10,000 months of simulated data. We constrain σ_t^2 to be greater than 10^{-7} – setting it to 10^{-7} if a shock drives it below that level.

B.3 Returns on zero-coupon consumption claims

We begin by examining returns on zero-coupon consumption claims. Define

$$PC_{n,t} = E_t^Q\left[\frac{C_{t+n}}{C_t}\right] \tag{67}$$

where Q is the pricing measure. $PC_{n,t}$ is thus the price of a claim to consumption on date $t+n$ scaled by current consumption. The return is then

$$R_{n,t+1} = \frac{PC_{n-1,t+1}C_{t+1}}{PC_{n,t}C_t} \tag{68}$$

Given that we have the pricing kernel from the model, it is straightforward to solve for $PC_{n,t}$ recursively, starting from the boundary condition that $PC_{0,t} = 1$.

As discussed in the main text, the approximation for the pricing kernel is

$$\Delta\mathbb{E}_{t+1} \log M_{t+1} = -\alpha\sigma_t\varepsilon_{c,t+1} + \frac{(\rho-\alpha)\,\theta}{1-\phi_x\theta}\sigma_t\varphi_e\varepsilon_{x,t+1} - \frac{\rho-\alpha}{1-\rho}k_1\frac{\theta}{1-v_1\theta}\sigma_w\varepsilon_{\sigma^2,t+1} \tag{69}$$

where, as discussed in the derivation of the pricing kernel under stochastic volatility, k_1 is the coefficient in the expression

$$E_t r_{w,t+1} = \bar{r} + \rho x_t + k_1\sigma_t^2 \tag{70}$$

Our analysis does not depend on any particular assumption about the structure of the expectation of the log pricing kernel. We therefore simply approximate

$$E_t \log M_{t+1} = \bar{m} + m_1\sigma_t^2 \tag{71}$$

That expression is also what is obtained in Bansal and Yaron's (2004) solution. We find $\bar{m}$ and m_1 as the values that best approximate our numerical solution (by minimizing the squared difference between $\bar{m}+m_1\sigma_t^2$ and the value in the numerical solution summed across the collocation points).

Finally, we also need a value for the coefficient k_1. As with $E_t \log M_{t+1}$, we obtain k_1 by simply regressing the values of $E_t r_{w,t+1} - \rho x_t$ from the numerical solution on a constant and σ_t^2.

The pricing equation is then,

$$pc_{n,t} = \log E_t \exp \left(\begin{array}{c} \bar{m} + m_1 \sigma_t^2 - \rho x_t - \alpha \sigma_t \varepsilon_{c,t+1} + \frac{(\rho-\alpha)\theta}{1-\phi_x \theta} \sigma_t \varphi_e \varepsilon_{x,t+1} - \frac{\rho-\alpha}{1-\rho} k_1 \frac{\theta}{1-v_1\theta} \sigma_w \varepsilon_{\sigma^2,t+1} \\ + pc_{n-1,t+1} + x_t + \sigma_t \varepsilon_{c,t+1} \end{array} \right) \tag{72}$$

where $pc_{n,t} = \log PC_{n,t}$. We then guess that prices can be expressed as

$$pc_{n,t} = \bar{p} + p_{x,n} x_t + p_{\sigma^2,n} \sigma_t^2 \tag{73}$$

This equation can be solved using standard methods to obtain

$$\bar{p}_n = \log \beta - \bar{r}_f + \bar{p}_{n-1} + p_{\sigma^2,n-1} (1 - v_1) \bar{\sigma}^2 + \frac{1}{2} \left(-\frac{\rho - \alpha}{1-\rho} k_1 \frac{\theta}{1 - v_1 \theta} + p_{\sigma^2,n-1} \right)^2 \sigma_w^2 \tag{74}$$

$$p_{x,n} = -\rho + p_{x,n-1} \phi_x + 1 \tag{75}$$

$$p_{\sigma^2,n} = -r_1 + p_{\sigma^2,n-1} v_1 + \frac{1}{2} (1-\alpha)^2 + \frac{1}{2} \left(\frac{(\rho - \alpha)\theta}{1 - \phi_x \theta} + p_{x,n-1} \right)^2 \varphi_e^2 \tag{76}$$

with the boundary conditions $\bar{p}_0 = p_{x,0} = p_{\sigma^2,0} = 0$.

We compare the risk premia implied by our approximation to those solved for numerically with the projection method by plotting Sharpe ratios across horizons calculated under our linear approximation and the numerical approximation. Figure A2 plots steady-state annualized Sharpe ratios (i.e. evaluated at $\sigma_t^2 = \bar{\sigma}^2$) for zero-coupon consumption claims with maturities from 1 to 240 months in the numerical solution to the model and also our log-linear approximation. The two series differ by less than 0.014 across all maturities. The root mean squared error is 0.0051. The linear approximation thus provides a highly accurate approximation to the risk premium for consumption claims and describes very well how the risk premia vary with maturity.

B.4 Hansen–Jagannathan distance

A standard measure of the distance between pricing kernels is the Hansen–Jagannathan (HJ; 1991) distance. For two pricing kernels, M_{t+1} and M'_{t+1}, the HJ distance is $std\left(M_{t+1} - M'_{t+1}\right)$. It is straightforward to show that the HJ distance is equal to the maximal difference in Sharpe ratios for an asset priced by the two kernels.

To examine how well the linearization approximates the numerically approximated pricing kernel, we calculate the HJ distance for a range of calibrations of the long-run risk model with different levels of persistence for expected consumption growth and volatility (ϕ_x and v_1). We hold φ_e and $\bar{\sigma}^2$ fixed across the simulations, which means that the unconditional standard deviation of x_t rises as ϕ_x rises. So the simulations with higher ϕ_x not only test robustness against higher persistence, they also test robustness against models where the state variable x_t moves farther away from its steady-state. Increasing dispersion in x_t also increases dispersion in the wealth/consumption ratio. So since our approximation is around a constant wealth/consumption ratio, the simulations with more persistent x_t provide a tougher test of the approximation.

We set σ_w in each simulation so that the unconditional standard deviation of σ_t^2 is unchanged from Bansal and Yaron's (2004) original calibration. We make that choice to ensure that σ_t^2 never falls below zero.

The table below reports the annualized HJ distance between the numerically approximated pricing kernel and the one obtained from our linear approximation scaled by the numerically derived HJ bound (the standard deviation of the pricing kernel divided by its expectation). That is, denoting the projection and linearized pricing kernels as M^{proj} and M^{linear}, the relative HJ distance is $\frac{std\left(M^{\text{proj}} - M^{\text{linear}}\right)}{std\left(M^{\text{proj}}\right)}$.

We report values of ϕ_x and v_1 in terms of the implied half-lives for x_t and σ_t^2.

As we would expect, the magnitude of the errors increases with the persistence of the state variables. The maximum relative HJ distance is 4 percent of the unconditional HJ bound. That is, across all possible assets in the economy, the linear approximation gets risk premia wrong by up to 4 percent. However, for the majority of calibrations, the size of the errors is relatively small. At Bansal and Yaron's (2004) original calibration, the relative HJ distance is only 2.7 percent. That is, risk premia deviate from their true value, for the most extreme possible asset, by only 2.7 parts in 100. The results in the table thus imply that for a broad range of calibrations, the HJ distance between our linear approximation and a numerical solution is reasonably small in relative terms.

x half-life	σ^2 half-life	Relative HJ distance
2.72	4.4	0.027
5	4.4	0.022
7.5	4.4	0.041
1.5	4.4	0.036
2.72	10	0.032
2.72	20	0.035
2.72	1.5	0.022

C Multiple priced variables and stochastic volatility

C.1 General result

The impulse response function in the multivariate case is denoted $\mathbf{g}_k = \mathbf{J}\Gamma^k$, where $\mathbf{g}_k$ is an $m \times n$ matrix whose $\{m, n\}$ element determines the effect of a shock to the nth element of $\boldsymbol{\varepsilon}_t$ on the mth element of $\mathbb{E}_t \bar{\mathbf{x}}_{t+k}$. The innovation to the SDF is then

$$\Delta \mathbb{E}_{t+1} m_{t+1} = -\left(\sum_{k=0}^{\infty} \mathbf{z}_k \mathbf{g}_k \right) \varepsilon_{t+1} \tag{77}$$

The price of risk for the jth element of $\boldsymbol{\varepsilon}$ is simply the jth element of $\sum_{k=0}^{\infty} \mathbf{z}_k \mathbf{g}_k$.

As before, we take the discrete Fourier transform of $\{\mathbf{g}_k\}$, defining

$$\tilde{\mathbf{G}}(\omega) \equiv \sum_{k=0}^{\infty} e^{-i\omega k}\mathbf{g}_k \tag{78}$$

Following the same steps as in section 2 and defining $\mathbf{G}(\omega) \equiv re\left(\tilde{\mathbf{G}}(\omega)\right)$, we arrive at

$$\begin{aligned}\sum_{k=0}^{\infty} \mathbf{z}_k \mathbf{g}_k \mathbf{b}_j &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbf{Z}(\omega)\,\mathbf{G}(\omega)\,\mathbf{b}_j d\omega & (79)\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_m \mathbf{Z}_m(\omega)\,\mathbf{G}_{m,j}(\omega)\,d\omega & (80)\end{aligned}$$

where

$$\mathbf{Z}(\omega) \equiv \mathbf{z}_0 + 2\sum_{k=1}^{\infty}\mathbf{z}_k \cos(\omega k) \tag{81}$$

and where $\mathbf{Z}_m(\omega)$ denotes the mth element of $\mathbf{Z}(\omega)$ and $\mathbf{G}_{m,j}(\omega)$ denotes the m,jth element of $\mathbf{G}(\omega)$. We thus have m different weighting functions, one for each of the priced variables. The m weighting functions each multiply n different impulse transfer functions, $\mathbf{G}_{m,j}(\omega)$. The price of risk for shock j depends on how it affects the various priced variables at all horizons.

C.2 Epstein–Zin with stochastic volatility

Using Result 2, we now extend the results on Epstein–Zin preferences to also allow for stochastic volatility, similarly to Campbell et al. (2015) and Bansal and Yaron (2004). We use the same log-normal and log-linear framework as above. The log stochastic discount factor under Epstein–Zin preferences is,

$$m_{t+1} = \frac{1-\alpha}{1-\rho}\log\beta - \rho\frac{1-\alpha}{1-\rho}\Delta c_{t+1} + \frac{\rho-\alpha}{1-\rho}r_{w,t+1} \tag{82}$$

where $r_{w,t+1}$ is the log return on a consumption claim on date $t+1$. Whereas we previously assumed that consumption growth was log-normal and homoskedastic, we now allow for time-varying volatility driven by a variable σ_t^2. We assume that σ_t^2 follows a linear, homoskedastic, and stationary process. We assume that log consumption growth is driven by a VMA process as in assumption 1, but that now the shocks ε_t have variances that scale linearly with σ_t^2.

It is then straightforward to show that expected returns on a consumption claim will follow

$$\mathbb{E}_t r_{w,t+1} = k_0 + \rho \mathbb{E}_t \Delta c_{t+1} + k_1 \sigma_t^2 \tag{83}$$

where k_0 and k_1 are constants that depend on the underlying process driving consumption growth. Using the Campbell–Shiller approximation, we can then write the innovation to the SDF as

$$\begin{aligned}
\Delta \mathbb{E}_{t+1} m_{t+1} &= -\alpha \Delta c_{t+1} - (\alpha - \rho) \Delta \mathbb{E}_{t+1} \sum_{j=1}^{\infty} \theta^j \Delta c_{t+1+j} && (84) \\
&\quad - \frac{\rho - \alpha}{1 - \rho} \Delta \mathbb{E}_{t+1} \theta k_1 \sigma_{t+1}^2 - \frac{\rho - \alpha}{1 - \rho} \Delta \mathbb{E}_{t+1} \sum_{j=1}^{\infty} \theta^j \theta k_1 \sigma_{t+j+1}^2 && (85)
\end{aligned}$$

The weighting functions for consumption growth and volatility are now

$$\begin{aligned}
Z_C^{EZ-SV}(\omega) &= \alpha + (\alpha - \rho) \sum_{j=1}^{\infty} \theta^j 2 \cos(\omega j) && (86) \\
Z_{\sigma^2}^{EZ-SV}(\omega) &= \theta k_1 \frac{\rho - \alpha}{1 - \rho} \left(1 + \sum_{j=1}^{\infty} \theta^j 2 \cos(\omega j) \right) && (87)
\end{aligned}$$

C.3 Epstein–Zin with time-varying higher moments

This section derives the result from the main text on the general case for Epstein–Zin preferences

We now guess that

$$E_t r_{w,t+1} = k_0 + \rho E_t \Delta c_{t+1} + k_1 \tilde{x}_t \tag{88}$$

Recall that according to the Campbell–Shiller approximation,

$$\Delta\mathbb{E}_{t+1}r_{w,t+1} = \sum_{j=0}^{\infty}\theta^j\Delta\mathbb{E}_{t+1}\Delta c_{t+j+1} - \sum_{j=1}^{\infty}\theta^j\Delta\mathbb{E}_{t+1}r_{w,t+j+1} \tag{89}$$

$$= \sum_{j=0}^{\infty}\theta^j\Delta\mathbb{E}_{t+1}\Delta c_{t+j+1} - \sum_{j=1}^{\infty}\theta^j\Delta\mathbb{E}_{t+1}\left(\rho\Delta c_{t+j+1} + k_1\tilde{x}_{t+j}\right) \tag{90}$$

$$= \sum_{j=0}^{\infty}\theta^j\sum_k g_{k,j}\varepsilon_{k,t+1} - \sum_{j=1}^{\infty}\theta^j\Delta\mathbb{E}_{t+1}\left(\sum_k\left(\rho g_{k,j} + k_1\tilde{g}_{k,j-1}\right)\varepsilon_{k,t+1}\right) \tag{91}$$

The pricing equation for the wealth portfolio is then

$$0 = \log E_t\exp\left(\frac{1-\alpha}{1-\rho}\log\beta - \rho\frac{1-\alpha}{1-\rho}\Delta c_{t+1} + \frac{1-\alpha}{1-\rho}r_{w,t+1}\right) \tag{92}$$

$$= \frac{1-\alpha}{1-\rho}\log\beta - \rho\frac{1-\alpha}{1-\rho}E_t\Delta c_{t+1} + \frac{1-\alpha}{1-\rho}E_t r_{w,t+1} \tag{93}$$

$$+\frac{1}{2}\log E_t\exp\left(-\rho\frac{1-\alpha}{1-\rho}\Delta\mathbb{E}_{t+1}\Delta c_{t+1} + \frac{1-\alpha}{1-\rho}\Delta\mathbb{E}_{t+1}r_{w,t+1}\right) \tag{94}$$

But given the assumption about how $\tilde{x}_t$ affects the distribution of $\varepsilon_{k,t+1}$, the final term above is linear in $\tilde{x}_t$, which confirms our guess for the form of the equation governing the expected return on the wealth portfolio.

We then have for the innovation in the SDF

$$\Delta\mathbb{E}_{t+1}m_{t+1} = -\rho\frac{1-\alpha}{1-\rho}\Delta\mathbb{E}_{t+1}\Delta c_{t+1} + \frac{\rho-\alpha}{1-\rho}\Delta\mathbb{E}_{t+1}r_{w,t+1} \tag{95}$$

$$= -\alpha\sum_k g_{k,0}\varepsilon_{k,t+1} + \frac{\rho-\alpha}{1-\rho}\sum_{j=1}^{\infty}\theta^j\left(\sum_k\left(\left(1-\rho\right)g_{k,j} - k_1\tilde{g}_{k,j-1}\right)\varepsilon_{k,t+1}\right) \tag{96}$$

So then the price of risk for any shock depends now on both the effects of the shock on consumption and also on the volatility process. If there were multiple volatility processes, then we would have multiple extra priced variables.

The time-domain weights for the consumption part are

$$z_{0,\Delta c} = \alpha \tag{97}$$

$$z_{j,\Delta c} = \theta^{j} (\alpha - \rho) \text{ for } j > 0 \tag{98}$$

for $\tilde{x}$, the weights are

$$z_{j,\tilde{x}} = k_1 \frac{\rho - \alpha}{1 - \rho} \theta^{j} \tag{99}$$

These are then rotated into the frequency domain using the same techniques as above.

D Predictability of volatility in consumption growth

In this section we examine whether the variables in our VAR – consumption growth and the two factors – are able to predict the volatility of future consumption growth. While there is certainly evidence that consumption growth is heteroskedastic (one way to find such evidence is to estimate an ARCH model on consumption growth) the key question for us is whether the state variables we examine are related to volatility.

We examine two tests of whether volatility in consumption growth is predicted by the lagged state variables: the Breusch–Pagan (1979) test and the Szroeter (1978) test. The Breusch–Pagan test, when all of the lagged state variables (with lags from 1 to 3) are allowed to potentially predict the variance of innovations, returns a p-value of 0.41. If only the first lag of the state variables is included, the p-value is 0.71. In other words, there is not significant evidence to reject the null that the volatility of consumption growth can be predicted.

We also examined a Szroeter (1978) test, which tests whether any of the lagged state variables individually predicts the variance of consumption growth. In that case, of the nine p-values, the smallest is 0.15 (which does not correct for multiple testing).

The two tests thus suggest that the state variables in the VAR are unable to predict the volatility of future consumption growth. While it may be the case that consumption growth volatility is predictable, the fact that these variables do not predict it means that their risk premia must depend on their effect on the conditional mean of consumption growth, rather than the conditional variance (ignoring the possibility that they predict higher moments like disaster risk). So even if stochastic volatility is priced, the pricing of the three state variables we examine will still reveal the pricing of fluctuations in expected consumption growth.

E Motivation for the bandpass basis from robust estimation

The bandpass specification can be obtained in equilibrium when investors use a robust estimation method for consumption dynamics. The full dynamic model of the economy is obviously difficult to estimate and summarize. There are numerous state variables, and the feedback between the various states and consumption itself may be complicated. Rather than try to actually estimate and process a full model of the economy when pricing assets, investors may summarize the effects of a particular shock on consumption growth by approximating its impulse transfer function with a step function that highlights the average power of the shock at meaningful ranges of frequencies. That way, rather than computing a full transfer function, which has an infinite number of degrees of freedom, they retain only the finite number of degrees of freedom required to define a step function.

Specifically, suppose that the true transfer functions are G_j, but that investors approximate them and price assets using step functions defined as

$$G_j^{Step}(\omega) = \begin{cases} \frac{1}{2\pi/32}\int_0^{2\pi/32} G_j(\kappa)\, d\kappa \text{ for } \omega \in [0, 2\pi/32) \\ \frac{1}{2\pi/8-2\pi/32}\int_{2\pi/32}^{2\pi/8} G_j(\kappa)\, d\kappa \text{ for } \omega \in [2\pi/32, 2\pi/8) \\ \frac{1}{\pi-2\pi/8}\int_{2\pi/8}^{\pi} G_j(\kappa)\, d\kappa \text{ for } \omega \in [2\pi/8, \pi] \end{cases} \tag{100}$$

Since investors do not perceive any variation in the transfer functions G_j^{Step} within the three frequency windows, variation in the weighting function, Z, in those windows is irrelevant – all that matters is its average value. In other words, if investors approximate the transfer function as a step function, then their behavior will be the same as if their weighting function Z were a step function.

More formally, suppose the true weighting function is some arbitrary $\tilde{Z}$, but investors measure risk using transfer functions that are step function approximations to the true transfer function. We then have:

$$\int_0^\pi \tilde{Z}(\omega) G_j^{Step}(\omega) d\omega = \int_0^\pi Z^{BP}(\omega;\mathbf{q}) G_j(\omega) d\omega \tag{101}$$

$$\text{where } Z^{BP}(\omega;\mathbf{q}) = \begin{cases} \frac{1}{2\pi/32}\int_0^{2\pi/32} \tilde{Z}(\kappa) d\kappa \text{ for } \omega \in [0, 2\pi/32) \\ \frac{1}{2\pi/8-2\pi/32}\int_{2\pi/32}^{2\pi/8} \tilde{Z}(\kappa) d\kappa \text{ for } \omega \in [2\pi/32, 2\pi/8) \\ \frac{1}{\pi-2\pi/8}\int_{2\pi/8}^{\pi} \tilde{Z}(\kappa) d\kappa \text{ for } \omega \in [2\pi/8, \pi] \end{cases} \tag{102}$$

So a model where investors have a weighting function $Z^{BP}(\omega;\mathbf{q})$ that is a step function is observationally equivalent to an alternative where they approximate transfer functions G_j as step functions. If the transfer functions that investors estimate are step functions, then risk prices may be calculated using a step function for Z, regardless of its true shape. Moreover, the steps in Z^{BP} correspond exactly to average risk prices in the three frequency windows.

In the end, then, the bandpass specification yields estimates of average risk prices in frequency windows and may be thought of as the result of investors estimating transfer functions G_j^{Step}. We show below that the step functions, G_j^{Step}, are far easier for investors to estimate than unrestricted functions, so we view the bandpass specification in the spirit of Campbell and Mankiw's (1989) estimation of the permanent income hypothesis in the presence of rule-of-thumb consumers. Similar to them, our findings suggest that a rule of thumb – in our case, the step function approximation – performs well.[39]

[39] We also note that approximating consumption dynamics in the frequency domain (rather than in the time domain) is the standard way to compress information in many fields of science. As a practical example,

F Details of the empirical analysis

F.1 Invariance of frequency-domain risk prices under rotations

Combining equations (51)–(54) from section (5.2.2), $\boldsymbol{\varepsilon}_t$ the risk prices for the shocks can be written in the time domain as

$$\mathbf{p} = \sum_{k=0}^{\infty} \mathbf{q}\mathbf{z}_k \mathbf{b}_1 \boldsymbol{\Phi}^k \tag{103}$$

$$\text{where } \mathbf{z}_k \equiv \frac{1}{\pi}\int_0^{\pi} \cos(\omega k)\left[Z_1(\omega), Z_2(\omega), Z_3(\omega)\right]' d\omega \tag{104}$$

$\mathbf{z}_k$ is the time-domain vector of basis functions. Note also that $\mathbf{b}_1\boldsymbol{\Phi}^k$ is the vector of IRFs of consumption growth to the reduced-form shocks $\boldsymbol{\varepsilon}_t$. Given the definition of $\mathbf{z}_k$ and using Result 1, the matrix $\mathbf{W}$ from equation (54) can be written as $\sum_{k=0}^{\infty} \mathbf{z}_k \mathbf{b}_1 \boldsymbol{\Phi}^k$.

Now suppose we considered a set of rotated shocks $\tilde{\boldsymbol{\varepsilon}}_t = \Theta\boldsymbol{\varepsilon}_t$ for some rotation matrix Θ. The estimated reduced-form risk prices for $\tilde{\boldsymbol{\varepsilon}}_t$, $\tilde{\mathbf{p}}$, will then have the property

$$\tilde{\mathbf{p}}\tilde{\boldsymbol{\varepsilon}}_t = \mathbf{p}\boldsymbol{\varepsilon}_t \tag{105}$$

$$\Rightarrow \tilde{\mathbf{p}}\Theta = \mathbf{p} \tag{106}$$

since the pricing kernel must be unchanged whether we examine the reduced-form innovations or a rotation of them.[40]

Furthermore, note that the IRFs for the rotated shocks are simply $\mathbf{b}_1\boldsymbol{\Phi}^k\Theta^{-1}$ (since $\Delta\mathbb{E}_{t+1}\Delta c_{t+k+1} = \mathbf{b}_1\boldsymbol{\Phi}^k\boldsymbol{\varepsilon}_t = \mathbf{b}_1\boldsymbol{\Phi}^k\Theta^{-1}\tilde{\boldsymbol{\varepsilon}}_t$). The rotation matrix for $\tilde{\boldsymbol{\varepsilon}}_t$ therefore becomes $\tilde{\mathbf{W}} = \sum_{k=0}^{\infty} \mathbf{z}_k \mathbf{b}_1 \boldsymbol{\Phi}^k\Theta^{-1} = \mathbf{W}\boldsymbol{\Theta}^{-1}$. So if we again take the reduced-form risk prices, $\tilde{\mathbf{p}}$, and

standard music, image and video compression, and noise-reduction procedures – whose objective is precisely to extract the most important components of each signal – use cosine transforms nearly identical to ours.

[40] That is, for the unrotated shocks, the asset pricing moments are $E[\exp(r_{it+1}) - \exp\left(r^f_{t+1}\right)] = \mathbf{p}\boldsymbol{\varepsilon}_{t+1} r_{i,t+1}$. For the rotated shocks, they are $E[\exp(r_{it+1}) - \exp\left(r^f_{t+1}\right)] = \tilde{\mathbf{p}}\tilde{\boldsymbol{\varepsilon}}_{t+1} r_{i,t+1}$. So the value of the objective function is the same with the rotated shocks when $\tilde{\mathbf{p}}\boldsymbol{\Theta} = \mathbf{p}$.

multiply them by the rotation matrix $\tilde{\mathbf{W}}^{-1}$, we obtain

$$\begin{aligned}
\tilde{\mathbf{p}}\tilde{\mathbf{W}}^{-1} &= \tilde{\mathbf{p}}\left(\mathbf{W}\boldsymbol{\Theta}^{-1}\right)^{-1} && (107)\\
&= \tilde{\mathbf{p}}\Theta\mathbf{W}^{-1} && (108)\\
&= \mathbf{p}\mathbf{W}^{-1} && (109)
\end{aligned}$$

So then whether we take the reduced form risk prices $\mathbf{p}$ and rotate them with $\mathbf{W}^{-1}$ or take a set of rotated risk prices $\tilde{\mathbf{p}}$ and rotate them with $\tilde{\mathbf{W}}^{-1}$, we obtain identical results.

F.2 Derivation of the asset pricing moment conditions

The derivation of the moments identifying the risk prices follows Campbell and Vuolteenaho (2004). Given the assumption of lognormality of all shocks, we can write:

$$\mathbb{E}_t r_{it+1} - r^f_{t+1} + \frac{1}{2}\sigma^2_{it} = -cov_t(m_{t+1}, r_{it+1}) \tag{110}$$

where $\sigma^2_{it} = Var_t(r_{it+1})$. We then note that

$$cov_t(m_{t+1}, r_{it+1}) = cov_t(\Delta\mathbb{E}_{t+1}m_{t+1}, r_{it+1}) = \mathbb{E}_t(\Delta\mathbb{E}_{t+1}m_{t+1}r_{it+1}) = \mathbb{E}_t(-\mathbf{q}\mathbf{W}\boldsymbol{\varepsilon}_{t+1}r_{it+1}) \tag{111}$$

Which implies

$$\mathbb{E}_t r_{it+1} - r^f_{t+1} + \frac{1}{2}\sigma^2_{it} = E_t(\mathbf{q}\mathbf{W}\boldsymbol{\varepsilon}_{t+1}r_{it+1}) \tag{112}$$

Since $\mathbb{E}_t r_{it+1} - r^f_{t+1} + \frac{1}{2}\sigma^2_{it} \approx \mathbb{E}_t[\exp(r_{it+1}) - \exp\left(r^f_{t+1}\right)]$, and taking unconditional expectations, we obtain

$$E[\exp(r_{it+1}) - \exp\left(r^f_{t+1}\right)] \approx E\left[\mathbf{q}\mathbf{W}\boldsymbol{\varepsilon}_{t+1}r_{it+1}\right] \tag{113}$$

F.3 Calculation of standard errors

The procedure in Hansen (2008) involves the following calculation. Define D to be the Jacobian of the moment conditions with respect to the parameters $[p_1', p_2']'$ (where $p_1 = vec\left(\hat{\boldsymbol{\Phi}}\right)$ and $p_2 = \hat{\mathbf{q}}$) partitioned in the two blocks of moments (where $D_{12} = 0$ since the VAR moments do not depend on $\hat{\mathbf{q}}$):

$$D = \begin{bmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{bmatrix}$$

Denote the weighting matrix for the VAR moments as W_1, and the weighting matrix for the asset pricing moments W_2. Finally, define

$$\begin{aligned} A_{11} &= D_{11}' W_1 \\ A_{22} &= D_{11}' W_2 \end{aligned}$$

Then the covariance matrix of $\hat{\mathbf{q}}$ is estimated as,

$$var(\hat{\mathbf{q}} - \mathbf{q}) = \frac{1}{T} \left\{ (A_{22} D_{22})^{-1} A_{22} \left[-D_{21} (A_{11} D_{11})^{-1} A_{11}, I \right] \right\}' S \left\{ (A_{22} D_{22})^{-1} A_{22} \left[-D_{21} (A_{11} D_{11})^{-1} A_{11}, I \right] \right\}$$

where the role played by the prespecified weighting matrices is clear from the terms A_{11} and A_{22}; the uncertainty about the parameters estimated in the first block comes through D_{11} and D_{21}. The matrix S is the covariance matrix of the asset pricing moments.

G Additional robustness tests

This section discusses a range of perturbations of the main model to examine the robustness of the main results.

G.1 Bootstrapped t-statistics

We compute bootstrapped t-statistics following suggestions in Efron and Tibshirani (1994). Specifically, in every bootstrap sample we calculate the t-statistic for each coefficient and then use the simulated distribution of the t-statistics to construct p-values for the test of whether the coefficients are different from zero.

Given a sample size of N, we take uniformly distributed draws from the set $\{1, 2, ..., N\}$ with replacement. The jth draw in bootstrap simulation i is denoted b_j^i. The ith simulated dataset is then the set of VAR residuals and test asset returns for observations $\left\{b_j^i\right\}_{j=1}^N$. To construct the set of state variables, we draw an initial value of the state variables randomly from the set of observations and then use the drawn innovations along with the point estimate for the feedback matrix, $\hat{\mathbf{\Phi}}$, to construct the full sample.

The estimation then proceeds on the simulated dataset exactly as it does on the true dataset. For each simulated sample we form t-statistics for the difference between the bootstrapped estimate of the coefficient and the point estimate. Suppose the empirically observed t-statistic in the main estimate for some coefficient k is equal to $\hat{t}_k > 0$. Then the bootstrapped p-value is twice the fraction of the simulated t-statistics at least as high as $\hat{t}_k$ (for a full description of the procedure, see Efron and Tibshirani, 1994)

The above procedure does not account for uncertainty in the estimation of the principal components for the FAVAR since Bai and Ng (2006) show that estimation error in the principal components is asymptotically negligible when $\sqrt{T}/N \to 0$ (see also the discussion in Ludvigson and Ng (2007)). But when considering the alternative specification in Table A5 that uses a cross-section of only nine time series to estimate the factors, this sampling uncertainty cannot be ignored.

Denote the variables used to calculate the principal components as $x_{i,t}$ for $i \in \{1, 2, ..., 9\}$. We proceed to account for uncertainty in estimating the principal components by resampling the $T \times N$ panel of observed variables $x_{i,t}$, and then re-estimating the factors in each sample, as in Ludvigson and Ng (2007). Denote the factors $f_{j,t}$, and the estimated coefficients on them $\hat{b}_{i,j}$. We then define the PC residuals as

$$\hat{e}_{i,t} \equiv x_{i,t} - \hat{b}_{1,t} f_{1,t} - \hat{b}_{2,t} f_{2,t} \tag{114}$$

As in Ludvigson and Ng (2007), we first estimate an AR(1) process on each individual PC residual $\hat{e}_{i,t}$:

$$\hat{e}_{i,t} = \rho_i \hat{e}_{i,t-1} + v_{i,t} \tag{115}$$

After the AR(1) specification is obtained and $\hat{\rho}_i$ is estimated for each i, $\hat{v}_{it}$ is resampled (preserving the cross-sectional correlation across different i) in each bootstrap sample. We then use the resampled AR(1) innovations to construct bootstrapped values of the individual errors e_{it}. Finally, those bootstrapped errors are added to $\hat{b}_{1,t} f_{1,t} + \hat{b}_{2,t} f_{2,t}$ to yield a bootstrapped sample of x_{it}. Principal components are then constructed using the bootstrapped sample of x_{it}. The remainder of the bootstrap procedure in this case (i.e. for consumption and returns) is otherwise identical to above.

G.2 Risk-sorted portfolios

The 25 Fama–French portfolios were originally constructed because their returns spanned a number of observed anomalies in the cross-section of excess returns. We would not necessarily expect them to have large spreads in their loadings on shocks to consumption growth at different horizons. In this section we therefore construct portfolios that are specifically designed to have a large spread in factor loadings.

In every quarter, we estimate factor loadings with respect to the low- and business-cycle frequency shocks (we refrain from also sorting on the high-frequency shocks to keep the portfolios relatively large and well diversified). The loadings are estimated on quarterly data over the previous 10 years. Stocks are then split in to three equally sized groups according to their loadings on the factors, and we construct nine portfolios by crossing the two groupings of loadings.

The low- and business-cycle frequency shocks are constructed using the bandpass specification. Specifically, we have

$$\Delta \mathbb{E}_{t+1} m_{t+1} = -\mathbf{qW}\boldsymbol{\varepsilon}_{t+1} \tag{116}$$

The rotated shocks are thus,

$$\mathbf{u}_{t+1} = \mathbf{W}\boldsymbol{\varepsilon}_{t+1} \tag{117}$$

And the low- and business-cycle frequency components are the first two elements of $\mathbf{u}$.

G.3 Results

Table A5 reports a range of alternative estimates of the risk prices.

First, we estimate our baseline specification (column 1 of Table 3) using annual data instead of quarterly data, motivated by recent evidence (e.g. Parker and Julliard 2005) that the consumption CAPM works better when looking at more time-aggregated data. The results with annual data are consistent with the ones obtained using quarterly data: low-frequency fluctuations are significantly priced.[41]

The second pair of columns uses two lags in the VAR, rather than the three suggested by cross-validation. The estimates are very close to those obtained with three lags, but they are no longer statistically significant.

The third pair of columns uses the optimal weighting matrix for the moments identifying the risk prices, which is derived by Hansen (2008). The optimal weighting matrix substantially shrinks the standard errors, but the point estimates are only minimally changed from our main results.

[41] We note that the shortest cycle we can identify with annual data is 2 years. Therefore, with annual data we cannot identify the price of risk for our "higher-than-business-cycle" frequency window.

Next, we calculate confidence intervals using the bootstrap procedure described above. The low-frequency risk prices remain highly significant, while risk prices for other frequencies are insignificant.

As described above, we also explore an alternative specification that extracts principal components from nine macro-financial data series as instead of the 131 series of Jurado et al. (2015): aggregate price/earnings and price/dividend ratios; the 10 year/3 month term spread; the Aaa–Baa corporate yield spread (default spread); the small-stock value spread; the unemployment rate minus its 8-year moving average; detrended short-term interest rate; the three-month Treasury yield rate; and Lettau and Ludvigson's (2001) *cay*. We compute the standard errors via bootstrap, with and without incorporating uncertainty in the estimation of the principal components.

Table A5 shows that even with the alternative method of constructing the factors for the FAVAR, and even taking into account uncertainty in the estimation of the factors, we continue to obtain highly significant coefficients on the low-frequency shocks to consumption. The point estimates are somewhat larger than in our main analysis, but not qualitatively different.

As a last extension, we attempt to estimate a version of the model with four instead of three frequency windows. In particular, we split the low-frequency window into one covering cycles lasting between 8 and 100 years and another covering cycles lasting more than 100 years. In order to estimate four risk prices we need four shocks, so we add a third principal component from the 131 data series to the FAVAR. Table A5 shows that in this case we obtain no results that are even close to significant and the standard errors are extremely large compared to the main results.

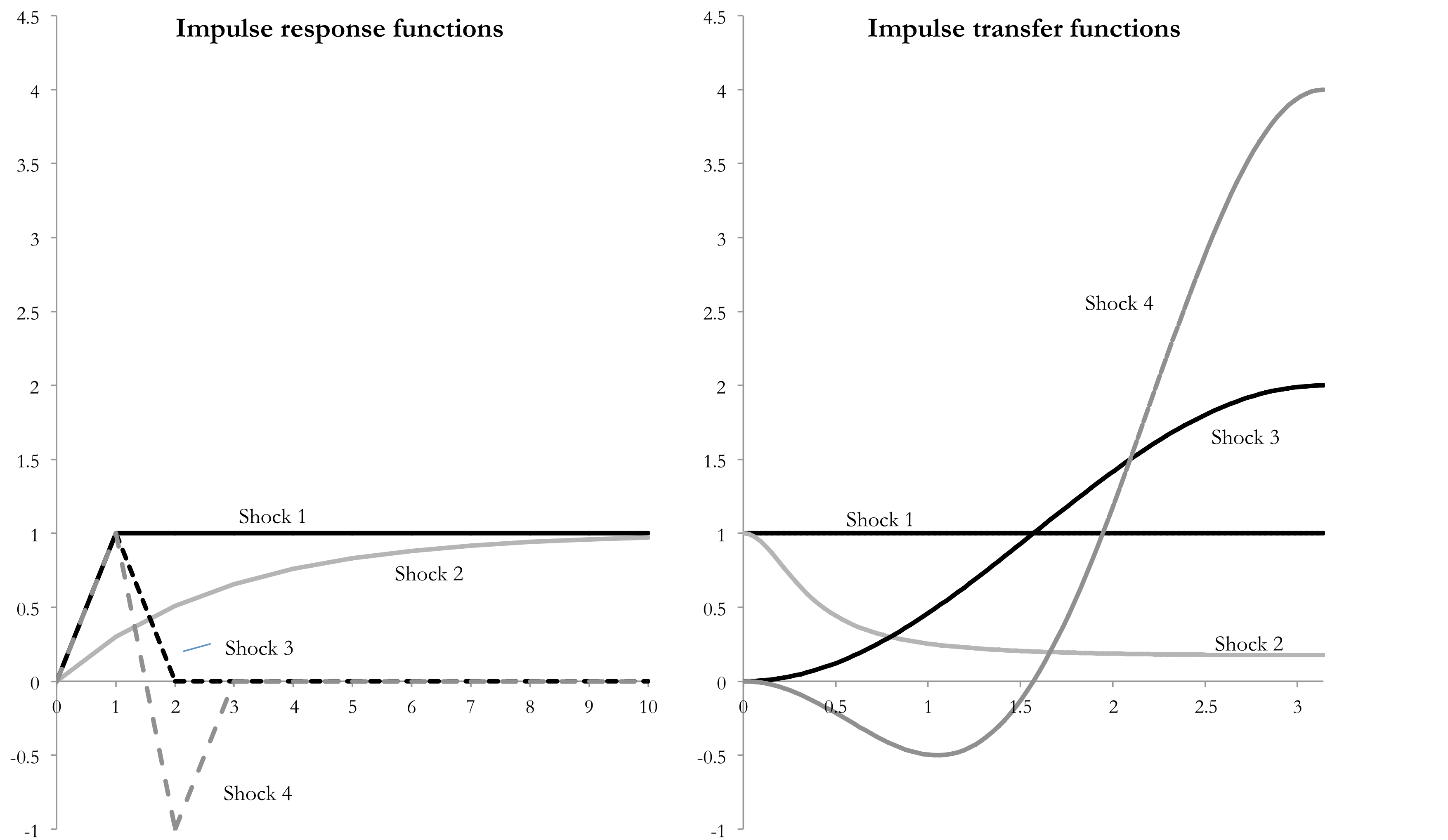


Figure 1. Impulse response functions and impulse transfer functions. The left panel plots responses of the level of consumption to four hypothetical shocks. The right-hand panel plots the cosine transforms of the corresponding shocks to consumption growth, which we refer to as the impulse transfer functions (plotted for frequencies 0 to π).

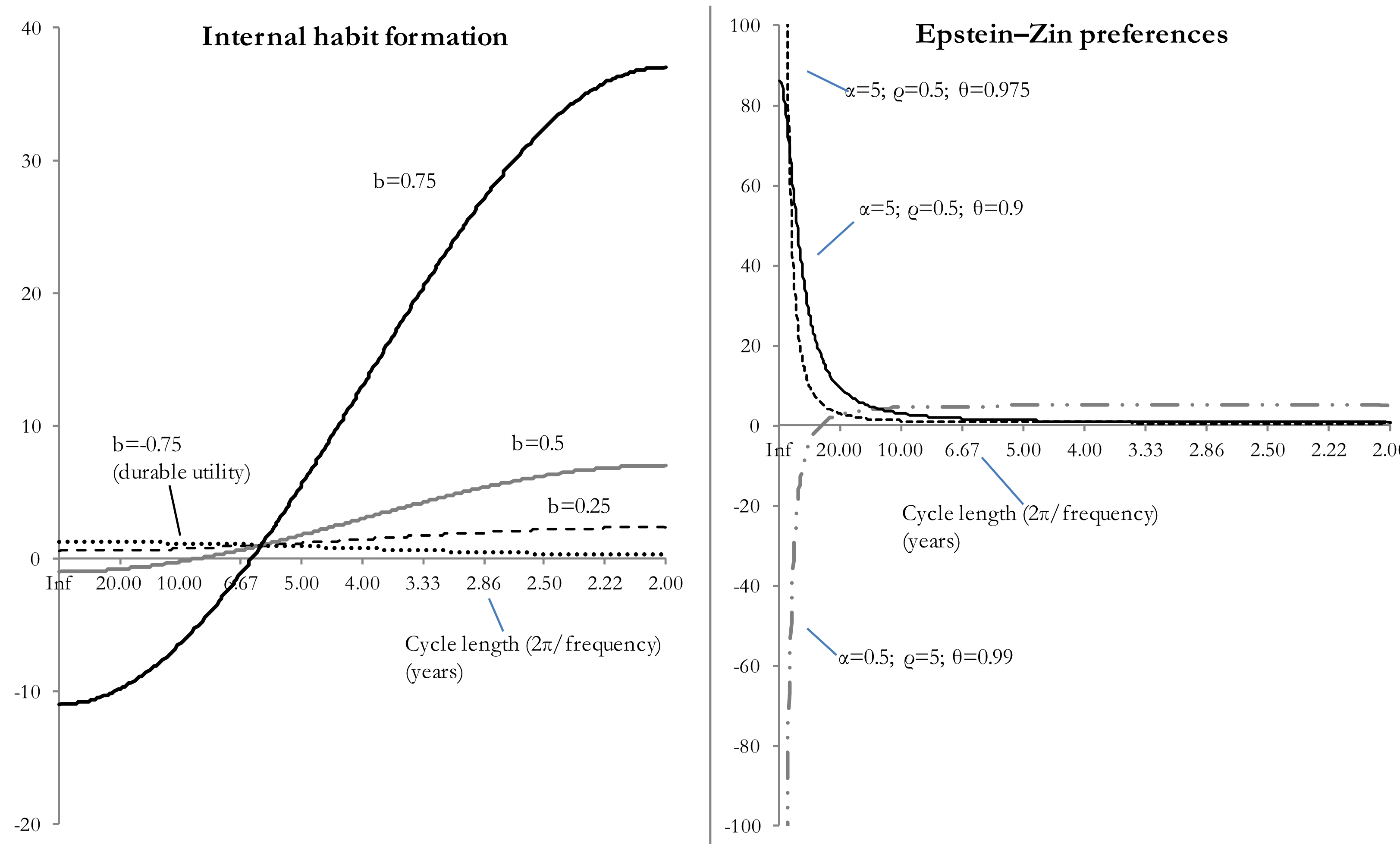


Figure 2. Theoretical spectral weighting functions. Plots of the spectral weighting function Z for various utility functions. The x-axis is the cycle length. In the left-hand panel, the parameter b determines the importance of the internal habit in the agent's utility function. In the right-hand panel, α is the coefficient of relative risk aversion; ρ is the inverse elasticity of intertemporal substitution; and θ is the discounting parameter closely related to the rate of pure time preference.

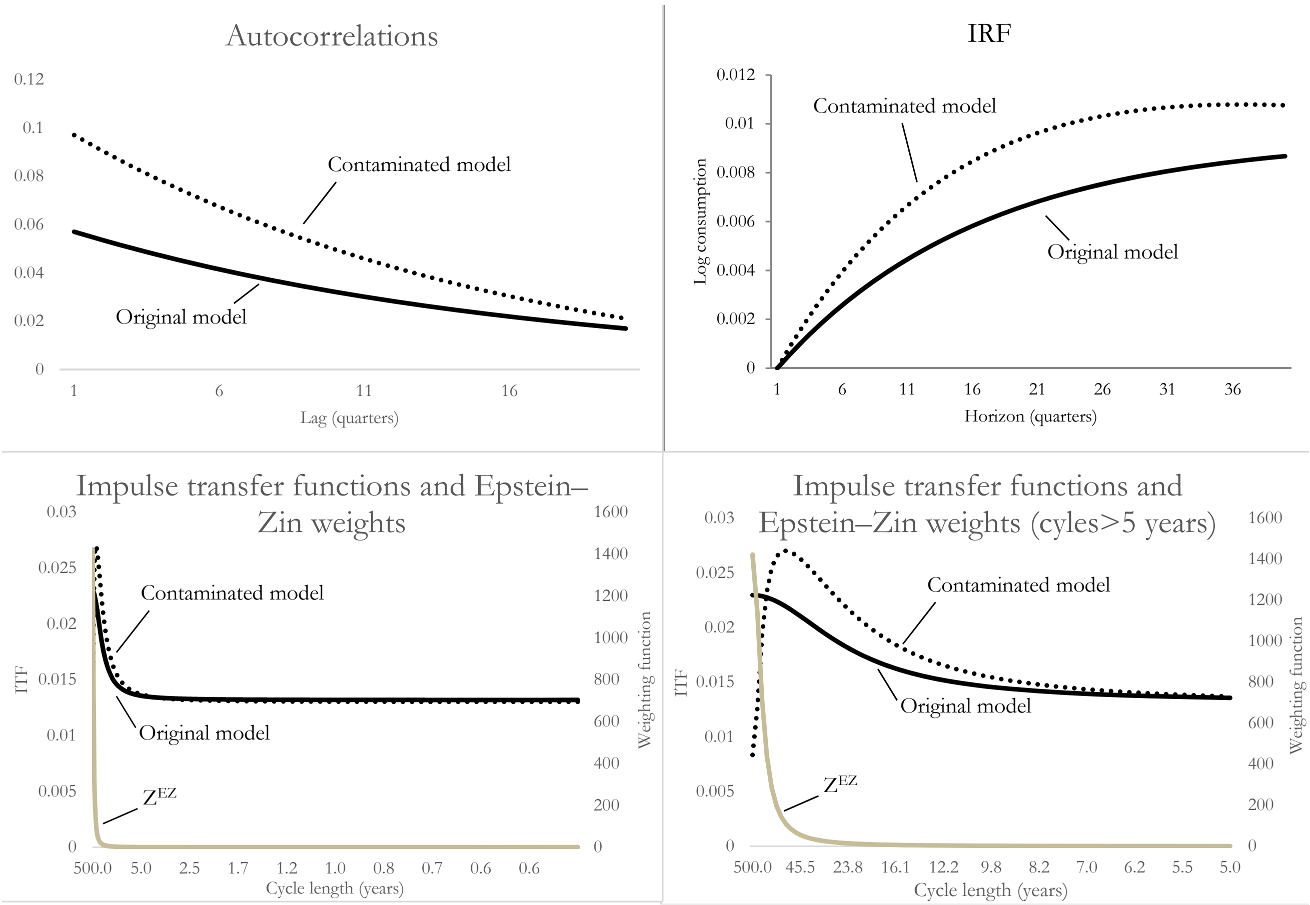
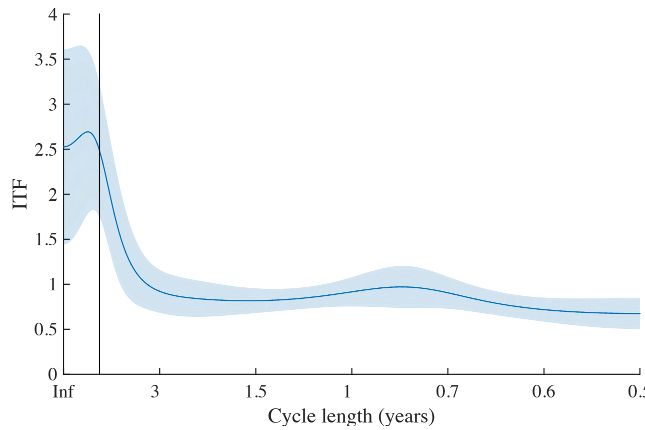
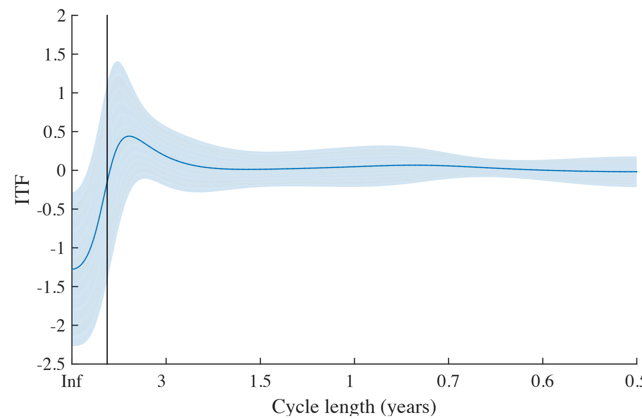


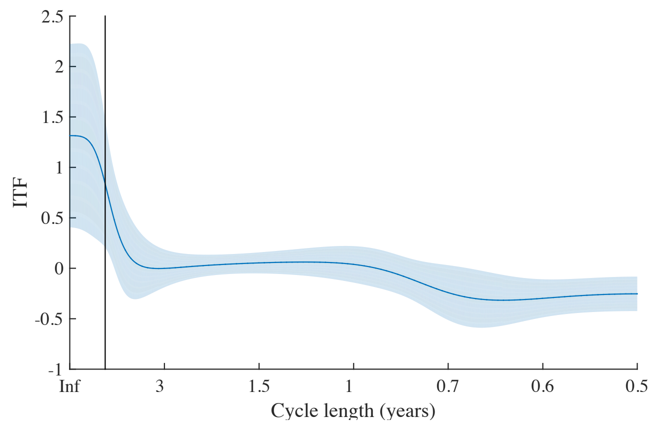
Figure 3. Characteristics of long-run risk model and contaminated version. In all panels, the solid line is the long-run risk model and the dotted line is the contaminated model. The top left panel plots autocorrelations of the consumption process. The top right panel plots IRFs for innovations to the persistent component of consumption growth, x_t. The bottom panels plot ITFs for innovations to the persistent component of consumption growth, x_t, with the right panel zooming in onto cycles longer than 5 years. The bottom panels also plot the weighting function for Epstein–Zin preferences, $Z^{EZ}(\omega)$.



(a) Shock to consumption growth



(b) Shock to first factor



(c) Shock to second factor

Figure 4. Estimated impulse transfer functions for consumption VAR. The figure shows the impulse transfer functions $G_j(\omega)$ estimated from a FAVAR with consumption growth and two principal components of a set of 131 macroeconomic variables. Each of the three panels shows the ITF for a different shock (note that shocks are not orthogonalized). Shaded regions represent 95-percent confidence intervals. The vertical line indicates cycles of 8 years, the break between business cycle frequencies (to the right of it) and "below-business-cycle" frequencies (to the left of it).

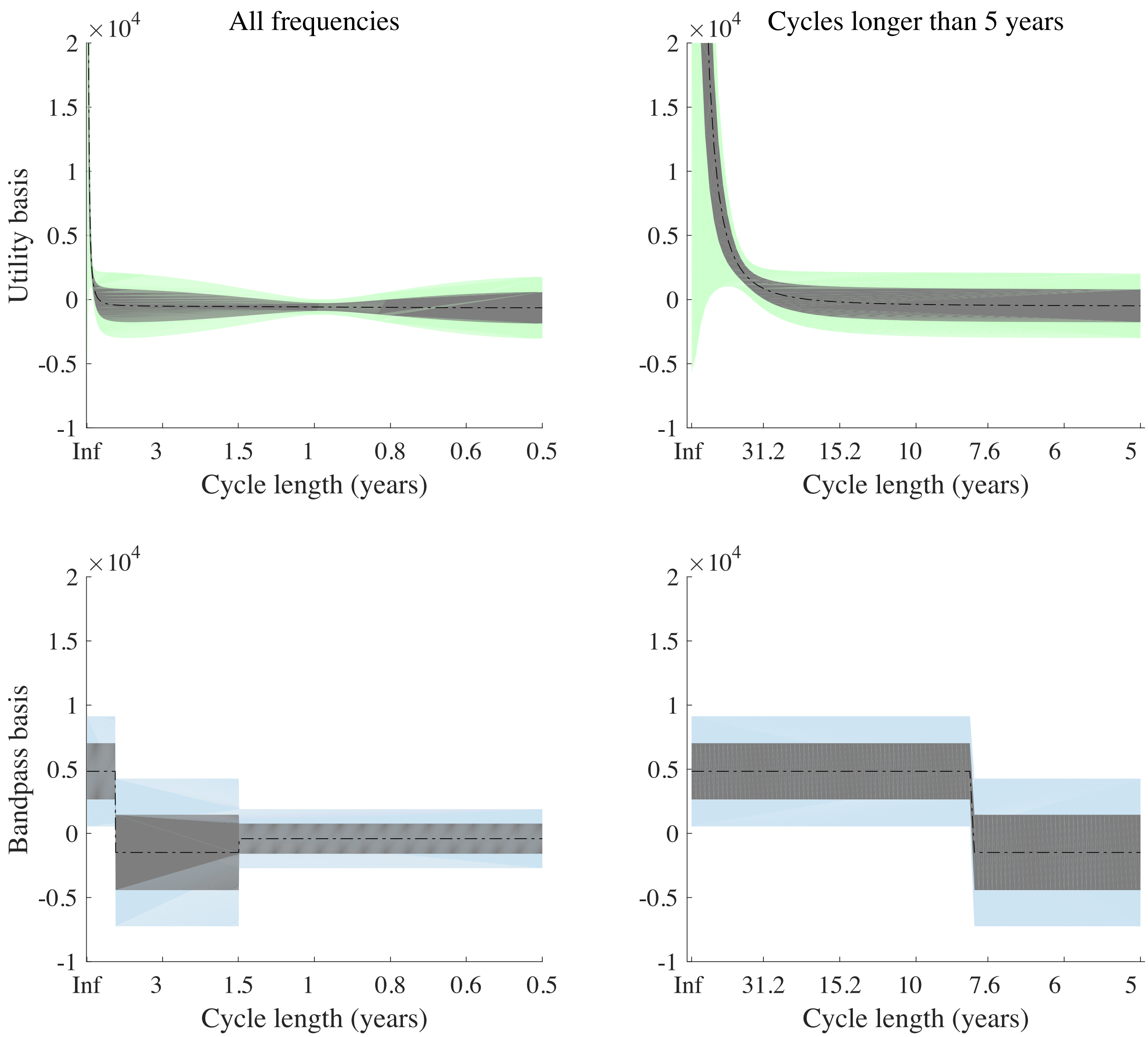


Figure 5. Estimated spectral weighting function. Estimated weighting function for consumption growth as the priced variable using the utility specification (top row) and the bandpass specification (bottom row). Risk prices are estimated using the 25 Fama–French portfolios. Light shaded areas denote 95-percent confidence regions. Dark shaded areas are 1 standard deviation confidence intervals. The utility specification uses a discount factor of 0.975 at the annual horizon. The x-axis gives the cycle length in years.

Epstein-Zin			Extremely low fr.	Low frequencies	Business Cycle	High frequencies	
α	ρ=1/EIS	θ	>210 years (%)	8 to 210 years (%)	1.5 to 8 years (%)	<1.5 years (%)	Median cycle
5	**0.5**	**0.975**	49.8	39.0	4.2	7.0	208.4 years
20	0.5	0.975	53.9	41.7	2.3	2.0	238.2 years
2.5	0.5	0.975	44.3	35.3	6.8	13.6	166.2 years
5	1	0.975	44.3	35.3	6.8	13.6	166.2 years
20	1	0.975	52.5	40.8	3.0	3.7	228.4 years
2.5	1	0.975	33.3	28.0	11.9	26.9	68.5 years
5	0.5	0.96	36.3	51.4	5.2	7.2	129.3 years
5	0.5	0.99	71.5	18.4	3.3	6.8	523.5 years
5	5	0.975	0.2	6.0	27.1	66.7	1.0 years
20	5	0.975	41.5	33.5	8.0	16.9	143.9 years
2.5	5	0.975	-54.8	-30.6	52.5	133.0	0.7 years

Internal Habit	Extremely low fr.	Low frequencies	Business Cycle	High frequencies	
b	>210 years (%)	8 to 210 years (%)	1.5 to 8 years (%)	<1.5 years (%)	Median cycle
0.25	0.1	2.3	13.9	83.6	0.5 years
0.5	-0.1	-1.9	-1.4	103.4	0.6 years
0.75	-0.2	-5.0	-12.3	117.6	0.6 years

Table 1. Calibration of the weight of Epstein-Zin and internal habit preferences in different frequency ranges. The table reports the fraction of the total weight (in percentage points) that different calibrations of Epstein–Zin preferences (top panel) and internal habits (bottom panel) assign to various frequency ranges. The table also reports the cycle length such that half of the pricing weight Z falls on either side of it (the median cycle).

Utility spec.		Shock 1		Shock 2		Shock 3	
Long-run		0.76	**	-0.61	**	0.64	**
	stderr	(0.27)		(0.25)		(0.23)	
	t-stat	2.83		-2.45		2.85	
Constant		1		0		0	
	stderr	-		-		-	
	t-stat	-		-		-	
Habit		0.19	***	-0.02		0.18	***
	stderr	(0.03)		(0.05)		(0.05)	
	t-stat	6.31		-0.45		3.53	

Bandpass spec.		Shock 1		Shock 2		Shock 3	
Long-run		0.16	***	-0.06	*	0.08	**
	stderr	(0.03)		(0.03)		(0.03)	
	t-stat	5.77		-1.68		2.85	
BC		0.29	***	0.04		0.02	
	stderr	(0.03)		(0.03)		(0.02)	
	t-stat	11.51		1.19		0.96	
High freq		0.55	***	0.02		-0.10	***
	stderr	(0.02)		(0.04)		(0.03)	
	t-stat	27.00		0.41		-3.42	

Table 2. Estimates of the rotation matrix W. The table reports the estimates of each element of the rotation matrix W, for both the bandpass and the utility specifications. * indicates significance at the 10-percent level, ** the 5-percent level, and *** the 1-percent level.

	Estimate	p-value		Estimate	p-value		Estimate	p-value		Estimate	p-value	
Epstein–Zin	556	0.07	*	183	0.26		210	0.18		729	0.10	*
Constant	-299	0.19		-226	0.05	*	-254	0.04	**	-795	0.09	*
Habit	62	0.96		827	0.17		787	0.22		-968	0.45	
Z_low	4837	0.03	**	2375	0.06	*	2526	0.03	**	4681	0.09	*
Z_BC	-1486	0.61		439	0.77		336	0.83		-3117	0.16	
Z_high	-413	0.72		-804	0.15		-820	0.18		-396	0.62	
p-value (difference test)		0.04			0.01			0.01			0.18	
Z_low	4386	0.01	***	2639	0.04	**	2796	0.01	**	5164	0.06	*
Z_BC and higher	-754	0.05	**	-358	0.27		-407	0.15		-1545	0.01	**
p-value (difference test)		0.01			0.06			0.02			0.04	
Portfolios used:												
25 BE and BE/ME	X			X			X			X		
49 Industries				X			X			X		
25 O.P. and Investment							X			X		
9 Risk-sorted										X		

Table 3. Estimates of the $Z(\omega)$ function. The table reports risk price estimates for the period 1962:1–2011:2 using quarterly data. The first set of rows presents the estimates of the coefficients for the utility specification. The second set of rows shows the estimates of the coefficients of the 3-window bandpass specification, i.e. the levels of the three steps (below-BC, BC, above-BC). The third set of rows shows the estimates for a 2-window bandpass specification (cycles below and above 8 years). For each bandpass specification, the table reports the p-value of a test for the difference in the coefficients. For the 3-window specification, the test is a chi2 test of the null hypothesis that the three coefficients are the same. For the 2-window specification, the test is a t-test for the difference between the two coefficients. Each column reports results using different set of portfolios, indicated at the bottom of the table: the 25 size and book-to-market sorted portfolios, 49 industry portfolios, 25 portfolios sorted by operating profitability and investment, and 9 portfolios double-sorted by their exposure to the long-run and medium-run shocks (corresponding to the first two risk prices reported in the table for each specification). * indicates significance at the 10-percent level, ** the 5-percent level, and *** the 1-percent level.

Appendix Tables and Figures

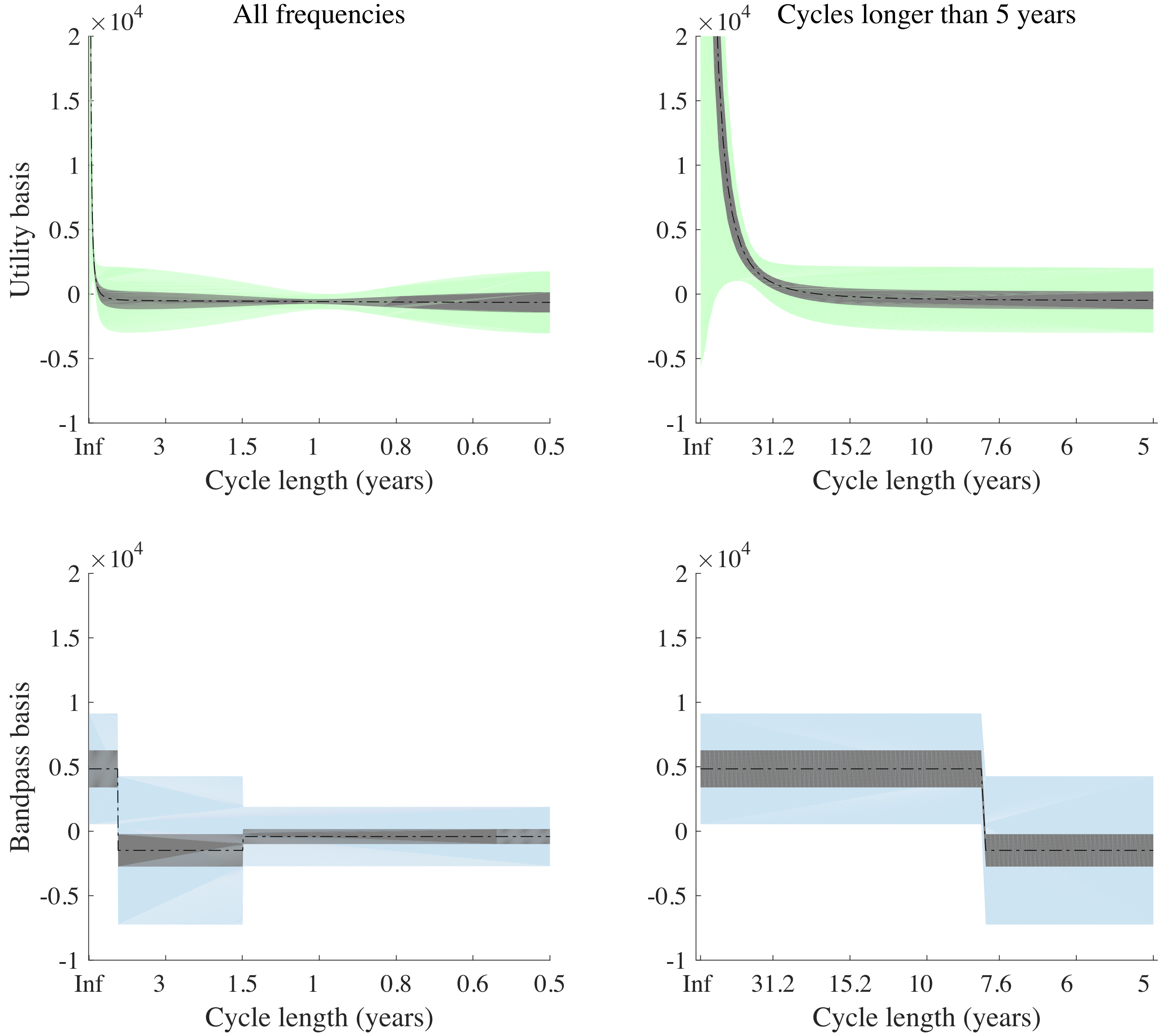


Figure A1. Estimated spectral weighting function, without VAR uncertainty. Estimated weighting function for consumption growth as the priced variable using the utility specification (top row) and the bandpass specification (bottom row). Risk prices are estimated using the 25 Fama–French portfolios. Light shaded areas denote 95-percent confidence regions. Dark shaded areas are 95-percent confidence intervals ignoring the estimation uncertainty of the VAR. The utility specification uses a discount factor of 0.975 at the annual horizon. The x-axis gives the cycle length in years.

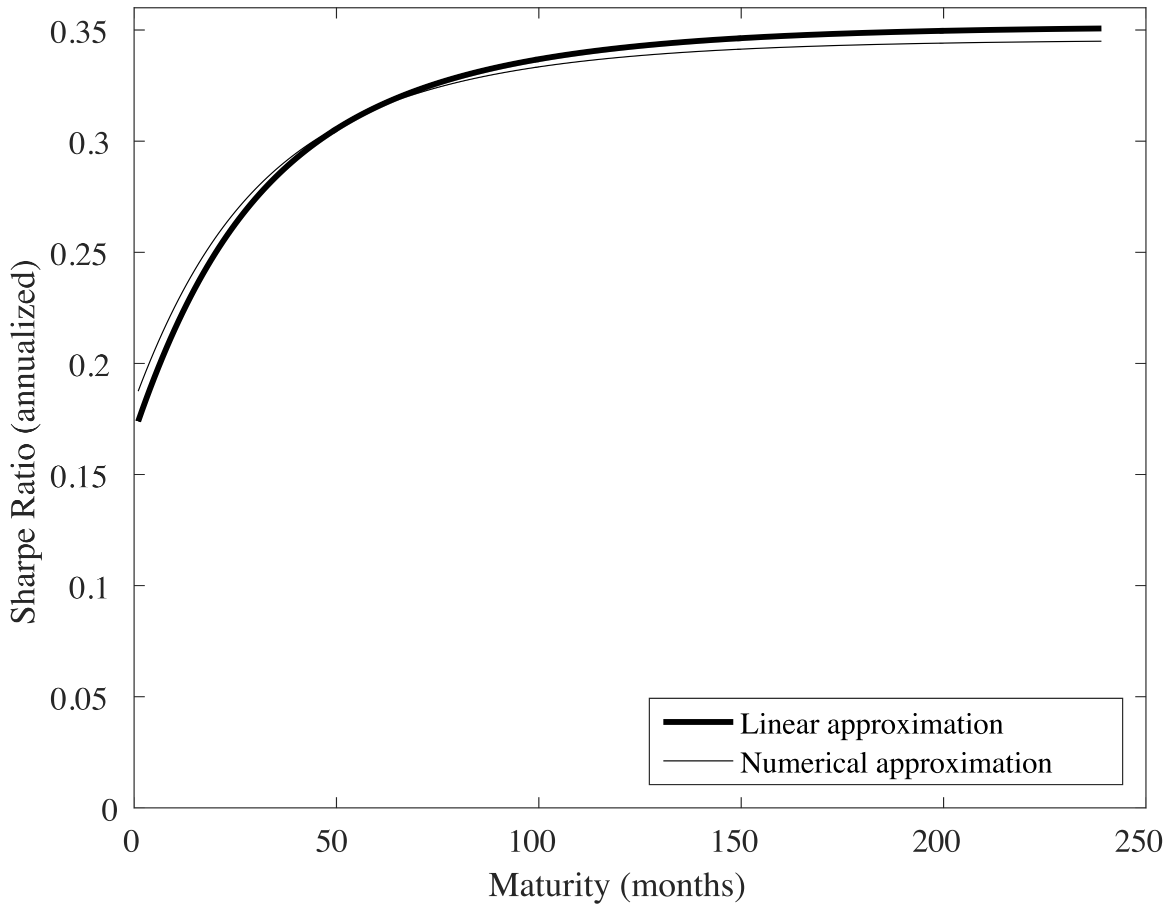


Figure A2. Sharpe ratios in exactly solved and log-linearized versions of the long-run risk model. The figure reports annualized Sharpe ratios for zero-coupon consumption claims of different maturity in the long-run risk model (Case II of Bansal and Yaron (2004)). The thin line uses 9th order projection methods to obtain the non-linear solution, while the thick line uses the log-linear approximation of the model as in Bansal and Yaron (2004).

	Lag 1			Lag 2			Lag 3		
	Cons.	Price	Cycle	Cons.	Price	Cycle	Cons.	Price	Cycle
Cons.	0.388 ***	-0.0487	0.365 ***	0.0652	0.112	-0.122	0.201 **	0.0581	-0.0389
	(4.97)	(-0.45)	(3.55)	(0.80)	(0.81)	(-1.21)	(2.55)	(0.50)	(-0.44)
Price	0.142 **	0.517 ***	0.21 ***	0.0141	0.111	-0.0223	0.205 ***	-0.101	-0.0254
	(2.29)	(6.08)	(2.58)	(0.22)	(1.02)	(-0.28)	(3.28)	(-1.10)	(-0.36)
Cycle	0.0740	-0.438 ***	0.265 ***	0.0814	-0.164	0.294 ***	0.0973	-0.0897	0.207 ***
	(1.26)	(-5.43)	(3.43)	(1.32)	(-1.58)	(3.88)	(1.64)	(-1.03)	(3.09)

Table A1. VAR estimates. VAR results for consumption growth and the two macroeconomic factors, with three lags. The sample is 1962:1–2011:2, quarterly. Standard errors are reported in brackets. * indicates significance at the 10-percent level, ** the 5-percent level, and *** the 1-percent level.

Fama-French 25 portfolios

Low-frequency loadings:

	Growth		2		3		4		Value		*Difference*	
Small	72.7	(18.7)	67.7	(15.2)	54.9	(13.5)	52.0	(12.9)	56.6	(14.6)	-16.1	(9.8)
2	66.7	(16.4)	52.7	(13.6)	50.3	(12.1)	50.8	(11.8)	56.6	(13.2)	-10.1	(10.0)
3	63.3	(14.8)	50.1	(12.2)	42.5	(11.0)	42.3	(11.0)	47.1	(11.9)	-16.2	(9.9)
4	55.0	(13.1)	47.0	(11.4)	42.4	(10.9)	40.1	(10.6)	47.8	(12.2)	-7.2	(9.4)
Large	37.6	(10.5)	28.5	(9.6)	25.1	(8.9)	27.0	(9.3)	32.2	(10.4)	-5.4	(8.4)
Difference	-35.1	(12.9)	-39.2	(10.4)	-29.8	(9.4)	-25.0	(8.7)	-24.4	(10.0)		

Business-cycle frequency loadings:

	Growth		2		3		4		Value		*Difference*	
Small	39.8	(9.8)	35.9	(8.0)	28.4	(7.2)	26.7	(6.9)	30.5	(7.7)	-9.3	(5.2)
2	34.3	(8.7)	26.2	(7.2)	24.9	(6.5)	24.5	(6.3)	28.6	(7.0)	-5.7	(5.3)
3	31.2	(7.9)	24.5	(6.5)	20.4	(5.9)	20.7	(5.9)	21.5	(6.4)	-9.7	(5.2)
4	26.9	(7.0)	22.7	(6.1)	21.3	(5.8)	18.6	(5.7)	23.9	(6.5)	-3.0	(5.0)
Large	18.9	(5.6)	13.4	(5.1)	12.9	(4.7)	13.9	(4.9)	16.8	(5.5)	-2.1	(4.4)
Difference	-20.9	(6.8)	-22.5	(5.5)	-15.5	(5.0)	-12.8	(4.6)	-13.8	(5.3)		

Table A2. Factor loadings for test portfolios. Each cell of each table is a factor loading for one of the portfolio returns with respect to either the low- or business-cycle frequency shock, for the 25 Fama–French portfolios. The numbers in parentheses are standard errors for the estimated factor loadings and their differences.

		(1)	p-value	(2)	p-value
Utility Spec.					
	Epstein–Zin	-2209	0.55	-7044	0.84
Consumption	Constant	-156	0.60		
	Habit	6338	0.40		
	Epstein–Zin	-2225	0.65	-18483	0.84
Volatility	Constant	-712	0.56	714	0.85
	Habit	7564	0.60	20483	0.85
Bandpass Spec.					
	Z_low	-40702	0.80	-4830	0.21
Consumption	Z_BC	19017	0.76		
	Z_high	-4247	0.67		
	Z_low	-66535	0.83	-19223	0.14
Volatility	Z_BC	33287	0.82	6793	0.26
	Z_high	-9286	0.82	-1095	0.61

Table A3. Model with stochastic volatility. The table estimates four models with stochastic volatility. In the first column, we estimate the model using the utility specification for the weighting function of consumption and volatility (top of the table), or using the bandpass specification for the weighting function of consumption and volatility (bottom of the table). In each of the two models estimated in the first column, the 6 parameters of the model (3 for the consumption weighting function and 3 for the volatility weighting function) are estimated using a factor-aumented VAR that includes observable real consumption growth, realized volatility of the S&P 500, and four principal components (macroeconomic factors) from Ludvigson and Ng (2007) and Jurado, Ludvigson and Ng (2015). The second column repeats the estimation but only includes the long-run component of the consumption weighting function, while leaving 3 parameters for the stochastic volatility weighting function. In this case, a 4-variable VAR is used, that uses real consumption growth, realized volatility of the S&P 500, and the first 2 principal components of the macroeconomic variables. * indicates significance at the 10-percent level, ** the 5-percent level, and *** the 1-percent level.

		Unrestricted	p-value		Restricted	p-value	
Utility Spec.							
Level	Epstein–Zin	671	0.08	*	333	0.15	
	Constant	-161	0.55		-481	0.01	***
	Habit	-261	0.87		455	0.59	
Interaction	Epstein–Zin	-857	0.03	**			
	Constant	-468	0.15		0.997	0.00	***
	Habit	2612	0.11				
Bandpass Spec.							
Level	Z_low	5656	0.04	**	5403	0.01	**
	Z_BC	-2034	0.56		-2545	0.38	
	Z_high	-73	0.96		561	0.65	
Interaction	Z_low	-5342	0.07	*			
	Z_BC	4911	0.15		-1.110	0.00	***
	Z_high	-2110	0.16				

Table A4. Model with time-varying risk premia. The table estimates four models with time-varying risk premia, conditional on the surplus consumption ratio. In the first column, we estimate the model in an unrestricted way, using lagged surplus consumption ratio as an instrument in the GMM estimation (standardized to have zero mean and unit varinace). For each of the utility specification (top) and bandpass specification (bottom), we report the coefficients on the three rotated shocks and those on the interaction between the lagged instrument and the rotated shocks. Negative estimates of the interacted coefficients indicate higher risk premia when the surplus consumption ratio is low, in the spirit of the Campbell-Cochrane (1999) habit model. * indicates significance at the 10-percent level, ** the 5-percent level, and *** the 1-percent level.

	Annual Data	p-value	Two lags	p-value	Altern. weighting	p-value	Bootstrap Results	p-value
Epstein–Zin	873	0.00 **	816	0.39	578	0.00 ***	556	0.01 ***
Constant	167	0.01	-331	0.09 *	-348	0.00 ***	-299	0.22
Habit	-1129	0.05	-529	0.85	154	0.81	62	1.00
Z_low	839	0.02 **	6998	0.38	5292	0.00 ***	4837	0.00 ***
Z_BC	-363	0.23	-2821	0.70	-1498	0.30	-1486	0.37
Z_high			64	0.98	-551	0.39	-413	0.99

	Altern. VAR No PC uncert.	p-value	Altern. VAR PC uncert.	p-value	4 Windows		p-value
Epstein–Zin	1111	0.00 ***	1111	0.00 ***			
Constant	-443	0.12	-443	0.12	Z>100yr	627805	0.45
Habit	1232	0.33	1232	0.33	Z_low	-66848	0.44
					Z_BC	14673	0.36
Z_low	8638	0.00 ***	8638	0.00 ***	Z_high	-3834	0.35
Z_BC	-3160	0.00 ***	-3160	0.00 ***			
Z_high	438	0.55	438	0.57			

Table A5. Robustness. The table reports alternative specifications and robustness results for the estimates of risk prices on different utility components (in the utility specification) or frequency groups (bandpass specification). The first set of results estimates the results as in Column 1 of Table 3, but using annual data. Since the minimum cycle discernible from annual data is 2 years, we cannot estimate the price of high-frequency fluctuations in the bandpass basis. The second set shows the results using two rather than three lags for the VAR. The third set computes standard errors using an alternative weighting matrix for the second stage of the sequential GMM procedure; in this case, the weighting matrix for the estimation of the risk prices from the cross section of portfolio returns depends not only on the moments of the asset pricing equations, but on the entire set of moment conditions, including the VAR moment conditions (see Hansen (2008)). The fourth set of results reports bootstrapped p-values, as described in the Appendix. The fifth set uses an alternative dataset to compute the VAR factors: principal components of 9 variables (aggregate price/earnings and price/dividend ratios; the 10 year/3 month term spread; the Aaa–Baa corporate yield spread (default spread); the small-stock value spread; the unemployment rate minus its 8-year moving average; detrended short-term interest rate; the three-month Treasury yield rate; and Lettau and Ludvigson's (2001) *cay*). p-values are computed via bootstrap, ignoring the sampling uncertainty in the construction of the principal components. The sixth set uses the same variables as in the fifth set, but accounts for sampling uncertainty in the principal component estimation. The seventh set estimates a bandpass specification replacing the low-frequency window with two separate ones, one covering cycles 10 to 100 years, one covering all cycles above 100 years. * indicates significance at the 10-percent level, ** the 5-percent level, and *** the 1-percent level.