An Anatomy of the Equity Premium

Paul Schneider†
February 3, 2016

Abstract

This paper introduces a decomposition of the market return in terms of higher-order realized, and option-implied risk aversion, connecting it to level, slope, and curvature of the implied volatility surface. Empirically, second-order risk aversion — loss aversion — explains most of the market return. Signals revealed by this risk anatomy provide predictive power out-of-sample for realized returns in particular for longer maturities. The decomposition also shows that compensation for disaster risk is not prominently featured in the market return. Furthermore it highlights that models with identically and independently distributed state variables are not suited to represent in particular longer-maturity returns.

1 Introduction

The equity premium as the expected profit from investing in the market is a central quantity to finance and economics for testing theories, but also to market participants who want to benefit from the expected growth of the economy with a simple trading strategy. In this paper I investigate a time-varying and model-free decomposition of the realized forward equity return

---

*I am grateful to Christian Julliard for raising the question whether the equity premium can be dissected into higher-order premia and to Anthony Neuberger and Christian Wagner for helpful discussions.

†University of Lugano and Swiss Finance Institute. Email: paul.schneider@usi.ch
in terms of economically appealing constituents that allow interpretation as realized variance, loss aversion, and temperance to assess the question which type of risk the market return compensates for. Conditional on S&P 500 option data I find that the main component of the equity return originates from aversion to losses, rather than variance, for short maturities. For longer maturities aversion to variance can be dominating, and for any maturity considered tail risk aversion is a third-order effect. I provide evidence also that the decomposition may help predicting the return and that models with i.i.d. state variables are inadequate modelling devices.

The early literature on the equity premium starts from simple decision-theoretic consumption-based models and tries to reconcile them with observed data. Mehra and Prescott (1985) note that it is very unlikely that realized equity returns are really generated by the expected-utility consumption-based model. Later, Julliard and Ghosh (2012) show that not even disaster risk can change this statement. Weil (1989) extends the utility specification and notes that neither non-expected recursive preferences can solve this equity premium puzzle and proposes in addition a new puzzle that also risk free interest rates do not behave conforming with the notion that agents would prefer early resolution of risk. Different preference specifications yield more success. Constantinides (1990) offers an explanation of equity return data via an extended habit utility model, and Benartzi and Thaler (1995) rationalize it with prospect theory. Many more extensions have been considered by the literature (Detemple and Murthy, 1994; Kogan et al., 2007; Lettau et al., 2008; Wachter, 2013; Bhamra and Uppal, 2014).

To investigate the trading strategy underlying the equity premium without a likely misspecified economic model, I take as given the realized equity return and start from viewing it purely as a signal, very much like the sound of a piano or a violin. To take the analogy to musical instruments further, certain frequencies are more relevant than others for different instruments. With this in mind I propose a set of basis functions which are more relevant a-priori economically to explain the equity return than others. These functions correspond to trading strategies with exposure to realized variance, realized loss aversion, and realized tail risk aversion, and possibly even
higher-order risk aversion. The transform technique I use for this purpose is model-free and therefore robust to misspecification errors. In finance, such transform techniques are usually used in connection with models to express unknown quantities through known quantities. Duffie et al. (2000), for example, use the known characteristic function of affine models to compute unknown densities and probabilities pertaining to them. Chen and Joslin (2012) extend this approach and develop tractable transform methods for certain well-behaved nonlinear functions of processes with known characteristic function. Martin (2013) makes use of the Fourier transform to study the behaviour of asset prices with Levy structure in the presence of disaster risk. Giglio and Dew-Decker (2015) investigate economic preferences through the lens of their frequency spectrum. In contrast, in this paper I apply transform techniques to known functions and the known (through option prices) forward-neutral distribution, to obtain a decomposition into known, but economically relevant quantities – a risk anatomy.

Using S&P 500 options data from 1990-2014 I perform this anatomy on market forward returns with maturities of 1, 3, 6, and 12 months. I find that the first three factors, realized variance, loss aversion and tail aversion, are sufficient to explain the equity return fully. This suggests that my particular choice of economic basis functions is indeed a reasonable one, given that the theory predicts that an infinite series may be needed. All factors are time-varying and not correlated too strongly. By far the greatest impact arises from the factor associated with loss aversion, in particular for shorter maturities, in line with the model-based predictions from Benartzi and Thaler (1995). I also perform predictive regressions and investigate the explained variation out-of-sample. Differently from Fama and French (2002), Campbell and Thompson (2008), Welch and Goyal (2008), Rapach et al. (2010), Dangl and Halling (2012) I do not use historic information for this exercise, but only prices of option portfolios at the time of forming the expectation, similar to Martin (2015). The decomposition promises some advantages over both Martin (2015) and Schneider and Trojani (2015b) for long maturities, while it performs worse than the aforementioned for the shorter 1 month and 3 month horizons. For all maturities but 3 months it is able to generate significant
predictive advantage over the expanding sample mean.

A decomposition of the equity return through the lens of the disaster model from Backus et al. (2011) shows that an i.i.d. specification of the factors driving the economy is generally inadequate. It also reveals that care must be taken to model rare crisis events in the evolution of consumption, despite the first two moments of the market return may be well matched, as the shape of the entire distribution is reflected in each factor of the risk anatomy.

The paper is organized as follows. Section 2.1 introduces the necessary notation and concepts, Section 2.2 develops the economic transform by projecting the equity return on certain economic contracts. Sections 3.1 and 3.2 work out some basic theoretical properties of the transform and contain the empirical study. Section 4 reviews the disaster model from Backus et al. (2011). Section 5 concludes, Appendix A develops economic Hellinger contracts, Appendix B reviews how to compute conditional moments from option portfolios, Appendix C contains proofs for the claims made in the text, Appendix D works out certain expressions needed for a disaster model, and Appendix E contains tables and figures supporting the empirics.

2 The Equity Premium

2.1 Preliminaries and Notation

The equity premium as the expected excess return of the market (here the S&P 500) is frequently considered in economic theory as a fundamental trading strategy whose response to risk attitudes of economic agents is studied. It is also considered a profitable trading strategy in itself, at least over longer time periods. Many investors are simply “long the market”, for example via ETF or futures contracts.

With \( S_t \) the S&P 500 spot price and with \( p_{t,T} \) denoting a zero coupon bond, the spot equity premium is the difference between the discrete returns
of the two instruments under the physical, natural, or time series measure $\mathbb{P}$

$$
\mathbb{E}_t^\mathbb{P} \left[ \frac{S_T - S_t}{S_t} \cdot \frac{1 - p_{t,T}}{p_{t,T}} \right] = \mathbb{E}_t^\mathbb{P} \left[ \frac{S_T}{S_t} - \frac{1}{p_{t,T}} \right].
$$

(1)

To avoid keeping track of the bond price and for technical reasons I consider subsequently the \textit{forward equity premium} which is defined as the expectation of the discrete return of a forward $F_{t,T}$ on the S&P 500 with maturity $T$ minus the discrete return of a forward bond

$$
\mathbb{E}_t^\mathbb{P} \left[ \frac{F_{T,T} - F_{t,T}}{F_{t,T}} \cdot \frac{1 - 1}{1} \right] = \mathbb{E}_t^\mathbb{P} \left[ \frac{F_{T,T} - F_{t,T}}{F_{t,T}} \right].
$$

(2)

The forward equity premium is therefore an excess return benchmarked to the risk-free market, but without explicit dependence on it.

Denote by $\mathbb{Q}$ a martingale measure under which the forward price $F_{t,T}$ for delivery of the S&P 500 at time $T$ is a martingale (the $T$-forward measure), and by

$$
R := \frac{F_{T,T}}{F_{t,T}}
$$

(3)

its gross return. For notational convenience I henceforth drop all time subscripts from returns and conditional expectations and develop all ideas for given points in time $t$ (now) and $T$ (tomorrow). With this convention, every quantity is to be understood conditionally on the information available at time $t$. By no arbitrage

$$
\mathbb{E}_t^\mathbb{Q} \left[ R \right] = 1,
$$

(4)

and hence $R$ can be seen to have a second nature as a density, say $d\mathbb{F}$, with respect to $\mathbb{Q}$

$$
R = \frac{d\mathbb{F}}{d\mathbb{Q}},
$$

(5)

in addition to being the S&P 500 forward gross return. There is another density defined through $R$ which is needed for the subsequent arguments, the Hellinger distribution with respect to $\mathbb{Q}$, defined as

$$
\frac{d\mathbb{H}}{d\mathbb{Q}} := \frac{R^{1/2}}{\mathbb{E}_t^\mathbb{Q} \left[ R^{1/2} \right]}.
$$

(6)
This density is termed one-half measure in Carr and Lee (2009). It takes a special place in that it is the probability measure to assess whether the forward-neutral distribution of an asset yields an implied volatility surface which is symmetric in log moneyness around the forward. Carr and Lee (2009) show that this is equivalent to an even characteristic function under the one-half measure. The idea originates from the desire to measure an asset distribution’s symmetry taking into account that prices are positive, that forward prices are martingales under the forward measure, and the presence of option prices along with the shape of the implied volatility surface. The notion of symmetry is intuitively and naturally connected to the notion of loss aversion in that a negatively skewed forward-neutral distribution suggests that negative deviations from the mean are more expensive than positive ones. The Black-Scholes model is one member of the class of put-call symmetric models.

**Example 2.1** (Black-Scholes is Put-Call-Symmetric (PCS)). In the Black-Scholes model, the log forward gross return is normally distributed

\[
\log R \sim Q \left( -\frac{1}{2} \sigma^2 (T - t), \sigma \sqrt{T - t} \right).
\]

From Theorem 2.5 in Carr and Lee (2009) an asset is PCS if the distribution of its log return is symmetric under \( \mathbb{H} \). For this purpose we investigate the characteristic function of \( \log R \) for \( u \in \mathbb{R} \)

\[
\mathbb{E}^\mathbb{H} \left[ e^{iu \log R} \right] = \frac{\mathbb{E}^Q \left[ e^{(1/2 + iu) \log R} \right]}{\mathbb{E}^Q \left[ R^{1/2} \right]} = \exp \left( -\frac{1}{2} \sigma^2 a^2 (T - t) \right), \quad (7)
\]

where \( i \) denotes the imaginary unit. The characteristic function is real and even, and hence the distribution of \( \log R \) is symmetric under \( \mathbb{H} \). Equivalently, the implied volatility surface of the Black-Scholes model as a constant is trivially symmetric.

The name *Hellinger distribution* which I use here is due to the one-half exponent which it shares with the Hellinger divergence from Schneider and Trojani (2015b). A related measure of which I make ample use below is the
inverse Hellinger distribution with respect to $Q$

$$
\frac{d \mathbb{H}}{d Q} := \frac{R^{-1/2}}{E^Q[R^{-1/2}]},
$$

(8)

In the next section I briefly introduce the mathematical foundation for a representation of $R$ in terms of functions that are related to realized measures of higher-order risk and to the financial concept of symmetry from Bates (1997), extended by Carr and Lee (2009). One advantage of the (inverse) Hellinger measure is that they are both measurable with respect to the implied volatility surface at time $t$ through the formula from Carr and Madan (2001), and the market gross return from time $t$ to $T$ reviewed in Appendix B. These are both observable quantities, computable from observable European put $P_{t,T}(K)$ and call prices $C_{t,T}(K)$ on the S&P 500 with strike price $K$ and maturity $T$.

2.2 The Equity Return on Inverse Hellinger Frequency

In this Section I derive an expression for realized equity returns in terms of option prices and a specific functional basis that is tailored for the economic context. Before introducing this basis I start from a formula at the heart of signal processing that requires some concepts and notation. Denote by $\mathcal{D} \subset \mathbb{R}^+$ the support of $R$ and define the Hilbert space $L^2_\Pi$ as the set of random variables $\mathcal{X}$ on $\mathcal{D}$ with finite norm

$$
\|X\|_{L^2_\Pi} := \sqrt{E^\Pi[X^2]} < \infty, \ X \in \mathcal{X},
$$

and corresponding expectation inner product

$$
\langle X, Y \rangle_{L^2_\Pi} := E^\Pi[XY] = \frac{E^Q[R^{-1/2}XY]}{E^Q[R^{-1/2}]}, \ \text{for} \ X, Y \in L^2_\Pi.
$$

(10)

For lighter notation I will suppress below the subscript $L^2_\Pi$ for the norm and the inner product and just write $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$. Suppose $R = \frac{d F}{d Q} \in L^2_\Pi, L^2_\Pi$
separable, then there exists a countable orthonormal basis \( \phi = \{ \phi_1, \phi_2, \ldots \} \) of \( \mathcal{L}_2^H \) such that we can write

\[
R^{(J)} := \sum_{i=0}^{J} c_i \phi_i, \text{ where } c_i = \langle R, \phi_i \rangle, \text{ and } \tag{12}
\]

\[
R = R^{(\infty)}, \tag{13}
\]

where the equals sign in (13) is to be understood in an \( \mathcal{L}_2^H \) sense. The basis \( \phi \) can be a function with multivariate argument, but in the present context it depends only on \( R \). From this construction \( c_i \) is known at time \( t \) (today), and \( \phi_i \) is only known at time \( T \) (tomorrow) through its dependence on \( R \). At time \( t \), \( c_i \) is therefore deterministic, while both \( R \) and \( \phi_i, i = 1, \ldots \) are random variables. Anticipating that \( \phi_0 = 1 \) we also have an expression for the simple, or discrete return as

\[
R^{(J)} - 1 = (c_0 - 1) + \sum_{i=1}^{J} c_i \phi_i. \tag{14}
\]

I now turn to put structure on \( \phi \). The additive functional form (12) is independent of the specific basis chosen. To find an economically meaningful one, I first introduce the sequence

\[
\tilde{B}_n(R) := \frac{2^{n+1}n! \left[ (-1)^n + R \right]}{4\sqrt{R}} - \frac{H_n(R)}{4\sqrt{R}}, \quad n = 0, 1, \ldots, \tag{15}
\]

with the \( n \)-th realized Hellinger contract \( H_n \) introduced by Schneider and Trojani (2015b) defined in Appendix A. They reflect economically appealing notions of higher-order risk aversion, linked to put-call symmetry and deviations thereof.

**Example 2.2** (Power Utility and Hellinger Contracts). Consider a power-

---

\(^1\)In terms of the inner product orthonormality is defined as the relation

\[
\langle \phi_i, \phi_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise}. \end{cases} \tag{11}
\]
utility economy, where the representative agent has inter-temporal marginal rate of substitution $R^{-\gamma}$ with $\gamma > 0$. Adopting this specification one can also look at the Hellinger measure with respect to the otherwise unknown physical $\mathbb{P}$ measure. For $u \in \mathbb{R}$ write

$$
\mathbb{E}^H[R^u] = \mathbb{E}^\mathbb{P}\left[ \frac{d\mathbb{H}}{d\mathbb{Q}} \frac{d\mathbb{Q}}{d\mathbb{P}} R^u \right] \propto \mathbb{E}^\mathbb{P}[R^{1/2+u-\gamma}] .
$$

(16)

From Carr and Lee (2009) we know that an asset is PCS if for every $u \in \mathbb{R}$

$$
\mathbb{E}^H[R^u] = \mathbb{E}^H[R^{-u}] .
$$

(17)

An expansion of the random variable in the $\mathbb{P}$ expectation from Eq. (16) in $u$ around $\gamma$ gives

$$
R^{1/2+u-\gamma} = \sqrt{R} + \sqrt{R} \log R(u - \gamma) + \frac{1}{2} \sqrt{R}(\log R)^2(u - \gamma)^2 \\
+ \frac{1}{6} \sqrt{R}(\log R)^3(u - \gamma)^3 + \cdots .
$$

(18)

Modulo delta-hedging (terms of order $R$ and constant) the first summand on the right hand side of Eq. (18) is $H_0$ (first-order risk aversion), the second $H_1$ (second-order risk aversion), the third $H_2$ (third-order risk aversion), and so forth. It is easy to see that the only way to guarantee condition 17 is to have risk neutrality ($\gamma = 0$ and hence $\mathbb{P} = \mathbb{Q}$) and for any odd summand

$$
0 = \mathbb{E}^\mathbb{Q}\left[ \sqrt{R} \log R \right] = \mathbb{E}^\mathbb{Q}\left[ \sqrt{R}(\log R)^3 \right] = \cdots .
$$

(19)

This example develops an analogy to Kraus and Litzenberger (1976) and Harvey and Siddique (2000), who relate the derivatives of the marginal rate of substitution to aversion against variance, prudence (loss aversion), temperance (tail aversion), to the implied volatility surface and its properties encoded through prices of Hellinger contracts. Within the Black-Scholes Example 2.1, Eq. (19) would indeed be valid for any odd power. The implied volatility surface under Black-Scholes is flat, and hence trivially PCS.

The distinction between loss aversion and tail risk here is meant in the
sense that loss aversion describes an aversion to negative realizations of the return, regardless of whether in the left tail of the distribution or near the mean, and tail risk describes realizations far out in the tails of the distribution for both positive and negative deviations.

Aside from the interpretation above in a power-utility economy, Hellinger contracts also have an information-theoretic grounding. The first contract $H_0$, for example, measures realized divergence of the sample path of $F$. The difference between implied and realized divergence is a generalized measure of aversion to variance, the variance risk premium. Furthermore, Hellinger contracts have an interpretation as realized power moments of log $R$ as

$$H_0(R) = 2(1 - 2\sqrt{R} + R) = \frac{1}{2}(\log R)^2 + O(\log R)^3,$$
$$H_1(R) = 4\left(R - \sqrt{R}\log R - 1\right) = \frac{1}{6}(\log R)^3 + O(\log R)^4,$$
$$H_2(R) = 16\left(R - \sqrt{R}\left(\frac{(\log R)^2}{4} + 2\right) + 1\right) = \frac{1}{12}(\log R)^4 + O(\log R)^5.$$

Finally, forward prices of $H_0, H_1,$ and $H_2$ allow the interpretation as level, slope and curvature of the implied volatility surface, respectively (Schneider and Trojani, 2015b). Normalizing with $\sqrt{R}$, a measure of realized variance, together with the additional delta-hedging in Eq. (15) is done for two reasons. Firstly the inner product $\langle R, \tilde{B}_i \rangle$ yields a forward-neutral expectation of Hellinger moments, and thus economically interpretable coefficients $c$. Secondly, $\tilde{B}_i$ are a basis of $L_\Pi^2$.

**Proposition 2.1** (Hellinger Basis of $L_\Pi^2$). Suppose $D$ is compact or the tails of log $R$ under $\Pi$ decay exponentially if it has unbounded support, then scaled Hellinger moments $\tilde{B}_n(R)$ are a basis of $L_\Pi^2$.  

The above proposition ensures that any function $f \in L_\Pi^2$ can be expressed in terms of realized and implied Hellinger moments. Below are the first few
basis functions with $r := \log R$

\[
\begin{align*}
\tilde{B}_0 &= 1, \\
\tilde{B}_1 &= r, \\
\tilde{B}_2 &= 8 + r^2, \\
\tilde{B}_3 &= r(24 + r^2), \text{ and} \\
\tilde{B}_4 &= 384 + 48r^2 + r^4.
\end{align*}
\]

Orthogonalizing with the Gram-Schmidt algorithm and defining the moment-generating function for a generic measure $\mathbb{M}$, assuming that it exists,

\[\Phi^\mathbb{M}(u) := \mathbb{E}^\mathbb{M}[e^{u \log R}],\]  \tag{20}

such that

\[
\mu_{u,v}^\mathbb{H} := \frac{\partial^v \Phi^\mathbb{H}(u)}{\partial u^v} = \mathbb{E}^\mathbb{H}[R^u (\log R)^v] = \frac{\mathbb{E}^\mathbb{Q}[R^{u-1/2}(\log R)^v]}{\mathbb{E}^\mathbb{Q}[R^{-1/2}]}, \tag{21}
\]

yields

\[
\begin{align*}
B_0 &= 1, \quad B_1 = r - \mu_{0,1}^\mathbb{H}, \\
B_2 &= r^2 - \mu_{0,2}^\mathbb{H} - \frac{\left(\mu_{0,3}^\mathbb{H} - \mu_{0,1}^\mathbb{H} \mu_{0,2}^\mathbb{H}\right) \left(r - \mu_{0,1}^\mathbb{H}\right)}{\mu_{0,2}^\mathbb{H} \left(\mu_{0,1}^\mathbb{H}\right)^2}, \tag{22}
\end{align*}
\]

orthogonal with respect to $\mathbb{H}$. Finally the orthonormal basis functions from Eq. (12) are

\[\phi_i := \frac{B_i(R)}{\|B_i(R)\|}.\]  \tag{23}

To emphasize the factor structure I also use below the notation

\[F_i := c_i \phi_i, \text{ so that} \]

\[R^{(n)} - 1 = (F_0 - 1) + F_1 + F_2 + \cdots + F_n.\]  \tag{24}

Each factor will be a product of implied Hellinger moments known at time $t$, and scaled, realized Hellinger moments known at time $T > t$. In the next Section I use this factor structure to interpret the market return in terms of (tail) risk and loss aversion.
3 An Anatomy of the Equity Premium

The forward price itself is a conditional expectation under the forward-neutral measure of the terminal spot value. As such it will be exposed to the (a)symmetry of this distribution. It is a remarkable feature of decomposition (12) that also the realized forward return, the future spot value divided by its forward price, depends on this distribution through the coefficients \( \langle R, \phi_i \rangle \) on \( \phi_i \). With these coefficients determined at time \( t \), and the gross return \( R \) realizing at time \( T > t \), this means that \( R \) depends on all higher-order forward-neutral moments, as well as the realizations of \((\log R)^n, n = 0, 1, 2, \ldots\)

Below I develop explicitly how these higher-order moments affect \( R \) starting from their most important contribution. For this purpose I introduce the notion of an order-\( J \) statement, where lower orders are more important quantitatively than higher orders with the rationale that \( R^{(J)} \) is closer to \( R^{(\infty)} = R \), the higher \( J \). Similar to a Taylor series expansion, the terms entering first are of higher importance.

**Definition 3.1** (Order-J Statements). A statement is of order \( J \), if it is made in terms of \( R^{(J)} \).

**Definition 3.2** (Skewed Economy). A quantity is said to be increasing in skew if it is smaller for negative, and larger for positive deviations from put-call-symmetry in the sense of Carr and Lee (2009) such that for every \( u \in \mathbb{R} \)

\[
\Phi^H(u) = \Phi^H(-u) \iff \text{market is PCS,} \tag{25}
\]
\[
\Phi^H(u) < \Phi^H(-u) \iff \text{neg. dev. from PCS,} \tag{26}
\]
\[
\Phi^H(u) > \Phi^H(-u) \iff \text{pos. dev. from PCS.} \tag{27}
\]

The below Eqs. show order-0 and order-1 equity returns.

\[
R^{(0)} = \mu_{1,0}, \text{ and} \tag{28}
\]
\[
R^{(1)} = \frac{\mu_{0,2} \mu_{1,0}^2 - \mu_{0,0} \mu_{1,1}^2}{\mu_{0,2}^2 - \left( \mu_{0,1}^2 \right)} + \frac{\mu_{1,1} - \mu_{0,1} \mu_{1,0}^2}{\mu_{0,2}^2 - \left( \mu_{0,1}^2 \right)} \log R. \tag{29}
\]
The expressions are collected here with respect to orders of \( \log R \), rather than the basis \( \phi_i \) to build intuition. The fact that the constant term in \( R^{(0)} \) is not the same as in \( R^{(1)} \) arises from the orthogonalization,\(^2\) but generally the change in coefficients from low order to high-order representations is small. Under no-arbitrage, the equity return is order-0 negatively skewed.

**Result 3.3** (Order-0 Equity Return). The \( \theta \)-order equity return

\[
R^{(0)} - 1 = \mu^{\Pi}_{1,0} - 1 < 0.
\]  

\( R^{(0)} \) tends to be increasing in skew for the data and extant models,\(^3\) and reflects the intuition that the forward-neutral distribution has a thicker left tail than the physical one. The next statement reveals a first glimpse at the connection of the market return to put-call symmetry.

**Result 3.4** (Order-1 Equity Return). The discrete order-1 equity return can be written

\[
R^{(1)} - 1 = \alpha + \beta \cdot \log R,
\]  

where \( \alpha \) and \( \beta \) have the familiar linear regression interpretation

\[
\beta = \frac{\text{Cov}^{\Pi}(R, \log R)}{\text{Var}^{\Pi} \log R}, \quad \alpha = (\mu^{\Pi}_{1,0} - 1) - \beta \mu^{\Pi}_{0,1},
\]  

where the regression slope \( \beta > 0 \).

Result 3.4 shows a CAPM-type equation with option-implied betas that capture market-conditional information in the spirit of Buss and Vilkov (2012).\(^4\) Eq. 32 also highlights very explicitly the difference between the

\(^2\)The first basis vector \( \phi_1 \) also has a constant part as can be seen from Eq. (22).

\(^3\)This arises from

\[
\mu^{\Pi}_{1,0} = \frac{1}{\mathbb{E}^{\Pi} [e^{-\log R}]},
\]  

The more left-skewed the distribution under \( \mathbb{H} \), the higher will be \( e^{-\log R} \), and thus the lower \( \mu^{\Pi}_{1,0} \). A sharp mathematical statement is not possible, however, in this case, since put-call symmetry does not say anything about the dispersion of a distribution, and the overall concentration of probability mass.

\(^4\)With the actual expression being different from theirs.
approach chosen in this paper and conventional statistical methods. Were
the measure used above the unconditional $\mathbb{P}$ measure, the expression would
correspond to a standard linear regression subject to measurement errors
and relying on asymptotics that averages of observations become their ex-
pectations for computing the coefficients $\alpha$ and $\beta$ fast enough. Using the
$\mathbb{H}$ measure instead, the coefficients become measurable conditionally from
option prices without resorting to asymptotics and stationarity of the data.
The practical implications are significant while the two concepts are the same
from a mathematical point of view and conceptually. If the market is put-
call symmetric then $\mu_{1,1}^{\mathbb{P}} = 0$, and $\text{Cov}^\mathbb{P}(R, \log R) = -\mu_{1,0}^{\mathbb{P}} \cdot \mu_{0,1}^{\mathbb{P}}$, such that
the entire expression depends just on the marginal moments of $\log R$ and $R$.
The relevance of PCS on the Hellinger representation of the equity return is
taken to a more general level by the next result.

**Result 3.5** (Put-Call Symmetry and Odd Hellinger Moments). *If the market
is PCS, the forward return depends only on even implied Hellinger moments.*

This statement is a generalization of the intuitive notion developed in
the power utility Example 2.2. If there is no loss aversion in the market, it
will also not be seen in equity returns. The power of this statement becomes
apparent in the empirical anatomy in Section 3.1 below, where this effect is
quantified.

A useful concept to assess the convergence of the series expansion is the
squared norm, or the energy of $R$

$$
\|R\|^2 = \mathbb{E}^\mathbb{P}[R^2] = \frac{\mathbb{E}^\mathbb{Q}[R^{3/2}]}{\mathbb{E}^\mathbb{Q}[R^{-1/2}]} = \sum_{i=0}^{\infty} \langle R, \phi_i \rangle^2 ,
$$

which is also benchmarked to PCS. This quantity is increasing in skew, with
$\|R\|^2 = 1$ in a put-call symmetric market, and consequently with a norm
greater (smaller) than 1 if the economy is positively (negatively) skewed.
Likewise we can compute the residual energy from

$$
\|R - R^{(J)}\|^2 = \|R\|^2 - \|R^{(J)}\|^2 = \sum_{i=J+1}^{\infty} \langle R, \phi_i \rangle^2 .
$$
With both $\|R\|$ and $\|R^{(J)}\|$ known from option prices, this expression can be used to determine a cut-off value for $J$ as formula (35) can be computed from option prices. Figure 1a shows the logarithm of $\|R - R^{(J)}\|^2$ for different option maturities for $J = 2$ over time. The time series suggest that $R^{(J)}$ is extremely close to $R$ for $J = 2$, and that we can safely ignore higher orders. This highlights the quality of the basis functions chosen as the infinite series (13) converges after the first three constituents.

3.1 A Decomposition of Realized Index Returns

The forward price of the simple return $E^Q [R - 1] = 0$ by no-arbitrage. As a consequence also the right-hand side of Eq. (13) has a forward price of zero. Given that the energy of the signal $R - 1$ is concentrated in the first few coefficients of the transform, it is reasonable to expect that the forward price of $R^{(J)} - 1$ will also be close to 0. Figure 1b shows that this is indeed the case. Deviations for the short maturity prices are virtually zero, and the longer-maturity ones are very small. With a forward price of zero I can take $R^{(J)}$ as the payoff from a trading strategy, just like entering into a forward contract on the market, with the difference that the $R^{(J)}$ strategy requires a portfolio of portfolio of options (one for each $\phi_i, 1 \leq i \leq J$), rather than two options to replicate the forward contract on $S_t$. Equipped with this tradeability property I can contrast the left-hand side of Eq. (12) with its right-hand side, and investigate the composition of the market return in terms of Hellinger portfolios, the joint forward price of which is zero.

Figure 2 shows the absolute value of realized forward market returns for maturities 1, 3, 6, and 12 months, along with the absolute value of its Hellinger decomposition for various orders. First-order risk aversion, or aversion to dispersion risk encoded by $|F_0 - 1|$, is time-varying and correlates with implied variance. It plays only a minor role for short maturities (Figures 2a and 2b), but makes up substantial parts for longer maturities (Figures 2c and 2d). In the late 1999’s it constitutes almost everything of the 1-year realized market return. In contrast, Figure 4 shows the corresponding decomposition through the lens of a Gaussian model introduced below in Section 4. The
shortcoming of this model becomes immediately visible through $|F_0 - 1|$, and in fact all coefficients $c_i$ being constant.\footnote{Were the volatility of the Gaussian model set at every point in time to implied volatility the coefficients would become time-varying, rendering the constant volatility assumption inconsistent. A Heston-type with its additional state variable would reflect this time variation and time-varying $c_i$ provided stochastic variance was time-varying.}

By far the biggest component across all maturities is the first factor $F_1$, which is associated with loss aversion. This important role comes almost naturally with its exposure to the asymmetry of $R - 1$. The factor $F_2$ associated with aversion to tail risk is visible only in periods of turmoil, and then more pronouncedly so for longer maturities. While compensation for tail risk is of high importance to premia on variance,\footnote{See, for example, \textit{Bollerslev and Todorov (2011)}} for instance, it is very intuitive that investors who take on linear market exposure are not so much affected by it.

A natural concern that arises with a factor decomposition (24) is the covariation between the factors. Optimally one would want as little of it as possible. Figure 7 shows the trajectories of the basis functions $\phi$ over time. The basis functions $\phi$ are constructed to be orthogonal with respect to the inverse Hellinger measure $\mathbb{H}$, but there is no reason to expect they should be orthogonal with respect to the physical measure $\mathbb{P}$ which has generated the time series. The results of an empirical assessment of their correlation can be seen in Table 1. Panel 1a shows correlations with a maximum of 76% between $\phi_1^{1m}$ and $\phi_2^{1m}$. This is an indication that the $\mathbb{P}$ measure seems to be quite far from $\mathbb{H}$, but not excessively so. The correlations of the factors $F$, the product of the time-varying coefficients $c_i$ and the basis functions $\phi_i$, are much lower with a maximum (across factors with the same maturity) of 65% between $F_1^{1m}$ and $F_2^{1m}$, and decreasing for longer maturities.

Below I investigate the question whether knowing the risk anatomy of $R$ makes a difference for predictions.

### 3.2 Does the Decomposition Improve Predictability?

Predictions of equity returns in the setting of this paper can be done through models for $\phi_i$ of decomposition (12), or directly. Benefits from the decompo-
sition could arise from differences in persistence between different $\phi_i$. With $i$ even, for example, they could inherit persistence from realized variance. Another potential source of predictability is the dependence of $\phi_i$ and $c_i$ on option prices, which reflect current market conditions and have been proven successful in empirical applications. At the same time, it is conceivable that forecasts from a multivariate model for $\phi_1, \phi_2, \ldots$ might increase the estimation error.

The linear model (32) would suggest the possibility of a persistent factor $F_2$. In Eq. (32), from $a - 1 \approx 0$ and $b \approx 1$ and using $R$ instead of $R^{(1)}$

$$R - \log R - 1 > 0.$$ 

This payoff is related to variance,\(^7\) which is known to be persistent. Figure 7 shows sample paths of the basis functions over time. With $\phi_0$ being identically one for any maturity, the time-varying $\phi_1$ in Panel 7a looks like white noise. Persistence is introduced mechanically into $\phi_1$ for higher maturities through overlapping returns. The trajectories for $\phi_2$ follow a similar pattern. The unpredictable nature of the basis functions over time is also reflected in their autocorrelation plots in Figure 9. The dynamics of the factors $F_i = c_i \phi_i$ show a slightly different behaviour. Figure 8 hints at more persistence, but the first-order factor $F_1$ still appears unpredictable according to its autocorrelation function, while the second-order factor $F_2$ shows significant autocorrelation at the first two lags (Panel 10a).\(^8\)

I take the conditional moments from Schneider and Trojani (2015a, ST) as a model for the expectation of the powers of $r$ in $\phi_i$ from Eq. (22) under the physical measure. These $\mathbb{P}$ moments are recovered from assumptions on expected profits of trading strategies together with a minimum variance projection of the pricing kernel onto the S&P 500, but otherwise model-free. They pertain to simply compounded returns, but can be expected to be close enough to the moments of log returns. Importantly, they are well-suited for an out-of-sample prediction, since they are based on option

\(^7\)It is realized entropy which also underlies the specification of the VIX² contract.

\(^8\)Evidence for predictability stemming from tail risk measures can also be found in Kelly and Jiang (2014) and Bollerslev et al. (2014).
data available at the time of the prediction only. I consider out-of-sample comparisons between 4 competing models according to the out-of-sample $R^2$ definition of Campbell and Thompson (2008). The first model is the sample average of past returns. The second is the decomposition (24) evaluated at the conditional moments from ST. The third is the first conditional moment from ST itself, and the fourth is the simple variance swap ($SVIX^2$) from Martin (2015), whose connection to the equity premium is motivated from an assumption on the covariation between the pricing kernel and the S&P 500 return. Table 2 shows that the decomposition is successful in predicting market forward returns out-of-sample only marginally for 1 month maturity, and marginally not for the 3 month maturity return. It is very successful for the longer maturities 6 and 12 months, however. The conditional first moment of ST and the $SVIX^2$ are very similar. It is therefore not surprising that the numbers are virtually identical. Here the decomposition does not add information for the shorter maturities 1 and 3 months, but significantly for the longer maturities 6 and 12 months. Variety in the persistence of the factors of the decomposition therefore does make a difference for predicting forward market returns, with values well within the plausible interval for $R^2$ in Ross (2005) and also in the range of Pettenuzzo et al. (2014), who use a much bigger information set.

4 Disaster Risk and Equity Returns

The anatomy of the market return in Section 3.1 is model-free, but lends itself also to be computed from models. This allows to confront the composition of model-implied equity returns to the observed model-free one. For this exercise, I employ the model from Backus et al. (2011), originating from Barro (2006). Its purpose is to connect a representative agent’s consumption with returns of the market portfolio emphasizing the role of disaster risk quantified through non-Gaussian innovations. This model is built upon power utility preferences with an i.i.d. factor structure, but reveals accessible insights due to its simplicity.

Time evolves discretely, and one period is normalized to one year.
Through the model’s i.i.d. structure, shorter or longer periods can be obtained through scaling the exponent of the characteristic function. The logarithm of consumption growth evolves according to

$$\log g_{t+1} = w_{t+1} + z_{t+1},$$  \hspace{1cm} (36)

where $w_{t+1} \overset{i.i.d.}{\sim} N(\mu, \sigma)$, and $z_{t+1}$ is determined through an independent Poisson counter $j$ with intensity $\omega$ and a conditional Normal distribution $z_{t+1} \mid j \overset{i.i.d.}{\sim} N(j\theta, \sqrt{j}\delta^2)$. This specification separates in simple terms “small” consumption risk (through $w$) from “big” (through $z$). The $\tau$-step-ahead moment-generating function of $\log g_{t+1}$ under the natural probability measure $\mathbb{P}$ reflects independent increments of log consumption growth

$$h(u) := \mathbb{E}^\mathbb{P}[e^{u \log g_{t+1}}] = \exp \left( \tau \left( \omega \left( e^{\frac{2u^2}{2} + \theta u} - 1 \right) + \frac{\sigma^2 u^2}{2} + \mu u \right) \right),$$ \hspace{1cm} (37)

where as before the expectation is taken to be conditional on time $t$ information. The corresponding cumulant generating function is $k(u) := \log h(u)$. The model is completed with dynamics for the one-period log pricing kernel and equity return, respectively,

$$\log m_{t+1} := \log \beta - \alpha \log g_{t+1}, \text{ and}$$

$$\log r_{t+1}^e := \lambda \log g_{t+1} - \log \beta - k(\lambda - \alpha),$$ \hspace{1cm} (38), (39)

where $\beta$ encodes the time preference rate of the representative agent, $\alpha$ her risk aversion, and $\lambda$ the leverage on the consumption asset. Appendix D shows the corresponding forward pricing kernel and forward market gross return $R$. As a first glance on the impact of big (compound Poisson) risk on the forward-neutral distribution of the market return, shutting down the non-Gaussian innovation $z_{t+1}$ in Eq. (36) gives PCS option prices identical to the Black-Scholes model from Example 2.1 with an even characteristic one-period function under $\mathbb{H}$. For $\omega > 0$ the implied volatility surface will be non-symmetric, independently of the values of $\theta$ and $\delta$. Denoting by
\[ r := \log R \] as before and setting \( \omega = 0 \) we can compute model-implied low-order representations as

\[ R^{(0)} = e^{-\frac{1}{2} \lambda^2 \sigma^2}, \quad \text{and} \quad R^{(1)} = R^{(0)} (1 + r + \lambda^2 \sigma^2). \quad (40) \]

From Eq. (40) the price of the tractability of the i.i.d. specification becomes apparent. The first factor \( c_0 \phi_0 = e^{-\frac{1}{2} \lambda^2 \sigma^2} \) is constant, in contrast to its data-implied counterpart in Figures 2 and 8. An obvious misspecification, less severe so for shorter-maturities, but quite sizeable for longer-maturities, where a great portion of the variation in market returns is due to this factor.

Next I check model-implied market returns against data-implied ones for three different parametrizations from Backus et al. (2011) in Table 3. These different parametrizations highlight the gap between option-implied distributions and macro-data-implied ones. There are two specifications (macro I and macro II) matching the unconditional properties of consumption growth, mean and standard deviation of the equity premium. The macro II specification also accommodates a very low-frequent but severe consumption disaster risk component. The third specification option uses parameters calibrated from option prices on the S&P 500 through the same model. Figure 3 shows that the densities of market returns are quite similar for the macro I and option specifications, in contrast to the density from the macro II specification, which is bimodal with significant probability mass on large negative losses.

Figure 4 highlights the drawback of an i.i.d. return model in that the leading term \( F_0 \) of the expansion is constant, whereas the model-free \( F_0 \) in Figure 2 is clearly time-varying. The effect of misspecification is increasing in maturity, as \( F_0 \) tends to be more important for long-maturity returns.

The Gaussian macro I, as well as the Gaussian-Poisson mixture specification

---

The expressions become much bigger and much less intuitive otherwise, for example, with \( \omega > 0 \) we have

\[ R^{(0)} = e^{\left(\omega e^{\frac{\sigma^2}{2}} - \sigma \lambda \left( e^{\frac{\lambda}{2} \left( (\lambda - 4n + 12e) - e^{\frac{\lambda}{2} \left( (2(\lambda - 4n + 12e) + 4e) - e^{\frac{\lambda}{2} \left( (2(\lambda + 4) + 4e) + e^{\lambda} \right)} - \frac{\sigma^2}{2} \right) \right)} \right) \right)}. \]
option calibrated to options in Figures 4 and 6 suggest a decomposition which is lacking in particular at modelling $F_0$ and $F_2$. For a maturity of 1 month, the macro I specification even exhibits a negative correlation of $F_2$ with its model-free counterpart (Table 4). The decomposition of risk compensation implied by the macro II parametrization in Figure 5 suggests a very unrealistic model. Most likely as a result of the distribution’s two modes (Figure 3), the $F_0$ and $F_2$ factors are too large by orders of magnitude. As a consequence, the representation of the risk structure of the equity return as a whole is visibly imprecise. Given that tail risk is a third-order contributor to market risk\textsuperscript{10} but a first-order contributor to model and parameter risk, its addition to a model for market returns may therefore not be imperative.

5 Conclusions

The forward market return is the payoff from entering into a forward contract on the S&P 500 divided by the forward price. The conditional expectation of this quantity is the forward equity premium, a central object of economic theory. The common understanding is that, provided there is risk aversion in the economy, it must be positive as a compensation for risk. Unfortunately it is impossible to make statements about this premium without a model. The literature concerned with making meaningful statements about the probability law determining this quantity with the least amount of assumptions has only started recently with Carr and Yu (2012), Borovička et al. (2014), Christensen (2014), Ross (2015), Martin (2015), Linetzky and Qin (2015b), Linetzky and Qin (2015a), assuming a combination of type of stochastic process driving the economy, Markovianity, stationarity and ergodicity, and the state space, or assumptions on the sign of expected profits from trading strategies themselves (Schneider and Trojani, 2015a).

This paper pursues a different goal by trying to understand the realized forward market return rather than its conditional expectation. The conceptual idea is to use economically interpretable basis functions in an economi-

\textsuperscript{10}The model-free factor $F_2$ associated with tail risk constitutes just a small portion of the realized market return (c.f. Figure 2).
cally relevant space to express this return. The resulting representation turns out to be attainable without the use of any model. The interpretations of the basis functions as realized variance, loss aversion and tail aversion follow from the thought experiment of imposing an expected utility representative agent on the economy with the implementation being model-free.

Using options data on the S&P 500 from 1990 to 2014 I confirm empirically that the economic basis functions are well-chosen, in that three suffice to explain the equity return completely. Its main driving force is realized loss aversion, only for longer maturities realized variance becomes important. The economic decomposition suggests also a promising extension to the literature on predicting the forward return for longer maturities. Disaster risk does not appear prominently in the market return, but modelling it may upset the return’s factor structure.

Appendices

A Hellinger Basis

From the divergence function

$$D_p(R) := \frac{R^p - 1 - p(R - 1)}{p(p - 1)}$$  (41)
we define the \( n - {th} \) Hellinger moment \( H_n(R) \) from

\[
\frac{\partial^n D_p(R)}{\partial p^n} = \frac{\partial^n}{\partial p^n} \left( \frac{1}{p - p^2} + \frac{(R - 1)}{1 - p} - \frac{R^p}{p - p^2} \right) \quad (42)
\]

\[
= n! \left( \left( \frac{1}{1 - p} \right)^{n+1} + \frac{(-1)^n}{p^{n+1}} \right) \quad (43)
\]

\[
+(R - 1)n! \left( \frac{1}{1 - p} \right)^{n+1} \quad (44)
\]

\[
- R^p \sum_{k=0}^{n} \binom{n}{k} (\log R)^{n-k} k! \left( \left( \frac{1}{1 - p} \right)^{k+1} + \frac{(-1)^k}{p^{k+1}} \right), \quad (45)
\]

as

\[
H_n(R) := \left. \frac{\partial^n D_p(R)}{\partial p^n} \right|_{p=1/2}
\]

\[
= 2^{n+1} n! \left( (-1)^n + R - \sqrt{R} \sum_{k=0}^{n} (\log R)^k \left( \frac{1}{2} \right)^{k+1} \frac{1}{k!} (1 + (-1)^{n+k}) \right) \quad (46)
\]

We also have for \( n \geq 2 \) the recursion

\[
\tilde{B}_n(R) = 4n(n-1)\tilde{B}_{n-2}(R) + (\log R)^n \quad (47)
\]

which can be easily seen from representation (46).

B Forward-Neutral Expectations from Option Portfolios

From the formula for the Lagrange remainder term of a \( C^2 \) function \( f \)

\[
f(x) = f(y) + f'(y)(x-y) + \int_y^y f''(K)(K-x)^+dK + \int_y^\infty f''(K)(x-K)^+dK. \quad (48)
\]
Evaluating this formula at \( x = F_{T,T} \) and \( y = F_{t,T} \) and taking forward-neutral expectations yields

\[
E^Q [f(F_{T,T})] = f(F_{t,T}) + \frac{1}{p_{t,T}} \left( \int_0^{F_{t,T}} f''(K) P_{t,T}(K) dK + \int_{F_{t,T}}^\infty f''(K) C_{t,T}(K) dK \right) \tag{49}
\]

C Proofs

C.1 Auxiliary Result

Lemma C.1 (Inverse Hellinger Moments and Put-Call Symmetry). The inverse Hellinger Moment \( \mu_{H_{u,0}} \) satisfies \( \mu_{H_{1/2,0}} < \mu_{H_{u,0}} < \mu_{H_{u,0}} \) for \( 1/2 < u < u^* < u \) and \( u < u^* < -1/2 \), as well as \( \mu_{H_{u,1}} < 0 \) for \( u < 1/2 \) and \( \mu_{H_{u,1}} > 0 \) for \( u > 1/2 \). Under PCS, the inverse Hellinger Moment \( \mu_{H_{u,0}} \) from Eq. (21) has its minimum at \( u = 1 \). This yields \( \mu_{H_{1,0}} < \mu_{H_{u,0}} < \mu_{H_{u,0}} \) for both \( 1 < u^* < u \) and \( u < u^* < 1 \). Furthermore, \( \mu_{H_{u,1}} < 0 \) for \( u < 1 \) and \( \mu_{H_{u,1}} > 0 \) for \( u > 1 \).

Proof. From no-arbitrage \( E^Q [e^{0 \log R}] = E^Q [e^{1 \log R}] = 1 \). The first claim follows from these points and the strict convexity and continuity of the exponential function together with the definition of \( \mu_{H_{u,0}} \). Another consequence of convexity is that \( E^Q [e^{p \log R}] < 1, p \in (0, 1) \). Under PCS we have from Carr and Lee (2009) that \( E^Q [e^{p \log R}] = E^Q [e^{(1-p) \log R}] \). This symmetry implies together with convexity that \( p = 1/2 \) gives the minimal point under PCS.

C.2 Proposition 2.1

Proof. If \( \mathcal{D} \) is unbounded, then denseness of the polynomial canonical basis follows from the exponential tails of \( \mu_{H_{u,0}} \) in \( r \) from Lemma 3.1 in Filipović et al. (2013). For compact \( \mathcal{D} \) it follows from the Stone-Weierstrass theorem. Eq. (47) shows that there is a linear map between the canonical \( \{1, r, r^2, r^3\} \) and the Hellinger basis \( \{\tilde{B}_0, \tilde{B}_1, \ldots\} \) which implies that the Hellinger basis is dense in \( \mathcal{D} \).\( \square \)
C.3 Result 3.3

Proof. We start with properties of inverse Hellinger moments. To prove claim 3.3, positivity of $R^{(0)}$ follows from the positivity of $R$. For the upper bound write

$$R^{(0)} = \langle R, \phi_0 \rangle \cdot \phi_0 = \langle R, 1 \rangle = \mu_{1,0} = \frac{\mathbb{E}^Q [R^{1/2}]}{\mathbb{E}^Q [R^{-1/2}]} < 1, \quad (50)$$

since $\mathbb{E}^Q [R^{-1/2}] > 1$ and $\mathbb{E}^Q [R^{1/2}] < 1$ from the convexity of the moment generating function.

C.4 Proof of Result 3.4

The “Hellinger CAPM” (32) is just a reformulation of Eq. (29). Positivity of $\beta$ follows from Schmidt (2003) which guarantees that the covariance of two increasing functions of a random variable is positive, and positivity of the variance.

$$\text{Cov}^\mathbb{H} (R, \log R) = \mathbb{E}^\mathbb{H} [R \log R] - \mathbb{E}^\mathbb{H} [R] \mathbb{E}^\mathbb{H} [\log R]$$

C.5 Proposition 3.5

We start with

Lemma C.2 (Odd-Order Log Returns in Inverse Hellinger Space). If the economy is put-call symmetric in the sense of Carr and Lee (2009), then the expression

$$\left\langle \frac{dF}{dQ}, \tilde{B}_n(R) \right\rangle = 0 \quad (51)$$

for $n$ odd.

Proof. Plugging in the definition (46) and computing the $\mathbb{H}$ moments reveals for $n$ odd

$$\left\langle \frac{dF}{dQ}, \tilde{B}_n(R) \right\rangle = \frac{1}{4} \cdot \frac{\mathbb{E}^Q [H_n(R)]}{\mathbb{E}^Q [R^{-1/2}]}, \quad (52)$$

and for $n$ even

$$\left\langle \frac{dF}{dQ}, \tilde{B}_n(R) \right\rangle = \frac{1}{4} \cdot \frac{2^{n+2} n! + \mathbb{E}^Q [H_n(R)]}{\mathbb{E}^Q [R^{-1/2}]}, \quad (53)$$
From Schneider and Trojani (2015b) we know that \( \mathbb{E}^Q [H_n(R)] = 0 \) for \( n \) odd under PCS, which proves the claim.

Now, from

\[
\mathbb{E}^Q [H_n(R)] \overset{PCS}{=} 0, \ n \ \text{odd},
\]

and the recursion Eq. (47), every odd-order power \( n \) of \( r \) has a representation in terms of odd-order \( \tilde{B}_n(R) \) and \( \tilde{B}_{n-2}(R) \). The orthogonalized basis \( \{B_0(R), B_1(R), \ldots\} \) is polynomial in \( r \), hence

\[
\left\langle \frac{d\mathbb{F}}{d\mathbb{Q}}, B_n(R) \right\rangle
\]

does not depend on odd \( i \)-order Hellinger moments for \( i = 0, \ldots, n \).

D Expressions for the Disaster Model

With the specification from Eqs. (36), (38), and (39), the forward pricing kernel for time \( \tau \geq t \) is given by

\[
\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{m_{t+\tau}}{\mathbb{E}^P [m_{t+\tau}]} = g_{t+\tau}^{-\alpha \tau} e \left( \alpha \mu - \frac{\omega^2 - 2}{2} - \omega \left( e^{\frac{1}{2} \alpha (\omega^2 - 2\theta)} - 1 \right) \right),
\]

and the forward market return over a period \( \tau \)

\[
R = g_{t+\tau}^{\lambda \tau} e \left( \frac{1}{2} \sigma \left( e^{\frac{1}{2} \alpha (\omega^2 - 2\theta)} - e^{\frac{1}{2} \sigma^2 (\alpha - \lambda)^2 + \theta (\lambda - \alpha)} \right) - \lambda \sigma^2 (\lambda - 2\alpha) - 2\lambda \mu \right).
\]

Model-implied Hellinger and inverse Hellinger measure changes follow from expressions (37), (38) and (39).
E Figures and Tables

Figure 1: Residual Energy. Panel (a) shows how much information is unexplained by option data according to Formula (35) after decomposing the equity return into the first three factors from Eq. (24). The log transformation accounts for the extremely small deviation between the true signal and its transform. Panel (b) shows the forward price of the true signal $R - 1$ and the forward prices of the decomposition with the first three economic basis functions from Eq. (23). The data producing both figures are European options on the S&P 500 from 1990 to 2014.
Figure 2: Equity Returns and Hellinger Decomposition. The panels show the factor decomposition (24) for different orders of the expansion. To gauge the magnitudes of the contributions they are depicted in absolute values. The factor decomposition can both over-, and undershoot the true return, but deviations for the higher-order representations are small (see also Figure 1). The data producing both figures are European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, 6, and 12 months.
Figure 3: Density Plots. This Figure shows densities of simply compounded (upper row) and continuously compounded (lower row) forward market returns from the Backus et al. (2011) model introduced in Section 4. The plots are obtained through Fourier inversion from expressions (38) and (39) using the parametrizations from Table 3.
Figure 4: Equity Returns and Hellinger Decomposition in Gaussian Model. The panels show the factor decomposition (24) using the model from Section 4 using the macro I parametrization in Table 3. To gauge the magnitudes of the contributions they are depicted in absolute values. The factor decomposition can both over-, and undershoot the true return. Forward prices entering the factors are computed from European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, 6, and 12 months.
Figure 5: Equity Returns and Hellinger Decomposition in Disaster Model. The panels show the factor decomposition (24) using the model from Section 4 using the macro II parametrization in Table 3. To gauge the magnitudes of the contributions they are depicted in absolute values. The factor decomposition can both over-, and undershoot the true return. Forward prices entering the factors are computed from European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, 6, and 12 months.
Figure 6: Equity Returns and Hellinger Decomposition in Disaster Model Calibrated to Options. The panels show the factor decomposition (24) using the model from Section 4 using the *option* parametrization in Table 3. To gauge the magnitudes of the contributions they are depicted in absolute values. The factor decomposition can both over-, and undershoot the true return. Forward prices entering the factors are computed from European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, 6, and 12 months.
Figure 7: Time-Varying Basis Functions. The panels show basis functions $\phi_i$ from Eq. (23) for orders $i = 0, 1,$ and 2. The data producing the panels are European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, 6, and 12 months.
Figure 8: Factors. The panels show factors $F_i$ from Eq. (24) for orders $i = 0, 1, \text{and } 2$. Factor $F_1$ is plotted on the left $y$-axis, while $F_0$ and $F_2$ are on the right $y$-axis. The data producing the panels are European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, 6, and 12 months.
Figure 9: Basis Function Autocorrelations. The panels show autocorrelations of basis functions $\phi_i$ from Eq. (23) for orders $i = 1$, and 2 (solid blue). The dotted red lines represent 95% confidence intervals. The data producing the panels are European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, and 12 months.
Figure 10: Factor Autocorrelations. The panels show autocorrelations of factors $F_i$ from Eq. (23) for orders $i = 1$, and 2 (solid blue). The dotted red lines represent 95% confidence intervals. The data producing the panels are European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, and 12 months.
### (a) Basis Function Correlations

<table>
<thead>
<tr>
<th>$\phi_1^m$</th>
<th>$\phi_2^m$</th>
<th>$\phi_3^m$</th>
<th>$\phi_4^m$</th>
<th>$\phi_6^m$</th>
<th>$\phi_1^{12m}$</th>
<th>$\phi_2^{12m}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
<td>54</td>
<td>24</td>
<td>45</td>
<td>29</td>
<td>36</td>
<td>24</td>
</tr>
<tr>
<td>44</td>
<td>30</td>
<td>32</td>
<td>30</td>
<td>27</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>74</td>
<td>37</td>
<td>56</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>52</td>
<td>24</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>74</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>47</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55</td>
</tr>
</tbody>
</table>

### (b) Factor Correlations

<table>
<thead>
<tr>
<th>$F_0^{1m}$</th>
<th>$F_1^{1m}$</th>
<th>$F_0^{3m}$</th>
<th>$F_1^{3m}$</th>
<th>$F_0^{6m}$</th>
<th>$F_1^{6m}$</th>
<th>$F_0^{12m}$</th>
<th>$F_1^{12m}$</th>
<th>$F_0^{12m}$</th>
<th>$F_1^{12m}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-12</td>
<td>24</td>
<td>97</td>
<td>-18</td>
<td>26</td>
<td>88</td>
<td>-23</td>
<td>-2</td>
<td>78</td>
<td>-25</td>
</tr>
<tr>
<td>62</td>
<td>-17</td>
<td>57</td>
<td>25</td>
<td>-19</td>
<td>47</td>
<td>28</td>
<td>-18</td>
<td>34</td>
<td>29</td>
</tr>
<tr>
<td>14</td>
<td>28</td>
<td>35</td>
<td>7.7</td>
<td>19</td>
<td>27</td>
<td>1.6</td>
<td>9.9</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-21</td>
<td>21</td>
<td>95</td>
<td>-28</td>
<td>-8.4</td>
<td>87</td>
<td>-30</td>
<td>-21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>-23</td>
<td>76</td>
<td>34</td>
<td>-21</td>
<td>57</td>
<td>42</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>14</td>
<td>59</td>
<td>5.9</td>
<td>16</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-30</td>
<td>-16</td>
<td>95</td>
<td>-31</td>
<td>-23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>-27</td>
<td>75</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-17</td>
<td>43</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-26</td>
<td>-20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1:** Factor and Basis Function Correlations. Table (a) shows correlations (in %) of the basis functions $\phi$ from Eq. (23). The basis functions are uncorrelated by construction under $\mathbb{F}$, the non-zero correlations stem from estimating them under the unconditional $\mathbb{P}$ measure. Table (b) show the correlations (in %) of the factors from Eq. (24) in the main text. Both tables are computed with SPX options in the years 1990-2014.
Table 2: Out-of-Sample $R^2$. The Table shows out-of-sample $R^2$ taken from Campbell and Thompson (2008) between two computing models 1 and 2 from the definition

$$R^2 := 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_{1,i})^2}{\sum_{i=1}^{n} (y_i - \hat{y}_{2,i})^2},$$

(58)

where $\hat{y}_{1,i}$ and $\hat{y}_{2,i}$ denote predictions based on the information up to time $i-1$ from models 1 and 2, respectively. In the first row the base model is the sample average (model 2), evaluated against the decomposition $E_{i-1}^P [R^{(2)}]$ at the conditional moments from Schneider and Trojani (2015a, ST) (model 1). In the second row the conditional second moment from ST is used as a baseline model (model 2), tested against the decomposition evaluated at the moments from ST. The third row uses the $SVIX^2$ from Martin (2015) as a base-line model. The data are European options written on the S&P 500 from January 1990 to January 2014.

<table>
<thead>
<tr>
<th></th>
<th>1m</th>
<th>3m</th>
<th>6m</th>
<th>12m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decom. vs Average</td>
<td>0.003</td>
<td>-0.004</td>
<td>0.042</td>
<td>0.032</td>
</tr>
<tr>
<td>Decom. vs Recov.</td>
<td>-0.003</td>
<td>-0.007</td>
<td>0.014</td>
<td>0.007</td>
</tr>
<tr>
<td>Decom. vs $SVIX^2$</td>
<td>-0.003</td>
<td>-0.007</td>
<td>0.014</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Table 3: Parameter Specification. This Table shows the parameter values used in Backus et al. (2011). For all three specifications leverage on the consumption asset $\lambda = 5.1$. The specifications imply an expected market forward return of 5.7%, 5%, and 5.7%, volatility of 18% (all three specs.) and a pricing kernel entropy of 4.9%, 4%, and 4.8% respectively for the macro I, macro II, and option specifications.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\omega$</th>
<th>$\theta$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>macro I</td>
<td>8.92</td>
<td>0.02</td>
<td>0.035</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>macro II</td>
<td>5.19</td>
<td>0.023</td>
<td>0.01</td>
<td>0.01</td>
<td>-0.3</td>
<td>0.15</td>
</tr>
<tr>
<td>option</td>
<td>8.7</td>
<td>0.0303</td>
<td>0.0253</td>
<td>1.3987</td>
<td>-0.0074</td>
<td>0.0191</td>
</tr>
</tbody>
</table>

Table 4: Model-Implied vs. Model-Free Factors. This Table shows correlations between factors implied by the Backus et al. (2011) model reviewed in Section 4 with the parameter values from Table 3 and the model-free factors from Section 3.

<table>
<thead>
<tr>
<th></th>
<th>$F^{1m}$</th>
<th>$F^{1m}$</th>
<th>$F^{1m}$</th>
<th>$F^{3m}$</th>
<th>$F^{3m}$</th>
<th>$F^{3m}$</th>
<th>$F^{12m}$</th>
<th>$F^{12m}$</th>
<th>$F^{12m}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>macro I</td>
<td>0</td>
<td>99</td>
<td>32</td>
<td>0</td>
<td>99</td>
<td>50</td>
<td>0</td>
<td>99</td>
<td>50</td>
</tr>
<tr>
<td>macro II</td>
<td>0</td>
<td>99</td>
<td>89</td>
<td>0</td>
<td>99</td>
<td>30</td>
<td>0</td>
<td>99</td>
<td>50</td>
</tr>
<tr>
<td>option</td>
<td>0</td>
<td>99</td>
<td>35</td>
<td>0</td>
<td>99</td>
<td>81</td>
<td>0</td>
<td>99</td>
<td>76</td>
</tr>
</tbody>
</table>
References


41
