An Anatomy of the Equity Premium*

Paul Schneider†

January 9, 2016

Abstract

This paper introduces a decomposition of the forward 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 forward market return. Signals revealed by this risk anatomy provide predictive power out-of-sample for realized forward returns in particular for longer maturities.

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 in terms of economically appealing constituents that allow interpretation as realized variance, loss aversion, and temperance to assess the question which kind of risk the return compensates for. Conditional on S&P 500 option data I find that the main component of the equity return originates from aversion

*I am grateful to Christian Julliard for raising the question whether the equity premium can be dissected into higher-order premia.

†University of Lugano and Swiss Finance Institute. Email:. paul.schneider@usi.ch

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.

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 (....literature).

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 originates from the factor associated with loss aversion, in particular for shorter maturities, in line with 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 appears to bear 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.

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 contracts. Sections 3.1 and 3.2 work out some basic theoretical properties of the transform and contain the empirical study. Section 5 concludes, Appendix A develops economic Hellinger contracts, Appendix B contains proofs for the claims made in the text, and Appendix C 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

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

To avoid keeping track of the bond price and for technical reasons I consider subsequently the *forward equity premium* which is defined as the expectation of the discrete return of a forward on the S&P 500 minus the discrete return of a forward bond

$$\mathbb{E}_t^{\mathbb{P}}\left[\frac{F_{T,T} - F_{t,T}}{F_{t,T}} - \frac{1-1}{1}\right] = \mathbb{E}_t^{\mathbb{P}}\left[\frac{F_{T,T} - F_{t,T}}{F_{t,T}}\right]. \tag{2}$$

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, and by

$$R := \frac{F_{T,T}}{F_{t,T}} \tag{3}$$

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

$$\mathbb{E}^{\mathbb{Q}}[R] = 1, \tag{4}$$

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

$$R = \frac{d\mathbb{F}}{d\mathbb{Q}}, \tag{5}$$

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

$$\frac{d\mathbb{H}}{d\mathbb{Q}} := \frac{R^{1/2}}{\mathbb{E}^{\mathbb{Q}}[R^{1/2}]}. \tag{6}$$

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 symmetry of the distribution of asset prices 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 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

$$\log R \overset{\mathbb{Q}}{\sim} N\left(-\frac{1}{2}\sigma^2(T-t), \sigma\sqrt{T-t}\right).$$

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 $u \in \mathbb{R}$

$$\mathbb{E}^{\mathbb{H}}\left[e^{\mathrm{i}u\log R}\right] = \frac{\mathbb{E}^{\mathbb{Q}}\left[e^{(1/2+\mathrm{i}u)\log R}\right]}{\mathbb{E}^{\mathbb{Q}}\left[R^{1/2}\right]} = \exp\left(-\frac{1}{2}\sigma^2 u^2 (T-t)\right), \tag{7}$$

where i denotes the imaginary unit. The characteristic function is real and even, and hence the distribution of $\log R$ is symmetric under $\mathbb{H}$. 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 $\mathbb{Q}$

$$\frac{d\overline{\mathbb{H}}}{d\mathbb{Q}} := \frac{R^{-1/2}}{\mathbb{E}^{\mathbb{Q}}\left[R^{-1/2}\right]}. \tag{8}$$

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 Carr and Lee (2009).

2.2 The Equity Return on 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 $\mathcal{L}^2_{\overline{\mathbb{H}}}$ as the set of random variables $\mathcal{X}$ on $\mathcal{D}$ with finite norm

$$\|X\|_{\mathcal{L}^2_{\overline{\mathbb{H}}}} := \sqrt{\mathbb{E}^{\overline{\mathbb{H}}}\left[X^2\right]} < \infty,\ X \in \mathcal{X}, \tag{9}$$

and corresponding expectation inner product

$$\langle X, Y \rangle_{\mathcal{L}^2_{\mathbb{H}}} := \mathbb{E}^{\overline{\mathbb{H}}}[XY] = \frac{\mathbb{E}^{\mathbb{Q}}\left[R^{-1/2}XY\right]}{\mathbb{E}^{\mathbb{Q}}\left[R^{-1/2}\right]},\ X, Y \in \mathcal{L}^2_{\mathbb{H}}. \tag{10}$$

For lighter notation I will suppress below the subscript $\mathcal{L}^2_{\mathbb{H}}$ for the norm and the inner product and just write $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$. Suppose $R = \frac{d\mathbb{F}}{d\mathbb{Q}} \in \mathcal{L}^2_{\mathbb{H}}$, $\mathcal{L}^2_{\mathbb{H}}$ separable, then there exists a countable orthonormal basis[1] $\phi = \{\phi_1, \phi_2, \ldots\}$ of $\mathcal{L}^2_{\mathbb{H}}$, 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 $\mathcal{L}^2_{\mathbb{H}}$ sense. The basis ϕ 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 ϕ_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 ϕ. The additive functional form (12) is independent of the specific basis chosen. To find an economically meaningful one, I first introduce the sequence

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

[1] In terms of the inner product orthonormality is defined as the relation

$$\langle \phi_i, \phi_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

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-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

$$\mathbb{E}^{\mathbb{H}}\left[R^u\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{H}}{d\mathbb{Q}}\frac{d\mathbb{Q}}{d\mathbb{P}}R^u\right] \propto \mathbb{E}^{\mathbb{P}}\left[R^{1/2+u-\gamma}\right]. \tag{16}$$

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

$$\mathbb{E}^{\mathbb{H}}\left[R^u\right] = \mathbb{E}^{\mathbb{H}}\left[R^{-u}\right]. \tag{17}$$

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

$$\begin{aligned} R^{1/2+u-\gamma} =& \sqrt{R} + \sqrt{R}\log R(u-\gamma) + \frac{1}{2}\sqrt{R}(\log R)^2(u-\gamma)^2 \\ &+ \frac{1}{6}\sqrt{R}(\log R)^3(u-\gamma)^3 + \cdots . \end{aligned} \tag{18}$$

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

$$0 = \mathbb{E}^{\mathbb{Q}}\left[\sqrt{R}\log R\right] = \mathbb{E}^{\mathbb{Q}}\left[\sqrt{R}(\log R)^3\right] = \cdots . \tag{19}$$

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 into aversion to 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.

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

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

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). Scaling with $\sqrt{R}$, which can be interpreted as realized variance, together with the additional delta-hedging in Eq. (15) is done for two reasons. Firstly the inner product $\left\langle R, \tilde{B}_i \right\rangle = \mathbb{E}^{\mathbb{Q}}\left[H_i(R)\right]$ yields a forward-neutral expectation of Hellinger moments, and thus economically interpretable coefficients c. Secondly, $\tilde{B}_i$ are a basis of $\mathcal{L}^2_{\overline{\mathbb{H}}}$.

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

The above proposition ensures that any function $f \in \mathcal{L}^2_{\overline{\mathbb{H}}}$ can be expressed in terms of realized and implied Hellinger moments. Below are the first few basis functions with $r := \log R$

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

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

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

such that

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

yields

$$B_0 = 1, \tag{22}$$

$$B_1 = r - \mu^{\overline{\mathbb{H}}}_{0,1}, \tag{23}$$

$$B_2 = r^2 - \mu^{\overline{\mathbb{H}}}_{0,2} - \frac{\left(\mu^{\overline{\mathbb{H}}}_{0,3} - \mu^{\overline{\mathbb{H}}}_{0,1}\mu^{\overline{\mathbb{H}}}_{0,2}\right)\left(r - \mu^{\overline{\mathbb{H}}}_{0,1}\right)}{\mu^{\overline{\mathbb{H}}}_{0,2} - \left(\mu^{\overline{\mathbb{H}}}_{0,1}\right)^2}, \tag{24}$$

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

$$\phi_i := \frac{B_i(R)}{\|B_i(R)\|}. \tag{25}$$

To emphasize the factor structure I also use below the notation

$$\begin{gathered}F_i := c_i \phi_i, \text{ so that}\\ R^{(n)} - 1 = (F_0 - 1) + F_1 + F_2 + \cdots + F_n.\end{gathered} \tag{26}$$

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 ϕ_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 $\log R$.

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 $u \in \mathbb{R}$

$$\Phi^{\mathbb{H}}(u) = \Phi^{\mathbb{H}}(-u) \Leftrightarrow \text{ market is PCS,} \tag{27}$$

$$\Phi^{\mathbb{H}}(u) < \Phi^{\mathbb{H}}(-u) \Leftrightarrow \text{ neg. dev. from PCS,} \tag{28}$$

$$\Phi^{\mathbb{H}}(u) > \Phi^{\mathbb{H}}(-u) \Leftrightarrow \text{ pos. dev. from PCS.} \tag{29}$$

The below Eqs. show order-1 and order-2 equity returns.

$$R^{(0)} = \mu_{1,0}^{\overline{\mathbb{H}}}, \text{ and} \tag{30}$$

$$R^{(1)} = \frac{\mu_{0,2}^{\overline{\mathbb{H}}}\mu_{1,0}^{\overline{\mathbb{H}}} - \mu_{0,1}^{\overline{\mathbb{H}}}\mu_{1,1}^{\overline{\mathbb{H}}}}{\mu_{0,2}^{\overline{\mathbb{H}}} - \left(\mu_{0,1}^{\overline{\mathbb{H}}}\right)^2} + \frac{\mu_{1,1}^{\overline{\mathbb{H}}} - \mu_{0,1}^{\overline{\mathbb{H}}}\mu_{1,0}^{\overline{\mathbb{H}}}}{\mu_{0,2}^{\overline{\mathbb{H}}} - \left(\mu_{0,1}^{\overline{\mathbb{H}}}\right)^2} \cdot \log R. \tag{31}$$

The expressions are collected here with respect to orders of $\log R$, rather than the basis ϕ_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 0-order equity return*

$$R^{(0)} - 1 = \mu_{1,0}^{\overline{\mathbb{H}}} - 1 < 0. \tag{32}$$

$R^{(0)}$ tends to be increasing in skew for the data and extant models,[3] and reflects a mechanical convexity correction between $R-1$ and $\log R$. 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*

$$R^{(1)} - 1 = \alpha + \beta \cdot \log R, \tag{34}$$

where α and β have the familiar linear regression interpretation

$$\beta = \frac{Cov^{\overline{\mathbb{H}}}(R, \log R)}{\mathbb{V}^{\overline{\mathbb{H}}}[\log R]}, \; \alpha = (\mu_{1,0}^{\overline{\mathbb{H}}} - 1) - \beta\mu_{0,1}^{\overline{\mathbb{H}}}, \tag{35}$$

[2] The first basis vector ϕ_1 also has a constant part as can be seen from Eq. (23).

[3] This arises from

$$\mu_{1,0}^{\overline{\mathbb{H}}} = \frac{1}{\mathbb{E}^{\mathbb{H}}\left[e^{-\log R}\right]}. \tag{33}$$

The more left-skewed the distribution under $\mathbb{H}$, the higher will be $e^{-\log R}$, and thus the lower $\mu_{1,0}^{\overline{\mathbb{H}}}$. 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.

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] The above Eq. 34 also highlights very explicitly the difference between the 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 expectations for computing the coefficients α and β fast enough. Using the $\overline{\mathbb{H}}$ 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 $\mu_{1,1}^{\overline{\mathbb{H}}} = 0$, and $Cov^{\overline{\mathbb{H}}}(R, \log R) = -\mu_{1,0}^{\overline{\mathbb{H}}} \cdot \mu_{0,1}^{\overline{\mathbb{H}}}$, 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

$$\|R\|^2 = \mathbb{E}^{\overline{\mathbb{H}}}\left[R^2\right] = \frac{\mathbb{E}^{\mathbb{Q}}\left[R^{3/2}\right]}{\mathbb{E}^{\mathbb{Q}}\left[R^{-1/2}\right]} = \sum_{i=0}^{\infty} \langle R, \phi_i \rangle^2 , \tag{36}$$

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

[4] With the actual expression being different from theirs.

greater (smaller) than 1 if the economy is positively (negatively) skewed. Likewise we can compute the residual energy from

$$\left\|R - R^{(J)}\right\|^2 = \|R\|^2 - \left\|R^{(J)}\right\|^2 = \sum_{i=J+1}^{\infty} \langle R, \phi_i \rangle^2 . \tag{37}$$

With both $\|R\|$ and $\left\|R^{(J)}\right\|$ known from option prices, this expression can be used to determine a cut-off value for J as formula (37) can be computed from option prices. Figure 1a shows the logarithm of $\left\|R - R^{(J)}\right\|^2$ 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 $\mathbb{E}^{\mathbb{Q}}[R - 1] = 0$ 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 the portfolio $R^{(J)}$ as the payoff from a trading strategy, just like entering into a forward contract on the market. 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 variance 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 2a and

2b). In the late 1999's it constitutes almost everything of the 1-year realized market return. 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, 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 (26) is the co-variation between the factors. Optimally one would want as little as possible. Figure 4 shows the trajectories of the basis functions ϕ over time. The basis functions ϕ are constructed to be orthogonal with respect to the inverse Hellinger measure $\overline{\overline{\mathbb{H}}}$, but there is no reason to expect they should be orthogonal with respect to the unconditional physical measure $\mathbb{P}$. 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 ϕ_1^{1m} and ϕ_2^{1m}. This is an indication that the $\mathbb{P}$ seems to be quite far from $\overline{\overline{\mathbb{H}}}$, but within reason from a modelling point of view. The correlations of the factors F, the product of the time-varying coefficients c_i and the basis functions ϕ_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.

The next section looks at the question whether knowing the 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 ϕ_i of decomposition (12), or directly. Benefits from the decomposition could arise from differences in persistence between different ϕ_i. With i even, for example, they could inherit persistence from realized variance. Another potential source of predictability is the dependence of ϕ_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 multiple models increase the estimation error.

The linear model (34) would suggest the possibility of a persistent factor F_2. In Eq. (34), from $a - 1 \approx 0$ and $b \approx 1$ and using R instead of $R^{(1)}$ gives

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

This payoff is related to variance,[5] which is known to be persistent. Figure 4 shows sample paths of the basis functions over time. With ϕ_0 being identically one for any maturity, the time-varying ϕ_1 in Panel 4a looks like white noise. Persistence is introduced mechanically into ϕ_1 for higher maturities through overlapping returns. The trajectories for ϕ_2 follow a similar pattern. The unpredictable nature of the basis functions over time is also reflected in their autocorrelation plots in Figure 6. The dynamics of the factors $F_i = c_i\phi_i$ show a slightly different behaviour. Figure 5 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 7a).

I take the conditional moments from Schneider and Trojani (2015a, ST) as a model for the expectation of the powers of r in ϕ_i from Eqs. (22), (23), (24) under the physical measure. They pertain to simply compounded returns, but can be expected to be close enough to the moments of log returns. In any case they are well-suited for an out-of-sample prediction, since they are based on option 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 (26) evaluated at the conditional moments from ST. The third is the conditional moment from ST itself, and the fourth is the simple variance swap ($SVIX^2$) from Martin (2015). Table 2 shows that the decomposition is more successful in predicting market forward returns out-of-sample only

[5] To be precise, it is realized entropy that also underlies the specification of the VIX^2 contract.

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 The Forward Market Return Through the Lens of a Model

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), 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 less ambitious goal in trying to understand the realized forward market return rather than its unattainable conditional expectation. The conceptual idea is to use economically interpretable basis functions in an economically relevant space to express this return. The resulting representation happens 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. The implementation is completely 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.

Appendices

A Hellinger Basis

From the divergence function

$$D_p(R) := \frac{R^p - 1 - p(R-1)}{p(p-1)} \tag{38}$$

we define the $n-th$ Hellinger moment $H_n(R)$ from

$$\frac{\partial^n D_p(R)}{\partial p^n} = \frac{\partial^n}{\partial p^n}\left(\frac{1}{p-p^2} + \frac{(R-1)}{1-p} - \frac{R^p}{p-p^2}\right) \tag{39}$$

$$= n!\left(\left(\frac{1}{1-p}\right)^{n+1} + \frac{(-1)^n}{p^{n+1}}\right) \tag{40}$$

$$+ (R-1)n!\left(\frac{1}{1-p}\right)^{n+1} \tag{41}$$

$$- R^p \sum_{k=0}^{n} \binom{n}{k} (\log R)^{n-k} k! \left(\left(\frac{1}{1-p}\right)^{k+1} + \frac{(-1)^k}{p^{k+1}}\right), \tag{42}$$

as

$$H_n(R) := \frac{\partial^n D_p(R)}{\partial p^n}|_{p=1/2}$$
$$= 2^{n+1} n! \left((-1)^n + R - \sqrt{R}\sum_{k=0}^{n} (\log R)^k \left(\frac{1}{2}\right)^k \frac{1}{k!}(1+(-1)^{n+k})\right) \tag{43}$$

We also have for $n \geq 2$ the recursion

$$\tilde{B}_n(R) = 4n(n-1)\tilde{B}_{n-2}(R) + (\log R)^n \tag{44}$$

which can be easily seen from representation (43).

B Proofs

B.1 Auxiliary Result

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

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

B.2 Proposition 2.1

Proof. If $\mathcal{D}$ is unbounded, then denseness of the polynomial canonical basis follows from the exponential tails of $\overline{\mathbb{H}}$ in r from Lemma 3.1 in Filipović et al. (2013). For compact $\mathcal{D}$ it follows from the Stone-Weierstrass theorem. Eq. (44) shows that there is a linear map between the canonical $\{1, r, r^2, r^3\}$ and the Hellinger basis $\left\{\tilde{B}_0, \tilde{B}_1, \ldots\right\}$ which implies that the Hellinger basis is dense in $\mathcal{D}$. □

B.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

$$R^{(0)} = \langle R, \phi_0\rangle \cdot \phi_0 = \langle R, 1\rangle = \mu^{\overline{\mathbb{H}}}_{1,0} = \frac{\mathbb{E}^{\mathbb{Q}}\left[R^{1/2}\right]}{\mathbb{E}^{\mathbb{Q}}\left[R^{-1/2}\right]} < 1, \tag{45}$$

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

B.4 Proof of Result 3.4

The "Hellinger CAPM" (34) is just a reformulation of Eq. (31). Positivity of β follows from Schmidt (2003) which guarantees that the covariance of two increasing functions of a random variable is positive, and positivity of the variance.

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

B.5 Proposition 3.5

We start with

Lemma B.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*

$$\left\langle \frac{d\mathbb{F}}{d\mathbb{Q}}, \tilde{B}_n(R) \right\rangle = 0 \tag{46}$$

for n odd.

Proof. Plugging in the definition (43) and computing the $\overline{\mathbb{H}}$ moments reveals for n odd

$$\left\langle \frac{d\mathbb{F}}{d\mathbb{Q}}, \tilde{B}_n(R) \right\rangle = \frac{1}{4} \cdot \frac{\mathbb{E}^{\mathbb{Q}}\left[H_n(R)\right]}{\mathbb{E}^{\mathbb{Q}}\left[R^{-1/2}\right]}, \tag{47}$$

and for n even

$$\left\langle \frac{d\mathbb{F}}{d\mathbb{Q}}, \tilde{B}_n(R) \right\rangle = \frac{1}{4} \cdot \frac{2^{n+2} n! + \mathbb{E}^{\mathbb{Q}}\left[H_n(R)\right]}{\mathbb{E}^{\mathbb{Q}}\left[R^{-1/2}\right]}, \tag{48}$$

From Schneider and Trojani (2015b) we know that $\mathbb{E}^{\mathbb{Q}}\left[H_n(R)\right] = 0$ for n odd under PCS, which proves the claim. □

Now, from

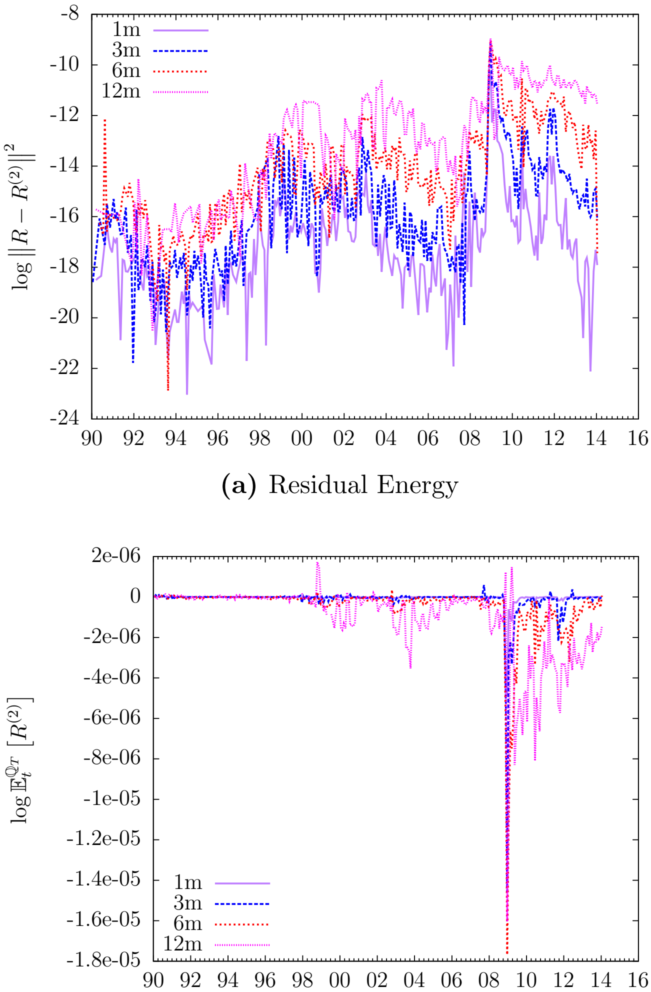
$$\mathbb{E}^{\mathbb{Q}}\left[H_n(R)\right] \overset{PCS}{=} 0, \text{ n odd }, \tag{49}$$

and the recursion Eq. (44), 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

$$\left\langle \frac{d\mathbb{F}}{d\mathbb{Q}}, B_n(R) \right\rangle \tag{50}$$

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

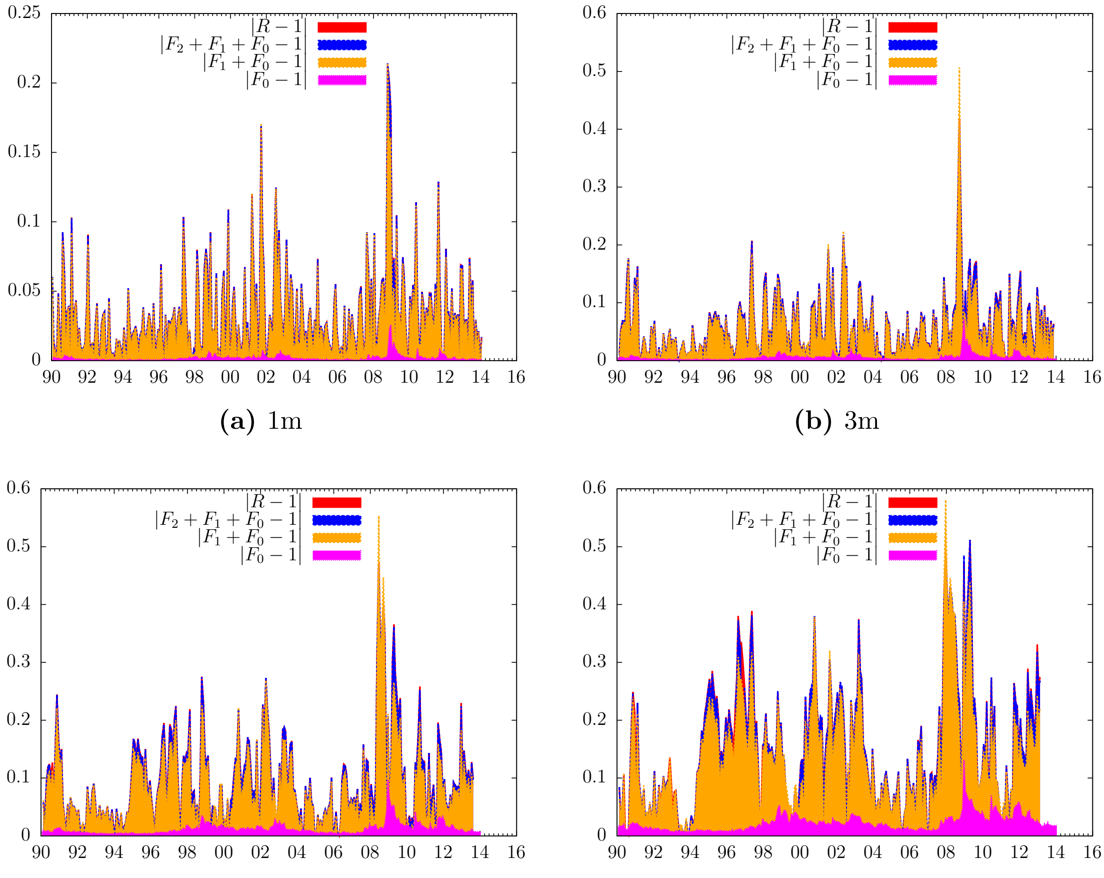
C Additional Figures and Tables



(a) Residual Energy

(b) Forward Price of Decomposition

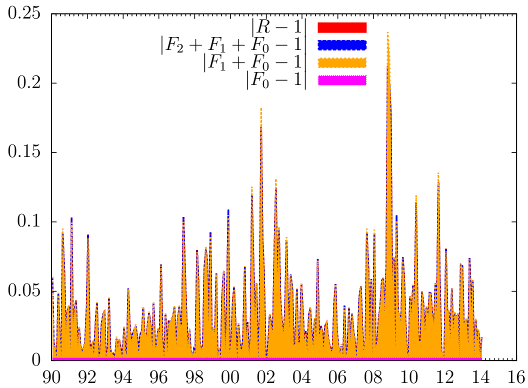
Figure 1: Residual Energy. Panel (a) shows how much information is unexplained by option data according to Formula (37) after decomposing the equity return into the first three factors from Eq. (26). The log transformation accounts for the extremely small deviation between the true signal and its transform. 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. (25). The data producing both figures are European options on the S&P 500 from 1990 to 2014.



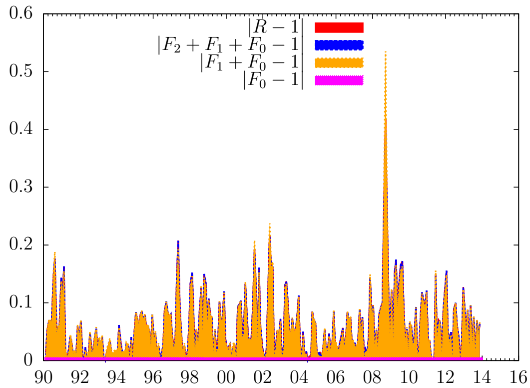
(a) 1m **(b)** 3m

(c) 6m **(d)** 12m

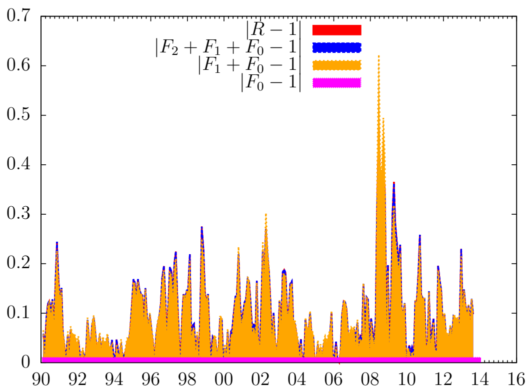
Figure 2: Equity Returns and Hellinger Decomposition. The panels show the factor decomposition (26) for different orders of the expansion. To gauge the magnitudes of the contributions they are depicted in absolute values. The factor decomposition can both over-, and undershoot the true return, but deviations 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.



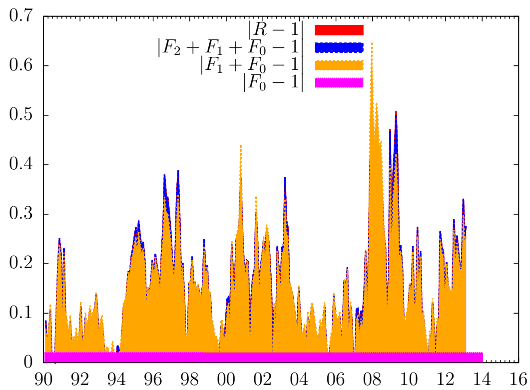
(a) 1m



(b) 3m



(c) 6m



(d) 12m

Figure 3: Equity Returns and Hellinger Decomposition in Black-Scholes Model. The panels show the factor decomposition (26) 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.

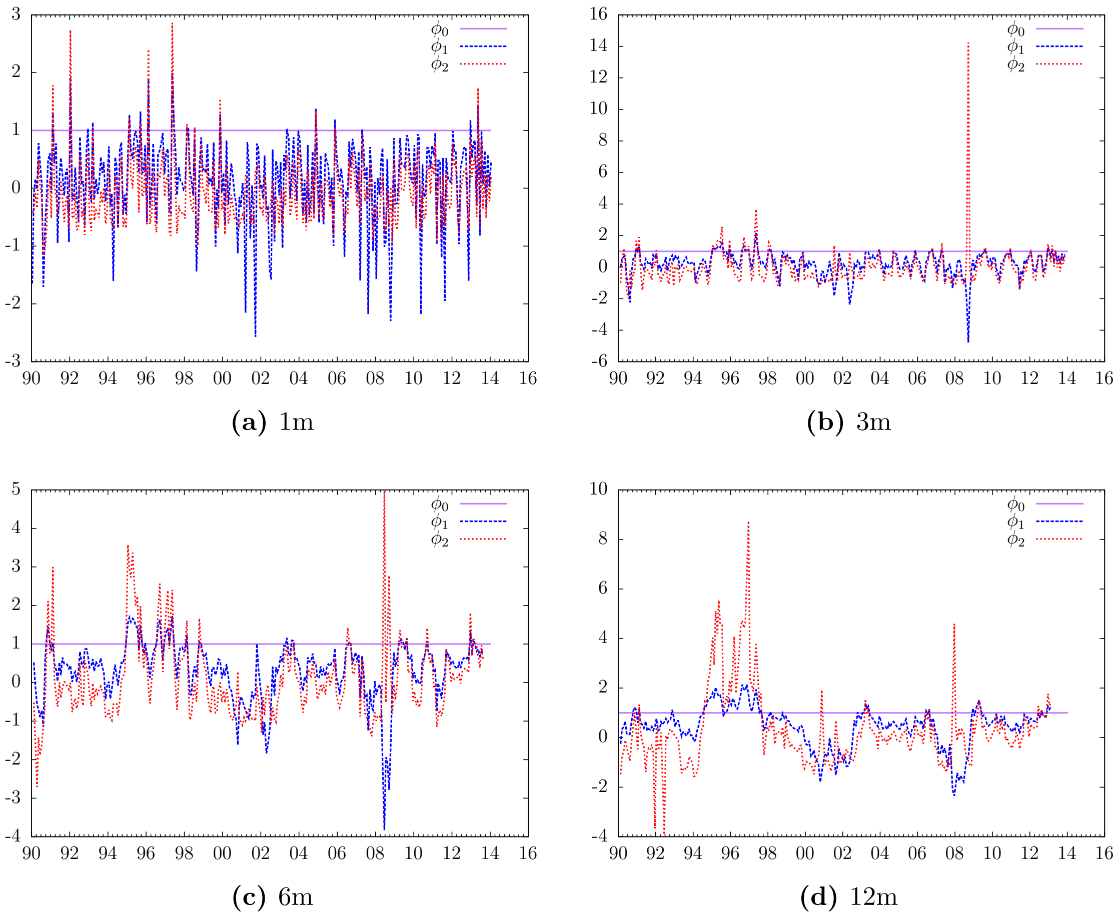
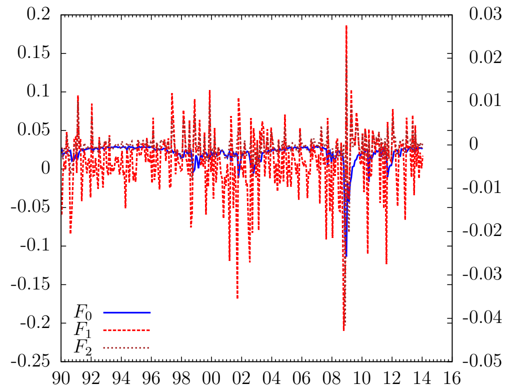
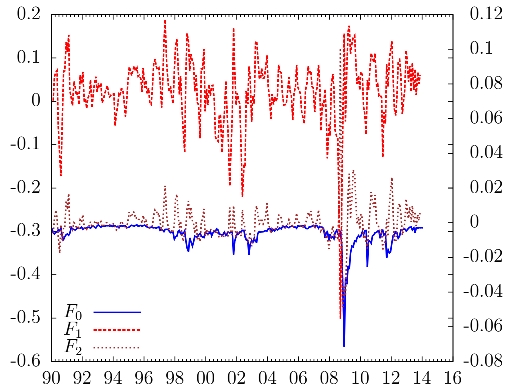


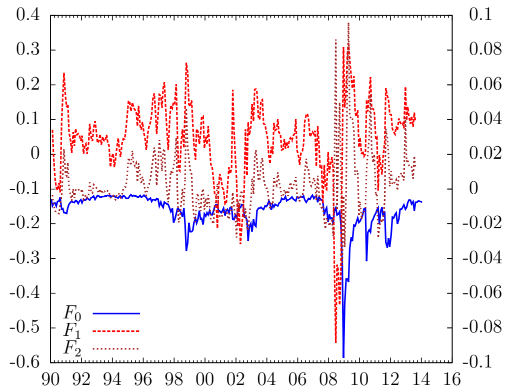
Figure 4: Time-Varying Basis Functions. The panels show basis functions ϕ_i from Eq. (25) for orders $i = 0, 1$, and 2. The data producing the panels are European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, 6, and 12 months.



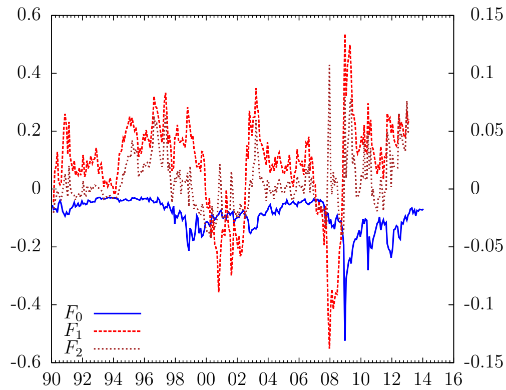
(a) 1m



(b) 3m

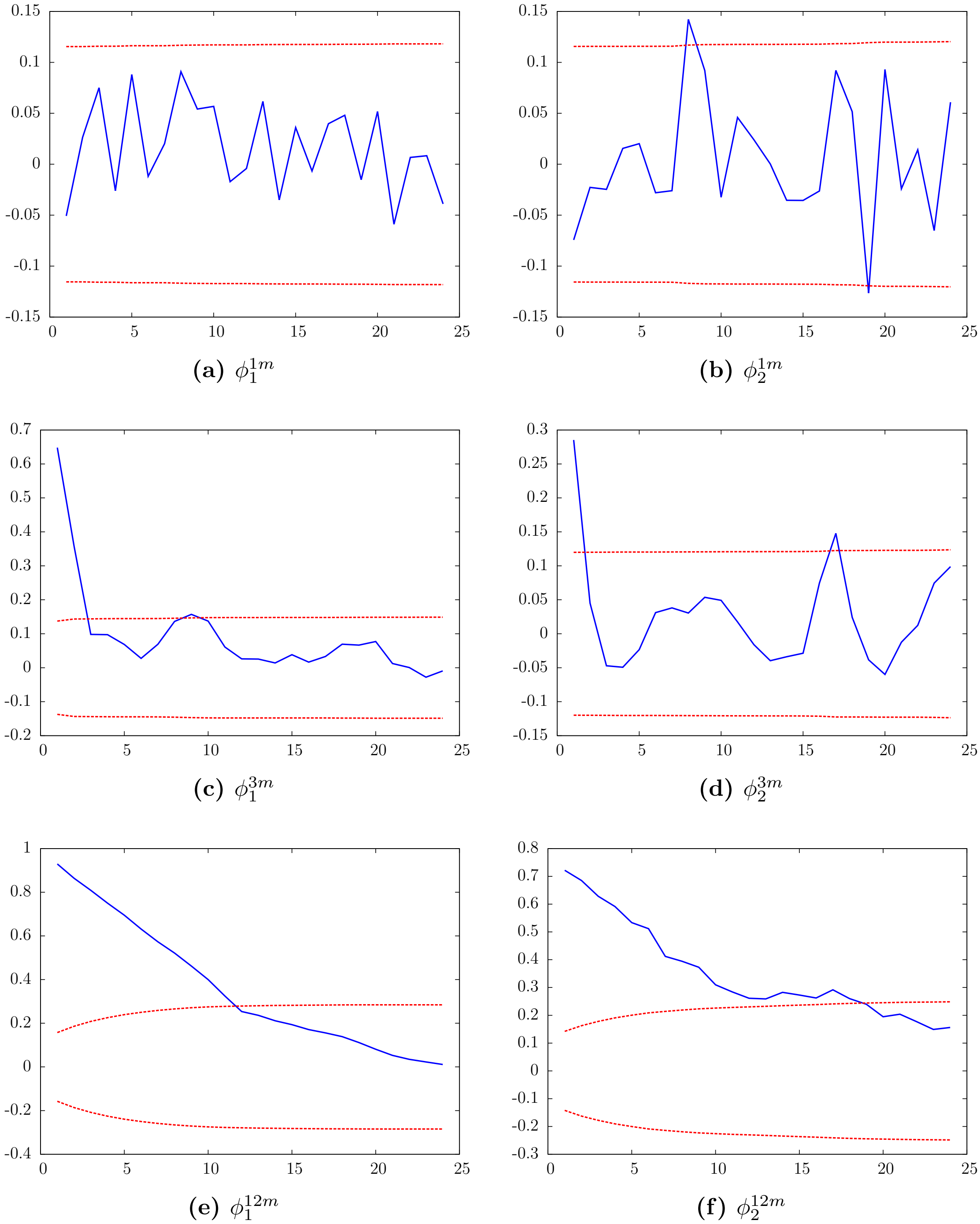


(c) 6m



(d) 12m

Figure 5: Factors. The panels show factors F_i from Eq. (26) for orders $i = 0, 1$, 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.



(a) ϕ_1^{1m} (b) ϕ_2^{1m} (c) ϕ_1^{3m} (d) ϕ_2^{3m} (e) ϕ_1^{12m} (f) ϕ_2^{12m}

Figure 6: Basis Function Autocorrelations. The panels show autocorrelations of basis functions ϕ_i from Eq. (25) 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.

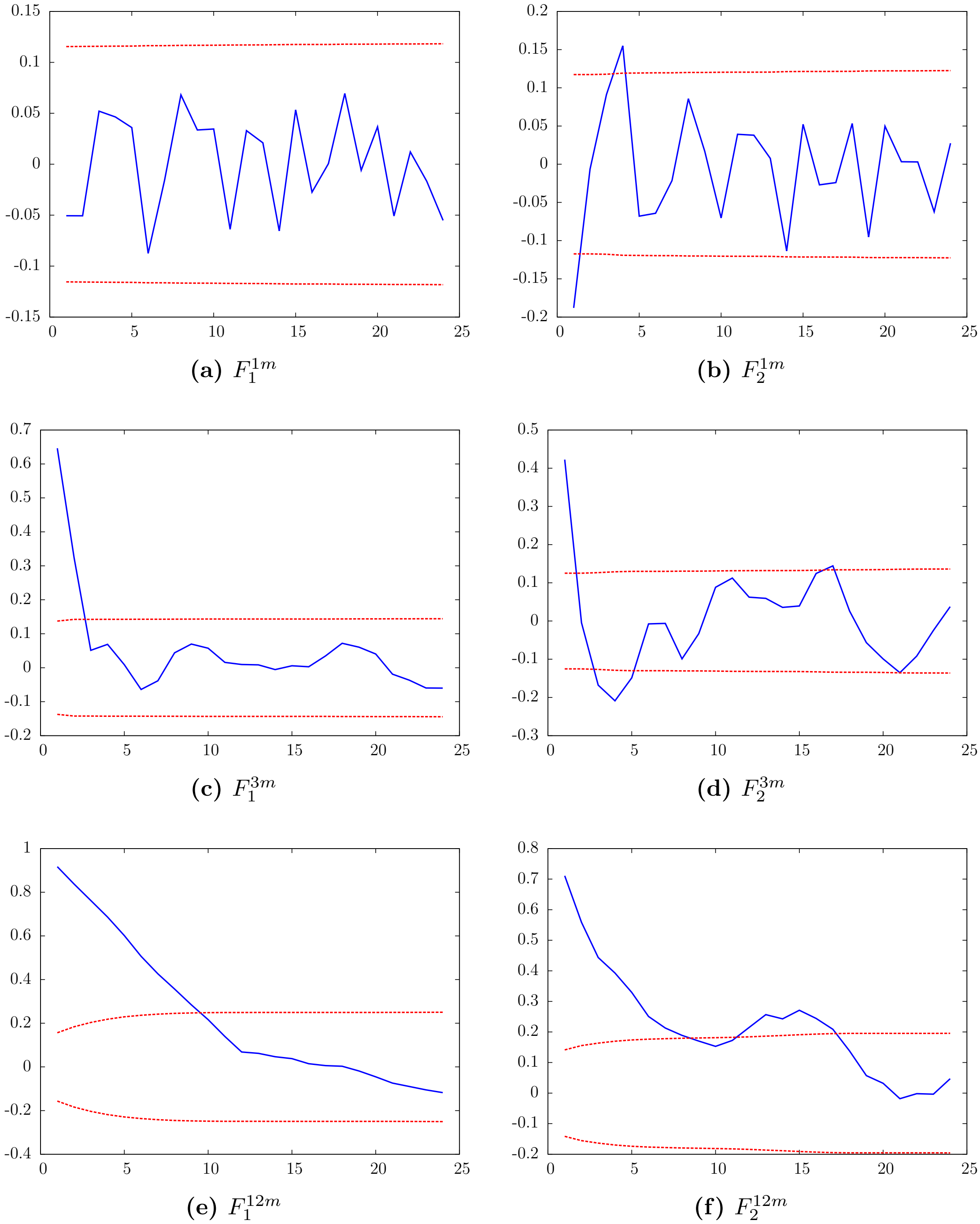


Figure 7: Factor Autocorrelations. The panels show autocorrelations of factors F_i from Eq. (25) for orders $i = 1$, and 2 (solid blue). The dotted red lines represent 95% confidence intervals. The data producing the panels are European options on the S&P 500 from 1990 to 2014 with maturities 1, 3, and 12 months.

	ϕ_2^{1m}	ϕ_1^{3m}	ϕ_2^{3m}	ϕ_1^{6m}	ϕ_2^{6m}	ϕ_1^{12m}	ϕ_2^{12m}
ϕ_1^{1m}	76	54	24	45	29	36	24
ϕ_2^{1m}		44	30	32	30	27	21
ϕ_1^{3m}			14	74	37	56	38
ϕ_2^{3m}				16	52	24	25
ϕ_1^{6m}					48	74	47
ϕ_2^{6m}						47	57
ϕ_1^{12m}							55

(a) Basis Function Correlations

	F_1^{1m}	F_2^{1m}	F_0^{3m}	F_1^{3m}	F_2^{3m}	F_0^{6m}	F_1^{6m}	F_2^{6m}	F_0^{12m}	F_1^{12m}	F_2^{12m}
F_0^{1m}	-12	24	97	-18	26	88	-23	-2	78	-25	-14
F_1^{1m}		62	-17	57	25	-19	47	28	-18	34	29
F_2^{1m}			14	28	35	7.7	19	27	1.6	9.9	19
F_0^{3m}				-21	21	95	-28	-8.4	87	-30	-21
F_1^{3m}					18	-23	76	34	-21	57	42
F_2^{3m}						13	14	59	5.9	16	13
F_0^{6m}							-30	-16	95	-31	-23
F_1^{6m}								42	-27	75	52
F_2^{6m}									-17	43	46
F_0^{12m}										-26	-20
F_1^{12m}											54

(b) Factor Correlations

Table 1: Correlations. Table (a) shows correlations (in %) of the basis functions ϕ from Eq. (25). The basis functions are uncorrelated by construction under $\overline{\mathbb{H}}$, 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. (26) in the main text. Both tables are computed with SPX options in the years 1990-2014.

	1m	3m	6m	12m
Decomp. vs Average	0.003	-0.004	0.042	0.032
Decomp. vs Recov.	-0.003	-0.007	0.014	0.007
Decomp. vs $SVIX^2$	-0.003	-0.007	0.014	0.007

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

$$R^2 := 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_{1,i})^2}{\sum_{i=1}^{n}(y_i - \hat{y}_{2,i})^2}, \tag{51}$$

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 we take as a base model the sample average (model 2) and evaluate the decomposition $\mathbb{E}^{\mathbb{P}}_{i-1}\left[R^{(2)}\right]$ at the conditional moments from Schneider and Trojani (2015a, ST) (model 1). In the second row with use the conditional moments from ST as a baseline model (model 2) and again test it against the decomposition evaluated at the moments from ST. The data are European options written on the S&P 500 from January 1990 to January 2014.

References

Benartzi, S. and R. H. Thaler (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *The Quarterly Journal of Economics*, 110, 73–92.

Borovička, J., L. P. Hansen, and J. Scheinkman (2014): "Misspecified Recovery," Working paper, New York University, University of Chicago, Princeton University and NBER.

Buss, A. and G. Vilkov (2012): "Measuring Equity Risk with Option-implied Correlations," *Review of Financial Studies*, 25, 3113–3140.

Campbell, J. Y. and S. B. Thompson (2008): "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies*, 21, 1509–1531.

Carr, P. and R. Lee (2009): "PUT-CALL SYMMETRY: EXTENSIONS AND APPLICATIONS," *Mathematical Finance*, 19, 523–560.

Carr, P. and J. Yu (2012): "Risk, Return, and Ross Recovery," *Journal of Derivatives*, 20, 38–59.

Chen, H. and S. Joslin (2012): "Generalized Transform Analysis of Affine Processes and Applications in Finance," *Review of Financial Studies*, 25, 2225–2256.

Christensen, T. (2014): "Nonparametric Stochastic Discount Factor Decomposition," Working paper, New York University.

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

Dangl, T. and M. Halling (2012): "Predictive regressions with time-varying coefficients," *Journal of Financial Economics*, 106, 157 – 181.

Duffie, D., J. Pan, and K. Singleton (2000): "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343–1376.

Fama, E. F. and K. R. French (2002): "The Equity Premium," *The Journal of Finance*, 57, 637–659.

Filipović, D., E. Mayerhofer, and P. Schneider (2013): "Transition Density Approximations for Multivariate Affine Jump Diffusion Processes," *Journal of Econometrics*, 176, 93–111.

Giglio, S. and I. Dew-Decker (2015): "Asset pricing in the frequency domain: theory and empirics," Working paper, University of Chicago and Northwestern University.

Harvey, C. and A. Siddique (2000): "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, 55, 1263–1295.

Julliard, C. and A. Ghosh (2012): "Can Rare Events Explain the Equity Premium Puzzle?" *Review of Financial Studies*, 25, 3037–3076.

Kraus, A. and R. H. Litzenberger (1976): "SKEWNESS PREFERENCE AND THE VALUATION OF RISK ASSETS," *The Journal of Finance*, 31, 1085–1100.

Linetzky, V. and L. Qin (2015a): "Long Term Risk: A Martingale Approach," Working paper, Northwestern University.

——— (2015b): "Positive Eigenfunctions of Markovian Pricing Operators: Hansen-Scheinkman Factorization, Ross Recovery and Long-Term Pricing," Working paper, Northwestern University.

Martin, I. (2013): "The Lucas Orchard," *Econometrica*, 81, 55–111.

——— (2015): "What is the Expected Return on the Market?" Working paper, London School of Economics and Political Science.

Mehra, R. and E. C. Prescott (1985): "The equity premium: A puzzle," *Journal of Monetary Economics*, 15, 145 – 161.

Pettenuzzo, D., A. Timmermann, and R. Valkanov (2014): "Forecasting stock returns under economic constraints," *Journal of Financial Economics*, 114, 517 – 553.

Rapach, D., J. Strauss, and G. Zhou (2010): "Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy," *Review of Financial Studies*, 23, 821–862.

Ross, S. (2005): *Neoclassical Finance*, Princeton Lectures in Finance, Princeton University Press.

——— (2015): "The Recovery Theorem," *Journal of Finance*, 70, 615–648.

Schmidt, K. (2003): "On the Covariance of Monotone Functions of a Random Variable," Working paper, Technische Universität Dresden.

Schneider, P. and F. Trojani (2015a): "(Almost) Model-Free Recovery," Working paper, University of Lugano, Boston University, University of Geneva, and Swiss Finance Institute.

——— (2015b): "Divergence and the Price of Uncertainty," Working paper, University of Lugano, Boston University, University of Geneva, and Swiss Finance Institute.

Weil, P. (1989): "The equity premium puzzle and the risk-free rate puzzle," *Journal of Monetary Economics*, 24, 401 – 421.

Welch, I. and A. Goyal (2008): "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21, 1455–1508.