Online appendix for "Asset pricing in the frequency domain: theory and empirics"

Ian Dew-Becker and Stefano Giglio

March 1, 2016

A Derivation of result 1

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

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

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

$$\begin{aligned} g_{j,k} &= g_{j,k} + g_{j,-k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{G}_j(\omega) \begin{pmatrix} \cos(\omega k) + i \sin(\omega k) \\ + \cos(-\omega k) + i \sin(-\omega k) \end{pmatrix} d\omega & (2) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{G}_j(\omega)\, 2\cos(\omega k)\, d\omega & (3) \end{aligned}$$

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

$$\sum_{k=0}^{\infty} z_k g_{j,k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} G_j(\omega) \left(z_0 + 2 \sum_{k=1}^{\infty} z_k \cos(\omega k)\right) d\omega \tag{4}$$

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

$$\Delta c_t = x_{t-1} + \sigma_{t-1}\varepsilon_{c,t} \tag{5}$$

$$x_t = \phi_x x_{t-1} + \varphi_e \sigma_{t-1}\varepsilon_{x,t} \tag{6}$$

$$\sigma_t^2 = v_1\left(\sigma_{t-1}^2 - \bar{\sigma}^2\right) + \bar{\sigma}^2 + \sigma_w \varepsilon_{\sigma,t} \tag{7}$$

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

$$PC_{n,t} = E_t^Q \left[\frac{C_{t+n}}{C_t} \right] \tag{8}$$

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

$$R_{n,t+1} = \frac{PC_{n-1,t+1} C_{t+1}}{PC_{n,t} C_t} \tag{9}$$

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

$$\Delta \mathbb{E}_{t+1} \log M_{t+1} = -\alpha \sigma_t \varepsilon_{c,t+1} + \frac{(\rho - \alpha)\,\theta}{1 - \phi_x \theta} \sigma_t \varphi_e \varepsilon_{x,t+1} - \frac{\rho - \alpha}{1 - \rho} k_1 \frac{\theta}{1 - v_1 \theta} \sigma_w \varepsilon_{\sigma^2,t+1} \tag{10}$$

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

$$E_t r_{w,t+1} = \bar{r} + \rho x_t + k_1 \sigma_t^2 \tag{11}$$

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

$$E_t \log M_{t+1} = \bar{m} + m_1 \sigma_t^2 \tag{12}$$

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,

$$pc_{n,t} = \log E_t \exp \left(\begin{array}{c} \bar{m} + m_1 \sigma_t^2 - \rho x_t - \alpha \sigma_t \varepsilon_{c,t+1} + \frac{(\rho-\alpha)\theta}{1-\phi_x \theta} \sigma_t \varphi_e \varepsilon_{x,t+1} - \frac{\rho-\alpha}{1-\rho} k_1 \frac{\theta}{1-v_1\theta} \sigma_w \varepsilon_{\sigma^2,t+1} \\ + pc_{n-1,t+1} + x_t + \sigma_t \varepsilon_{c,t+1} \end{array} \right) \tag{13}$$

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

$$pc_{n,t} = \bar{p} + p_{x,n} x_t + p_{\sigma^2,n} \sigma_t^2 \tag{14}$$

This equation can be solved using standard methods to obtain

$$\bar{p}_n = \log \beta - \bar{r}_f + \bar{p}_{n-1} + p_{\sigma^2,n-1} (1 - v_1) \bar{\sigma}^2 + \frac{1}{2} \left(-\frac{\rho - \alpha}{1-\rho} k_1 \frac{\theta}{1 - v_1 \theta} + p_{\sigma^2,n-1} \right)^2 \sigma_w^2 \tag{15}$$

$$p_{x,n} = -\rho + p_{x,n-1} \phi_x + 1 \tag{16}$$

$$p_{\sigma^2,n} = -r_1 + p_{\sigma^2,n-1} v_1 + \frac{1}{2} (1-\alpha)^2 + \frac{1}{2} \left(\frac{(\rho-\alpha)\theta}{1-\phi_x \theta} + p_{x,n-1} \right)^2 \varphi_e^2 \tag{17}$$

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 $\frac{std\left(M^{\text{proj}}-M^{\text{linear}}\right)}{std(M^{\text{proj}})}$.

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 ε_t on the mth element of $\mathbb{E}_t \bar{\mathbf{x}}_{t+k}$. The innovation to the SDF is then

$$\Delta \mathbb{E}_{t+1} m_{t+1} = -\left(\sum_{k=0}^{\infty} \mathbf{z}_k \mathbf{g}_k\right) \varepsilon_{t+1} \tag{18}$$

The price of risk for the jth element of ε 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

$$\tilde{\mathbf{G}}(\omega) \equiv \sum_{k=0}^{\infty} e^{-i\omega k} \mathbf{g}_k \tag{19}$$

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

$$\sum_{k=0}^{\infty} \mathbf{z}_k \mathbf{g}_k \mathbf{b}_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{Z}(\omega) \mathbf{G}(\omega) \mathbf{b}_j d\omega \tag{20}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_m \mathbf{Z}_m(\omega) \mathbf{G}_{m,j}(\omega) d\omega \tag{21}$$

where

$$\mathbf{Z}(\omega) \equiv \mathbf{z}_0 + 2\sum_{k=1}^{\infty} \mathbf{z}_k \cos(\omega k) \tag{22}$$

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,

$$m_{t+1} = \frac{1-\alpha}{1-\rho}\log\beta - \rho\frac{1-\alpha}{1-\rho}\Delta c_{t+1} + \frac{\rho-\alpha}{1-\rho}r_{w,t+1} \tag{23}$$

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

$$\mathbb{E}_t r_{w,t+1} = k_0 + \rho\mathbb{E}_t\Delta c_{t+1} + k_1\sigma_t^2 \tag{24}$$

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

$$\Delta\mathbb{E}_{t+1}m_{t+1} = -\alpha\Delta c_{t+1} - (\alpha-\rho)\,\Delta\mathbb{E}_{t+1}\sum_{j=1}^{\infty}\theta^j\Delta c_{t+1+j} \tag{25}$$

$$-\frac{\rho-\alpha}{1-\rho}\Delta\mathbb{E}_{t+1}\theta k_1\sigma_{t+1}^2 - \frac{\rho-\alpha}{1-\rho}\Delta\mathbb{E}_{t+1}\sum_{j=1}^{\infty}\theta^j\theta k_1\sigma_{t+j+1}^2 \tag{26}$$

The weighting functions for consumption growth and volatility are now

$$Z_C^{EZ-SV}(\omega) = \alpha + (\alpha - \rho)\sum_{j=1}^{\infty}\theta^j 2\cos(\omega j) \tag{27}$$

$$Z_{\sigma^2}^{EZ-SV}(\omega) = \theta k_1 \frac{\rho - \alpha}{1-\rho}\left(1 + \sum_{j=1}^{\infty}\theta^j 2\cos(\omega j)\right) \tag{28}$$

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

$$E_t r_{w,t+1} = k_0 + \rho E_t \Delta c_{t+1} + k_1 \tilde{x}_t \tag{29}$$

Recall that according to the Campbell–Shiller approximation,

$$\begin{aligned}
\Delta \mathbb{E}_{t+1} r_{w,t+1} &= \sum_{j=0}^{\infty}\theta^j \Delta \mathbb{E}_{t+1}\Delta c_{t+j+1} - \sum_{j=1}^{\infty}\theta^j \Delta \mathbb{E}_{t+1} r_{w,t+j+1} && (30)\\
&= \sum_{j=0}^{\infty}\theta^j \Delta \mathbb{E}_{t+1}\Delta c_{t+j+1} - \sum_{j=1}^{\infty}\theta^j \Delta \mathbb{E}_{t+1}\left(\rho \Delta c_{t+j+1} + k_1 \tilde{x}_{t+j}\right) && (31)\\
&= \sum_{j=0}^{\infty}\theta^j \sum_k g_{k,j}\varepsilon_{k,t+1} - \sum_{j=1}^{\infty}\theta^j \Delta \mathbb{E}_{t+1}\left(\sum_k \left(\rho g_{k,j} + k_1 \tilde{g}_{k,j-1}\right)\varepsilon_{k,t+1}\right) && (32)
\end{aligned}$$

The pricing equation for the wealth portfolio is then

$$\begin{aligned}
0 &= \log E_t \exp\left(\frac{1-\alpha}{1-\rho}\log\beta - \rho\frac{1-\alpha}{1-\rho}\Delta c_{t+1} + \frac{1-\alpha}{1-\rho} r_{w,t+1}\right) && (33)\\
&= \frac{1-\alpha}{1-\rho}\log\beta - \rho\frac{1-\alpha}{1-\rho}E_t\Delta c_{t+1} + \frac{1-\alpha}{1-\rho}E_t r_{w,t+1} && (34)\\
&\quad + \frac{1}{2}\log E_t \exp\left(-\rho\frac{1-\alpha}{1-\rho}\Delta \mathbb{E}_{t+1}\Delta c_{t+1} + \frac{1-\alpha}{1-\rho}\Delta \mathbb{E}_{t+1} r_{w,t+1}\right) && (35)
\end{aligned}$$

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

$$\begin{aligned}
\Delta\mathbb{E}_{t+1}m_{t+1} &= -\rho\frac{1-\alpha}{1-\rho}\Delta\mathbb{E}_{t+1}\Delta c_{t+1}+\frac{\rho-\alpha}{1-\rho}\Delta\mathbb{E}_{t+1}r_{w,t+1} &(36)\\
&= -\alpha\sum_{k}g_{k,0}\varepsilon_{k,t+1}+\frac{\rho-\alpha}{1-\rho}\sum_{j=1}^{\infty}\theta^{j}\left(\sum_{k}\left(\left(1-\rho\right)g_{k,j}-k_{1}\tilde{g}_{k,j-1}\right)\varepsilon_{k,t+1}\right) &(37)
\end{aligned}$$

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

$$z_{0,\Delta c} = \alpha \quad (38)$$

$$z_{j,\Delta c} = \theta^{j}\left(\alpha-\rho\right) \text{ for } j>0 \quad (39)$$

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

$$z_{j,\tilde{x}} = k_{1}\frac{\rho-\alpha}{1-\rho}\theta^{j} \quad (40)$$

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

$$G_j^{Step}(\omega) = \begin{cases} \frac{1}{2\pi/32}\int_0^{2\pi/32} G_j(\kappa)\, d\kappa \text{ for } \omega \in [0, 2\pi/32) \\ \frac{1}{2\pi/8-2\pi/32}\int_{2\pi/32}^{2\pi/8} G_j(\kappa)\, d\kappa \text{ for } \omega \in [2\pi/32, 2\pi/8) \\ \frac{1}{\pi-2\pi/8}\int_{2\pi/8}^{\pi} G_j(\kappa)\, d\kappa \text{ for } \omega \in [2\pi/8, \pi] \end{cases} \tag{41}$$

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:

$$\int_0^{\pi} \tilde{Z}(\omega)\, G_j^{Step}(\omega)\, d\omega = \int_0^{\pi} Z^{BP}(\omega; \mathbf{q})\, G_j(\omega)\, d\omega \tag{42}$$

$$\text{where } Z^{BP}(\omega; \mathbf{q}) = \begin{cases} \frac{1}{2\pi/32}\int_0^{2\pi/32} \tilde{Z}(\kappa)\, d\kappa \text{ for } \omega \in [0, 2\pi/32) \\ \frac{1}{2\pi/8-2\pi/32}\int_{2\pi/32}^{2\pi/8} \tilde{Z}(\kappa)\, d\kappa \text{ for } \omega \in [2\pi/32, 2\pi/8) \\ \frac{1}{\pi-2\pi/8}\int_{2\pi/8}^{\pi} \tilde{Z}(\kappa)\, d\kappa \text{ for } \omega \in [2\pi/8, \pi] \end{cases} \tag{43}$$

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.[1]

F Details of the empirical analysis

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

The risk prices for the shocks can be written in the time domain as

$$\mathbf{p} = \sum_{k=0}^{\infty} \mathbf{q}\mathbf{z}_k \mathbf{b}_1 \mathbf{\Phi}^k \tag{44}$$

$$\text{where } \mathbf{z}_k \equiv \frac{1}{\pi}\int_0^{\pi} \cos(\omega k)\left[Z_1(\omega), Z_2(\omega), Z_3(\omega)\right]' d\omega \tag{45}$$

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

Now suppose we considered a set of rotated shocks $\tilde{\varepsilon}_t = \Theta\varepsilon_t$ for some rotation matrix Θ.

[1]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, 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.

The estimated reduced-form risk prices for $\tilde{\varepsilon}_t$, $\tilde{\mathbf{p}}$, will then have the property

$$\begin{aligned} \tilde{\mathbf{p}}\tilde{\varepsilon}_t &= \mathbf{p}\varepsilon_t && (46) \\ &\Rightarrow \tilde{\mathbf{p}}\Theta = \mathbf{p} && (47) \end{aligned}$$

since the pricing kernel must be unchanged whether we examine the reduced-form innovations or a rotation of them.[2]

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

$$\begin{aligned} \tilde{\mathbf{p}}\tilde{\mathbf{W}}^{-1} &= \tilde{\mathbf{p}}\left(\mathbf{W}\mathbf{\Theta}^{-1}\right)^{-1} && (48) \\ &= \tilde{\mathbf{p}}\Theta\mathbf{W}^{-1} && (49) \\ &= \mathbf{p}\mathbf{W}^{-1} && (50) \end{aligned}$$

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:

$$\mathbb{E}_t r_{it+1} - r_{t+1}^f + \frac{1}{2}\sigma_{it}^2 = -cov_t(m_{t+1}, r_{it+1}) \tag{51}$$

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

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

$$cov_t(m_{t+1}, r_{it+1}) = cov_t(\Delta\mathbb{E}_{t+1} m_{t+1}, r_{it+1}) = \mathbb{E}_t(\Delta\mathbb{E}_{t+1} m_{t+1} r_{it+1}) = \mathbb{E}_t(-\mathbf{qW}\varepsilon_{t+1} r_{it+1}) \tag{52}$$

Which implies

$$\mathbb{E}_t r_{it+1} - r_{t+1}^f + \frac{1}{2}\sigma_{it}^2 = E_t(\mathbf{qW}\varepsilon_{t+1} r_{it+1}) \tag{53}$$

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

$$E[\exp(r_{it+1}) - \exp\left(r_{t+1}^f\right)] \approx E\left[\mathbf{qW}\varepsilon_{t+1} r_{it+1}\right] \tag{54}$$

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 $p_1 = vec\left(\hat{\mathbf{\Phi}}\right)$ 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

$$\hat{e}_{i,t} \equiv x_{i,t} - \hat{b}_{1,t} f_{1,t} - \hat{b}_{2,t} f_{2,t} \tag{55}$$

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

$$\hat{e}_{i,t} = \rho_i \hat{e}_{i,t-1} + v_{i,t} \tag{56}$$

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

$$\Delta\mathbb{E}_{t+1} m_{t+1} = -\mathbf{q}\mathbf{W}\varepsilon_{t+1} \tag{57}$$

The rotated shocks are thus,

$$\mathbf{u}_{t+1} = \mathbf{W}\varepsilon_{t+1} \tag{58}$$

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.[3]

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.

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.

[3] 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.

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.

Appendix Tables and Figures

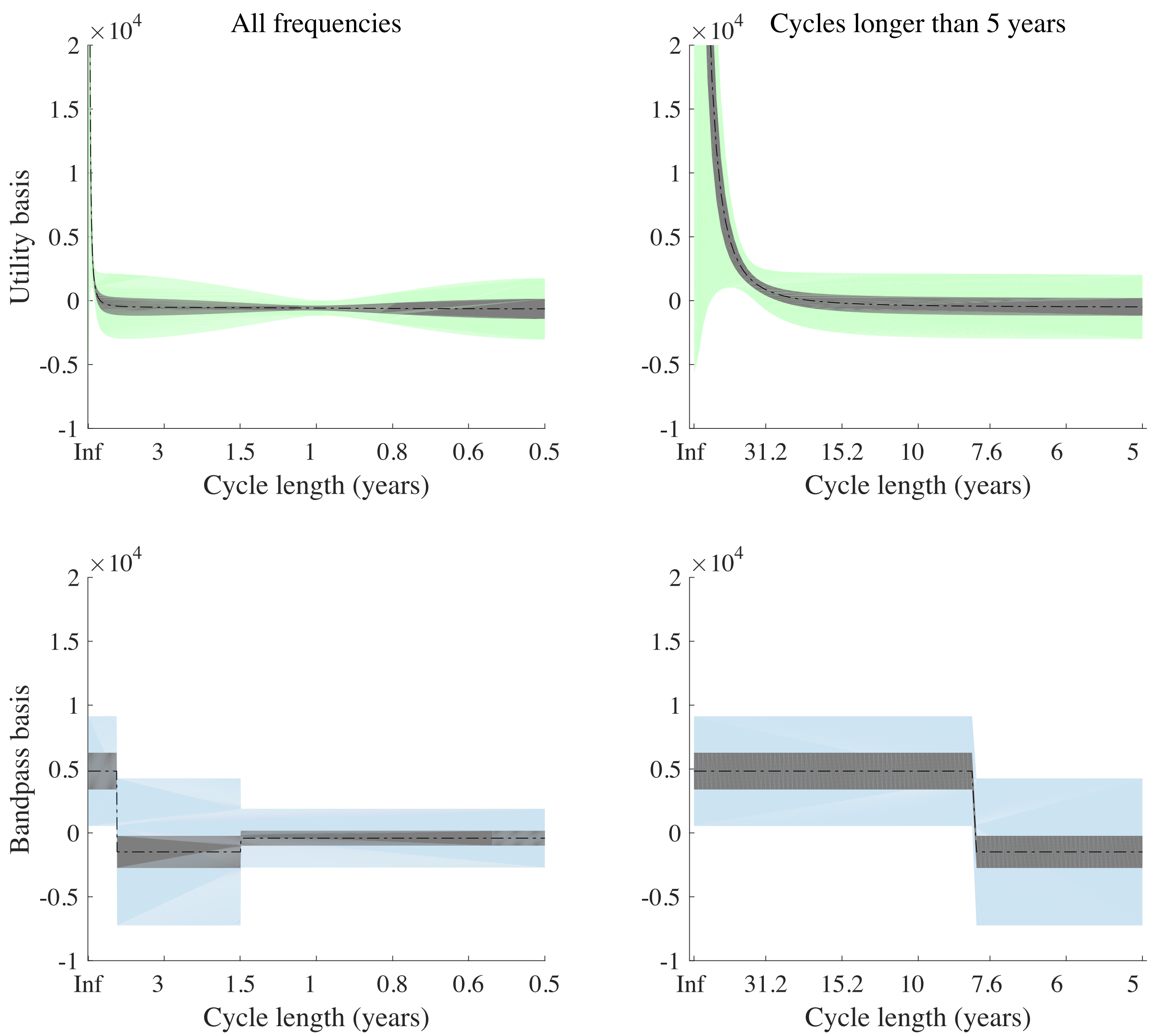

Figure A1: **Estimated spectral weighting function, without VAR uncertainty**. Estimated weighting function for consumption growth as the priced variable using the utility specification (top row) and the bandpass specification (bottom row). Risk prices are estimated using the 25 Fama–French portfolios. 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.

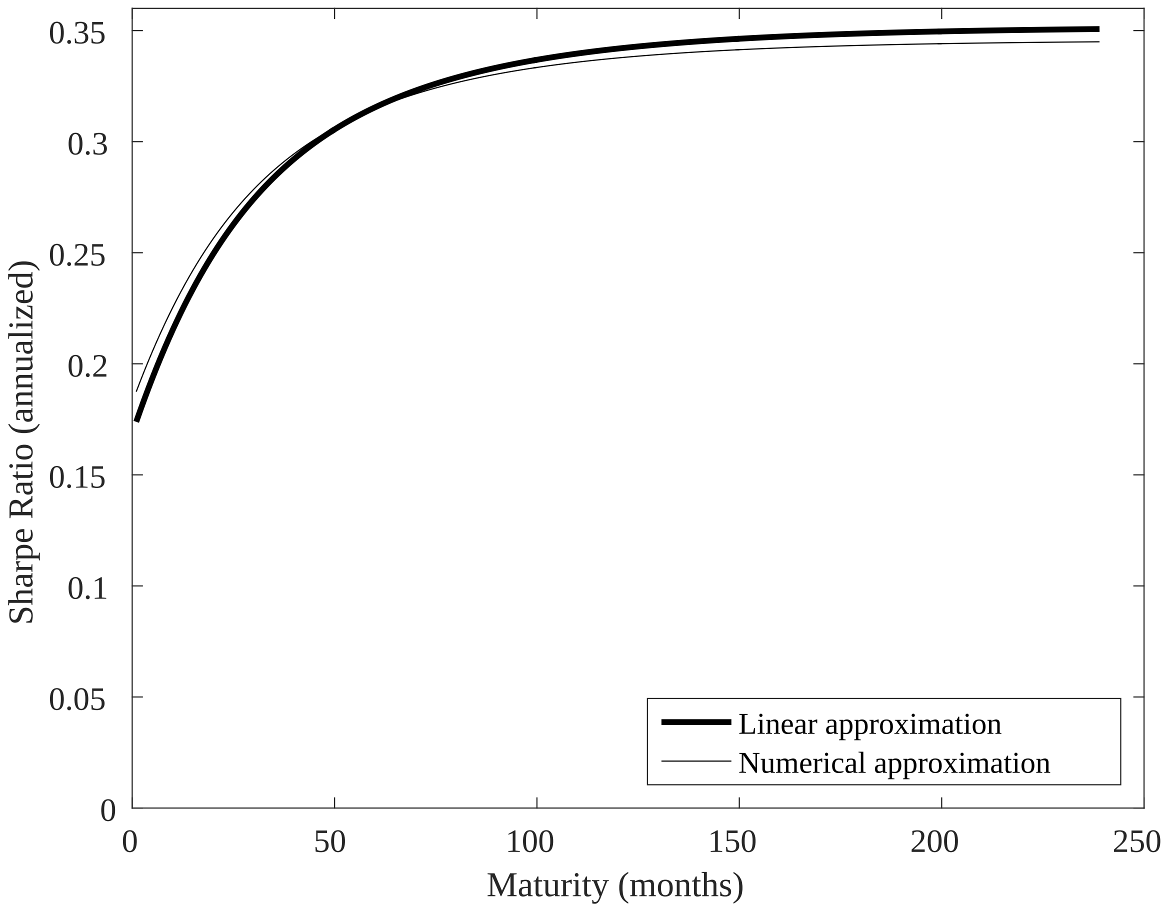

Figure A2: **Sharpe ratios in exactly solved and log-linearized versions of the long-run risk model**. The figure reports annualized Sharpe ratios for zero-coupon consumption claims of different maturity in the long-run risk model (Case II of Bansal and Yaron (2004)). The thin line uses 9th order projection 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.							
	Epstein–Zin	671	0.08	*	333	0.15	
Level	Constant	-161	0.55		-481	0.01	***
	Habit	-261	0.87		455	0.59	
	Epstein–Zin	-857	0.03	**			
Interaction	Constant	-468	0.15		0.997	0.00	***
	Habit	2612	0.11				
Bandpass Spec.							
	Z_low	5656	0.04	**	5403	0.01	**
Level	Z_BC	-2034	0.56		-2545	0.38	
	Z_high	-73	0.96		561	0.65	
	Z_low	-5342	0.07	*			
Interaction	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.