Modeling Conditional Factor Risk Premia Implied by Index Option Returns*

Mathieu Fournier — HEC Montréal
Kris Jacobs — University of Houston
Piotr Orłowski — HEC Montréal

June 3, 2021

Abstract

We propose a novel factor model for option returns. Option exposures are modeled nonparametrically and factor risk premia may vary non-linearly with states. The model allows for estimation of factor risk premia and factor exposures using regressions, with minimal assumptions on the dynamics of factors and/or option returns. We implement the model using index option returns. The model explains expected option returns across moneyness and maturities, and its hedging performance is impressive. We obtain estimates of the average risk premia on the market index, the market variance, as well as factors associated with tail risk and intermediary risk, and we characterize the time variation in these risk premia. The signs of the average risk premia are consistent with economic intuition for all factors and risk premia spike during crises. The magnitudes of the risk premia on the market and variance factors are reasonable and they have the expected sign throughout the sample.

JEL Classification: G10; G12; G13.

Keywords: Factor models; option-implied risk premia; time-varying exposures; machine learning.

*For helpful comments we thank Daniel Andrei, Patrick Augustin, Sebastien Betermier, Mark Carhart, Laurent Barras, Giorgio De Santis, Christian Dorion, Patrick Gagliardini, Christian Gourieroux, Alexandre Jeanneret, Gerardo Manzo, Dmitriy Muravyev, Neil Pearson, Guillaume Roussellet, Valeri Sokolovski, and seminar participants at McGill University, Kepos Capital, HEC Montréal, and CREST. The authors gratefully acknowledge financial support from Bank of Canada, CDI (formally IFSID). Correspondence to: Mathieu Fournier, HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, QC H3T 2A7. E-mail: mathieu.fournier@hec.ca.

1 Introduction

Because of the state-contingent nature of their payoffs, options are highly informative about state prices and the price of risk. Traditionally, the index and stock option literature has studied option prices. This contrasts with the literature on the underlying securities, which is almost exclusively based on returns. Moreover, while much of the study of stock returns is focused on factor models that analyze the market-wide risks determining these returns, the literature on option returns is more focused on option characteristics. While the literature on factor models of option returns has been growing, it is still very limited and we do not know much about the properties of risk premia implied by option returns.[1]

A potential explanation for the limited progress on this topic is that existing methodologies are developed to estimate risk premia from stock returns and may not be suitable for option returns. Option prices, returns, and exposures exhibit pronounced non-linear variation across time and option characteristics, which is hard to capture using standard methods for stock returns. This paper proposes a factor model specifically designed to estimate conditional risk premia and option exposures based on large panels of option returns. The framework allows for multiple factors and is sufficiently flexible to accommodate any factor of interest. It is able to provide reliable estimates of the relative importance of various sources of systematic risk in determining the expected return on options.

The traditional approach to estimating risk premiums implied by options adopts parametric models, typically featuring jump, volatility, and tail risks.[2] These studies provide important insights about equity, variance, jump, and tail risk premia. However, their parametric structure constrains the number and type of factors that can be simultaneously accommodated. These models are therefore silent on the exposure of options beyond return, variance, and jump risks, even though the literature suggests the pres-

[1]See Jones (2006), Israelov and Kelly (2017), and Horenstein, Vasquez, and Xiao (2018) for notable exceptions.

[2]For seminal contributions to this literature, see, among others, Pan (2002), Eraker, Johannes, and Polson (2003), Carr and Wu (2009), Bates (2008), Broadie, Chernov, and Johannes (2009), Todorov (2010), Bollerslev, Todorov, and Xu (2015), and Andersen, Fusari, and Todorov (2015a,b).

ence of other risks, such as liquidity and intermediary risks.[3] Moreover, the structure of the factor risk premia in parametric models is very restrictive.

The option pricing literature is also well aware of existing problems with the measurement of option exposures to risk factors, not only over time but also across moneyness and maturity.[4] Duarte, Jones, and Wang (2019) show that measurement errors in option exposures in a Black-Scholes economy induce large biases in the estimation of expected option returns, which in turn contaminate risk premium estimates. Israelov and Kelly (2017) document that state-of-the-art affine models also produce counterfactual predictions about expected option returns. Developing models that are better suited to the study of option returns is thus of paramount interest.

We model option returns' factor exposures using nonparametric functions of instruments such as moneyness, maturity, and conditional volatility. This nonparametric approach allows for rich patterns in model exposures, not only across maturity and moneyness, but also over time. This approach is supported by a growing literature showing that non-linear stochastic discount factors help explain the cross-section of expected stock and option returns, as well as another literature demonstrating that non-affine option pricing models provide a better fit to the cross-section of option prices.[5]

To estimate the model, we develop a two-step procedure. In a first step, it jointly estimates option return exposures to all factors. Taking these exposures as given, we then estimate the conditional risk premia associated with the factors. We illustrate the value of our approach using more than 22 years of daily data on index option returns with various maturity and moneyness. The resulting option return panel has more than one million observations, but we find that the estimation of exposures and risk premia is fast, which

[3] Christoffersen, Goyenko, Jacobs, and Karoui (2018b) show that illiquidity impacts the expected returns on delta-hedged equity options. Barras and Malkhozov (2016), Fan, Imerman, and Dai (2016), and Gârleanu, Pedersen, and Poteshman (2009) study the impact of demand shocks on options.

[4] This issue is more critical for option returns, but also poses a problem when studying the cross-section of stock returns. For discussions, see Shanken (1990), Jagannathan and Wang (1996), Ghysels (1998), Lustig and van Nieuwerburgh (2005), Lewellen and Nagel (2006), Nagel and Singleton (2011), Buss and Vilkov (2012), and Bollerslev, Li, and Todorov (2016).

[5] On non-linear stochastic discount factors, see, among others, Dittmar (2002), Aït-Sahalia and Duarte (2003), Bakshi, Madan, and Panayotov (2010), Chabi-Yo (2012), and Schneider and Trojani (2019). On non-affine models, see for instance Jones (2003), who develops a non-affine stochastic volatility model to study index option prices, and Eraker and Wang (2015), who develop a non-linear model of the variance premium.

proves that our framework is practical for large-scale empirical exercises. Our main model contains three factors: the market return and changes in its variance, as well as gamma (the squared market return). We use an affine risk premia specification as the benchmark. We find that while a non-affine specification generates more variation in conditional risk premia, particularly during periods of high uncertainty, an affine one also performs well and generates substantial variation in risk premiums. Moreover, it is more successful in delivering intuitively plausible signs for the risk premia. To instrument option exposures and factor risk premia, we use the conditional market variance, which we proxy using the squared VIX and the square of the Tail Volatility index of Bollerslev et al. (2015). We also estimate models with additional factors: the intermediary capital factor of He, Kelly, and Manela (2017) and the tail factor in Bollerslev et al. (2015).

Overall, the model performs well. Its explanatory power in the cross-section of options with different moneyness and maturity varies between 92% and 100%. This fit is impressive because it is based on daily returns, whereas most existing studies use less noisy monthly returns. One of our main empirical contributions is a characterization of the relative contribution of various economic factors in explaining the cross-section of option returns. In the benchmark model, market return risk is the main driver of index option return risk and explains between 47% and 73% of index option return total variance. Variance risk is also an important driver, with R-squareds between 19% and 54%. The contribution of gamma risk to overall option return variation is small compared to these two factors. The contribution of market return risk and market variance risk continues to dominate when we add the intermediary and tail factors to the model.

The impact of factors strongly differs across the cross-section of options. SPX puts are much more sensitive to market variance risk than calls, and calls are more sensitive to market return risk. As expected, the tail factor is mainly important for OTM puts. The intermediary factor on the other hand is more important for calls than for puts. These results suggest that the index option market seems to be somewhat segmented, consistent with Constantinides, Jackwerth, and Perrakis (2009) and Bakshi, Crosby, and Gao (2019), who show that index call options are priced differently from index put options.

We provide estimates of conditional and unconditional factor exposures. Unconditionally, fitted exposures to option pricing factors have the correct sign, their magnitude is plausible, and they behave as expected across moneyness. The term-structure of exposures to variance risk implied by index option returns contrasts in some dimensions with the predictions of standard affine models. To illustrate the value of the estimated conditional exposures, we compare the hedging performance of our model with standard benchmarks. An investor who uses our model instead of Black and Scholes (1973) or Heston (1993) realizes considerable hedging gains. This is important for financial institutions and traders who must continuously manage the risk exposures of their positions and strategies.

The model also provides estimates of conditional factor risk premiums. In the benchmark model, the signs of the average risk premia are intuitively plausible for all three factors: positive for the market premium and negative for the variance and gamma premiums. The average market risk premium is 9.1%, close to the sample average. The average of the variance risk premium is −22.1%. Moreover, the model-implied equity and variance risk premia have the expected positive and negative sign every day. Their variation over time seems plausible and they peak at the height of the financial crisis in 2008.

The signs and magnitudes of the market and variance risk premiums are robust when including additional factors in the model. The intermediary factor based on He et al. (2017) carries a positive risk premium of 17.3%, consistent with intuition and similar in magnitude to the estimate in He et al. (2017). However, the percentage of variation in option returns explained by this factor is modest compared to the role played by the market and variance factors. The same is true for the tail factor from Bollerslev et al. (2015). As expected, it carries a negative risk premium, but its role in explaining option returns is much smaller than that of the market and variance factors.

restrictions. A critical difference is that our approach is based on option returns rather than prices. Existing studies typically specify the physical and risk-neutral dynamics of the risk factors and then estimate the model by combining option prices with the underlying stock return data, because the underlying security returns are required to identify the physical dynamics and risk premium parameters. Unlike option prices, option returns embed information about both physical and risk-neutral probability measures, which allows us to identify conditional risk premia using option information only. Moreover, in the existing literature risk premiums are typically estimated for conditional moments of returns, whereas we can estimate risk premiums for other factors too.

Our model and empirical results are related to various strands of the literature on options and empirical asset pricing. Gagliardini, Ossola, and Scaillet (2016) develop a factor model with linear specification of exposures and risk premia to study the risks priced in large panels of individual stock returns. Kelly, Pruitt, and Su (2018) develop a factor model of stock returns with principal component factors and instrumented exposures. We adapt the panel framework of Gagliardini et al. (2016) to option returns while allowing for instrumented option exposures in the spirit of Kelly et al. (2018).

Our work is also closely related to a burgeoning literature that uses principal component and non-linear factor decomposition methods for modeling option returns. Jones (2006) develops a factor model of option returns with factor innovations that are non-linear transformations of economic variables. Israelov and Kelly (2017) model index option returns using a principal components approach applied to implied volatilities. Horenstein et al. (2018) study the factor structure in equity option returns based on the principal components of delta-hedged equity option returns. These approaches provide an impressive explanatory power for option returns, but they are not able to identify the underlying determinants of option returns. We complement these studies by providing a dynamic factor model of option returns which allows us to measure the relative importance of various economic factors for the cross-section of option returns.

The paper proceeds as follows. Section 2 provides the theoretical framework. Section 3 discusses the estimation strategy. Section 4 discusses the data, Section 5 presents the empirical results for the benchmark model, and Section 6 presents robustness results. Finally, Section 7 concludes.

2 The Model

We first present the model. We then discuss the relationship between our approach and affine option pricing models, using the Heston (1993) stochastic volatility model as an example. This exercise clarifies the similarities and differences between our framework and existing approaches to the analysis of option prices.

2.1 A Dynamic Factor Model of Option Returns

Our objective is to estimate option exposures to risk factors and factor risk premia while imposing as few parametric assumptions as possible. We denote by O_t the price of an option written on the underlying asset S_t observed at time t with a given exercise price, K_O, expiration date, τ_O, and payoff function (i.e., call or put). We focus primarily on the modeling of deleveraged option returns, however, our framework can just as easily be applied to raw option returns. The fundamental consequence of deleveraging, or lack thereof, is the change in the relative importance of returns on options with different characteristics in the estimation of the model. We choose to deleverage the returns because it brings the average returns on all options to a similar magnitude, and hence gives them similar weights in the estimation process. Another advantage of working with deleveraged returns is that it leads to easy to understand specification of normalized sensitivities of options to underlying risk factors. To deleverage option returns without relying on parametric assumptions about option deltas, we scale option returns by the ratio of the option price to the price of the underlying asset. The deleveraged return on option O from t to $t + \Delta t$ is thus defined as:

$$r^O_{t+\Delta t} \equiv \left(\frac{\Delta O_{t+\Delta t}}{O_t} \right) \cdot \frac{O_t}{S_t}, \tag{1}$$

where $\Delta O_{t+\Delta t} \equiv O_{t+\Delta t} - O_t$ denotes the option price increment over the period Δt.

We assume that $r^O_{t+\Delta t}$ can be characterized by a dynamic factor structure of the form

$$r^O_{t+\Delta t} = \gamma^O_t + \sum_l \beta^O_{l,t} \cdot f_{l,t+\Delta t} + \varepsilon^O_{t+\Delta t}, \tag{2}$$

where $f_{l,t+\Delta t}$ denotes the factor l's realization between t and $t+\Delta t$, M is the number of factors, and γ_t^O and $\beta_{l,t}^O$ are the option's conditional intercept and conditional loading on factor l, respectively.[6] The model residual $\varepsilon_{t+\Delta t}^O$ satisfies $E_t^P\left[\varepsilon_{t+\Delta t}^O\right]=0$, where $E_t^P\left[\cdot\right]$ denotes the conditional expectation under the physical probability measure. We do not explicitly model heteroskedasticity and autocorrelation in the residuals, but we account for their presence when conducting inference about the estimated parameters.

To provide more intuition for our approach, consider an economy in which deleveraged option returns are consistent with the Heston (1993)'s model and $\Delta t \to 0^+$. In this case, deleveraged option returns follow a two-factor structure with the return on the asset, $f_{s,t+\Delta t}=\frac{\Delta S_{t+\Delta t}}{S_t}$, and changes in asset's variance, $f_{v,t+\Delta t}=\Delta v_{t+\Delta t}$, as risk factors.

To allow for rich dynamics of the option exposures, we model the $\beta_{l,t}^O$ as non-parametric functions of option characteristics (i.e., time to maturity and moneyness) and instruments. Option exposures to a given risk factor thus vary non-linearly across maturity and moneyness. They also vary over time as the instruments fluctuate. We discuss the specification of exposures in more detail in Section 2.2 below.

According to the first fundamental theorem of asset pricing, absence of arbitrage opportunities guarantees the existence of a stochastic discount factor (SDF), ξ_t, and a risk-neutral probability measure, Q, such that the price of any option, O_t, is given by

$$O_t = E_t^P\left[\frac{\xi_{t+\Delta t}}{\xi_t}O_{t+\Delta t}\right] = E_t^Q\left[\frac{O_{t+\Delta t}}{1+r_t}\right]. \tag{3}$$

In the previous expression, $O_{t+\Delta t}$ is the option value at time $t+\Delta t$, r_t is time-t conditional risk-free rate for Δt period of time, and $E_t^Q\left[\cdot\right]$ is the conditional expectation under the risk-neutral measure. The no-arbitrage condition is imposed in our model by setting $\gamma_t^O = -\sum_l \beta_{l,t}^O E_t^Q[f_{l,t+\Delta t}]$ in Equation 2. Equation (3) in turn implies that the risk premium on the deleveraged option return is given by

$$E_t^P\left[r_{t+\Delta t}^O\right] - E_t^Q\left[r_{t+\Delta t}^O\right] = -Cov_t\left(\frac{\xi_{t+\Delta t}}{\xi_t}\cdot(1+r_t)\,;r_{t+\Delta t}^O\right). \tag{4}$$

[6]To apply the framework to raw option returns, simply multiply the deleveraged return (1) by S_t/O_t and note that this changes the functional form of both the conditional intercept, and the conditional factor loading.

Conditioning on the factors, we assume that $\varepsilon^O_{t+\Delta t}$ in equation (2) is not priced and satisfies $E^Q_t\left[\varepsilon^O_{t+\Delta t}\right] = 0$.[7] The factor structure in equation (2), equation (4), and $E^Q_t\left[\varepsilon^O_{t+\Delta t}\right] = 0$ together imply that the conditional risk premium of deleveraged option returns satisfies

$$E^P_t\left[r^O_{t+\Delta t}\right] - E^Q_t\left[r^O_{t+\Delta t}\right] = \sum_l \beta^O_{l,t} \cdot \left(E^P_t\left[f_{l,t+\Delta t}\right] - E^Q_t\left[f_{l,t+\Delta t}\right]\right). \tag{5}$$

The conditional risk premium on option O's deleveraged return is the sum of each conditional exposure multiplied by the factor's conditional risk premium.

Building on Gagliardini et al. (2016), we allow the physical and risk-neutral conditional expectation of factors to be time-varying. For a given measure m, the conditional expectation of $f_{l,t+\Delta t}$ is a linear transformation of the q by 1 vector $\boldsymbol{g}_{l,t}$ of instruments.[8] This gives:

$$E^m_t\left[f_{l,t+\Delta t}\right] \equiv \boldsymbol{\mu}^{m\prime}_l \boldsymbol{g}_{l,t}, \tag{6}$$

for $m = P, Q$ and where $\boldsymbol{\mu}^m_l$ is a q by 1 vector of conditional expectation parameters.[9] Equation (6) in turn implies that factor risk premia are given by

$$E^P_t\left[f_{l,t+\Delta t}\right] - E^Q_t\left[f_{l,t+\Delta t}\right] = \left(\boldsymbol{\mu}^P_l - \boldsymbol{\mu}^Q_l\right)' \boldsymbol{g}_{l,t} \equiv \boldsymbol{\lambda}'_l \boldsymbol{g}_{l,t}, \tag{7}$$

where $\boldsymbol{\lambda}_l = \boldsymbol{\mu}^P_l - \boldsymbol{\mu}^Q_l$ denotes the vector of risk premium parameters. The conditional expectation and risk premium specifications (6) and (7) are quite general and nest standard asset pricing models.

When the change in the underlying asset variance is the factor of interest and $f_{v,t+\Delta t} = \Delta v_{t+\Delta t}$, standard affine models imply that the factor's conditional expectations (6) and conditional variance risk premium (7) are affine in the level of variance. In those models, the variance risk premium is given by $E^P_t\left[\Delta v_{t+\Delta t}\right] - E^Q_t\left[\Delta v_{t+\Delta t}\right] = \lambda_v v_t \Delta t$, where λ_v

[7] If $Cov_t\left(\frac{\xi_{t+\Delta t}}{\xi_t} \cdot (1 + r_t); \varepsilon^O_{t+\Delta t}\right) = 0$ (i.e., $\varepsilon^O_{t+\Delta t}$ is not priced) then we have $E^Q_t\left[\varepsilon^O_{t+\Delta t}\right] = E^P_t\left[\varepsilon^O_{t+\Delta t}\right] = 0$ in the absence of arbitrage opportunities.

[8] Throughout the text, we use bold characters or letters to differentiate vectors or matrices from scalars.

[9] When required, we use the notation $\boldsymbol{x}'$ to indicate that we are taking the transpose of the vector or a matrix $\boldsymbol{x}$.

is a scalar and v_t is the underlying asset's return spot variance.[10] More generally, our framework can accommodate non-affine specifications. For instance, when $\boldsymbol{g}_{v,t}$ contains non-linear transformations of v_t, the model's conditional P- and Q-expectations and risk premium dynamics for $f_{v,t+\Delta t} = \Delta v_{t+\Delta t}$ become non-linear in v_t.

To estimate option exposures and factor risk premium dynamics, we exploit the time-series and cross-section of deleveraged excess option returns, $R^O_{t+\Delta t} \equiv r^O_{t+\Delta t} - E^Q_t\left[r^O_{t+\Delta t}\right]$. Empirically, deleveraged excess option returns correspond to the difference between deleveraged option returns and the leverage-adjusted risk-free rate, $\frac{O_t}{S_t} r_t$. We thus have $R^O_{t+\Delta t} = r^O_{t+\Delta t} - \frac{O_t}{S_t} r_t$ given that $E^Q_t\left[r^O_{t+\Delta t}\right] = \frac{O_t}{S_t} E^Q_t\left[\frac{\Delta O_{t+\Delta t}}{O_t}\right] = \frac{O_t}{S_t} r_t$ under the risk-neutral measure. To derive the model prediction for $R^O_{t+\Delta t}$, we first take the risk-neutral expectation of (2) and subtract the result obtained from (2). We get

$$R^O_{t+\Delta t} = \sum_l \left[-E^Q_t\left[f_{l,t+\Delta t}\right] \cdot \beta^O_{l,t} + f_{l,t+\Delta t} \cdot \beta^O_{l,t}\right] + \varepsilon^O_{t+\Delta t} \quad (8)$$

$$= \sum_l \left[-\boldsymbol{\mu}_l^{Q'} \boldsymbol{g}_{l,t} \cdot \beta^O_{l,t} + f_{l,t+\Delta t} \cdot \beta^O_{l,t}\right] + \varepsilon^O_{t+\Delta t}, \quad (9)$$

where we replace $E^Q_t\left[f_{l,t+\Delta t}\right]$ with $\boldsymbol{\mu}_l^{Q'} \boldsymbol{g}_{l,t}$ given (6) to obtain the second equation. Equation 9 embeds the no-arbitrage condition and states that the conditional intercept of the excess option return equation is equal to the sum of Q-expectations of risk factors multiplied by factor sensitivities.

2.2 Option Return Exposures

Regardless of their distributional assumptions, option pricing models all specify that option exposures to risk factors are non-linear functions of the model's state variables option characteristics,[11] such as, e.g., maturity, moneyness, and payoff structure, and thus the options are differentially exposed to systematic risks. Moreover, the relation between

[10] This specification is used, among others, in Heston (1993), Pan (2002), Jones (2003), Bates (2008), Egloff, Leippold, and Wu (2010), Gourier (2014), and Bégin, Dorion, and Gauthier (2020) among others.

[11] Consider for example the Black and Scholes (1973) and Heston (1993) models. In both models, the delta and the vega depend non-linearly on time to maturity and on the state variables (S_t in Black and Scholes (1973) and $\{S_t, v_t\}$ in Heston (1993), where S_t denotes the underlying asset price and v_t is the asset's stochastic variance).

exposures and option characteristics is highly non-linear even in the most parsimonious frameworks such as the Black and Scholes (1973) model. The relationship between option exposures and option characteristics may also vary with the state of the economy and the level of uncertainty.

To allow for rich cross-sectional and temporal variations in risk loadings, we adopt a non-parametric approach. We model exposures to a given risk factor as non-linear functions of a $p \times 1$ vector of option characteristics, $\boldsymbol{c}_{O,t}$, and a $s \times 1$ vector of factor-specific instruments, $\boldsymbol{i}_{l,t}$. For example, in our benchmark model, the relevant option characteristics are time to maturity measured in years, $\tau_{O,t}$, and moneyness, $k_{O,t} = K_O/S_t$. We thus have $\boldsymbol{c}_{O,t} \equiv [\tau_{O,t}\ k_{O,t}]'$ and $\boldsymbol{i}_{l,t} = v_t$. An option's maturity mechanically decreases with time, which introduces an additional source of time variation in exposures captured by $\tau_{O,t}$. In turn, $k_{O,t}$ captures the impact of S_t on option exposure. The vector of factor-specific instruments may contain additional signals (state variables) that are informative about option exposures. For instance, many option pricing modelspredict that option exposures are function of the level of conditional variance. In the single-instrument case when $i_{l,t}$ is a scalar, this is consistent with imposing $i_{l,t} = v_t$, and we indeed do so in our benchmark specification.

We denote by $\boldsymbol{z}_{O,l,t} \equiv \left[\boldsymbol{c}'_{O,t}\ \boldsymbol{i}'_{l,t}\right]'$ the $(p+s) \times 1$ vector of all the signals used to model exposures. Armed with this notation, the risk loading of option O on factor l satisfies

$$\beta^O_{l,t} \equiv \beta_l\left(\boldsymbol{z}_{O,l,t}\right). \tag{10}$$

Fluctuations in exposures across time and options are entirely induced by the vector of signals $\boldsymbol{z}_{O,l,t}$. In contrast, the non-linear mapping between the signals and exposures, $\beta_l(\cdot)$, is unique for each risk factor.

We model $\beta_l(\cdot)$ using basis expansion theory. Basis expansions allow for transformation of a set of signals into any (smooth) non-linear transformation. While a wide variety of basis expansion methods are available, we model $\beta_l(\cdot)$ using the thin plate regression spline basis (TPRS basis henceforth). The TPRS basis is able to parsimoniously account for multidimensional signals, and transforms the vector $\boldsymbol{z}_{O,l,t}$ as follows

$$\boldsymbol{z}_{O,l,t} \underset{TRPS}{\longrightarrow} [1 \quad \boldsymbol{z}_{O,l,t} \quad \varphi_1(\boldsymbol{z}_{O,l,t}) \quad \cdots \quad \varphi_h(\boldsymbol{z}_{O,l,t})]' \equiv \boldsymbol{\phi}_{O,l,t}, \tag{11}$$

where each element $\varphi_j(\boldsymbol{z}_{O,l,t})$ with $j \in \{1, ..., h\}$ is a non-linear scalar transformation of $\boldsymbol{z}_{O,l,t}$. Appendix A contains the details regarding the specification of $\varphi_j(\cdot)$ in the context of the TPRS basis transforms. From equation (11), note that $\boldsymbol{\phi}_{O,l,t}$ is a $k \times 1$ vector composed of 1, $\boldsymbol{z}_{O,l,t}$, and h non-linear basis functions such that $k = 1 + (p + s) + h$. We further elaborate on the choice of h below in the empirical analysis.

Based on the basis expansion (11), the risk loading $\beta^O_{l,t}$ satisfies

$$\beta^O_{l,t} = \beta_l(\boldsymbol{z}_{O,l,t}) = \boldsymbol{\phi}'_{O,l,t}\boldsymbol{b}_l = \sum_{j=1}^{k} \phi^j_{O,l,t} b^j_l, \tag{12}$$

where $\boldsymbol{b}_l$ is a $k \times 1$ vector of factor-specific parameters capturing the basis representation of $\beta^O_{l,t}$ across O and t.[12] Thus, exposures to a given risk factor are linear ($\boldsymbol{b}_l$) in the (non-linear) basis expansions ($\boldsymbol{\phi}_{O,l,t}$). Given that $\boldsymbol{\phi}_{O,l,t}$ is fixed once the basis expansion method is chosen, it can be treated as a vector of explanatory variables.

This approach to the modeling of option exposures builds on Kelly et al. (2018) who develop a framework for modeling instrumented exposures. They model the exposures of stocks as power functions of stock characteristics. In their empirical application, they analyze a restricted version of their model with linear stock exposures. We extend their framework by allowing exposures to be functions of basis expansions of option characteristics. In particular, the TPRS basis we choose has many optimal properties for modeling scalar-valued functions of multiple variables, such as smoothness and parsimony. As we argue below, the use of basis expansions is particularly important for adequately measuring option exposures, given the nature of the security.

Combining equation (9) with the definition of exposures (12) above implies the fol-

[12] Asymptotically, when the dimension of the basis expansion h increases at a slower rate than the sample size, then as h and the sample size become infinitely large, $\beta^O_{l,t} = \beta_l(\boldsymbol{z}_{O,l,t}) = \boldsymbol{\phi}'_{O,l,t}\boldsymbol{b}_l$ converges almost surely to the true (unobserved) option exposure.

lowing factor representation of deleveraged excess option returns

$$R^O_{t+\Delta t} = \sum_l \left[-\boldsymbol{\mu}_l^{Q'} \boldsymbol{g}_{l,t} \cdot \boldsymbol{\phi}'_{O,l,t} \boldsymbol{b}_l + f_{l,t+\Delta t} \cdot \boldsymbol{\phi}'_{O,l,t} \boldsymbol{b}_l \right] + \varepsilon^O_{t+\Delta t}, \tag{13}$$

where $\boldsymbol{\mu}_l^{Q'} \boldsymbol{g}_{l,t} \cdot \boldsymbol{\phi}'_{O,l,t} \boldsymbol{b}_l$ corresponds to $E_t^Q \left[f_{l,t+\Delta t} \right] \beta^O_{l,t}$ and $f_{l,t+\Delta t} \cdot \boldsymbol{\phi}'_{O,l,t} \boldsymbol{b}_l$ to $f_{l,t+\Delta t} \beta^O_{l,t}$.

In summary, compared to existing studies, our methodological contribution is to develop a factor model of option returns with rich option exposures and factor risk premium dynamics. We extend the panel approach of Gagliardini et al. (2016) to options by instrumenting option exposures in the spirit of Kelly et al. (2018).

2.3 Relation to Affine Option Pricing Models

We next discuss how our model extends standard affine option pricing frameworks by allowing for more general exposure and risk premium dynamics. To illustrate this, we consider a simple example: The stochastic volatility (SV) model of Heston (1993). Note that this example can easily be generalized to more general models considered in the literature, such as the stochastic volatility with jump (SVJ) case.

In the Heston (1993) model, two factors determine the prices of options. The first factor is the underlying asset return, which follows a geometric Brownian motion. We have:

$$\begin{gathered} \frac{dS_{t+dt}}{S_t} = (r + \lambda_s v_t)\, dt + \sqrt{v_t} \cdot dW^{s,P}_{t+dt} \quad \text{under} \quad P \\ \frac{dS_{t+dt}}{S_t} = r dt + \sqrt{v_t} \cdot dW^{s,Q}_{t+dt} \quad \text{under} \quad Q, \end{gathered} \tag{14}$$

under the physical and risk-neutral measures, respectively, where λ_s denotes the return risk premium parameter, v_t the stochastic variance, and $dW^{s,\cdot}_{t+dt}$ denotes the Brownian motion that drives return dynamics. The second factor is the asset's stochastic variance. The P- and Q-dynamics for the spot variance are given by:

$$\begin{gathered} dv_{t+dt} = \kappa^P \left(\theta^P - v_t \right) dt + \sigma \sqrt{v_t} \cdot dW^{v,P}_{t+dt} \quad \text{under} \quad P \\ dv_{t+dt} = \kappa^Q \left(\theta^Q - v_t \right) dt + \sigma \sqrt{v_t} \cdot dW^{v,Q}_{t+dt} \quad \text{under} \quad Q, \end{gathered} \tag{15}$$

where κ^P and $\kappa^Q = \kappa^P + \lambda_v$ are the mean reversion speed parameters, λ_v is the variance risk premium parameter, and θ^P and $\theta^Q = \frac{\kappa^P \theta^P}{\kappa^Q}$ are the unconditional variances. Moreover, σ is the volatility of variance parameter, and $dW_{t+dt}^{v,\cdot}$ denotes the Brownian motion driving variance dynamics such that $dW_{t+dt}^{s,\cdot} \cdot dW_{t+dt}^{v,\cdot} = \rho dt$, where ρ captures the leverage effect.

Equations (14) and (15) have important implications for the conditional moments of the factors. Using the notation introduced in Sections 2.1 and 2.2, we can express the physical and risk-neutral expectations of the factors in Heston (1993) as

$$\begin{aligned} E_t^m\left[\frac{dS_{t+dt}}{S_t}\right] &= \begin{bmatrix} r \\ \lambda_s \mathbf{1}_{\{m=P\}} \end{bmatrix}' \cdot \begin{bmatrix} 1 \\ v_t \end{bmatrix} dt = \boldsymbol{\mu}_s^{m,H\prime} \boldsymbol{g}_{s,t}^H \\ E_t^m\left[dv_{t+dt}\right] &= \begin{bmatrix} \kappa^m \theta^m \\ -\kappa^m \end{bmatrix}' \cdot \begin{bmatrix} 1 \\ v_t \end{bmatrix} dt = \boldsymbol{\mu}_v^{m,H\prime} \boldsymbol{g}_{v,t}^H, \end{aligned} \tag{16}$$

for $m = P, Q$ where $\mathbf{1}_{\{m=P\}}$ is the indicator function that takes the value 1 when $m = P$ and 0 otherwise, and where $\boldsymbol{\mu}_s^m$, $\boldsymbol{\mu}_v^m$, $\boldsymbol{g}_{s,t}$ and $\boldsymbol{g}_{v,t}$ are implicitly defined in equation (16).[13] Again using the notaion from Sections 2.1 and 2.2, the conditional factor risk premia in Heston (1993) are given by

$$\begin{aligned} E_t^P\left[\frac{dS_{t+dt}}{S_t}\right] - E_t^Q\left[\frac{dS_{t+dt}}{S_t}\right] &= \left(\boldsymbol{\mu}_s^{P,H} - \boldsymbol{\mu}_s^{Q,H}\right)' \boldsymbol{g}_{s,t}^H = \lambda_s v_t dt \\ E_t^P\left[dv_{t+dt}\right] - E_t^Q\left[dv_{t+dt}\right] &= \left(\boldsymbol{\mu}_v^{P,H} - \boldsymbol{\mu}_v^{Q,H}\right)' \boldsymbol{g}_{v,t}^H = \lambda_v v_t dt, \end{aligned} \tag{17}$$

respectively. Together, equations (16) and (17) provide the restrictions imposed by the Heston (1993) model's assumptions on the dynamics (6) and (7) in our modeling framework. In stark contrast with affine models which fully specify factor dynamics, our setup does not require any assumption about the factors' conditional distribution aside from the functional form of the conditional expectations of the factors.

Building on Broadie et al. (2009), one can show that the Heston (1993) model implies the following (instantaneous) dynamics of deleveraged excess option returns as defined in

[13]For instance, we have $\boldsymbol{g}_{s,t} = \boldsymbol{g}_{v,t} = [1 \quad v_t]' \, dt$.

(8) and (9)

$$\begin{aligned} R^O_{t+dt} &= \sum_{l=s,v} \left[-E^Q_t \left[f_{l,t+dt} \right] \cdot \beta^{O,H}_{l,t} + f_{l,t+dt} \cdot \beta^{O,H}_{l,t} \right] \\ &= \sum_{l=s,v} \left[-\boldsymbol{\mu}^{Q,H'}_l \boldsymbol{g}^H_{l,t} \cdot \beta^{O,H}_{l,t} + f_{l,t+dt} \cdot \beta^{O,H}_{l,t} \right], \end{aligned} \tag{18}$$

where $f_{s,t+dt} = \frac{dS_{t+dt}}{S_t}$ and $f_{v,t+dt} = dv_{t+dt}$ are the factors, and where $\beta^{O,H}_{s,t} \equiv \frac{\partial O_t}{\partial S_t}$ and $\beta^{O,H}_{v,t} \equiv \frac{\partial O_t}{\partial v_t} \frac{1}{S_t}$ denote the delta and leverage-adjusted vega in the Heston (1993) model. Clearly, the return dynamic (18) is equivalent to equations (8) and (9) when $\Delta t \to dt$ in the limit.

In summary, the existing option literature provides important insights about the pricing of volatility and jump risks. However, the nature and number of factors that existing models can accommodate is limited given the complex parametric nature of these models. In contrast, our framework can be used to study the pricing of return- or volatility-based or other factors, provided that the factors are observable (or can be constructed).

3 Estimation Strategy

Our estimation strategy is designed to handle large unbalanced panels, which is critical for studying the cross-section of SPX option returns.[14] We develop a simple three-stage OLS estimation approach which combines features of Fama and MacBeth (1973) Fama-MacBeth (1973) with standard cross-sectional panel regressions. Our approach adapts various results and methodologies originally developed to study stock returns to our setup.[15] We first explain our strategy for estimating exposures. We then discuss the estimation of conditional risk premia.

[14]The unbalanced nature of option panels comes from two sources. First, options of different moneyness and maturity are introduced at different points in time. Second, the return observation of a given option may be missing on some days but not others.

[15]For discussions about the estimation of cross-sectional asset pricing models for stocks, see, Nagel and Singleton (2011) and the surveys by Goyal and Saretto (2009) Goyal (2012) and Nagel and Singleton (2011) Nagel (2013).

3.1 Estimating Option Exposures

Recall that combining equations (6), (7), and (9) with the definition of exposures (12) yields equation (13). Because the explanatory variables $\boldsymbol{g}_{l,t}\boldsymbol{\phi}'_{O,l,t}$ are pre- and post-multiplied by the vectors of parameters, $\boldsymbol{\mu}_l^Q$ and $\boldsymbol{b}_l$, equation (13) is non-linear in the model parameters. Without further simplifications, the return dynamics (13) therefore cannot be estimated by OLS. To circumvent this problem, the following proposition presents an alternative factor representation of deleveraged excess option returns which allows for OLS estimation.

Proposition 1 *The dynamics of deleveraged excess option return in equation (13) admits the following factor structure*

$$R^O_{t+\Delta t} = \boldsymbol{\phi}^{g'}_{O,t}\boldsymbol{b}^Q + \boldsymbol{\phi}^{f'}_{O,t+\Delta t}\boldsymbol{b} + \varepsilon^O_{t+\Delta t}, \tag{19}$$

with

$$\begin{aligned} \boldsymbol{\phi}^{g'}_{O,t}\boldsymbol{b}^Q &= \textstyle\sum_l -E^Q_t\left[f_{l,t+\Delta t}\right]\cdot\beta^O_{l,t} \\ \boldsymbol{\phi}^{f'}_{O,t+\Delta t}\boldsymbol{b} &= \textstyle\sum_l f_{l,t+\Delta t}\cdot\beta^O_{l,t}, \end{aligned} \tag{20}$$

where $\boldsymbol{\phi}^g_{O,t} \equiv \left[\boldsymbol{\phi}^{g'}_{O,1,t}\cdots\boldsymbol{\phi}^{g'}_{O,M,t}\right]'$ *is a* $Mkq\times 1$ *vector composed of* $\boldsymbol{\phi}^g_{O,l,t} \equiv \boldsymbol{g}_{l,t}\otimes\boldsymbol{\phi}_{O,l,t}$ *for which* $\otimes$ *denotes the Kronecker product,* $\boldsymbol{\phi}^f_{O,t+\Delta t} \equiv \left[\boldsymbol{\phi}^{f'}_{O,1,t+\Delta t}\ \cdots\ \boldsymbol{\phi}^{f'}_{O,M,t+\Delta t}\right]'$ *is an* $Mk\times 1$ *vector composed of* $\boldsymbol{\phi}^f_{O,l,t+\Delta t} \equiv f_{l,t+\Delta t}\otimes\boldsymbol{\phi}_{O,l,t}$*. The parameter vectors are given by*

$$\boldsymbol{b} \equiv \left[\boldsymbol{b}'_1\cdots\boldsymbol{b}'_M\right]' \quad \textit{and} \quad \boldsymbol{b}^Q \equiv \left[\boldsymbol{b}^{Q'}_1\cdots\boldsymbol{b}^{Q'}_M\right]', \tag{21}$$

where $\boldsymbol{b}^Q_l \equiv -\boldsymbol{\mu}^Q_l\otimes\boldsymbol{b}_l$ *is a vector of dimension* $qk\times 1$*, which corresponds to the product of* $\boldsymbol{b}_l$*, the* $k\times 1$ *vector of parameters defining options' exposures to the* l^{th}*-factor in (12), and* $\boldsymbol{\mu}^Q_l$*, the* $q\times 1$ *vector of parameters defining the* l^{th}*-factor's risk-neutral conditional expectation.*

Proof. By the properties of the Kronecker product, equation (13) is equivalent to

$$\begin{aligned}
R^O_{t+\Delta t} &= \sum_l \left[(\boldsymbol{g}_{l,t} \otimes \boldsymbol{\phi}_{O,l,t})' \left(-\boldsymbol{\mu}^Q_l \otimes \boldsymbol{b}_l \right) + (f_{l,t+\Delta t} \otimes \boldsymbol{\phi}_{O,l,t})' \boldsymbol{b}_l \right] + \varepsilon^O_{t+\Delta t} \\
&= \sum_l \left[\boldsymbol{\phi}^{g'}_{O,l,t} \boldsymbol{b}^Q_l + \boldsymbol{\phi}^{f'}_{O,l,t+\Delta t} \boldsymbol{b}_l \right] + \varepsilon^O_{t+\Delta t} \\
&= \boldsymbol{\phi}^{g'}_{O,t} \boldsymbol{b}^Q + \boldsymbol{\phi}^{f\,'}_{O,t+\Delta t} \boldsymbol{b} + \varepsilon^O_{t+\Delta t}.
\end{aligned}$$

■

Note that $\boldsymbol{\phi}^g_{O,t} = \boldsymbol{g}_{l,t} \otimes \boldsymbol{\phi}_{O,l,t}$ and $\boldsymbol{\phi}^f_{O,t+\Delta t} = f_{l,t+\Delta t} \otimes \boldsymbol{\phi}_{O,l,t}$ are the model's explanatory variables. As illustrated by equation (20), $\boldsymbol{\phi}^{g'}_{O,t} \boldsymbol{b}^Q$ corresponds to the sum of conditional exposures multiplied by the factors' conditional risk-neutral expectation. In contrast, $\boldsymbol{\phi}^{f'}_{O,t+\Delta t} \boldsymbol{b}$ captures the total impact of the risk factor realizations on the returns on option O (i.e., the sum of exposures multiplied by factor realizations).

Using data on risk factors, option characteristics, and the instruments, we first construct measures of $\boldsymbol{g}_{l,t}$, $\boldsymbol{\phi}_{O,l,t}$, and $f_{l,t+\Delta t}$ for all factors l, options O, and date t. Following the results in Proposition 1, the vectors $\boldsymbol{\phi}^g_{O,t}$ and $\boldsymbol{\phi}^f_{O,t+\Delta t}$ can then be obtained. Conditioning on $\boldsymbol{\phi}^g_{O,t}$ and $\boldsymbol{\phi}^f_{O,t+\Delta t}$ for all dates and options, equation (19) can be used to derive the OLS estimator of $\boldsymbol{b}$ and $\boldsymbol{b}^Q$. The following proposition presents the result.

Proposition 2 (First-stage estimator) *The OLS estimator of* $\boldsymbol{B} \equiv \left[\left(\boldsymbol{b}^Q \right)' \quad \boldsymbol{b}' \right]'$ *is*

$$\widehat{\boldsymbol{B}} \equiv \boldsymbol{Q}^{-1}_{\phi} \frac{1}{T} \sum_t \frac{1}{N_t} \sum_O \boldsymbol{\phi}^{fg}_{O,t+\Delta t} R^O_{t+\Delta t} \,, \tag{22}$$

where

$$\boldsymbol{Q}_\phi \equiv \frac{1}{T} \sum_t \frac{1}{N_t} \sum_O \boldsymbol{\phi}^{fg}_{O,t+\Delta t} \left(\boldsymbol{\phi}^{fg}_{O,t+\Delta t} \right)' \,,$$

and $\boldsymbol{\phi}^{fg}_{O,t+\Delta t} = \left[\left(\boldsymbol{\phi}^g_{O,t} \right)' \quad \left(\boldsymbol{\phi}^f_{O,t+\Delta t} \right)' \right]'$.[16]

[16] Depending on model specification, the matrix of first-stage regressors $\boldsymbol{\Phi}$ defined by stacking the transposes of $\boldsymbol{\phi}^{fg}_{O,t+\Delta t}$ may not have full rank. This situation can arise if the factor-instrument vectors $\boldsymbol{i}_{l,t}$ have common elements across different factors. In this case, one must drop the redundant columns in $\boldsymbol{\Phi}$ before estimating the model. This in turn impacts the dimension of $\boldsymbol{b}^Q$ which decreases to reflect the fact that $E^Q_t \left[f_{l,t+\Delta t} \right] \beta^O_{l,t}$ may no longer be identified for each risk factor separately. However, dropping redundant columns in $\boldsymbol{\Phi}$ does not impact the identification of the vector $\boldsymbol{b}$ which contains relevant information about option exposures.

Proof. The proof follows from standard OLS results applied to the factor representation in Proposition 1. ■

Under conditions stated in Assumption 1(i) in Appendix B, the estimator $\widehat{\boldsymbol{B}}$ is unbiased and exhibits, for a fixed basis dimension k, the following limiting behavior.

Proposition 3 (Asymptotic normality of first-stage estimator) *For fixed basis size k, and under assumption 1(i) in Appendix B,*

$$\sqrt{T}\left(\widehat{\boldsymbol{B}} - \boldsymbol{B}\right) \sim N(0, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{B}}}),$$

with

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{B}}} \equiv \boldsymbol{Q}_{\phi}^{-1}\boldsymbol{\Sigma}_{\phi\varepsilon}\boldsymbol{Q}_{\phi}^{-1},$$

and

$$\boldsymbol{\Sigma}_{\phi\varepsilon} \equiv V\left[\frac{1}{\sqrt{T}}\sum_{t}\frac{1}{N_t}\sum_{O}\boldsymbol{\phi}^{fg}_{O,t+\Delta t}\varepsilon^{O}_{t+\Delta t}\right].$$

Inference about model parameters requires to estimate the covariance matrix $\boldsymbol{\Sigma}_{\phi\varepsilon}$. In Proposition 5 in Appendix B we give the appropriate estimator $\widehat{\boldsymbol{\Sigma}}_{\phi\varepsilon}$ which is robust to heteroskedasticity, autocorrelation, cross-correlation and cross-sectional correlation in $\varepsilon^{O}_{t+\Delta t}$.

With the estimated exposures on hand, we now turn to the estimation of risk premia.

3.2 Estimating Factor Risk Premia

No arbitrage condition for traded risk factors implies that their Q-expectation must be equal to the risk-free rate. One way to impose this condition is to estimate risk premia for traded and non-traded risk factors separately and to express traded factors in excess returns. In our empirical set-up, the market risk factor is the only traded factor while the remaining factors are usually not tradable. Separating the risk premium estimation of traded and non-traded factors has also the advantage of improving the precision of risk premia estimation when working with option returns because of the fact that the second stage regressors – products of option sensitivities and risk premia predictors – are strongly correlated in practice.

Our strategy thus consists in estimating the market risk premium in the first step by regressing the market-return projection of option returns on market risk premium predictors. We then remove the exposures of option returns to the market factor from option returns and regress market-hedged option returns on non-traded factors to obtain the remaining risk premia parameters in the second step.

In what follows, we use the factor subscript $l = 1$ for the market risk factor (i.e., $f_{1,t+\Delta t}$ denotes the excess return on the S&P 500 index) and denotes by Ω (i.e., $l \neq 1$) the set of non-market factors.

Step 1: Estimation of the Market Risk Premium

To estimate risk premia, we first need to express deleveraged excess option return as a linear function of risk premium parameters. To this end, we define the projection of the excess option return on the realization of the market risk factor,

$$\widehat{R}^{O\Delta}_{t+\Delta t} \equiv \widehat{\beta}^{O}_{1,t} \left(f_{1,t+\Delta t} - E^{Q}_{t} \left[f_{1,t+\Delta t} \right] \right), \tag{23}$$

where $E^{Q}_{t}\left[f_{1,t+\Delta t}\right] = 0$ because $f_{1,t+\Delta t}$ is the excess return on the S&P 500 index. Adding and subtracting $\beta^{O}_{1,t} E^{P}_{t}\left[f_{1,t+\Delta t}\right]$ to $\widehat{R}^{O\Delta}_{t+\Delta t}$ yields

$$\widehat{R}^{O\Delta}_{t+\Delta t} = \widehat{\beta}^{O}_{1,t} \left(E^{P}_{t}\left[f_{1,t+\Delta t}\right] - E^{Q}_{t}\left[f_{1,t+\Delta t}\right] \right) + \widehat{\beta}^{O}_{l,t} \left(f_{l,t+\Delta t} - E^{P}_{t}\left[f_{1,t+\Delta t}\right] \right) = \widehat{\beta}^{O}_{1,t} \lambda_{1,t} + \eta^{\Delta}_{O,t+\Delta t}\,, \tag{24}$$

where $\lambda_{1,t} = \boldsymbol{g}'_{1,t} \boldsymbol{\lambda_1}$ is the conditional market risk premium and where the error term $\eta^{\Delta}_{O,t+\Delta t} \equiv \beta^{O}_{1,t}\left(f_{1,t+\Delta t} - E^{P}_{t}\left[f_{1,t+\Delta t}\right]\right)$ has zero P-expectation by construction. Given that $\widehat{\beta}^{O}_{1,t}\lambda_{1,t} = \left(\widehat{\beta}^{O}_{1,t}\boldsymbol{g}_{1,t}\right)' \boldsymbol{\lambda}_1$, we obtain the following regression equation, which is linear in market premium parameters $\boldsymbol{\lambda}_1$,

$$\widehat{R}^{O\Delta}_{t+\Delta t} = \left(\widehat{\beta}^{O}_{1,t}\boldsymbol{g}_{1,t}\right)' \boldsymbol{\lambda}_1 + \eta^{\Delta}_{O,t+\Delta t} = \widehat{\boldsymbol{d}}^{\Delta\prime}_{O,t} \boldsymbol{\lambda}_1 + \eta^{\Delta}_{O,t+\Delta t}\,. \tag{25}$$

In the previous equation, $\widehat{\boldsymbol{d}}_{O,t}^{\Delta\prime} \equiv \widehat{\beta}_{1,t}^{O} \boldsymbol{g}_{1,t}'$ is a vector of dimension $s \times 1$ which contains the product of the estimated option loadings on the market excess return with market risk premium predictors. It is worth noting that $\eta_{O,t+\Delta t}^{\Delta}$ is orthogonal to the first-stage estimation error $\varepsilon_{t+\Delta t}^{O}$ and to $\widehat{\boldsymbol{d}}_{O,t}^{\Delta}$ which is important for inference purposes.

The regression equation (25) can be straightforwardly estimated by OLS.[17] However, and as pointed out by Shanken (1992), such approach introduces one potential problem. More precisely, note that measurement errors in market loadings impact both sides of the equation (25) since $\beta_{l,t}^{O}$ impacts $\widehat{R}_{t+\Delta t}^{O\Delta}$ and $\widehat{\boldsymbol{d}}_{O,t}^{\Delta}$ simultaneously. While the presence of measurement error on the right-hand side of equation (25) is not a problem in itself, its presence in the left-hand side of the regression equation results in an asymptotic bias of the standard OLS estimator.[18] With this in mind, the following proposition presents the asymptotic behavior of the de-biased OLS estimator of the vector of market risk premium parameters.

Proposition 4 (Limiting behavior of de-biased estimator of $\boldsymbol{\lambda}_1$) *Under Assumptions 2(iii)-2(iv) in Appendix B, with the unbiased estimator of market premia parameters $\widehat{\boldsymbol{\lambda}}_1^{\Psi}$ given in Proposition 8 in Appendix B, $\boldsymbol{Q}_{\widehat{\boldsymbol{d}}^{\Delta}} \equiv \frac{1}{T}\sum_t \frac{1}{N_t}\sum_O \widehat{\boldsymbol{d}}_{O,t}^{\Delta}\left(\widehat{\boldsymbol{d}}_{O,t}^{\Delta}\right)'$, and the components of $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}_1}$ given in Proposition 9 in Appendix B, the unbiased estimator of market premia parameters follows the asymptotic distribution*

$$\sqrt{T}\left(\widehat{\boldsymbol{\lambda}}_1^{\Psi} - \boldsymbol{\lambda}_1\right) \sim N(0, \boldsymbol{\Sigma}_{\boldsymbol{\lambda}_1}),$$

with

$$\boldsymbol{\Sigma}_{\boldsymbol{\lambda}_1} \equiv \boldsymbol{Q}_{\widehat{\boldsymbol{d}}^{\Delta}}^{-1}\left[\boldsymbol{\Sigma}_{\boldsymbol{d}^{\Delta}\boldsymbol{\eta}^{\Delta}} + \boldsymbol{\Sigma}_{\boldsymbol{d}^{\Delta}\widehat{\boldsymbol{b}}^{\Delta}} + \boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}} + 2\boldsymbol{C}_{\widehat{\boldsymbol{b}}^{\Delta}\widehat{\boldsymbol{d}}^{\Delta}} + \boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}\eta^{\Delta}}\right]\boldsymbol{Q}_{\widehat{\boldsymbol{d}}^{\Delta}}^{-1},$$

and where the estimators of $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}_1}$ is provided in Proposition 10 in Appendix B.

[17] For completeness, we present the expression of the OLS estimator of $\boldsymbol{\lambda}_1$ in Proposition 7 in Appendix B.

[18] Similarly to Shanken (1992), we do not explicitly account for the errors-in-variables bias by assuming that the matrix $\boldsymbol{Q}_{\widehat{\boldsymbol{d}}^{\Delta}}$ defined in Proposition 4 converges to its probability limit.

Step 2: Estimation of Risk Premia for Non-Market Factors

We now describe our strategy to estimate the risk premium parameters of the remaining factors. The estimation and inference procedures described below follow closely the methodology used to estimate market risk premium. First, let us define the regressand used to estimate the remaining risk premia which corresponds to the projection of option returns on non-market risk factors and their Q-expectations. We have

$$\widehat{R}^{O\Omega}_{t+\Delta t} \equiv \sum_{l\in\Omega} -\widehat{\beta}^{O}_{l,t} E^{Q}_{t}\left[f_{l,t+\Delta t}\right] + \sum_{l\in\Omega} \widehat{\beta}^{O}_{l,t} f_{l,t+\Delta t} \, . \tag{26}$$

Adding and subtracting $\sum_{l\in\Omega} \widehat{\beta}^{O}_{l,t} E^{P}_{t}\left[f_{l,t+\Delta t}\right]$ from (26), we can express the non-market component of the option return as a linear equation of risk premium parameters

$$\begin{aligned}
\widehat{R}^{O,\Omega}_{t+\Delta t} &= \sum_{l\in\Omega} \widehat{\beta}^{O}_{l,t} \left(E^{P}_{t}\left[f_{l,t+\Delta t}\right] - E^{Q}_{t}\left[f_{l,t+\Delta t}\right] \right) + \eta^{\Omega}_{O,t+\Delta t} \\
&= \sum_{l\in\Omega} \widehat{\beta}^{O}_{l,t} \left(\boldsymbol{g}'_{l,t} \boldsymbol{\lambda}_{l} \right) + \eta^{\Omega}_{O,t+\Delta t} \\
&= \widehat{\boldsymbol{d}}^{\Omega\prime}_{O,t} \boldsymbol{\lambda}_{\Omega} + \eta^{\Omega}_{O,t+\Delta t} \, ,
\end{aligned} \tag{27}$$

where $\eta^{\Omega}_{O,t+\Delta t} \equiv \sum_{l\in\Omega} \left(f_{l,t+\Delta t} - E^{P}_{t}\left[f_{l,t+\Delta t}\right] \right) \beta^{O}_{l,t}$ with $E^{P}_{t}\left[\eta^{\Omega}_{O,t+\Delta t}\right] = 0$. Note that we used the definition of the factor risk premium (7) to obtain the second equality from the first, and define $\widehat{\boldsymbol{d}}^{\Omega}_{O,t} \equiv [\ \widehat{\beta}^{O}_{2,t} \boldsymbol{g}'_{2,t} \cdots \widehat{\beta}^{O}_{M,t} \boldsymbol{g}'_{M,t}\ ]'$ and $\boldsymbol{\lambda}_{\Omega} \equiv [\boldsymbol{\lambda}'_{2} \cdots \boldsymbol{\lambda}'_{M}]'$ to obtain (27) from the second equality. In (27), $\widehat{\boldsymbol{d}}^{\Omega}_{O,t}$ corresponds to the $(M-1)\, s \times 1$ vectors of non-market exposures times non-market return risk premium instruments and $\boldsymbol{\lambda}_{\Omega}$ is the vector of risk premium parameters.

Note that equation (27) does not allow for an intercept because it is derived under the assumption of absence of arbitrage opportunities (i.e., it assumes the model prices all assets perfectly). Similarly to testing the CAPM, for example, allowing for an intercept can provide insights about model misspecification. It is straightforward to include an intercept in (27) by augmenting $\widehat{\boldsymbol{d}}^{\Omega}_{O,t}$ with a vector of 1 and by adding the respective intercept parameter to $\boldsymbol{\lambda}_{\Omega}$. The intercept will absorb the part of expected option returns that is not explained by the factors in the set Ω.

Similarly to Step 1, the standard OLS estimator of $\boldsymbol{\lambda}_\Omega$ is biased due to measurement error in exposures. The bias adjustment for $\boldsymbol{\lambda}_\Omega$ is given in Appendix B.2.2.

4 Data

Our empirical analysis relies on five publicly available datasets.

We use the S&P 500 index (SPX) (obtained from OptionMetrics) to proxy for the market portfolio, while excluding dividend payments due to the fact that the options are written on the ex-dividend index value. We use daily deleveraged index option returns as defined in equation (1), computed from end-of-day SPX option mid-quotes obtained from OptionMetrics. In this dataset, we only retain options with maturity ranging from 1 to 6 months. We further restrict the sample to put options with moneyness (K/S_t) between 0.80 and 1.025 and call options with moneyness between 0.975 and 1.15. This allows us to span the entire moneyness spectrum without relying on in-the-money options, which are less liquid. From the resulting sample, we filter out options with zero bid prices and ask prices five times larger than the bid price. Data on risk-free rates are also obtained from OptionMetrics.

We use the square of the VIX index to proxy for index variance. Thus, we set Δv_t to the daily increment in the square of the VIX index (which we obtain from the CBOE website[19]), which we denote VAR. For instrumenting the risk premia, we use the VIX_t^2 and tail variance TVAR_t, the square of the Tail Volatility index of Bollerslev et al. (2015), which we obtain from the Tail Index website of Andersen and Todorov.[20] One of the models we analyze in the robustness section includes the intermediary factor of He et al. (2017) and intermediary capital ratio which we obtain from Manela's website.[21]

We report separate results for puts and calls by moneyness and maturity buckets throughout our empirical analysis. For calls, we use the K/S_t cutoffs of (0.975;1.02], (1.02;1.07] and (1.07;1.15] to form moneyness buckets of at-the-money (ATM), out-of-

[19]https://www.cboe.com/tradable_products/vix/vix_historical_data/
[20]https://tailindex.com/volatilitymethodology.html
[21]http://apps.olin.wustl.edu/faculty/manela/

the-money (OTM), and deep-out-of-the-money (DOTM) options respectively. For ATM puts, we use (0.975;1.02], while for OTM and DOTM puts we use (0.9;0.975] and (0.8;0.9] respectively. We use a wider range of moneyness for SPX puts because of the higher liquidity of deep-out-of-the-money puts relative to calls. For each option type and maturity bucket, we further report separate results for options with maturity of 1-2, 2-3, and 3-6 months. This provides additional insights about the term structure of expected option returns and option sensitivities.

Table 1 presents the sample average, standard deviation, Sharpe ratio, skewness, and kurtosis of daily deleveraged and daily delta-hedged deleveraged option returns. The average return, standard deviation, and Sharpe ratio are annualized assuming 252 trading days in a year. For delta-hedged returns, we delta-hedge one long position in each deleveraged option using our model deltas on a daily basis. On average, deleveraged call returns are positive across maturity and moneyness buckets, while put returns are negative. Average put returns are substantially larger (in absolute value) than call returns. Puts load negatively on the market return premium, which is positive, and positively on the variance risk premium, which is negative. The market return and variance risk premia thus do not cancel each other out for puts. Calls on the other hand are positively exposed to market return and variance risks. The lower magnitude of average call returns is thus partly due to their positive exposure to two factors with risk premia of opposite sign.

Unlike deleveraged returns, both put and call delta-hedged excess returns are negative on average. This is consistent with the existing literature, see for instance Bakshi, Kapadia, and Madan (2003).[22] Table 1 also indicates that both unhedged and delta-hedged deleveraged put returns have a higher Sharpe ratio (in absolute value), are less skewed, and have lower kurtosis than call options. The magnitude of the Sharpe ratio of both calls and puts increases (in absolute value) and their standard deviation decreases after delta-hedging. As we discuss in more detail below, this is due to the low signal-to-noise ratio of market return risk compared to the high signal-to-noise ratio of market variance

[22]This finding is also supported by a related literature on negative variance risk premia, see e.g. Coval and Shumway (2001), Buraschi and Jackwerth (2001), Carr and Wu (2009)).

risk.

Comparing the returns across maturity and moneyness buckets, we observe a number of salient patterns. First, put options with shortest maturities (1-2 months) earn more negative returns (before and after delta-hedging) than put options with longer maturities. In call options, we only observe such a pattern, albeit weaker, in delta-hedged returns. Return volatility changes across option type, moneyness and maturity alongside average returns. Overall, before delta-hedging, put options offer larger (in absolute value) annualized Sharpe ratios ranging from -0.58 to -0.38 while call options deliver very low Sharpe ratios (with a maximum of 0.31, but typically below 0.15). After delta-hedging, the Sharpe ratios of both put and call options increase in magnitude, however those of put options are much greater than those of call options (between -1.15 and -0.88 for puts vs -0.44 and -0.15 for calls). Finally, the Sharpe ratios of delta-hedged short-maturity options are significantly lower than of options with longer maturities, in all moneyness buckets. This pattern does not robustly occur in call options. A cursory comparison of option returns across moneyness buckets suggests that the returns on out of the money options are smaller in magnitude than returns on at the money options. This pattern, however, is a mechanical consequence of our chosen method of deleveraging of the returns shown in equation (1).

[Table 1 about here.]

5 Empirical Results

Our objectives are twofold. First, we aim to quantify option return exposures to various systematic risk factors. Second, taking these estimated exposures as given, we aim to estimate the factors' conditional risk premiums.

We first discuss our benchmark model specification. We then present results on model fit and discuss the decomposition of index option returns based on risk factors. We discuss the cross-section of index option exposures to different factors and assess the contribution of each factor to index option risks and expected returns. Finally, we present our estimates of factor risk premiums and discuss how the model estimates compare to well-known benchmark models.

5.1 The Benchmark Model Specification

Our benchmark features three risk factors that are deeply rooted in option pricing theory. We denote the first ($l = 1$) factor by MKT. It corresponds to the market return risk factor and is measured using daily SPX returns without dividends. The second ($l = 2$) factor captures aggregate variance risk and we denote it by VAR. It is measured using daily changes in SPX conditional variance, proxied for by the square of the VIX index.

The third ($l = 3$) factor, GAM, is the gamma risk factor, which is proxied for by daily squared SPX returns. GAM controls for the fact that the return on an option is a non-linear function of the return on the market factor, and the magnitude of GAM's contribution to explaining realized option returns depends on the return holding period. Omitting GAM from the model could lead to an important omitted variable bias. As we will discuss below, the GAM factor accounts for a relatively small percentage of the variation in option returns, but its inclusion improves model performance in some dimensions.

Importantly, the GAM factor is distinct from the VAR factor. While VAR is the change in the level of aggregate variance risk, GAM is the squared SPX return, so that $E_t[\text{GAM}_{t+\Delta t}] = v_t + E_t[\text{MKT}_{t+\Delta t}]^2$. The realization of $\text{GAM}_{t+\Delta t}$ is thus on average close to v_t, the level of aggregate variance prevailing at time t. In our sample, the correlation of $\text{GAM}_{t+\Delta t}$ and $v_t = \text{VIX}_t^2$ is approximately 60%, while the correlations of $\text{GAM}_{t+\Delta t}$ with $\text{MKT}_{t+\Delta t}$ and $\text{VAR}_{t+\Delta t}$ are close to zero.

Given the set of factors, we need to specify the P- and Q-conditional expectations of each factor by choosing the vector of conditional expectation parameters $\boldsymbol{\mu}_l^m$ for $m = P, Q$ along with the vector of risk premium instruments $\boldsymbol{g}_{l,t}$. Recall that $\boldsymbol{\mu}_l^m$ for $m = P, Q$ combined with $\boldsymbol{g}_{l,t}$ define $E_t^m\left[f_{l,t+\Delta t}\right] = \boldsymbol{\mu}_l^{m\prime} \boldsymbol{g}_{l,t}$ and the dynamics of the conditional risk premium, $E_t^P\left[f_{l,t+\Delta t}\right] - E_t^Q\left[f_{l,t+\Delta t}\right]$.

Although the model allows for the presence of nonlinear risk premia, our benchmark analysis focuses on a specification with a linear risk premium.[23] For the MKT factor, we consider the vector of physical expectation parameters $\boldsymbol{\mu}_1^P = [\ \mu_{0,1}^P\ \mu_{1,1}^P\ \mu_{2,1}^P\ ]'$ and set $\boldsymbol{g}_{1,t}$ equal to $[1 \quad \text{VIX}_t^2 \quad \text{TVAR}_t]$. Note that the no-arbitrage condition requires $E_t^Q [f_{MKT,t+\Delta t}] = r_t$ which eliminates the need to specify its risk-neutral expectation. Our choice of instruments for the market risk premium is based on an extensive literature which considers the relationship between market and variance risk premia (see e.g., Bollerslev and Todorov, 2011, Martin, 2017), and the fact that the tail risk measure TVAR_t is demonstrated to be a predictor of the equity premium (Bollerslev et al., 2015). For VAR, we set $\boldsymbol{\mu}_2^m = [\ \mu_{0,2}^m\ \mu_{1,2}^m\ ]'$ for $m = P, Q$ and $\boldsymbol{g}_{\text{VAR},t} = [1 \quad \text{VIX}_t^2]$. We set $\boldsymbol{\mu}_3^m = \mu_{0,3}^m$ and $\boldsymbol{g}_{3,t} = 1$ for GAM.

Finally, we need to specify the option exposures. Recall that the exposure of option O to a given risk factor corresponds to the TPRS-basis expansion of the vector $\boldsymbol{z}_{O,l,t} = \left[\boldsymbol{c}_{O,t}' \ \boldsymbol{i}_{l,t}'\right]'$ where $\boldsymbol{c}_{O,t} \equiv [\tau_{O,t}\ k_{O,t}]'$ is the vector consisting of option O 's moneyness and time to maturity, and $\boldsymbol{i}_{l,t}$ (resp. $i_{l,t}$) is the vector of instruments (resp. the single instrument) used to model option exposures to factor l. We follow option pricing theory and set $i_{l,t}$ to $v_t = \text{VIX}_t^2$ for all factors.[24] Finally, we set the TPRS basis order h in equation (11) used to estimate exposures equal to twenty for all factors. Henceforth, consistent with the literature, we refer to market return risk exposures as deltas, to market variance exposures as vegas, and to gamma risk exposures as gammas.

5.2 Unconditional Factor Exposures

We estimate our benchmark model on individual (i.e., not portfolios) SPX deleveraged option returns, following the three-stage estimation methodology outlined in Section 3. The existing literature about the relative contribution of economic factors to option returns is still rather limited. This is partly because option pricing models are usually fitted on prices, and the return implications of these models have received less attention. More-

[23]For discussions of non-linear pricing kernels and risk premia see, among others, Dittmar (2002), Jones (2003), Chabi-Yo (2011), Christoffersen, Heston, and Jacobs (2013), and Schneider and Trojani (2019).

[24]In alternative factor specifications studied in Section 6, the exposures to the HKM factor are instrumented with $\boldsymbol{i}_t = \left[\text{VIX}_t^2\ \text{ICR}_t\right]$ where ICR_t is the inverse capital ratio in He et al. (2017), while the exposures to the TVAR factor are instrumented with the level of tail variance, TVAR_t.

over, standard parametric models are usually silent about the exposures of options to factors other than return, volatility and jump risks. Nevertheless, the literature suggests that additional factors, absent from no-arbitrage reduced-form models, could drive option returns (see e.g., Gârleanu et al., 2009, He et al., 2017). We investigate the importance of several such factors in a series of robustness exercises in Section 6.1 below, and find that their contribution to realized and, especially, expected option returns is modest.

Our objective is to provide new insights into the unconditional exposures and factor decomposition of option returns through the lens of the model.

[Table 2 about here.]

[Figure 1 about here.]

Table 2 presents the results from the first stage estimation. In Panel A we report the sample average (unconditional) exposures of options by moneyness and maturity buckets. To obtain the unconditional exposures for each bucket, we first average the daily exposures of all options in a given bucket to obtain one observation for each bucket on each day and then we average the daily exposures of each bucket over the full sample. The results in Table 2 indicate that the unconditional exposures have the sign predicted by no-arbitrage option pricing models and are consistent with economic intuition, for both calls and puts. For instance, the deltas, vegas, and gammas of calls decrease for all maturities with increasing moneyness (K/S_t). For puts, the deltas increase towards zero as moneyness decreases, and the vegas and gammas decrease. ATM calls have deltas ranging from 0.41 to 0.47 while ATM puts have deltas ranging from -0.43 to -0.40. Note that these numbers correspond to sample averages within buckets which contain options with different strikes and maturities. When considering puts and calls with $K/S_t = 1$, we observe that the estimated deltas are close to -0.5 and 0.5, respectively.

Figure 1 plots exposures to each factor, computed at the the sample median values of the instruments. We then compute the cross-section of exposures for calls and puts by varying moneyness over the respective ranges (i.e., $K/S_t \in [0.8, 1.0]$ for puts and $K/S_t \in [1.0, 1.15]$ for calls), using maturities of 1, 3, and 6 months, respectively. Model-implied exposures display rich patterns across moneyness depending on the factor, the type of option (i.e., puts or calls), and maturity. The delta of ATM calls is around 0.5 and decreases monotonically as moneyness increases. The delta of DOTM puts is close to zero and decreases as K/S_t increases. Model deltas thus behave similarly to deltas implied by standard reduced-form models.

In contrast, the behavior of fitted exposures to market variance risk (VAR) is not always trivial. For ATM puts, vegas display an inverted term-structure in the sense that vega decreases with maturity. This observation on the term-structure of model vegas for ATM puts is inconsistent with the predictions of standard affine models such as Black and Scholes (1973) and Heston (1993) models, which predict that vega increases monotonically with maturity for ATM options. Model prediction for the vega of ATM calls is broadly consistent with standard affine models. Finally, the term-structure of gammas for ATM calls and puts is downward sloping when $K/S_t = 1$,. This is also consistent with standard option pricing models.

5.3 A Variance Decomposition of Option Returns

Panel B of Table 2 presents the R^2 contribution of each factor to the total variance of option returns. To compute the contribution of each factor to the R^2, we first use the average of daily exposures based on all options in a given bucket. We then multiply this average by the daily factor realizations to get the total impact of a given factor on option returns for each bucket on each day. Finally, we compute the Shapley-Owen R^2 for each bucket.[25] The Shapley-Owen R^2 conveniently allows to isolate the contribution of each factor in explaining deleveraged option returns while accounting for potential co-movements between factors. Another advantage of this measure is that it is invariant to the order in which the factors are included in the model.

The total R^2 for each option bucket is reported in the rightmost column of the Put and Call option sub-panels, and can be used to assess model fit. From these results we

[25] The Shapley-Owen value of R^2 is defined as the average marginal contribution of a variable towards R^2 across all possible model specifications and orders of variable inclusion. As such, this measure does not depend on the order in which the regressors are included in the model, and the total contribution of the regressors in the full sample is equal to the R^2 of the complete regression model. For more information see e.g. Huettner and Sunder (2012).

conclude that the fit obtained by our benchmark model is impressive. The R^2 ranges between 92% and 99% across buckets for calls, and between 98% to 108% for puts.[26] The lowest R^2 is obtained for long-maturity DOTM calls.

The R^2 contributions of market return and variance risk far exceed that of gamma risk for both calls and puts. Overall, market risk is more important than variance risk, but for OTM and DOTM puts, the contributions are relatively close. The relative contribution of market risk is more important for calls than for puts. It ranges from 64% to 73% for calls, and from 47% to 61% for puts. The contribution of variance risk to option return R^2 ranges from 35% to 54% for puts and from 23% to 27% for calls. The contribution of gamma risk to option return R^2 is small, and ranges from 2% to 6% for puts and from 1% to 9% for calls. Thus, two strongly correlated factors (MKT and VAR) explain more than 90% of the variation in index option returns across all option types, maturity and moneyness buckets.

5.4 Conditional Exposures and Hedging Performance

Option market participants such as banks, pension funds, and hedge-fund portfolio managers actively trade index options and need to quantify the conditional exposures of their positions to relevant risks for the purpose of asset allocation and risk management. With this in mind, we now assess the hedging performance of the model based on conditional factor exposures.

Table 3 presents the hedging performance of our model relative to the Black and Scholes (1973) and Heston (1993) models. We first analyze the results for the former, presented in Panel A. Each day, we hedge one long position in each deleveraged option by subtracting the product of the conditional exposures and the factor realizations from deleveraged option returns. Conditional exposures are computed from either the Black and Scholes (1973) or our model. For each maturity/moneyness bucket and each day, we first calculate the average hedged return for all options in that bucket. Using the time series of daily average hedged returns for each bucket, we then compute the full sample

[26]Note that the Shapley-Owen R^2 exceeds 100% in a number of instances. This can occur when return fit is evaluated on a subset of the estimation sample.

variance of hedged returns for each model and report the variance ratio between the two models. Because the Black and Scholes (1973) hedged option return variance is in the numerator, a ratio greater than 1.0 indicates that the variance of hedged returns implied by our model is smaller than the variance of hedged returns implied by the Black-Scholes model (1973).

[Table 3 about here.]

The results are consistent across calls and puts. First, not surprisingly the performance of the Black and Scholes (1973) model relative to our model deteriorates substantially when we add factors. Second, the variance factor improves hedging more than the gamma factor. Third, while the gamma factor in itself leads to a modest improvement in hedging (the third columns), it substantially improves hedging when combined with the market and variance factors (the fourth columns). This illustrates the importance of accounting for factor dependence. In our model, exposures are estimated jointly and directly account for factor co-movements. Note that factor exposures implied by conventional models correspond to the partial derivative of the option price with respect to the factor, and therefore do not account for factor dependence. Fourth, the improvements provided by our model over Black-Scholes are consistently higher for call options than for put options.

When considering the delta-hedging strategy, the ratios are lower than one for the majority of buckets for put options. Note that this finding does not indicate that Black and Scholes (1973) deltas are more precise than the deltas estimated with our model. Instead, it is a direct consequence of the fact that our estimated exposures account for factor co-movement.[27]

While the Black and Scholes (1973) model is a useful first benchmark, given the importance of the variance factor in our empirical specification it is useful to compare

[27]The MKT and VAR factors are (negatively) correlated, and this fact is taken into account in estimation, while in the Black and Scholes (1973) model volatility is constant, and the delta is simply a partial derivative of the option price. As a consequence, the Black and Scholes model tends to over-hedge option returns, in the sense that an investor willing to hold options delta-hedged with Black-Scholes greeks in order to earn compensation for variance risk would see their investment performance suffer because of removing too much underlying return variation.

its hedging performance with that of the Heston (1993) model. Panel B of Table 3 presents these results, where the Heston model is estimated using particle filtering. The results in Panel B of Table 3 confirm those of Panel A in the sense that our model also outperforms the Heston (1993) model. Not surprisingly, the differences in performance are less dramatic compared to the Black-Scholes case due to the improved modeling of the variance factor.

Overall, we conclude that the economic magnitude of the gains from hedging using our model instead of the Black and Scholes (1973) model are considerable, especially when it comes to hedging variance risk, where our model's hedge is an order of magnitude better than the Black-Scholes greeks hedge, which in certain cases increases the variance of hedged returns.

5.5 How Much does Each Factor Contribute to Expected Returns?

To provide additional insights into expected option returns, Panel A of Table 4 presents results on a factor decomposition of expected deleveraged option returns. In columns 1-3 of each sub-panel of Panel A, we report the contribution of each factor to expected put and call option returns, respectively. Column 4 presents the model prediction for total expected returns, which corresponds to the sum of columns 1 through 3. Consistent with the results on factor exposures in Table 2, market return risk is the main determinant of expected option returns for both calls and puts, followed by exposure to market variance risk. Gamma risk plays a relatively minor role. The model predictions for deleveraged option expected return in the fourth columns are similar to the sample return average reported in Table 1. Note of course that the model provides expected return predictions at each point in time in addition to these averages. We examine these risk premia below.

To further assess the model's ability to adequately capture option return dynamics (conditional expected option returns), we regress daily deleveraged option returns against the model predictions using the full-sample time series of daily option returns.[28] We

[28] The expected returns in Panel B of Table 4 include an intercept term which is identical for all options,

report the results in Panel B of Table 4. Impressively, when put options with all maturities and moneyness levels are considered together, the point estimates of the intercept and slope of the regression are 0.05 and 1.00, respectively. They are slightly off for calls, at 0.42 and 1.13, but these values are not statistically significantly different from 0.0 and 1.0, respectively. This is strong evidence that our model correctly captures the conditional dynamics of option returns. Even more strikingly, the slope coefficients obtained are all positive and close to one across buckets for put options. For call options, this holds across maturity buckets, however in the analysis of moneyness buckets we see that our model struggles with certain groups of options: ATM and OTM (but not DOTM) calls with medium to long maturity. Nevertheless, in no case are the estimated loadings statistically different from 1.0 at standard significance levels. This result further supports the model's ability to reliably extract the information embedded in option return dynamics.

[Table 4 about here.]

5.6 Implied Factor Risk Premia

Index options are highly informative about the dynamics of risk premia because they embody market expectations for different states of nature. However, the existing literature has mainly focused on the study of variance and jump risk premia implied by index options through the lens of affine no-arbitrage pricing models. Relative to these frameworks, our flexible factor approach to the modeling of option returns allows us to back out risk premium dynamics associated with all relevant factors. Importantly, because we use option returns, we are also able to better extract the conditional premium for market return risk thanks to the tight no-arbitrage relationships between the options and the index return.

Table 5 illustrate the implications of our model for factor risk premia embedded in index option returns. Panel A reports the risk premium parameters and corresponding p-values, as well as the F-statistics for the significance of premia on each factor. Panel B

and thus not reported in Panel A. The inclusion of this intercept term in the estimation of risk premia is discussed in Section 3.2.

presents various summary statistics on daily risk premia, including the average, median, standard deviation, skewness, and kurtosis of the time series of daily conditional risk premium estimates for each factor. Because the risk premium for the GAM-factor is constant in our benchmark model, we do not report higher-order statistics for this factor.

[Table 5 about here.]

Table 5 indicates that the parameter estimates reported in Panel A result in economically meaningful conditional risk premia. The loading of the market risk premium on each instrument is positive. This result is consistent with the idea that as the state of the economy deteriorates and systematic volatility and tail risks increase, the expected market return increases. The loading of the VAR risk premium on VIX_t^2 is negative, indicating that the conditional variance risk premium decreases and becomes more negative as volatility increases. This is consistent with a wealth of empirical evidence and it is a model feature in many reduced form option pricing models as well as more structural models. Furthermore, our VAR premium specification is confirmed by the value of the F-statistic (p-value 0.049).[29]

Panel B of Table 5 provides additional insights about the distribution of daily conditional risk premia. The average market risk premium implied by our model is 9.1% and the variance risk premium is -22.1%. The risk premium associated with gamma risk is negligible.

Daily risk premium estimates of are highly leptokurtic. Impressively, the annualized market and variance risk premium have the anticipated sign and their magnitudes are reasonable, despite the fact that the sample includes both the bursting of the dot-com bubble and the financial crisis.

To further evaluate the plausibility of the average level of our estimated aggregate variance premium we confront it with other recent evidence. Note that the model-implied variance risk premium captures the impact of variance changes in the level of aggregate

[29] In contrast, the market risk premium's F-statistic is not statistically significant. We therefore also estimated a specification with a constant market risk premium. In unreported results, we find that this specification is statistically significant, but significantly hampers the model's ability to fit the dynamics of option returns.

variance on deleveraged option returns and as such, one needs to be careful when comparing it with the implications of no-arbitrage reduced-form option pricing models. For example, in Heston's model discussed in Section 2.3 the counterpart of our premium is $\frac{1}{dt}\left(E^P[dv_t] - E^P[E_t^Q[dv_t]]\right) = -\kappa^Q(\theta^Q - \theta^P)$. Using the estimates of Christoffersen, Fournier, and Jacobs (2018a), for example, yields a much smaller risk premium of approximately -2.5% per year. Two other interesting reference points that provide variance risk premia from asset returns are available. The first one is Cheng (2020), who reports an average VIX return premium of about -2.51% monthly using VIX futures data, which translates into a -30% annual premium. Ang, Hodrick, Xing, and Zhang (2006) report a -1% per-year unconditional premium on the VIX based on cross-sectional regressions with ΔVIX as a factor. The three estimates of the variance premium differ a great deal which emphasises the difficulty of obtaining reliable and interpretable estimates. Overall, our result is conceptually closest to Cheng (2020) because his study, similarly to ours, is focused on directly analysing the returns on derivatives, and our estimate is of a similar order of magnitude. We thus conclude that our estimate of the variance risk premium is plausible.

[Figure 2 about here.]

Figure 2 complements the results in Table 5 by plotting the time-series of annualized conditional risk premia, computed by annualizing daily estimates assuming 252 days in a year. We conclude that the estimated market and variance risk premiums not only have the anticipated signs on average in our sample, but that the signs are plausible at each point in time. This is particularly impressive given that we compute these risk premia at the daily frequency and that we use potentially noisy individual option returns in estimation. Figure 2 indicates that the risk premia are substantially higher during the financial crisis period and the sovereign debt crisis. We conclude that the cross-section of index option returns contain valuable information about market return risk premium dynamics.

6 Additional Results and Robustness Analysis

Tables 6 through 8 report on several additional results and robustness exercises. Our benchmark model includes three factors, the market return MKT, the market variance VAR, and the squared market return GAM. We first report on additional factors. We focus on the HKM factor from He et al. (2017) and the TVAR factor from Bollerslev et al. (2015). Subsequently we investigate the robustness of our results to changes in the specification of the functional form of the risk premia. Our third robustness exercise reports on the choice of proxy for the variance factor. The fourth robustness exercise reports on different sample periods.

[Table 6 about here.]

[Table 7 about here.]

6.1 Alternative Factors

Our benchmark model includes three factors: The market return MKT, the market variance VAR, and the squared market return GAM. Based on our prior intuition, and as validated by the empirical results, the MKT and VAR factors ought to be included in any model of expected option returns because regardless of the other model factors, MKT and VAR will always account for the lion's share of the risk premia and return variation.

The third factor, the squared market return (GAM) does not account for a large share of expected option returns, and only marginally impacts the risk premium estimation for the other two factors. To demonstrate this, column (2) in Panel B of Table 6 reports results for the two-factor model with MKT and VAR as factors. The sample period and other aspects of model specification are identical to those of the benchmark model. A comparison with the results in Panel B of Table 4, which are repeated in column (1), shows that the risk premium on the MKT factor is somewhat larger (11.2% vs. 9.1%), while the risk premium on the VAR factor is somewhat smaller in absolute value (−20.0% vs. −22.1%).

The three-factor specification shows its superiority over the two-factor specification in the analysis of the accuracy of the conditional expected return prediction (untabulated results available from the authors upon request).

Next we report on additional factors. We focus on two interesting factors studied in the recent literature: the HKM factor from He et al. (2017) and the TVAR factor, i.e. the daily increment in the left-tail variance of Bollerslev et al. (2015). The HKM factor captures the importance of intermediaries in asset markets in general and derivatives markets in particular, see for instance Gârleanu et al. (2009), He and Krishnamurthy (2013), He et al. (2017). The TVAR factor builds on an extensive literature that shows that volatility factors are not sufficient to characterize the cross-section of option prices, and that jumps and tail risk explain a substantial amount of the variation in option prices. The specifications of factor premia for HKM and TVAR are as follows. For HKM we follow equation (12) in He et al. (2017) and set the premium to be an affine function of the inverse capital ratio of financial intermediaries. For TVAR we mimic our specification for other factors and set the premium to be an affine function of the level of tail variance.

Columns (3) and (4) of Panel B of Table 6 report on the risk premia associated with these two factors. In both cases we include HKM and TVAR as the fourth factor in a model with MKT, VAR, and GAM as the first three factors. As mentioned above, MKT and VAR are indispensable factors, GAM helps with accurate return prediction, and including HKM and TVAR by themselves gives them the best chance to make a big impact on the model. The findings are very interesting. First, both TVAR and HKM have the expected sign on average, negative in the case of TVAR and positive in the case of HKM. However, while HKM has the expected sign 100.0% of the time, the corresponding number for TVAR is only 44.0%. Note also that the MKT and VAR factors once again have the expected sign almost all the time. The average estimate of the HKM risk premium is 32.8%, similar to the estimate in He et al. (2017). The negative premium on TVAR (−7.1%) is rather small compared to the premium on VAR. We are not aware of existing studies that provide benchmark estimates of risk premiums for this factor.

Table 7 reports on option loadings, the Shapley-Owen R^2 values (Huettner and Sunder,

2012) and the contribution of the HKM and TVAR factors to expected returns for options with different moneyness and maturity. The contributions of both factors to expected returns are an order of magnitude smaller than the contributions of MKT and VAR.

Moreover, it is clear that HKM and TVAR affect the cross-section of options very differently. For HKM, call options with maturity above two months load on the factor positively, while call options with short maturities load negatively. For put options, this maturity pattern is only present in DOTM puts). HKM has significantly more explanatory power than TVAR across all options, and contributes more to explaining variation in call options than put options. TVAR only has limited explanatory power for short-maturity DOTM options.

Figure 3 complements Tables 6 and 7 by plotting the variation over time for the entire 1996-2019 sample period for TVAR, and for the 2000-2019 sample period for which HKM is available. As in Figure 2, we break the sample in three sub-periods and use different scaling on the Y-axes because of the very high premia during the financial crisis. For both factors, the absolute value of the risk premium is by far the largest in the financial crisis, but the time-series behavior of the risk premia is very different outside of the 2008-2009 period. The time series of the HKM risk premium is much more persistent and has relatively few outliers. Not surprisingly, the risk premium associated with the TVAR factor is very transitory and contains many very large outliers.

[Figure 3 about here.]

6.2 Alternative Model Specifications

While we use a linear specification for the benchmark model, our model allows for alternative assumptions on the functional form of the risk premia. Columns (2)-(4) of Panel A of Table 6 report on the robustness of the risk premium estimates to the model specification. All three specifications use the three-factor model with the MKT, VAR, and GAM factors, but we only report on MKT and VAR to save space. As discussed previously, the risk premia associated with the GAM factor are relatively small regardless of the specifications. For completeness, the results in column (3) of Panel A report the estimates from a specification that drops TVAR_t as a predictor of the market premium.

First consider the constant risk premium specification in column (2) of Table 6. These results suggest that the constant risk premium specification is misspecified. This is easily verified by further inspecting the performance of this model, such as the expected option returns (tabulated for the benchmark model in Table 5). We do not report these results to save space. The average risk premium associated with the MKT factor is relatively similar to the benchmark result, but the risk premium associated with the VAR factor is much closer to zero. In contrast, the VAR and MKT risk premiums in the linear and quadratic cases are relatively similar to the benchmark case. The signs of the market and variance premia are consistent with economic intuition in all specifications.

6.3 The Estimation Sample Period

Figure 2 clearly indicates the highly time-varying nature of the risk premia, and specifically the enormous increase in risk premia during the financial crisis. It is therefore natural to wonder how the financial crisis affects model estimates. We therefore report on three subsamples. The 1996-2007 sample does not include the crisis and is dominated by the period of great moderation after the burst of the dot-com bubble. The 2008-2019 contains the key events of the Great Financial Crisis. The sample denoted "Apr 2009-2019" starts in April 2009, the first month after the DJIA reached its lowest post-crisis level. All results are based on the three-factor model with MKT, VAR, and GAM with the baseline risk premium specification

Panel D of Table 6 shows that the estimates of the MKT premium for the 2008-2019 sample are similar to those of the longer benchmark sample. However, both risk premiums are larger in absolute value for the post-crisis Apr 2009-2019 sample, and the VAR premium is significantly more negative both samples that do not include the period prior to 2008. The pre-crisis 1996-2007 sample yields estimates of the MKT and VAR premia that are much closer to zero. In all samples that contain the financial crisis the premia estimates have the anticipated sign at every point in time, but this is not the case in the subsamples that exclude the crisis period. Finally, in all samples except for the post-crisis sample, the average MKT premium is close to the average realized excess return on the S&P 500 index for the relevant period.

6.4 Alternative Proxies for the Variance Factor

Consistent with our priors, the benchmark model indicates that both the MKT and VAR factors are associated with large risk premia. This result is robust to the inclusion of additional factors and the sample period used in estimation. It is also robust to model specification.

While the measurement of the MKT factor is straightforward, the same cannot be said about the VAR factor. We believe our choice of the squared VIX as a variance factor is a reasonable one, because it is a well-known measure of variance. Moreover, we avoid critiques of data-mining because the VIX series is given, in contrast to many measures of variance used in the literature that are often implemented using ancillary assumptions. While most of the resulting measures of variance are highly correlated, there are at least two very important issues with respect to measurement that need to be addressed in a robustness analysis. The first issue is the choice between a measure of the physical versus the risk-neutral variance. The VIX is a measure of the risk-neutral variance and therefore already includes a risk premium. It is worth investigating if measures of physical variance lead to similar results. Second, most variance estimates include a diffusive as well as a jump component, and it is useful to verify if removing the jump component changes inference on variance risk premium estimates.

To answer these questions, we re-estimate our baseline three-factor model using alternative proxies for the variance factor. The four proxy factors are constructed from daily increments in measures of the physical variance of index returns. The first proxy, labeled SPOT, is based on the non-parametric volatility estimate of Todorov (2019). While this quantity is obtained from option prices, it measures the physical volatility of the Brownian component of index returns. The second proxy, labeled GARCH, is based on the filtered conditional variance series from a GJR-GARCH(1,1) model (Glosten, Jagannathan, and Runkle, 1993) estimated on our sample of index returns. The remaining two proxies are realized variance (RV) and integrated quadratic variation (IQV), which are obtained from intraday high-frequency return data. While RV is an estimate of the total instantaneous variance of index returns, IQV is an estimate of the Brownian component of return variance, i.e. it estimates the same concept as the SPOT proxy.[30] Thus, GARCH and RV contain both the continuous and jump variance components while SPOT and IQV only contain the continuous component. We report on four aspects of using alternative variance proxies: (a) the time-series properties of the proxies, (b) the estimates of the equity and variance-proxy premia, (c) their explanatory power for option returns, and (d) the accuracy of the resulting expected option returns.

Table 8 contains a summary of our findings. The time-series properties of the variance proxies are summarised in Panel A. SPOT and GARCH are highly correlated with our baseline VAR factor and have similar volatility-of-volatility as well as kurtosis.

RV and IQV are markedly different from VAR and SPOT. Their correlation with VAR and SPOT is much lower, ranging from 17% (with SPOT) to 32% (with VAR).[31] The vol-of-vol and kurtosis levels are much higher than those of the VAR proxy. This high variability of RV and IQV suggests they are noisy estimates of aggregate asset market variance. Finally, GARCH is highly correlated with VAR and SPOT, but its high positive skewness is similar to that of the RV and IQV proxies.

Panel B of Table 8 reports summary statistics of risk premia estimates for the MKT factor and the various proxies for the variance factor. The time-series properties of the MKT premium are virtually unchanged when VAR is replaced by SPOT or GARCH. In contrast, with RV and IQV, the estimate of the MKT premium is much lower on average, becomes negatively skewed with heavier tails, and, most importantly, is positive only 68% (86%) of the time with RV (IQV). The properties of the premium on the aggregate variance factor are similar for VAR, SPOT and GARCH (with averages of −51.6% for SPOT and −36.8% for GARCH) but much larger (in absolute value) for IQV and RV

[30]The data for SPOT were obtained from www.tailindex.com. The estimates of realized variance and integrated quadratic used for RV and IQV were obtained from https://realized.oxford-man.ox.ac.uk.

[31]Note that the correlation coefficients reported in Panel A of Table 8 are calculated from daily increments in the variance proxy series. The correlations of the levels of the five variance proxy series are much higher and range from X to X.

(−62.8% and −664.0%, respectively). However, it is only for the SPOT proxy that the conditional variance premium estimate is negative at each point in time, as is the case for the benchmark VAR proxy.

Panel C of Table 8 reports the explanatory power of the various variance factor proxies for option returns. GARCH, RV and IQV only explain between 2% and 3% of option return variation, in contrast with VAR which explains 23% (36%) of call (put) return variation. SPOT accounts for 10% (24%) of call (put) return variation. Note that while the average sensitivities of the option returns to VAR and SPOT are similar, the estimated sensitivities to the remaining proxies are an order of magnitude higher. This fact, combined with the low explanatory power and high vol-of-vol suggests that these factors may be imprecise estimates of aggregate market variance.

Finally, Panel D of Table 8 uses Mincer-Zarnowitz regressions to examine how well the alternative variance proxies fit the conditional dynamics of option returns. The results suggest that SPOT captures the dynamics of put options rather well. IQV captures the conditional dynamics of both puts and calls, but fails to match the average return level. GARCH and RV fail on both measures of return fit.

We conclude that alternatives to our proxy for VAR in our benchmark model, the daily increment in VIX^2, lead to similar conclusions about the dynamics of risk premia, but VAR delivers results that are more closely aligned with economic priors. Variance factors recovered from option prices contain information that is not available in variance factors exclusively based on low-frequency (GARCH) or high-frequency (IQV and RV) return data. Finally, the implications of our approach are robust to using different proxies of aggregate asset variance, provided the proxies are not very noisy, as is the case with RV for example.

[Table 8 about here.]

7 Conclusion

We propose and implement a novel dynamic factor model for option returns. Our framework is very general. Option exposures to a given factor are modeled nonparameterically using parsimonious basis expansions of moneyness, maturity, and conditional volatility. The model provides estimates of the conditional risk premium of various economic factors and the time-varying exposure of options with different moneyness and maturity to those factors.

We illustrate the value of our approach with an implementation on a 22-year sample of returns on index option calls and puts with various moneyness and maturity. Our benchmark model contains a market factor, a factor capturing variance risk, and a gamma factor. We also investigate the intermediary risk factor of He et al. (2017) and a tail risk factor (Bollerslev et al., 2015). The signs of the average risk premia of the factors are consistent with economic intuition, the magnitudes are plausible, and the premia spike during crises. The risk premiums on the market and variance factors are by far the largest, and they have the expected sign on every day in the sample. The model explains expected option returns well across moneyness and maturities, and its hedging performance is impressive.

Several issues are left for future research. First, it may be useful to extend the model by allowing for a more refined aggregate jump factor (see e.g., Bates, 2008, Bollerslev and Todorov, 2011, Kelly, Lustig, and Van Nieuwerburgh, 2016) or to account for frictions in option markets (see e.g., Bollen and Whaley, 2004, Bates, 2003, Gârleanu et al., 2009). Second, estimating the model by combining index option and equity option returns data may provide many additional insights as well as improved conditional risk premium estimates (see e.g. Bégin et al., 2020, Duarte et al., 2019). Finally, the model can be used to estimate and analyze the conditional risk premia implied by other asset classes, such as corporate bonds (see e.g., Bai, Bali, and Wen, 2019).

Appendices

A Thin Plate Regression Spline Basis (TPRS basis)

Our approach to the modeling of option exposures builds on Thin Plate Regression Spline basis (TPRS-basis). Basis expansion theory is based on the idea that any sufficiently well-behaved and smooth function can be expressed as a linear combination of basis functions. It is important to bear in mind that the theoretical expression of a given basis expansion does not depend on whether or not realizations of the function to be estimated are directly observable. Any basis expansion is a combination of basis functions and loadings parameters. The particular form of the basis functions composing a given basis expansion is invariant with respect to the function to be approximated. Thus, one only needs a valid measurement equation for estimating the parameter loadings of the basis expansion whether or not the measurement equation is directly or indirectly related to the function of interest. In our case, the functions of interest to be estimated are option exposures and the measurement equation used to estimate these exposures (i.e., the basis parameters) corresponds to the dynamics of deleveraged option returns (2.13) of the paper, which are indirectly related to option exposures.

Splines are functions constructed by joining sections of simpler functions that belong to a specific family (e.g., cubic polynomials). The points at which functions meet are usually referred to as knots. Ultimately, the functions are joined in such a way so that the resulting target function and its derivatives up to a given order are continuous.

For illustrative purposes, let us denote by $\tilde{\beta}_l\left(\mathbf{z}_{O,l,t}\right)$ the true exposure of option O to factor l given the vector of signals $\mathbf{z}_{O,l,t}$. Once a basis of a given dimension κ is chosen, the value of $\tilde{\beta}_l\left(\cdot\right)$ at each of the N_l (i.e., $N_l \equiv \sum_{O,t}$) observations in $\{\mathbf{z}_{O,l,t}\}_{\forall O,t} \equiv Z_l$ can be approximated as a linear combination of the κ basis functions. For $\mathbf{z} \in Z_l$, we have

$$\tilde{\beta}_l\left(\mathbf{z}\right) \approx \beta_l\left(\mathbf{z}\right) \equiv \sum_{j=1}^{\kappa} \phi^j\left(\mathbf{z}\right) b_l^j. \tag{AO.1}$$

The literature differentiates between smoothing splines and regression splines. In smoothing splines, all the elements in Z_l are used as knots and $\kappa \geq N_l$. The econometrician thus needs to rely on penalization methods such as ridge regressions for estimating b_l^j. In regression splines, κ is a choice variable set by the econometrician to some relatively low number as to avoid overfitting. In regression splines, estimation does not require penalization methods which in turn makes inference easier. In some regression spline methods, the econometrician may also need to select the κ points in the domain of the function $\tilde{\beta}_l(\cdot)$ that will serve as knots.

In our context, an appropriate basis expansion should achieve two key objectives. As we discuss in Section 2.2 of the main paper, option pricing theory suggests that using a multidimensional vector $\mathbf{z}_{O,l,t}$ to estimate exposures is important. As a result, a good basis expansion for option exposures needs to efficiently handle a multidimensional vector of signals. Second, the basis expansion chosen should avoid overfitting, be computationally efficient, and limit ad-hoc specifications including ad-hoc choices of knots. To meet these objectives, we choose the Thin Plate Regression Spline Basis (TPRS-basis) which was originally developed by Wood (2003). The TPRS-basis expansion extends Thin Plate Smoothing Splines (TPSS) by allowing for a low dimensional basis representation of the function of interest (i.e., low κ). TPSS in turn is a multivariate extension of Cubic Smoothing Splines (CSS). We first introduce the foundations of the TPSS-basis. We then extend our discussion to the TPRS-basis.

Both CSS and TPSS builds on the notion of *roughness* of the estimated function. In the univariate case when the dimension of z is 1, the roughness of β_l can be measured by

$$\int_{z\in D} \left[\frac{d^m \beta_l(z)}{(dz)^m}\right]^2 dz, \tag{AO.2}$$

where D is the domain of β_l. Intuitively, equation (AO.2) measures the integrated variance of the m^{th} derivative of β_l. The notion of *roughness* as defined by equation (AO.2) is thus inversely related to the level of smoothness of β_l. From a fitting perspective, excessive *roughness* of the estimated function is indicative of the degree of overfitting. The CSS basis expansion is obtained by solving a functional optimization problem by minimizing the sum of squared basis residuals $\sum_{\mathbf{z}\in Z_l}\left(\tilde{\beta}_l(z)-\beta_l(z)\right)^2$ over all possible well-behaved functions $\beta_l(\cdot)$ subject to the constraint (AO.2) when z is a scalar and m is set to 2 in (AO.2). TPSS extends the CSS algorithm by allowing for multivariate signals and by providing a general solution of the mean square problem taking into consideration all derivatives of β_l in (AO.2) of the m^{th}-order which is also function of cross-derivatives in the multivariate $\mathbf{z}$ case.[32] The optimal solution of this problem gives the following TPSS-basis approximation of the function $\widetilde{\beta}_l:\mathbb{R}^d\to\mathbb{R}$

$$\tilde{\beta}_l(\mathbf{z})=\sum_{s=1}^{M}\omega_{s,l}\cdot\nu_s(\mathbf{z})+\sum_{u=1}^{N_l}\delta_{u,l}\cdot f_{m,d}(\mathbf{z},\mathbf{z}_u), \tag{AO.3}$$

where $\nu_s(\cdot):\mathbb{R}^d\to\mathbb{R}$ and $f_{m,d}(\cdot,\mathbf{z}_u):\mathbb{R}^d\to\mathbb{R}$ are basis functions which exact definitions are provided below, and $\omega_{s,l}$ and $\delta_{u,l}$ are the basis parameters to be estimated. Equation (AO.3) can be rewritten in matrix form as

$$\widetilde{\beta}_l=\mathbf{V}\cdot\mathbf{w}_l+\mathbf{F}\cdot\delta_l+\varepsilon^{TPSS}, \tag{AO.4}$$

where $\mathbf{V}$ is an $N_l\times M$ matrix with elements $V_{ij}=\nu_j(\mathbf{z}_i)$ for $i=1,\ldots,N$ and $j=1,\ldots,M$, and $\mathbf{F}$ is an $N_l\times N_l$ matrix with elements $F_{ij}=f_d(\mathbf{z}_i,\mathbf{z}_j)$ for $i,j=1,\ldots,N_l$. In the previous equation, $\mathbf{w}_l$ and δ_l are two vectors of coefficients and $\widetilde{\beta}_l$ is the $N_l\times 1$ vector of true exposures.[33] Moreover, $M=\frac{(m+d-1)!}{d!(m-1)!}$ where $m>d/2$ is an integer that defines the order of the derivatives taken into account for the "*roughness*" penalization to obtain the TPSS solution.

In equations (AO.3) and (AO.4), $\nu_s(\cdot)$ are all the monomials of the elements of $\mathbf{z}$ and of order smaller than or equal to $m-1$.[34] Recall that our benchmark specification uses $\mathbf{z}_{O,l,t}=[k_{O,l,t}\ \tau_{O,l,t}\ v_t]$ and thus $d=3$. In addition, we consider the case of $m=2$ (i.e.,

[32] In the two predictors case with $d=2$ and with roughness measured using the second derivative ($m=2$), we have $\int\int\left(\left[\frac{d^2\beta_l(\mathbf{z})}{(dz_1)^2}\right]^2+\left[\frac{d^2\beta_l(\mathbf{z})}{dz_1dz_2}\right]^2+\left[\frac{d^2\beta_l(\mathbf{z})}{(dz_2)^2}\right]^2\right)dz_1dz_2.$

[33] To ensure the uniqueness of the coefficients, it is necessary that $\sum_{u=1}^{N}\delta_{u,l}\nu_s(\mathbf{z}_u)=0$ for all s such that $\mathbf{V}'\delta_l=\mathbf{0}_{M\times 1}$. This restriction in turn guarantees that the penalized least square criterion $\varepsilon^{TPSS'}\varepsilon^{TPSS}$ is positive definite and admits a unique solution (i.e., the vectors $\mathbf{w}_l$ and δ_l are unique).

[34] Monomials are polynomials with just one term in each dimension of a given order. For example when $\mathbf{z}$ is of dimension 3, one monomial of $\mathbf{z}=[z_1\ z_2\ z_3]$ is $z_1^3z_2^2z_3^{74}$ for instance.

quadratic "*roughness*" penalization) which together with d implies that $M = 4$. In this case, we have $\nu_1(\mathbf{z}_{O,l,t}) = 1$, $\nu_2(\mathbf{z}_{O,l,t}) = k_{O,l,t}$, $\nu_3(\mathbf{z}_{O,l,t}) = \tau_{O,l,t}$, and $\nu_4(\mathbf{z}_{O,l,t}) = v_t$.[35] The N_l functions $f_{m,d}(\cdot, \mathbf{z}_u)$ are then defined as

$$f_{m,d}(\mathbf{z}, \mathbf{z}_u) = \begin{cases} \theta_{e,d} \, ||\mathbf{z} - \mathbf{z}_u||^{2m-d} \ln(||\mathbf{z} - \mathbf{z}_u||) & \text{for } d \text{ even} \\ \theta_{o,d} \, ||\mathbf{z} - \mathbf{z}_u||^{2m-d} & \text{for } d \text{ odd} \end{cases}, \tag{AO.5}$$

for $\mathbf{z} \in Z_l$, and where $||\cdot||$ denotes the Euclidian norm. In the previous definition, the constants $\theta_{e,d}$ and $\theta_{o,d}$ are given by

$$\begin{aligned} \theta_{e,d} &= \tfrac{(-1)^{m+1+d/2}}{2^{2m-1}\pi^{d/2}(m-1)!(m-d/2)!} \\ \theta_{o,d} &= \tfrac{\Gamma(d/2-m)}{2^{2m}\pi^{d/2}(m-1)!}. \end{aligned} \tag{AO.6}$$

There are two important drawbacks of the TPSS-basis. First, when the dimension d of $\mathbf{z}$ is greater than 2 the evaluation of the basis functions becomes computationally expensive for large data panels. Second, penalized estimation becomes necessary which in turn makes inference challenging. Both problems are alleviated by the TPRS-basis. The TPRS-basis chooses a reduced-dimension basis among all the functions $f_{m,d}(\cdot, \mathbf{z}_u)$ while keeping $\nu_s(\cdot)$ unchanged. As a result, the TPRS-basis is much less computationally demanding than TPSS-basis and its low-dimensional nature allows for OLS estimation of basis parameters without relying on penalization methods as the dimension reduction is already built into the algorithm.

An *ad-hoc* TPRS-basis can be constructed by first choosing κ distinct points (knots) from Z_l, $\mathbf{z}_1$ through $\mathbf{z}_\kappa$ which then define the κ basis functions $f_{m,d}(\cdot, \mathbf{z}_u)$, $u = 1, \ldots, \kappa$. In a hypothetical situation where $\widetilde{\beta}_l$ is directly observed, a regression model for $\widetilde{\beta}_l$ can be formulated based on equation (AO.3) for $\mathbf{z}_{O,l,t} \in Z_l$ as

$$\widetilde{\beta}_l(\mathbf{z}_{O,l,t}) = \sum_{s=1}^{M} w_{s,l} \cdot \nu_s(\mathbf{z}_{O,l,t}) + \sum_{u=1}^{\kappa} \delta_{\kappa,u,l} \cdot f_{m,d}(\mathbf{z}_{O,l,t}, \mathbf{z}_u) + \varepsilon_{O,l,t}^{TPRS}. \tag{AO.7}$$

Defining the $N_l \times \kappa$ matrix $\widetilde{\mathbf{F}}_\kappa$ with element $\widetilde{F}_{ij} = f_d(\mathbf{z}_i, \mathbf{z}_u)$ with $i = 1, \ldots, N_l$ where

[35] Note that the functions $\nu_s(\cdot)$ have a *roughness* equal to 0 by construction and thus are not penalized.

a given $\mathbf{z}_i$ is a particular realization of $\mathbf{z}_{O,l,t}$ in Z_l and $j = 1, \ldots, \kappa$. One can rewrite equation (AO.7) as

$$\widetilde{\beta}_l = \mathbf{V} \cdot \mathbf{w}_l + \widetilde{\mathbf{F}}_\kappa \cdot \delta_{\kappa,l} + \varepsilon^{TPRS}. \tag{AO.8}$$

Because an inappropriate choice of the κ knots may significantly impacts model fit, Wood (2003) develops a general algorithm that built on an optimal basis derived from the spectral approximation of $\mathbf{F}$ while imposing the uniqueness restriction $\sum_{u=1}^{N_l} \delta_{u,l} \nu_s(\mathbf{z}_u) = 0$, $\mathbf{V}'\delta_l = \mathbf{0}_{M\times 1}$.

For a given basis dimension κ, the goal of the TPRS-algorithm is to construct based on Z_l, $\kappa - M$ basis functions which, together with $\nu_s(\cdot)$, summarizes the information from the complete TPSS-basis represented by $\mathbf{F}$. The guiding principle is that the approximating low-dimensional basis $\widetilde{\mathbf{F}}_\kappa$ should allow for a minimal change in model fit from the full TPSS basis as described in Wood (2003). The construction starts with the eigen decomposition of the matrix $\mathbf{F} = \mathbf{U}\mathbf{D}\mathbf{U}'$ where $\mathbf{D}$ is the diagonal matrix of eigenvalues and $\mathbf{U}$ contains the eigenvectors. The eigenvalues in $\mathbf{D}$ are sorted so that $|D_{i,i}| \geq |D_{i+1,i+1}|$, and the eigenvectors in $\mathbf{U}$ have the corresponding order. The spectral approximation of $\mathbf{F}$ of rank κ satisfies

$$\widehat{\mathbf{F}}_\kappa = \mathbf{U}_\kappa \mathbf{D}_\kappa \mathbf{U}_\kappa', \tag{AO.9}$$

where the columns of $\mathbf{U}_\kappa$ contain the first κ eigenvectors associated with the first κ eigenvalues and $\mathbf{D}_\kappa$ is the diagonal matrix of the first κ eigenvalues. For uniqueness of the TPRS-basis representation, one still need to impose $\mathbf{V}'\delta_l = \mathbf{0}_{M\times 1}$. To this end, note that the κ columns of the $N_l \times \kappa$ matrix $\mathbf{U}_\kappa$ form an orthogonal basis. The dimension reduction of the TPRS-basis constraints the TPSS-basis coefficient vector according to $\delta_l = \mathbf{U}_\kappa \delta_{\kappa,l}$. We thus have

$$\begin{aligned}\widetilde{\beta}_l &= \mathbf{V} \cdot \mathbf{w}_l + \widehat{\mathbf{F}}_\kappa \cdot \mathbf{U}_\kappa \delta_{\kappa,l} + \varepsilon^{TPRS} \\ &= \mathbf{V} \cdot \mathbf{w}_l + \mathbf{U}_\kappa \mathbf{D}_\kappa \cdot \delta_{\kappa,l} + \varepsilon^{TPRS},\end{aligned} \tag{AO.10}$$

under the constraint $\mathbf{V}'\delta_l = \mathbf{V}'\mathbf{U}_\kappa \delta_{\kappa,l} = \mathbf{0}_{M\times 1}$. Up to this point, the TPRS-basis replaced $\mathbf{F}$ with $\mathbf{U}_\kappa \mathbf{D}_\kappa$ of dimension $N_l \times \kappa$ while keeping the basis functions $\nu_s(\cdot)$ of dimension M

unchanged. This implies that the total dimension of the basis equal to $\kappa+M$. To arrive at the final dimension κ, the TPRS-basis exploits the constraint $\mathbf{V}'\mathbf{U}_\kappa\delta_{\kappa,l} = \mathbf{0}_{M\times 1}$ to restrict possible parameter vectors to a lower-dimensional space. To do this, TPRS-basis finds an orthogonal column basis $\mathbf{Q}_{\kappa-M}$ which solves $\mathbf{V}'\mathbf{U}_\kappa\mathbf{Q}_{\kappa-M} = \mathbf{0}_{M\times(\kappa-M)}$. Then TPRS-basis represents the coefficient vector $\delta_{\kappa,l}$ as $\delta_{\kappa,l} = \mathbf{Q}_{\kappa-M}\delta_{\kappa-M,l}$ where $\mathbf{V}'\mathbf{U}_\kappa\mathbf{Q}_{\kappa-M}\delta_{\kappa-M,l} = \mathbf{0}_{M\times 1}$ is satisfied for any $\delta_{\kappa-M,l}$ by construction. A $\mathbf{Q}_{\kappa-M}$ can then be easily found from the QR decomposition of $\mathbf{U}_\kappa'\mathbf{V}$. It is given by the final $\kappa - M$ columns of the orthogonal factor $\mathbf{Q}$.[36] Armed with $\mathbf{Q}_{\kappa-M}$, the TPRS-basis representation can be written (without constraints) as

$$\begin{aligned}
\widetilde{\beta}_l &= \mathbf{V}\cdot\mathbf{w}_l + \mathbf{U}_\kappa\mathbf{D}_\kappa\mathbf{U}_\kappa'\cdot\mathbf{U}_\kappa\mathbf{Q}_{\kappa-M}\delta_{\kappa-M,l} + \varepsilon^{TPRS} \\
&= \mathbf{V}\cdot\mathbf{w}_l + \mathbf{U}_\kappa\mathbf{D}_\kappa\mathbf{Q}_{\kappa-M}\cdot\delta_{\kappa-M,l} + \varepsilon^{TPRS} \\
&= \mathbf{V}\cdot\mathbf{w}_l + \mathbf{F}_{\kappa-M}\cdot\delta_{\kappa-M,l} + \varepsilon^{TPRS},
\end{aligned} \tag{AO.11}$$

where we have used the fact that $\mathbf{U}_\kappa'\cdot\mathbf{U}_\kappa$ is equal to the identity matrix to obtain the second equality, defined $\mathbf{F}_{\kappa-M} \equiv \mathbf{U}_\kappa\mathbf{D}_\kappa\mathbf{Q}_{\kappa-M}$ in the third, and where $\delta_{\kappa-M,l}$ is the matrix of TPRS-basis parameters of dimension $N_l\times(\kappa-M)$. Each row i of the matrix is constructed such that, after multiplying it by the parameter vector $\delta_{\kappa-M}$, it approximates the i-th row of the full basis matrix $\mathbf{F}$ multiplied by the parameter vector $\delta_l = \mathbf{U}_\kappa\mathbf{Q}_{\kappa-M}\delta_{\kappa-M,l}$. Each element $\mathbf{F}_{\kappa-M,ij}$ of the matrix $\mathbf{F}_{\kappa-M}$ is a linear combination of the i-th row of the rank-κ spectral approximation of $\mathbf{F}$ where the coefficients of the linear combination are given by the j-th column of $\mathbf{U}_\kappa\mathbf{Q}_{\kappa-M}$.

We can re-write equation (AO.11) as equation (2.12) in the paper. We have

$$\beta_l\left(\mathbf{z}_{O,l,t}\right) = \begin{bmatrix} \mathbf{V} \\ \mathbf{F}_{\kappa-M} \end{bmatrix}' \cdot \begin{bmatrix} \mathbf{w}_l \\ \delta_{\kappa-M,l} \end{bmatrix} = \phi_{O,l,t}'\mathbf{b}_l = \sum_{j=1}^{k}\phi_{O,l,t}^j b_l^j. \tag{AO.12}$$

To implement the TRPS-basis expansion, we use the `R` package `mgcv` version `1.8.26`

[36]The QR decomposition of a matrix $\boldsymbol{X}$ factorizes $\boldsymbol{X} = \boldsymbol{QR}$ where $\boldsymbol{Q}$ is an orthogonal matrix and $\boldsymbol{R}$ is an upper triangular matrix.

which directly builds on Wood (2003) and Wood (2017).

B Estimation and inference

B.1 Estimation of factor loadings

Recall, with the definitions from Proposition 1 in the main text, we can write

$$R^O_{t+\Delta t} = \left(\boldsymbol{\phi}^g_{O,t}\right)^\intercal \boldsymbol{b}^Q + \left(\boldsymbol{\phi}^f_{O,t+\Delta t}\right)^\intercal \boldsymbol{b} + \varepsilon^O_{t+\Delta t}.$$

In this section we present the results for statistical inference in the first-stage estimation. The results are given for a fixed order k of the TPRS basis. We account for time-series and cross-sectional correlation in the residuals $\varepsilon^O_{t+\Delta t}$.

Assumptions 1 *In what follows, we maintain the following assumptions:*

(i) $E\left[\varepsilon^O_{t+\Delta t}|\phi^{fg}_{O,t+\Delta t}\right] = 0$

(ii) $\varepsilon^O_{t+\Delta t}$ *is spatially strong mixing, as in Driscoll and Kraay (1998).*

Assumption (i) means that, in particular, which options are actively traded on a given day does not depend on instrument and factor realizations.[37]

Proposition 5 (Estimator of $\boldsymbol{\Sigma}_{\phi\varepsilon}$) *Define* $\boldsymbol{h}_t \equiv \frac{1}{N_t}\sum_O \boldsymbol{\phi}^{fg}_{O,t+\Delta t}\varepsilon^O_{t+\Delta t}$ *and set S, a positive integer lag order. Under assumption (ii),* $\boldsymbol{\Sigma}_{\phi\varepsilon}$ *can be estimated with*

$$\widehat{\boldsymbol{\Sigma}}_{\phi\varepsilon} \equiv \frac{1}{T}\sum_t \boldsymbol{h}_t\boldsymbol{h}_t^\intercal + \frac{1}{T}\sum_{s=1}^{S}\sum_{t=s+1}^{T} w_s\left(\boldsymbol{h}_t\boldsymbol{h}_{t-l}^\intercal + \boldsymbol{h}_{t-l}\boldsymbol{h}_t^\intercal\right),$$

and, $w_s = 1 - \frac{s}{S+1}$*, which is the Newey and West (1987) estimator applied to the time series of* $\boldsymbol{h}_t$*.*

With these ingredients we present the immediate result which allows for inference about the statistical significance of a given factor for explaining deleveraged excess option returns. Because option exposures $\widehat{\beta}^O_{l,t}$ are linear transformations of the estimated vector

[37] Arguably, this assumption would be tenuous if we chose to use the whole option surface in our estimation. However, in our dataset we limit the strike ranges and option maturities in such a way as to eliminate this possible source of bias.

of parameters $\widehat{\boldsymbol{b}}_l$, the asymptotic distribution of $\beta^O_{l,t}$ inherits the distributional properties of $\widehat{\boldsymbol{b}}_l$.

Proposition 6 *With deleveraged excess option returns satisfying* (19) *and option exposures defined by* (12)*, the asymptotic distribution of option O's exposure to factor l on day t satisfies*

$$\widehat{\beta}^O_{l,t} = \boldsymbol{\phi}'_{O,l,t}\widehat{\boldsymbol{b}}_l \sim N\left(\boldsymbol{\phi}'_{O,l,t}\boldsymbol{b}_l; \boldsymbol{\phi}'_{O,l,t}\boldsymbol{\Sigma}_{\widehat{\boldsymbol{b}}_l}\boldsymbol{\phi}_{O,l,t}\right), \tag{28}$$

where $\boldsymbol{b}_l$ and $\boldsymbol{\Sigma}_{\boldsymbol{b}_l}$ are the appropriate subvector and submatrix of $\boldsymbol{B}$ and $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{B}}}$, respectively. Moreover, the joint hypothesis test that the parameters in $\boldsymbol{b}_l$ are jointly equal to zero (i.e., factor l does not influence the dynamics of option returns) can be tested by the following Wald statistic:

$$\widehat{\boldsymbol{b}}'_l\widehat{\boldsymbol{\Sigma}}^{-1}_{\widehat{\boldsymbol{b}}_l}\widehat{\boldsymbol{b}}_l \sim \boldsymbol{\chi}^2_{\boldsymbol{k}}. \tag{29}$$

Proof. Definition (12) combined with equations (**??**), (**??**), and standard results for linear models gives (28) and (29). ■

B.2 Estimation of risk premia

In this section we present the results allowing for inference about risk premia parameters. Due to multicollinearity problems which we observed in our data (with strongly correlated MKT and VAR factors), it is useful to split the standard second-stage estimation into two independent sub-stages: the premium for tradable factors (MKT), and the premium for non-tradable factors. Inference in the two sub-stages is identical after accounting for the different regressors carried over from the first stage. We first lay out the complete set of results for the estimation of premia on tradable factors in Section B.2.1. Then, in Section B.2.2, we indicate what changes when the results of Section B.2.1 are applied to non-tradable factors. Our results are fundamentally analogous to those obtained by Shanken (1992) for two-stage estimation of risk premia.[38]

We rely on the following assumptions.

[38] Shanken's results also allow for estimating the premia on tradable and non-tradable factors separately, albeit in the case of an unconditional factor model with constant risk premia this approach benefits from the fact that premia on tradable factors can be estimated by taking sample averages of the factor realizations.

Assumptions 2 *In what follows, we maintain the following assumptions:*

(iii) $E\left[\eta^{\Delta}_{O,t+\Delta t}|\phi^{fg}_{O,t+\Delta t}\right]=0$

(iv) $\eta^{\Delta}_{O,t+\Delta t}$ *is spatially strong mixing, as in Driscoll and Kraay (1998).*

B.2.1 Premia on tradable factors

In this sub-stage we first project the realized option returns on the tradable factors, and then regress these projections on the predictors of risk premia for tradable factors. Without loss of generality assume that factors 1 through D are tradable. We denote the projection as

$$R^{O\Delta}_{t+\Delta t}\equiv\sum_{k=1}^{D}\beta^{O}_{kt}f_{kt+\Delta t}\,.$$

Let $\boldsymbol{d}^{\Delta}_{O,t}\equiv\begin{bmatrix}\beta^{O}_{1t}\boldsymbol{g}^{\intercal}_{1t} & \ldots & \beta^{O}_{Dt}\boldsymbol{g}^{\intercal}_{Mt}\end{bmatrix}$, and let $\widehat{\boldsymbol{d}}^{\Delta}_{O,t}$ be its sample analogue, i.e., with $\widehat{\beta}^{O}_{kt}$ replacing β^{O}_{kt}. Then the following relationship holds between the projection of option returns on the tradable factors and the risk premia predictors,

$$R^{O\Delta}_{t+\Delta t}=\left(\boldsymbol{d}^{\Delta}_{O,t}\right)^{\intercal}\boldsymbol{\lambda}_{\Delta}+\eta^{\Delta}_{O,t+\Delta t}\,.$$

In fact, with the true factor loadings unknown, we estimate $\boldsymbol{\lambda}_{\Delta}$ based on the feasible version of the above equation,

$$\widehat{R}^{O\Delta}_{t+\Delta t}\equiv\sum_{k=1}^{D}\widehat{\beta}^{O}_{kt}f_{kt+\Delta t}=\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\right)^{\intercal}\boldsymbol{\lambda}_{\Delta}+\widetilde{\eta}^{\Delta}_{O,t+\Delta t}\,. \tag{30}$$

In the above regression equation, the quantities on both sides contain measurement noise, which has to be accounted for in subsequent inference. There are two consequences of the measurement noise: Asymptotic bias, and increases in the variance of the estimated coefficients.

Proposition 7 (Estimator of $\boldsymbol{\lambda}_{\Delta}$) *The OLS estimator of $\boldsymbol{\lambda}_{\Delta}$ is given by*

$$\widehat{\boldsymbol{\lambda}}_{\Delta}\equiv\boldsymbol{Q}^{-1}_{\widehat{d}^{\Delta}}\frac{1}{T}\sum_{t}\frac{1}{N_t}\sum_{O}\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\cdot\left(\boldsymbol{\phi}^{f_{\Delta}}_{O,t+\Delta t}\right)^{\intercal}\widehat{\boldsymbol{b}}_{\Delta}\,,$$

where $\widehat{\boldsymbol{b}}_\Delta$ *is the subvector of* $\widehat{\boldsymbol{b}}$ *which contains the loadings on the tradable factors,* $\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}$ *is the subvector of the vector* $\boldsymbol{\phi}^{f}_{O,t+\Delta t}$ *with entries related to tradable factors, and* $\boldsymbol{Q}_{\widehat{\boldsymbol{d}}^\Delta} \equiv \frac{1}{T}\sum_t \frac{1}{N_t}\sum_O \widehat{\boldsymbol{d}}^\Delta_{O,t}\left(\widehat{\boldsymbol{d}}^\Delta_{O,t}\right)^\intercal$.

Corollary 1 (Estimation error in $\widehat{\boldsymbol{\lambda}}_\Delta$) *The following decomposition of estimation error in* $\widehat{\boldsymbol{\lambda}}_\Delta$ *holds:*

$$\sqrt{T}\left(\widehat{\boldsymbol{\lambda}}_\Delta - \boldsymbol{\lambda}_\Delta\right) = \boldsymbol{Q}^{-1}_{\widehat{\boldsymbol{d}}^\Delta}\frac{1}{\sqrt{T}}\sum_t \frac{1}{N_t}\sum_O \Big[\boldsymbol{d}^\Delta_{O,t}\eta^\Delta_{O,t+\Delta t} + \boldsymbol{d}^\Delta_{O,t}\left(\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\right)^\intercal\left(\widehat{\boldsymbol{b}}_\Delta - \boldsymbol{b}_\Delta\right)$$
$$+\left(\widehat{\boldsymbol{d}}^\Delta_{O,t} - \boldsymbol{d}^\Delta_{O,t}\right)\left(\boldsymbol{d}^\Delta_{O,t}\right)^\intercal \lambda_\Delta + \left(\widehat{\boldsymbol{d}}^\Delta_{O,t} - \boldsymbol{d}^\Delta_{O,t}\right)\eta^\Delta_{O,t+\Delta t}$$
$$+\left(\widehat{\boldsymbol{d}}^\Delta_{O,t} - \boldsymbol{d}^\Delta_{O,t}\right)\left(\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\right)^\intercal\left(\widehat{\boldsymbol{b}}_\Delta - \boldsymbol{b}_\Delta\right)\Big].$$

Furthermore, all terms in the decomposition, except for the last one, are zero-expectation.

Proof. The decomposition follows from observing that $\widehat{\boldsymbol{d}}^\Delta_{O,t} = \boldsymbol{d}^\Delta_{O,t} + \left(\widehat{\boldsymbol{d}}^\Delta_{O,t} - \boldsymbol{d}^\Delta_{O,t}\right)$ and $\widehat{\boldsymbol{b}}_\Delta = \boldsymbol{b}_\Delta + \left(\widehat{\boldsymbol{b}}_\Delta - \boldsymbol{b}_\Delta\right)$ in the definition of $\widehat{\boldsymbol{\lambda}}_\Delta$. That the first four terms of the decomposition are zero expectation stems from (a) the fact that the first stage estimates are unbiased, and (b) the fact that the first-stage estimation error is independent of factor realizations (by Assumption (i)). ■

To derive further results it is useful to define the following matrix:

$$\widetilde{\boldsymbol{D}}_{O,t} \equiv \begin{bmatrix} \boldsymbol{\phi}^\intercal_{1O,t}\otimes\boldsymbol{g}^\intercal_{11t} & 0 & 0 & \dots & 0 \\ \vdots & & & & \\ \boldsymbol{\phi}^\intercal_{1O,t}\otimes\boldsymbol{g}^\intercal_{1pt} & 0 & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & \dots & \boldsymbol{\phi}^\intercal_{dO,t}\otimes\boldsymbol{g}^\intercal_{dpt} & \dots & 0 \\ \vdots & & & & \\ 0 & \dots & 0 & \dots & \boldsymbol{\phi}^\intercal_{DO,t}\otimes\boldsymbol{g}^\intercal_{DPt} \end{bmatrix}.$$

It is used to represent $\boldsymbol{d}^\Delta_{O,t}$ as $\widetilde{\boldsymbol{D}}_{O,t}\boldsymbol{b}_\Delta$, where $\boldsymbol{b}_\Delta$ is the $Dk \times 1$ vector of the parameters of factor sensitivities and $\widetilde{\boldsymbol{D}}_{O,t}$ is of size $DP \times Dk$.[39]

[39] In the interest of simplicity we assume we have D factors and a k-dimensional basis expansion, and

Proof. The proof follows directly from Corollary 1 and the observation that the final term in the decomposition of estimation noise is only a function of the estimation error in the factor loading parameters $\boldsymbol{b}_\Delta$:

$$\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t} - \boldsymbol{d}^{\Delta}_{O,t}\right)\left(\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\right)^{\intercal}\left(\widehat{\boldsymbol{b}}_1 - \boldsymbol{b}_\Delta\right) = \widetilde{\boldsymbol{D}}_{O,t}\left(\widehat{\boldsymbol{b}}_\Delta - \boldsymbol{b}_\Delta\right)\left(\widehat{\boldsymbol{b}}_\Delta - \boldsymbol{b}_\Delta\right)^{\intercal}\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\,.$$

■

Corollary 1 implies that the estimator is asymptotically biased due to the fact that measurement error is present on both sides in the regression equation. In what follows, we first state a result about the asymptotic bias term.

Proposition 8 (Unbiased estimator of $\boldsymbol{\lambda}_\Delta$)

$$\widehat{\boldsymbol{\lambda}}^{\Psi}_{\Delta} \equiv \widehat{\boldsymbol{\lambda}}_\Delta - \boldsymbol{\Psi}_\Delta$$

is an unbiased estimator of the vector of risk premia parameters. The asymptotic bias term is given by

$$\Psi_\Delta \equiv \left[plim\ \boldsymbol{Q}_{\widehat{\boldsymbol{d}}^\Delta}\right]^{-1}\ plim\ \left[\frac{1}{T}\sum_t \frac{1}{N_t}\sum_O \widetilde{\boldsymbol{D}}_{O,t}\left(\widehat{\boldsymbol{b}}_\Delta - \boldsymbol{b}_\Delta\right)\left(\widehat{\boldsymbol{b}}_\Delta - \boldsymbol{b}_\Delta\right)^{\intercal}\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\right],$$

and it is estimated by

$$\widehat{\Psi}_\Delta = \boldsymbol{Q}^{-1}_{\widehat{\boldsymbol{d}}^\Delta}\frac{1}{T}\sum_t \frac{1}{N_t}\sum_O \widetilde{\boldsymbol{D}}_{O,t}\widehat{V}_1\left[\widehat{\boldsymbol{b}}_\Delta\right]\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\,,$$

where $\widehat{V}_1\left[\widehat{\boldsymbol{b}}_\Delta\right]$ is the submatrix of $\widehat{V}_1\left[\widehat{\boldsymbol{B}}\right]$ that contains the covariance matrix of $\widehat{\boldsymbol{b}}_\Delta$.

Coming back to Corollary 1, we state a result about the variance of the estimation error.

Proposition 9 (Variance of estimation error) *The variance of estimation error consists of five components, four of which are due to measurement error arising from first-*

P risk premia instruments per factor. It is straightforward to generalize the result to a setting with an arbitrary number of instruments per factor.

stage estimation.

$$V\left[\boldsymbol{Q}_{\widehat{\boldsymbol{d}}^{\boldsymbol{\Delta}}}\sqrt{T}\left(\widehat{\boldsymbol{\lambda}}_{\Delta}-\boldsymbol{\lambda}_{\Delta}-\boldsymbol{\Psi}_{\Delta}\right)\right]=\boldsymbol{\Sigma}_{\boldsymbol{d}^{\boldsymbol{\Delta}}\boldsymbol{\eta}^{\boldsymbol{\Delta}}}+\boldsymbol{\Sigma}_{\boldsymbol{d}^{\Delta}\widehat{\boldsymbol{b}}^{\Delta}}+\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}}+2\boldsymbol{C}_{\widehat{\boldsymbol{b}}^{\Delta}\widehat{\boldsymbol{d}}^{\Delta}}+\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}\eta^{\Delta}}\,.$$

The components are

$$\begin{aligned}
\boldsymbol{\Sigma}_{\boldsymbol{d}^{\boldsymbol{\Delta}}\boldsymbol{\eta}^{\boldsymbol{\Delta}}} &\equiv V\left[\frac{1}{\sqrt{T}}\sum_t\frac{1}{N_t}\sum_O \boldsymbol{d}^{\Delta}_{O,t}\eta^{\Delta}_{O,t+\Delta t}\right],\\
\boldsymbol{\Sigma}_{\boldsymbol{d}^{\Delta}\widehat{\boldsymbol{b}}^{\Delta}} &\equiv V\left[\frac{1}{\sqrt{T}}\sum_t\frac{1}{N_t}\sum_O \boldsymbol{d}^{\Delta}_{O,t}\left(\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\right)^{\intercal}\left(\widehat{\boldsymbol{b}}_{\Delta}-\boldsymbol{b}_{\Delta}\right)\right],\\
\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}} &\equiv V\left[\frac{1}{\sqrt{T}}\sum_t\frac{1}{N_t}\sum_O \left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}-\boldsymbol{d}^{\Delta}_{O,t}\right)\left(\boldsymbol{d}^{\Delta}_{O,t}\right)^{\intercal}\lambda_{\Delta}\right],\\
\boldsymbol{C}_{\widehat{\boldsymbol{b}}^{\Delta}\widehat{\boldsymbol{d}}^{\Delta}} &\equiv E\left[\frac{1}{T}\sum_t\frac{1}{N_t}\sum_O \left(\boldsymbol{d}^{\Delta}_{O,t}\right)^{\intercal}\widehat{\boldsymbol{\lambda}}_{\Delta}\cdot\boldsymbol{d}^{\Delta}_{O,t}\left(\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\right)^{\intercal}\left(\widehat{\boldsymbol{b}}_{\Delta}-\boldsymbol{b}_{\Delta}\right)\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}-\boldsymbol{d}^{\Delta}_{O,t}\right)^{\intercal}\right],\\
\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}\eta^{\Delta}} &\equiv V\left[\frac{1}{\sqrt{T}}\sum_t\frac{1}{N_t}\sum_O \left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}-\boldsymbol{d}^{\Delta}_{O,t}\right)\eta^{\Delta}_{O,t+\Delta t}\right].
\end{aligned}$$

Proof. Note that

$$\begin{aligned}
V\left[\boldsymbol{Q}_{\widehat{\boldsymbol{d}}^{\boldsymbol{\Delta}}}\sqrt{T}\left(\widehat{\boldsymbol{\lambda}}_{\Delta}-\boldsymbol{\lambda}_{\Delta}-\boldsymbol{\Psi}_{\Delta}\right)\right] = V\Bigg[&\frac{1}{\sqrt{T}}\sum_t\frac{1}{N_t}\sum_O\Big[\boldsymbol{d}^{\Delta}_{O,t}\eta^{\Delta}_{O,t+\Delta t}+\boldsymbol{d}^{\Delta}_{O,t}\left(\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\right)^{\intercal}\left(\widehat{\boldsymbol{b}}_{\Delta}-\boldsymbol{b}_{\Delta}\right)\\
&+\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}-\boldsymbol{d}^{\Delta}_{O,t}\right)\left(\boldsymbol{d}^{\Delta}_{O,t}\right)^{\intercal}\lambda_{\Delta}+\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}-\boldsymbol{d}^{\Delta}_{O,t}\right)\eta^{\Delta}_{O,t+\Delta t}\\
&+\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}-\boldsymbol{d}^{\Delta}_{O,t}\right)\left(\boldsymbol{\phi}^{f_\Delta}_{O,t+\Delta t}\right)^{\intercal}\left(\widehat{\boldsymbol{b}}_{\Delta}-\boldsymbol{b}_{\Delta}\right)\Big]\Bigg].
\end{aligned}$$

Then recall that $\boldsymbol{\Sigma}_{\boldsymbol{d}^{\boldsymbol{\Delta}}\boldsymbol{\eta}^{\boldsymbol{\Delta}}}$ is the variance of standard OLS score, which is conditionally independent from the other components. $\boldsymbol{\Sigma}_{\boldsymbol{d}^{\Delta}\widehat{\boldsymbol{b}}^{\Delta}}$ and $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}}$ both arise from estimation error in $\widehat{\boldsymbol{b}}_{\Delta}$, and thus are correlated, with $\boldsymbol{C}_{\widehat{\boldsymbol{b}}^{\Delta}\widehat{\boldsymbol{d}}^{\Delta}}$ the covariance matrix. $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}\eta^{\Delta}}$ arises from the multiplicative contribution of the estimation error in $\widehat{\boldsymbol{b}}_{\Delta}$ and the regression error $\eta^{O}_{t=\Delta t}$; these terms have zero covariance with the preceding terms due to the conditional independence of the first-stage estimation noise and second-stage regression error $\eta^{O}_{t=\Delta t}$.

■

Proposition 10 (Estimators of estimation noise variance) *The consistency of* $\widehat{\boldsymbol{b}}^{\Delta}$

implies that we can replace $\boldsymbol{d}^{\Delta}$ *with* $\widehat{\boldsymbol{d}}^{\Delta}$ *where applicable. Under assumption (iv), the following are consistent estimators of estimation noise variance.*

(a) $\boldsymbol{\Sigma}_{\boldsymbol{d}^{\Delta}\boldsymbol{\eta}^{\Delta}}$. *Define* $\boldsymbol{h}_t \equiv \frac{1}{N_t}\sum_O \widehat{\boldsymbol{d}}^{\Delta}_{O,t}\eta^{\Delta}_{O,t+\Delta t}$ *and set* S*, a positive integer lag order.* $\boldsymbol{\Sigma}_{\boldsymbol{d}^{\Delta}\boldsymbol{\eta}^{\Delta}}$ *can be estimated with*

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{d}^{\Delta}\boldsymbol{\eta}^{\Delta}} \equiv \frac{1}{T}\sum_t \boldsymbol{h}_t\boldsymbol{h}_t^{\intercal} + \frac{1}{T}\sum_{s=1}^{S}\sum_{t=s+1}^{T} w_s\left(\boldsymbol{h}_t\boldsymbol{h}_{t-l}^{\intercal} + \boldsymbol{h}_{t-l}\boldsymbol{h}_t^{\intercal}\right),$$

and $w_s = 1 - \frac{s}{S+1}$*, which is the Newey and West (1987) estimator applied to the time series of* $\boldsymbol{h}_t$*.*

(b) $\boldsymbol{\Sigma}_{\boldsymbol{d}^{\Delta}\widehat{\boldsymbol{b}}^{\Delta}}$. *The estimator of the variance component is given by*

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{d}^{\Delta}\widehat{\boldsymbol{b}}^{\Delta}} \equiv \frac{1}{T^2}\left[\sum_t \frac{1}{N_t}\sum_O \widehat{\boldsymbol{d}}^{\Delta}_{O,t}\left(\boldsymbol{\phi}^{f_{\Delta}}_{O,t+\Delta t}\right)^{\intercal}\right]\widehat{V}_1\left[\widehat{\boldsymbol{b}}_{\Delta} - \boldsymbol{b}_{\Delta}\right]\left[\sum_t \frac{1}{N_t}\sum_O \boldsymbol{\phi}^{f_{\Delta}}_{O,t+\Delta t}\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\right)^{\intercal}\right],$$

where $\widehat{V}_1\left[\widehat{\boldsymbol{b}}_{\Delta} - \boldsymbol{b}_{\Delta}\right]$ *is the submatrix of* V_1 *defined in Proposition 3 that corresponds to the loadings on tradable factors.*

(c) $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}}$. $\widehat{\boldsymbol{\lambda}}^{\Psi}_{\Delta}$ *is consistent, hence replace* $\left(\boldsymbol{d}^{\Delta}_{O,t}\right)^{\intercal}\boldsymbol{\lambda}_{\Delta}$ *with* $\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\right)^{\intercal}\widehat{\boldsymbol{\lambda}}^{\Psi}_{\Delta}$*. Then use the fact that* $\widehat{\boldsymbol{d}}^{\Delta}_{O,t} - \boldsymbol{d}^{\Delta}_{O,t} = \widetilde{\boldsymbol{D}}_{O,t}\left(\widehat{\boldsymbol{b}}_{\Delta} - \boldsymbol{b}_{\Delta}\right)$*. The estimator is given by*

$$\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{d}}^{\Delta}} = \frac{1}{T^2}\left(\sum_t \frac{1}{N_t}\sum_O\left(\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\right)^{\intercal}\widehat{\boldsymbol{\lambda}}^{\Psi}_{\Delta}\right)\cdot\widetilde{\boldsymbol{D}}_{O,t}\right)\widehat{V}_1\left[\widehat{\boldsymbol{b}}_{\Delta} - \boldsymbol{b}_{\Delta}\right]\left(\sum_t \frac{1}{N_t}\sum_O\left(\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\right)^{\intercal}\widehat{\boldsymbol{\lambda}}^{\Psi}_{\Delta}\right)\cdot\widetilde{\boldsymbol{D}}^{\intercal}_{O,t}\right).$$

(d) $\boldsymbol{C}_{\widehat{\boldsymbol{b}}^{\Delta}\widehat{\boldsymbol{d}}^{\Delta}}$. $\widehat{\boldsymbol{\lambda}}^{\Psi}_{\Delta}$ *is consistent, hence replace* $\left(\boldsymbol{d}^{\Delta}_{O,t}\right)^{\intercal}\boldsymbol{\lambda}_{\Delta}$ *with* $\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\right)^{\intercal}\widehat{\boldsymbol{\lambda}}^{\Psi}_{\Delta}$*. The estimator is given by*

$$\widehat{\boldsymbol{C}}_{\widehat{\boldsymbol{b}}^{\Delta}\widehat{\boldsymbol{d}}^{\Delta}} \equiv \frac{1}{T^2}\left(\sum_t \frac{1}{N_t}\sum_O\left(\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\right)^{\intercal}\widehat{\boldsymbol{\lambda}}^{\Psi}_{\Delta}\cdot\widehat{\boldsymbol{d}}^{\Delta}_{O,t}\left(\boldsymbol{\phi}^{f_{\Delta}}_{O,t+\Delta t}\right)^{\intercal}\right)\widehat{V}_1\left[\widehat{\boldsymbol{b}}_{\Delta} - \boldsymbol{b}_{\Delta}\right]\left(\sum_t \frac{1}{N_t}\sum_O \widetilde{\boldsymbol{D}}_{O,t}\right).$$

(e) $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}\eta^{\Delta}}$. *With* $\widehat{\boldsymbol{d}}^{\Delta}_{O,t} - \boldsymbol{d}^{\Delta}_{O,t} = \widetilde{\boldsymbol{D}}_{O,t}\left(\widehat{\boldsymbol{b}}_{\Delta} - \boldsymbol{b}_{\Delta}\right)$*, define* $\widehat{V}_1^{1/2}$ *as the principal matrix square root of* $\widehat{V}_1\left[\widehat{\boldsymbol{b}}_{\Delta}\right]$*, and define* $\boldsymbol{h_t} \equiv \frac{1}{N_t}\sum_O \eta^{\Delta}_{O,t+\Delta t}\widetilde{\boldsymbol{D}}_{O,t}\widehat{V}_1^{1/2}$*. Set* S*, a positive* *integer lag order.* $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{d}}^{\Delta}\eta^{\Delta}}$ *can be estimated with*

$$\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{d}}^{\Delta}\eta^{\Delta}} \equiv \frac{1}{T}\sum_{t}\boldsymbol{h}_t\boldsymbol{h}_t^{\intercal} + \frac{1}{T}\sum_{s=1}^{S}\sum_{t=s+1}^{T} w_s\left(\boldsymbol{h}_t\boldsymbol{h}_{t-l}^{\intercal} + \boldsymbol{h}_{t-l}\boldsymbol{h}_t^{\intercal}\right) ,$$

and $w_s = 1 - \frac{s}{S+1}$*, which is the Newey and West (1987) estimator applied to the time series of* $\boldsymbol{h}_t$*.*

B.2.2 Premia on non-tradable factors

The results in this section closely follow the preceding section, with the exception that the projection of option returns on first-stage variables contains the $-E^Q[R_{O,t+\Delta t}]$ component:

$$R^{O\Omega}_{t+\Delta t} \equiv \sum_{k=1}^{M}\left(f_{k,t+\Delta t} - E^Q[f_{k,t+\Delta t}]\right)\beta^O_{k,t} = \left(\boldsymbol{\phi}^{gf_\Omega}_{O,t+\Delta t}\right)^{\intercal}\boldsymbol{B}_\Omega\,,$$

where Ω denotes the set of the non-tradable factors, $\boldsymbol{\phi}^{gf_\Omega}_{O,t+\Delta t}$ is the subvector of the vector $\boldsymbol{\phi}^{gf}_{O,t+\Delta t}$ associated with the non-tradable factors and the corresponding risk-premia predictors.

Let $\boldsymbol{d}^{\Omega}_{O,t} \equiv \begin{bmatrix}\beta^O_{1t}\boldsymbol{g}^{\intercal}_{1,t} & \dots & \beta^O_{Dt}\boldsymbol{g}^{\intercal}_{M,t}\end{bmatrix}$, and let $\widehat{\boldsymbol{d}}^{\Omega}_{O,t}$ be its sample analogue, i.e., with $\widehat{\beta}^O_{kt}$ replacing β^O_{kt} and $\widehat{\gamma}^O_t$ replacing γ^O_t. Then the following relationship holds between the projection of option returns on the tradable factors and the risk premia predictors,

$$R^{O\Omega}_{t+\Delta t} = \left(\boldsymbol{d}^{\Omega}_{O,t}\right)^{\intercal}\boldsymbol{\lambda}_\Omega + \eta^{\Omega}_{O,t+\Delta t}\,.$$

In fact, with the true factor loadings unknown, we estimate $\boldsymbol{\lambda}_\Omega$ based on the feasible version of the above equation,

$$\widehat{R}^{O\Omega}_{t+\Delta t} \equiv \left(\boldsymbol{\phi}^{gf_\Omega}_{O,t+\Delta t}\right)^{\intercal}\widehat{\boldsymbol{B}}_\Omega = \left(\boldsymbol{d}^{\Omega}_{O,t}\right)^{\intercal}\boldsymbol{\lambda}_\Omega + \widetilde{\eta}^{\Omega}_{O,t+\Delta t}\,. \tag{31}$$

Inference in the case of non-tradable factors follows the results of Section B.2.1 with $\boldsymbol{d}^{\Omega}_{O,t}$ replacing $\boldsymbol{d}^{\Delta}_{O,t}$, $\widehat{\boldsymbol{B}}_\Omega$ replacing $\widehat{\boldsymbol{b}}_\Delta$, and $\boldsymbol{\phi}^{gf_\Omega}$ replacing $\boldsymbol{\phi}^{gf_\Delta}$.

References

Yacine Aït-Sahalia and Jefferson Duarte. Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, 116(1-2):9–47, 2003.

Torben G. Andersen, Dobrislav Dobrev, and Ernst Schaumburg. Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93, 2012. Recent Advances in Panel Data, Nonlinear and Nonparametric Models: A Festschrift in Honor of Peter C.B. Phillips.

Torben G. Andersen, Nicola Fusari, and Viktor Todorov. Parametric Inference and Dynamic State Recovery From Option Panels. *Econometrica*, 83(3):1081–1145, may 2015a.

Torben G. Andersen, Nicola Fusari, and Viktor Todorov. The risk premia embedded in index options. *Journal of Financial Economics*, 117(3):558–584, sep 2015b.

Andrew Ang, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. The cross-section of volatility and expected returns. *Journal of Finance*, 2006.

Jennie Bai, Turan G. Bali, and Quan Wen. Common risk factors in the cross-section of corporate bond returns. *Journal of Financial Economics*, 2019.

Gurdip Bakshi, Nikunj Kapadia, and Dilip Madan. Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *Review of Financial Studies*, 16(1):101–143, jan 2003.

Gurdip Bakshi, Dilip Madan, and George Panayotov. Returns of claims on the upside and the viability of U-shaped pricing kernels. *Journal of Financial Economics*, 2010.

Gurdip S. Bakshi, John Crosby, and Xiaohui Gao. The Dark Matter in Equity Index Volatility Dynamics: Assessing the Economic Rationales for Unspanned Risks. *SSRN Electronic Journal*, 2019.

Laurent Barras and Aytek Malkhozov. Does variance risk have two prices? Evidence from the equity and option markets. *Journal of Financial Economics*, 121(1):79–92, jul 2016.

David S. Bates. Empirical option pricing: A retrospection. *Journal of Econometrics*, 116 (1-2):387–404, 2003.

David S Bates. The market for crash risk. *Journal of Economic Dynamics and Control*, 32(7):2291–2321, 2008.

Jean-François Bégin, Christian Dorion, and Geneviève Gauthier. Idiosyncratic Jump Risk Matters: Evidence from Equity Returns and Options. *The Review of Financial Studies*, 33(1):155–211, jan 2020.

Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

Nicolas P. B. Bollen and Robert E. Whaley. Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance*, 59(2):711–753, 2004.

Tim Bollerslev and Viktor Todorov. Tails, fears, and risk premia. *The Journal of Finance*, 66(6):2165–2211, 2011.

Tim Bollerslev, Viktor Todorov, and Lai Xu. Tail risk premia and return predictability. *Journal of Financial Economics*, 118(1):113–134, oct 2015.

Tim Bollerslev, Sophia Zhengzi Li, and Viktor Todorov. Roughing up Beta: Continuous versus Discontinuous Betas and the Cross Section of Expected Stock Returns. *Journal of Financial Economics*, 120(3):464–490, jun 2016.

Mark Broadie, Mikhail Chernov, and Michael Johannes. Understanding index option returns. *Review of Financial Studies*, 2009.

A Buraschi and J C Jackwerth. The price of a smile: Hedging and spanning in option markets. *Review of Financial Studies*, 14:495–527, 2001.

Adrian Buss and Grigory Vilkov. Measuring equity risk with option-implied correlations, 2012.

Peter Carr and Liuren Wu. Variance Risk Premiums. *Review of Financial Studies*, 22 (3):1311–1341, mar 2009.

Fousseni Chabi-Yo. Explaining the idiosyncratic volatility puzzle using stochastic discount factors. *Journal of Banking and Finance*, 35(8):1971–1983, 2011.

Fousseni Chabi-Yo. Pricing kernels with stochastic skewness and volatility risk. *Management Science*, 58(3):624–640, 2012.

Ing-Haw Cheng. Volatility markets underreacted to the early stages of the COVID-19 pandemic. *Review of Asset Pricing Studies, Forthcoming*, 2020.

Peter Christoffersen, Steven Heston, and Kris Jacobs. Capturing Option Anomalies with a Variance-Dependent Pricing Kernel. *The Review of Financial Studies*, 26(8):1963–2006, 06 2013.

Peter Christoffersen, Mathieu Fournier, and Kris Jacobs. The Factor Structure in Equity Options. *The Review of Financial Studies*, 31(2):595–637, feb 2018a.

Peter Christoffersen, Ruslan Goyenko, Kris Jacobs, and Mehdi Karoui. Illiquidity Premia in the Equity Options Market. *The Review of Financial Studies*, 31(3):811–851, mar 2018b.

George M. Constantinides, Jens Carsten Jackwerth, and Stylianos Perrakis. Mispricing of S&P 500 index options. *Review of Financial Studies*, 2009.

Joshua D. Coval and Tyler Shumway. Expected option returns. *The Journal of Finance*, 56(3):983–1009, jun 2001.

Robert F Dittmar. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *Journal of Finance*, 57(1):369–403, 2002.

John C. Driscoll and Aart C. Kraay. Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics*, 80(4):549–560, nov 1998.

Jefferson Duarte, Christopher S Jones, and Junbo L. Wang. Very Noisy Option Prices and Inferences Regarding Option Returns, 2019.

Daniel Egloff, Markus Leippold, and Liuren Wu. The term structure of variance swap rates and optimal variance swap investments. *Journal of Financial and Quantitative Analysis*, 2010.

Bjørn Eraker and Jiakou Wang. A non-linear dynamic model of the variance risk premium. *Journal of Econometrics*, 2015.

Bjørn Eraker, Michael Johannes, and Nicholas Polson. The Impact of Jumps in Equity Index Volatility and Returns. *Journal of Finance*, 58(3):1269–1300, 2003.

Eugene F. Fama and James D. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636, 1973.

Jianqing Fan, Michael B. Imerman, and Wei Dai. What Does the Volatility Risk Premium Say About Liquidity Provision and Demand for Hedging Tail Risk? *Journal of Business and Economic Statistics*, 2016.

Patrick Gagliardini, Elisa Ossola, and Olivier Scaillet. Time-Varying Risk Premium in Large Cross-Sectional Equity Data Sets. *Econometrica*, 84(3):985–1046, 2016.

Nicolae Gârleanu, Lasse Heje Pedersen, and Allen M. Poteshman. Demand-Based Option Pricing. *Review of Financial Studies*, 22(10):4259–4299, oct 2009.

Eric Ghysels. On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance*, 1998.

Lawrence R. Glosten, Ravi Jagannathan, and David E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801, 1993.

Elise Gourier. Pricing of idiosyncratic equity and variance risks. *Working Paper*, 2014.

Amit Goyal and Alessio Saretto. Cross-section of option returns and volatility. *Journal of Financial Economics*, 94(2):310–326, 2009.

Zhiguo He and Arvind Krishnamurthy. Intermediary asset pricing. *American Economic Review*, 103(2):732–70, April 2013.

Zhiguo He, Bryan Kelly, and Asaf Manela. Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 2017.

Steven L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 1993.

Alex R. Horenstein, Aurelio Vasquez, and Xiao Xiao. Common Factors in Equity Option Returns. *SSRN Electronic Journal*, nov 2018.

Frank Huettner and Marco Sunder. Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values. *Electronic Journal of Statistics*, 6:1239–1250, 2012.

Roni Israelov and Bryan T. Kelly. Forecasting the Distribution of Option Returns. *SSRN Electronic Journal*, 2017.

Ravi Jagannathan and Zhenyu Wang. The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, 51(1):3–53, 1996.

Christopher S. Jones. The dynamics of stochastic volatility: evidence from underlying and options markets. *Journal of Econometrics*, 116:181–224, 2003.

Christopher S. Jones. A Nonlinear Factor Analysis of S&P 500 Index Option Returns. *The Journal of Finance*, 61(5):2325–2363, oct 2006.

Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh. Too-systemic-to-fail: What option markets imply about sector-wide government guarantees. *American Economic Review*, 106(6):1278–1319, June 2016.

Bryan T. Kelly, Seth Pruitt, and Yinan Su. Instrumented Principal Component Analysis. *SSRN Electronic Journal*, 2018.

Jonathan Lewellen and Stefan Nagel. The conditional {CAPM} does not explain asset pricing anomalies. *Journal of Financial Economics*, 82(2):289–314, nov 2006.

Hanno Lustig and Stijn van Nieuwerburgh. Housing collateral, consumption insurance and risk premia: an empirical perspective. *Journal of Finance*, 60(3):1167–1219, 2005.

Ian Martin. What is the Expected Return on the Market?*. *The Quarterly Journal of Economics*, 132(1):367–433, feb 2017.

Stefan Nagel and Kenneth J. Singleton. Estimation and Evaluation of Conditional Asset Pricing Models. *Journal of Finance*, 2011.

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

Jun Pan. The Jump-risk Premia Implicit in Options: Eevidence from an iItegrated Time-series Study. *Journal of Financial Economics*, 63(1):3–50, 2002.

Pauol Schneider and Fabio Trojani. (Almost) Model-Free Recovery. *The Journal of Finance*, 74(1):323–370, feb 2019.

Jay Shanken. Intertemporal asset pricing: An empirical investigation. *Journal of Econometrics*, 45:99–120, 1990.

Jay Shanken. On the Estimation of Beta-Pricing Models. *Review of Financial Studies*, 5 (1):1–33, jan 1992.

Viktor Todorov. Variance Risk-Premium Dynamics: The Role of Jumps. *Review of Financial Studies*, 23(1):345–383, jan 2010.

Viktor Todorov. Nonparametric spot volatility from options. *The Annals of Applied Probability*, 29(6):3590–3636, dec 2019.

Simon N Wood. Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1):95–114, 2003.

Simon N. Wood. *Generalized additive models: An introduction with R*. Routledge, 2017. ISBN 9781498728348.

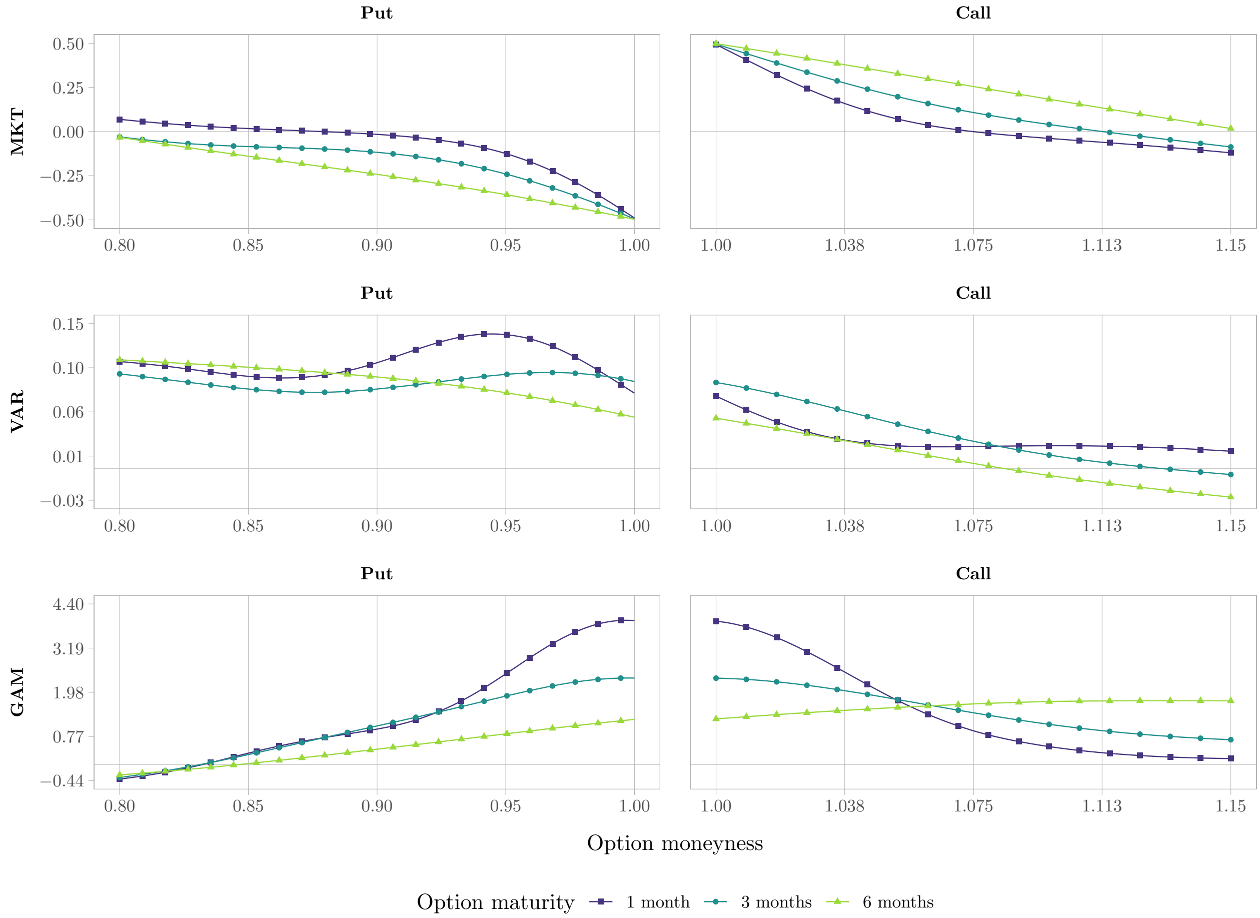

Figure 1: Option exposures

Model exposures plotted against moneyness for calls and puts, separately, for each factor. Exposures are calculated by setting the instruments to their sample medians and varying moneyness and maturity. In each panel, we consider three maturities of 1, 3, and 6 months, respectively.

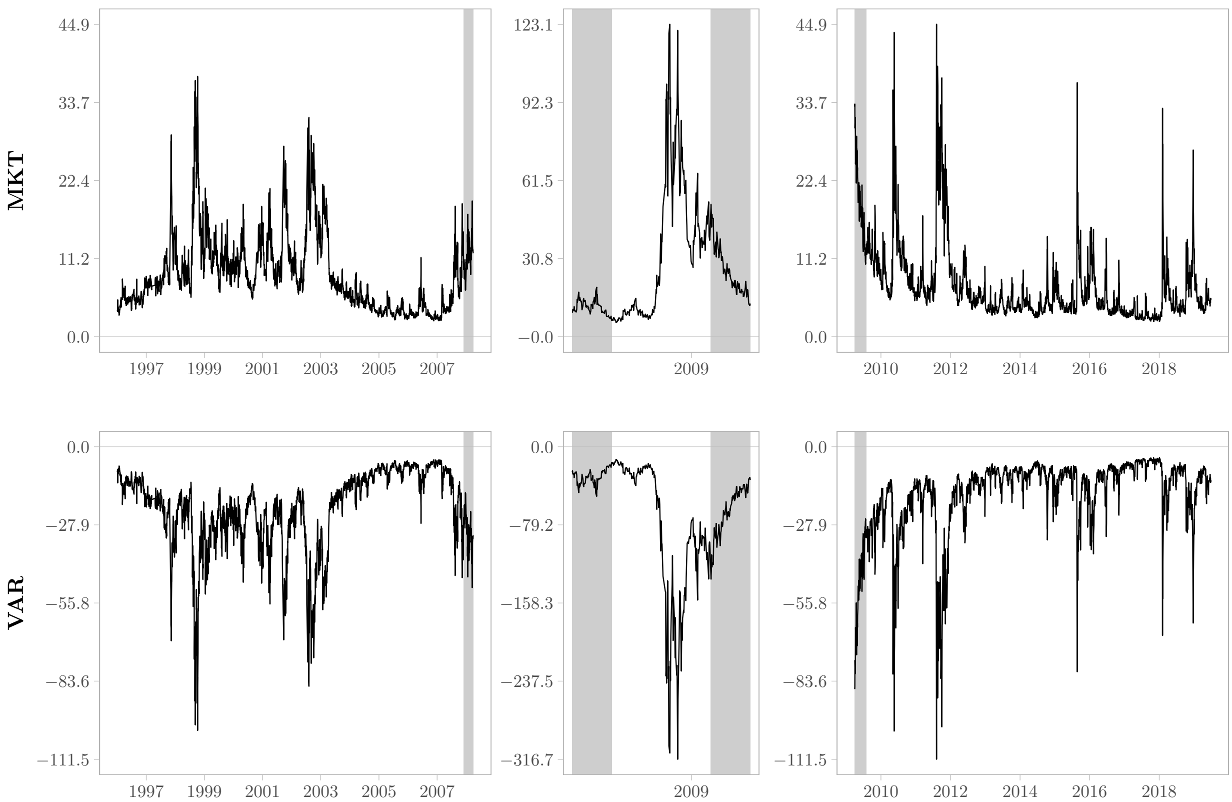

Figure 2: Time series of risk premia

Daily time series of annualized (252 days per year) risk premia implied by the benchmark model specification. The sample period is from January 3rd, 1996 to June 27th, 2019.

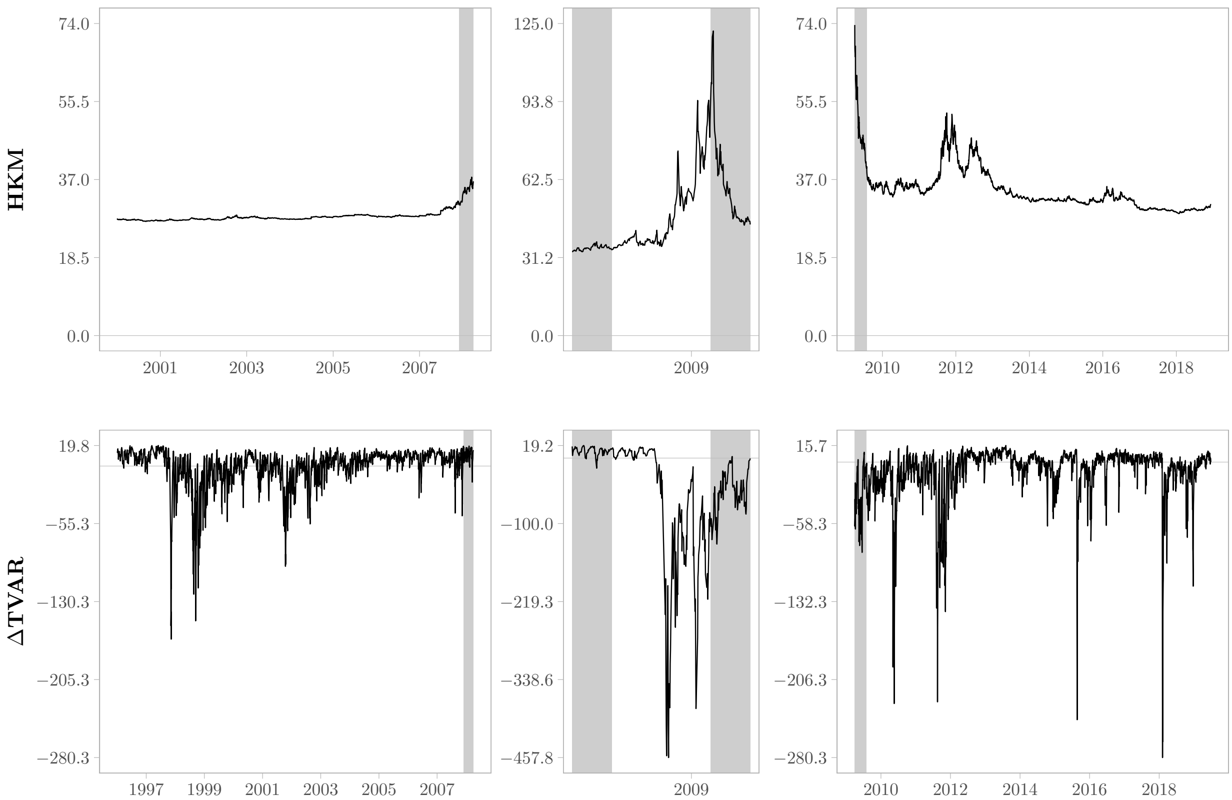

Figure 3: Time series of risk premia on additional factors

We plot the daily time series of annualized risk premia on the HKM and ΔTVAR factors implied by our model. The risk premium on HKM is an affine function of the squared inverse capital ratio at the end of each day, as in equation (12) in He et al. (2017). The risk premium on ΔTVAR is an affine function of the level of TVAR_t at the end of each day. The premia are annualized by multiplying by the factor fo 252 trading days. For ΔTVAR, the sample period is from January 3^{rd}, 1996 to June 27^{th}, 2019. For HKM, the sample period starts on January 2^{nd}, 2000.

Table 1: Descriptive statistics of excess option returns

The summary statistics are calculated based on daily returns on S&P 500 index options between January 1996 and June 2019. Expected returns (ER), standard deviations (SD) and Sharpe Ratios (SR) are expressed in percent per annum, with 252 trading days per year. Option returns are averaged within each day and bucket before calculating the descriptive statistics. The remaining quantities are Skewness (Sk) and Kurtosis (K). Delta-hedged returns are calculated with deltas obtained from our baseline factor model specification described in Section 5. For at the money (ATM) options, $K/S_t \in (0.975, 1.02]$. Out of the money (OTM) calls (puts), $K/S_t \in (1.02, 1.07]$ ($K/S_t \in (0.9, 0.975]$). Deep out of the money (DOTM) calls (puts), $K/S_t \in (1.07, 1.15]$ ($K/S_t \in (0.8, 0.9]$).

Panel A: Put options

	Unlevered excess returns					Unlevered and delta-hedged excess returns				
τ group	ER	SD	SR	Sk	K	ER	SD	SR	Sk	K
Moneyness: All										
1-2M	−3.14	6.18	−0.51	1.63	25.28	−2.08	1.99	−1.05	2.35	35.99
2-3M	−2.71	6.50	−0.42	1.29	19.57	−1.83	2.06	−0.89	1.89	32.32
3-6M	−2.81	6.85	−0.41	1.13	17.72	−1.74	2.25	−0.77	1.34	32.98
All	−2.89	6.51	−0.44	1.33	20.51	−1.89	2.10	−0.90	1.80	33.99
Moneyness: ATM										
1-2M	−4.64	8.76	−0.53	1.26	14.01	−2.69	2.43	−1.11	2.03	22.51
2-3M	−3.35	8.79	−0.38	0.97	11.65	−2.03	2.30	−0.88	1.59	15.42
3-6M	−3.69	8.91	−0.41	0.96	12.27	−1.85	2.26	−0.82	1.33	14.90
All	−3.91	8.82	−0.44	1.06	12.65	−2.19	2.33	−0.94	1.68	18.13
Moneyness: OTM										
1-2M	−3.22	5.53	−0.58	1.68	26.80	−2.25	1.96	−1.15	2.74	30.39
2-3M	−2.96	6.17	−0.48	1.41	20.42	−2.00	2.10	−0.95	2.42	29.10
3-6M	−2.88	6.69	−0.43	1.08	15.72	−1.89	2.22	−0.85	1.78	21.60
All	−3.02	6.15	−0.49	1.35	20.07	−2.04	2.10	−0.97	2.26	26.53
Moneyness: DOTM										
1-2M	−1.54	2.69	−0.57	2.24	64.79	−1.31	1.45	−0.90	2.01	102.52
2-3M	−1.80	3.33	−0.54	1.95	44.46	−1.47	1.72	−0.85	1.44	84.17
3-6M	−1.87	4.09	−0.46	1.43	30.05	−1.49	2.28	−0.65	0.93	60.36
All	−1.74	3.42	−0.51	1.78	42.07	−1.42	1.86	−0.77	1.29	80.23

Panel B: Call options

	Unlevered excess returns					Unlevered and delta-hedged excess returns				
τ group	ER	SD	SR	Sk	K	ER	SD	SR	Sk	K
Moneyness: All										
1-2M	0.57	5.33	0.11	1.60	30.99	−0.75	2.00	−0.38	2.67	41.66
2-3M	0.56	5.81	0.10	1.27	25.25	−0.67	1.94	−0.34	2.09	34.82
3-6M	0.86	6.11	0.14	0.55	17.13	−0.35	1.94	−0.18	1.51	24.15
All	0.67	5.76	0.12	1.07	23.48	−0.59	1.96	−0.30	2.10	33.82
Moneyness: ATM										
1-2M	0.96	7.54	0.13	0.91	14.75	−1.07	2.40	−0.44	2.25	30.02
2-3M	0.81	7.84	0.10	0.74	13.50	−0.85	2.35	−0.36	1.77	27.45
3-6M	0.57	7.86	0.07	0.03	8.93	−0.34	2.29	−0.15	1.39	21.03
All	0.78	7.74	0.10	0.55	12.31	−0.76	2.35	−0.32	1.83	26.54
Moneyness: OTM										
1-2M	0.62	4.56	0.14	2.43	45.69	−0.78	1.94	−0.40	3.01	46.77
2-3M	0.54	5.38	0.10	1.69	30.99	−0.65	1.94	−0.33	2.10	31.67
3-6M	1.88	6.06	0.31	1.09	20.65	−0.34	1.96	−0.17	1.16	17.01
All	1.02	5.37	0.19	1.59	29.53	−0.59	1.95	−0.30	2.08	31.63
Moneyness: DOTM										
1-2M	0.07	2.21	0.03	7.71	218.80	−0.37	1.50	−0.24	2.82	57.33
2-3M	0.32	2.95	0.11	4.56	100.40	−0.50	1.35	−0.37	2.71	49.07
3-6M	0.09	3.71	0.02	2.14	51.18	−0.36	1.50	−0.24	2.38	41.55
All	0.15	3.04	0.05	3.67	88.56	−0.41	1.45	−0.28	2.63	49.53

Table 2: Option exposures

In Panel A we report the sample averages of option exposures to the factors. Exposures are averaged within each day and bucket before calculating the time-series averages. In Panel B we report Shapley-Owen values of R^2 (Huettner and Sunder, 2012) from regressions of realized option returns on the products of option exposures and factor realizations.*

Panel A: Average option exposures

	Put options			Call options		
τ group	MKT	VAR	GAM	MKT	VAR	GAM
Moneyness: All						
1-2M	−0.20	0.10	1.99	0.21	0.05	1.92
2-3M	−0.24	0.09	1.42	0.26	0.05	1.73
3-6M	−0.28	0.09	1.06	0.30	0.05	1.35
All	−0.24	0.09	1.49	0.25	0.05	1.67
Moneyness: ATM						
1-2M	−0.40	0.10	3.30	0.41	0.07	3.20
2-3M	−0.42	0.10	2.26	0.44	0.08	2.37
3-6M	−0.43	0.09	1.53	0.47	0.08	1.52
All	−0.42	0.09	2.37	0.44	0.08	2.38
Moneyness: OTM						
1-2M	−0.16	0.12	1.93	0.17	0.04	1.91
2-3M	−0.22	0.10	1.45	0.24	0.05	1.85
3-6M	−0.27	0.09	1.26	0.31	0.05	1.36
All	−0.22	0.10	1.55	0.24	0.05	1.71
Moneyness: DOTM						
1-2M	−0.03	0.08	0.74	0.02	0.03	0.49
2-3M	−0.08	0.08	0.53	0.07	0.03	0.90
3-6M	−0.12	0.09	0.38	0.11	0.01	1.17
All	−0.08	0.08	0.55	0.07	0.02	0.86

Panel B: Factor explanatory power for realized option return

	Put options				Call options			
τ group	MKT	VAR	GAM	Tot R^2	MKT	VAR	GAM	Tot R^2
Moneyness: All								
1-2M	0.60	0.36	0.04	1.01	0.73	0.23	0.03	0.99
2-3M	0.61	0.36	0.03	1.00	0.71	0.24	0.03	0.98
3-6M	0.61	0.36	0.02	0.99	0.70	0.23	0.02	0.95
All	0.61	0.36	0.03	1.00	0.72	0.23	0.02	0.97
Moneyness: ATM								
1-2M	0.58	0.37	0.04	0.99	0.71	0.25	0.03	0.99
2-3M	0.60	0.35	0.03	0.99	0.71	0.25	0.02	0.98
3-6M	0.61	0.35	0.02	0.98	0.71	0.24	0.01	0.97
All	0.60	0.36	0.03	0.99	0.71	0.24	0.02	0.98
Moneyness: OTM								
1-2M	0.54	0.42	0.05	1.01	0.68	0.27	0.05	0.99
2-3M	0.56	0.39	0.04	0.99	0.69	0.25	0.03	0.98
3-6M	0.57	0.39	0.03	0.99	0.70	0.24	0.02	0.96
All	0.56	0.40	0.03	0.99	0.70	0.24	0.03	0.97
Moneyness: DOTM								
1-2M	0.48	0.54	0.06	1.08	0.64	0.23	0.09	0.96
2-3M	0.50	0.48	0.05	1.03	0.65	0.25	0.08	0.99
3-6M	0.47	0.51	0.04	1.02	0.69	0.19	0.04	0.92
All	0.48	0.51	0.04	1.03	0.69	0.21	0.05	0.95

* Note that the Shapley-Owen R^2 values can sum to more than one when evaluated on a subset of the original estimation sample.

Table 3: Factor model hedging performance vs Black-Scholes and Heston

The table reports ratios of variance of unlevered option returns hedged with Black-Scholes (Panel A) and Heston (1993) Greeks and with estimated $\widehat{\beta}$ coefficients, $\frac{V\left[R^O_{t+\Delta t}-\text{H/BS hedge}\right]}{V\left[R^O_{t+\Delta t}-\text{Estimated hedge}\right]}$. Values above (below) 1 imply that the given option pricing model model hedges away less (more) than our estimates. In each column we report, respectively, the results for: 1. Delta hedging, 2. Delta hedging followed by Vega hedging, 3. Delta hedging followed by Gamma hedging, 4. Delta hedging followed by Vega hedging, followed by Gamma hedging. We define the BS Vega as $V \equiv \partial O/\partial \sigma^2_{IV}$, where O denotes the option price.

Panel A: Black-Scholes model

	Put options				Call options			
τ group	Δ	$\Delta+V$	$\Delta+\Gamma$	$\Delta+V+\Gamma$	Δ	$\Delta+V$	$\Delta+\Gamma$	$\Delta+V+\Gamma$
Moneyness: ATM								
1-2M	0.83	2.59	0.91	4.58	0.88	3.95	1.21	7.87
2-3M	0.81	5.08	0.80	7.22	1.06	8.80	1.28	13.32
3-6M	0.74	9.03	0.71	10.91	1.22	15.44	1.35	19.23
All	0.79	5.82	0.79	7.96	1.05	10.02	1.29	14.37
Moneyness: OTM								
1-2M	0.97	2.64	1.07	4.79	1.05	3.57	1.29	7.23
2-3M	1.02	5.44	1.04	7.67	1.12	7.42	1.21	11.16
3-6M	1.10	10.04	1.08	12.37	1.27	13.91	1.30	17.24
All	1.03	5.94	1.07	8.37	1.14	8.15	1.27	12.04
Moneyness: DOTM								
1-2M	0.79	2.60	0.84	4.66	0.82	4.59	1.22	9.60
2-3M	0.78	4.91	0.75	7.28	1.03	9.75	1.29	15.09
3-6M	0.77	9.92	0.74	12.05	1.26	17.41	1.42	22.23
All	0.78	6.19	0.77	8.59	1.04	11.54	1.31	16.94
Moneyness: All								
1-2M	0.50	2.15	0.52	3.11	0.49	4.08	0.93	7.40
2-3M	0.48	3.94	0.44	5.36	0.93	11.33	1.46	17.31
3-6M	0.37	5.69	0.34	6.59	1.06	15.06	1.35	18.34
All	0.43	4.61	0.40	5.71	0.83	12.00	1.23	15.96

Panel B: Heston model with filtered states

	Put options					Call options				
τ group	Δ	V	Γ	$\Delta+V$	$\Delta+V+\Gamma$	Δ	V	Γ	$\Delta+V$	$\Delta+V+\Gamma$
Moneyness: ATM										
1M-2M	1.06	0.93	1.00	1.95	2.34	1.28	1.21	1.05	1.85	2.26
2M-3M	1.09	0.84	1.00	2.02	2.20	1.33	1.27	1.03	1.85	1.99
3M-6M	1.12	0.83	0.99	1.93	2.00	1.29	1.29	1.04	1.72	1.78
All	1.09	0.87	1.00	1.97	2.18	1.30	1.26	1.04	1.81	2.01
Moneyness: OTM										
1M-2M	0.73	1.24	1.02	1.88	2.37	1.72	1.32	1.10	1.75	2.50
2M-3M	0.78	1.00	1.00	1.90	2.16	1.84	1.36	1.05	1.74	1.94
3M-6M	0.88	0.96	0.99	1.91	1.94	1.53	1.40	1.02	1.42	1.47
All	0.79	1.04	1.00	1.90	2.12	1.68	1.37	1.05	1.61	1.87
Moneyness: DOTM										
1M-2M	0.58	3.33	1.05	2.39	2.51	2.36	1.55	1.27	1.35	2.33
2M-3M	0.49	1.97	1.01	1.24	1.23	2.34	1.58	1.12	1.24	1.41
3M-6M	0.41	1.73	0.99	0.97	0.96	1.86	1.74	1.05	1.02	1.08
All	0.48	2.03	1.01	1.30	1.31	2.09	1.66	1.11	1.14	1.37
Moneyness: All										
1M-2M	0.80	1.05	1.01	1.99	2.41	1.53	1.26	1.08	1.75	2.36
2M-3M	0.76	0.93	1.00	1.81	1.97	1.67	1.32	1.04	1.68	1.84
3M-6M	0.72	0.93	0.99	1.64	1.67	1.51	1.38	1.04	1.41	1.46
All	0.76	0.97	1.00	1.80	1.97	1.57	1.32	1.05	1.59	1.81

Table 4: Expected returns on options

Expected unlevered option returns in percent per annum. In Panel A we present the decomposition expected returns due to the exposure to each factor, the sample average of $\beta^O_{lt} \cdot \lambda_{lt}$. In Panel B we present Mincer-Zarnowitz regressions of realized returns on expected returns,

$$R_{O,t,t+\Delta_t} = \alpha + \beta E_t[R_{O,t,t+\Delta_t}] + \nu_{O,t} \, .$$

In column p-val α we report the p-value of the test of the hypothesis that $\alpha = 0$. In column p-val β we report the p-value of the test of the hypothesis that $\beta = 1$.

Panel A: Decomposition of expected returns (percent per annum)

	Put options				Call options			
τ group	(1) MKT	(2) VAR	(3) GAM	(4) Tot	(1) MKT	(2) VAR	(3) GAM	(4) Tot
Moneyness: All								
1-2M	−2.05	−1.87	−0.12	−4.03	2.40	−1.29	−0.11	1.00
2-3M	−2.29	−1.82	−0.08	−4.19	2.68	−1.33	−0.10	1.26
3-6M	−2.50	−1.95	−0.06	−4.51	2.88	−1.07	−0.08	1.73
All	−2.28	−1.88	−0.09	−4.25	2.66	−1.23	−0.10	1.33
Moneyness: ATM								
1-2M	−3.69	−2.03	−0.19	−5.92	4.03	−1.62	−0.19	2.22
2-3M	−3.75	−1.99	−0.13	−5.88	4.16	−1.77	−0.14	2.25
3-6M	−3.85	−1.86	−0.09	−5.80	4.27	−1.65	−0.09	2.53
All	−3.77	−1.96	−0.14	−5.87	4.15	−1.68	−0.14	2.33
Moneyness: OTM								
1-2M	−1.88	−2.11	−0.11	−4.10	2.17	−1.19	−0.11	0.87
2-3M	−2.22	−1.91	−0.08	−4.21	2.61	−1.32	−0.11	1.18
3-6M	−2.54	−2.01	−0.07	−4.63	3.05	−1.18	−0.08	1.79
All	−2.21	−2.01	−0.09	−4.32	2.61	−1.23	−0.10	1.28
Moneyness: DOTM								
1-2M	−0.58	−1.45	−0.04	−2.07	0.84	−1.04	−0.03	−0.23
2-3M	−0.87	−1.55	−0.03	−2.45	1.14	−0.85	−0.05	0.24
3-6M	−1.12	−1.99	−0.02	−3.13	1.33	−0.39	−0.07	0.87
All	−0.86	−1.67	−0.03	−2.55	1.11	−0.75	−0.05	0.31

Panel B: Mincer-Zarnowitz regressions

	Put options				Call options			
τ group	$\widehat{\alpha}$	p-val $\widehat{\alpha} = 0$	$\widehat{\beta}$	p-val $\widehat{\beta} = 1$	$\widehat{\alpha}$	p-val $\widehat{\alpha} = 0$	$\widehat{\beta}$	p-val $\widehat{\beta} = 1$
Moneyness: All								
1-2M	0.17	0.99	1.02	1.00	0.77	0.71	1.09	1.00
2-3M	−0.52	0.95	0.81	0.96	−0.86	0.81	1.30	0.98
3-6M	0.34	0.98	1.13	0.98	0.92	0.82	1.07	1.00
All	0.05	1.00	1.00	1.00	0.42	0.92	1.13	1.00
Moneyness: ATM								
1-2M	0.28	0.99	1.08	1.00	1.22	0.84	0.78	0.99
2-3M	−2.44	0.77	0.39	0.79	0.05	1.00	−0.01	0.77
3-6M	1.32	0.88	1.37	0.88	5.42	0.50	−0.68	0.75
All	−0.17	1.00	0.97	1.00	1.55	0.79	0.35	0.91
Moneyness: OTM								
1-2M	0.14	1.00	1.03	1.00	0.53	0.85	1.46	0.97
2-3M	−0.35	0.98	0.91	0.99	−1.28	0.75	1.97	0.91
3-6M	−0.14	1.00	0.98	1.00	−0.31	0.99	2.10	0.88
All	−0.01	1.00	0.99	1.00	0.04	1.00	1.68	0.93
Moneyness: DOTM								
1-2M	−0.11	0.99	0.72	0.93	0.90	0.24	0.95	1.00
2-3M	0.60	0.89	1.32	0.93	−0.10	0.99	3.70	0.27
3-6M	0.37	0.96	1.14	0.98	−0.07	1.00	1.20	0.99
All	0.17	0.98	0.99	1.00	0.46	0.52	1.49	0.94

Table 5: Factor premia estimates

Estimates of risk premia parameters (Panel A) and summary statistics (Panel B) of the time-series of risk premia for the baseline model specification. The sample runs from January 1996 until June 2019. p-values are reported in parentheses. Coefficient standard errors and F-statistics are robust to general autocorrelation, cross-correlation and cross-sectional correlation patterns between options and account for the measurement error in option betas (see Appendix B for details).

Panel A: Estimates of λ

$E_t[\text{MKT}] = \lambda_0^{(1)} + \lambda_1^{(1)} \cdot VIX_t^2 + \lambda_2^{(1)} \cdot TVAR_t$			
λ_0	λ_1	λ_2	F-stat
1.62e-05	0.00692	0.00423	2.67
(0.966)	(0.735)	(0.953)	(0.446)

$E_t[\text{VAR}] = \lambda_0^{(2)} + \lambda_1^{(2)} \cdot VIX_t^2$			
λ_0	λ_1	λ_2	F-stat
5.36e-06	−0.0192	–	6.02
(0.993)	(0.549)	–	(0.049)

$E_t[\text{GAM}] = \lambda_0^{(3)}$			
λ_0	λ_1	λ_2	F-stat
−2.3e-06	–	–	0.19
(0.666)	–	–	(0.666)

Panel B: Summary statistics of risk premia

Factor	Average (%)	Median (%)	Std. dev. (%)	Skewness	Kurtosis	Exp. sign (%)
MKT	9.1	6.8	0.57	5.0	40.8	100.0
VAR	−22.1	−16.0	1.45	−4.9	40.5	100.0
GAM	−0.1					

Table 6: Factor premia in alternative model specifications and sample periods

Summary statistics of factor risk premia with alternative premia specifications (Panel A), with additional factors (Panel B), under different basis dimension choices (Panel C), and in subsamples (Panel D). The results for our baseline model are repeated in the first column of each panel. The average premium and its standard deviation are reported in percent per annum. The row denoted "Exp. sign" reports the fraction, in percent, of the sample period in which the factor premium is of the expected sign, i.e., positive in the case of MKT and HKM, and negative in the case of VAR and TVAR. The baseline sample is from 1996 until 2019, except for the model with the HKM factor, where the sample starts in 2000.

Panel A: Alternative premia specifications

	(1) Baseline		(2) 3F Constant		(3) 3F Semi-VIX		(4) 3F Quadratic	
	MKT	VAR	MKT	VAR	MKT	VAR	MKT	VAR
Average (%)	9.1	−22.1	7.9	−8.0	10.5	−17.3	7.9	−23.1
Std. dev. (%)	0.6	1.5			0.8	1.8	0.7	1.9
Skewness	5.0	−4.9			4.8	−4.8	11.7	−7.5
Kurtosis	40.8	40.5			38.4	47.5	183.6	86.8
Exp. sign (%)	100.0	100.0	100.0	100.0	100.0	89.4	100.0	100.0

Panel B: Additional factors

	(1) Baseline		(2) 2F		(3) 3F + HKM			(4) 3F + TVAR		
	MKT	VAR	MKT	VAR	MKT	VAR	HKM	MKT	VAR	TVAR
Average (%)	9.1	−22.1	11.2	−20.0	6.2	−22.2	32.8	12.2	−20.5	−7.1
Std. dev. (%)	0.6	1.5	0.9	1.3	0.4	1.9	0.5	0.9	1.1	2.1
Skewness	5.0	−4.9	4.7	−4.7	5.1	−4.9	4.1	4.9	−5.0	−5.4
Kurtosis	40.8	40.5	38.0	38.0	40.1	38.1	29.4	40.8	41.1	47.4
Exp. sign (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	44.0

Panel C: Robustness to basis dimension

	(1) Baseline ($k = 20$)		(2) $k = 10$		(3) $k = 30$	
	MKT	VAR	MKT	VAR	MKT	VAR
Average (%)	9.1	−22.1	9.3	−24.4	9.2	−23.0
Std. dev. (%)	0.6	1.5	0.6	1.6	0.6	1.7
Skewness	5.0	−4.9	4.9	−4.9	5.0	−4.9
Kurtosis	40.8	40.5	40.2	40.5	40.8	40.5
Exp. sign (%)	100.0	100.0	100.0	100.0	100.0	100.0

Panel D: Subsamples

	(1) Baseline		(2) 1996–2007		(3) 2008–2019		(4) Apr 2009–2019	
	MKT	VAR	MKT	VAR	MKT	VAR	MKT	VAR
Average (%)	9.1	−22.1	3.4	−4.4	10.7	−36.4	20.1	−31.6
Std. dev. (%)	0.6	1.5	0.7	0.6	0.8	3.3	2.2	2.3
Skewness	5.0	−4.9	2.2	−1.5	4.5	−4.5	2.7	−2.6
Kurtosis	40.8	40.5	10.1	6.4	30.1	29.9	12.4	11.4
Exp. sign (%)	100.0	100.0	52.3	60.7	100.0	100.0	74.0	98.8

Table 7: Option exposures and expected returns with additional factors

Scaled option exposures ($\beta \times 10^2$), Shapley-Owen values of R^2 (Huettner and Sunder, 2012) and contribution to expected returns ("Exp. ret.") of additional factors in alternative model specifications. The presented estimates are from four-factor models that add the HKM or ΔTVAR factor to the baseline specification. The contribution to expected return is calculated as the sample average of the products of the exposures and premia and is reported in percent per annum. Exposures and returns are averaged within each day (and bucket) before calculating the time-series averages. HKM is the intermediary risk factor of He et al. (2017) with the risk premium specified as in equation (12) therein, i.e., an affine function of the square of the intermediary leverage ratio; daily data on the HKM factor was obtained from the website of Asaf Manela. ΔTVAR is the daily increment in the square of the Tail Volatility factor of Bollerslev et al. (2015); daily data on Tail Volatility was obtained from the TailIndex website

	Put options						Call options					
	HKM			TVAR			HKM			TVAR		
Maturity	$\beta \times 10^2$	R^2	Exp. ret.	$\beta \times 10^2$	R^2	Exp. ret.	$\beta \times 10^2$	R^2	Exp. ret.	$\beta \times 10^2$	R^2	Exp. ret.
Moneyness: All												
1-2M	0.07	0.130	0.03	0.231	0.004	−0.11	0.27	0.169	0.09	0.944	0.006	−0.16
2-3M	−0.15	0.063	−0.04	0.012	0.003	−0.07	0.48	0.145	0.16	1.055	0.003	−0.18
3-6M	−0.33	0.003	−0.10	−1.757	0.007	0.01	0.67	0.138	0.22	4.230	0.006	−0.34
All	−0.14	0.047	−0.04	−0.514	0.000	−0.05	0.47	0.147	0.16	2.100	0.003	−0.23
Moneyness: ATM												
1-2M	0.24	0.142	0.09	−1.519	0.005	−0.03	0.47	0.182	0.16	−0.547	0.001	−0.07
2-3M	0.02	0.101	0.02	−2.979	0.016	0.06	0.28	0.160	0.10	−1.871	0.004	0.01
3-6M	−0.03	0.077	0.00	−0.776	0.000	−0.03	0.26	0.170	0.09	0.734	0.004	−0.10
All	0.08	0.105	0.03	−1.740	0.006	0.00	0.34	0.166	0.12	−0.550	0.000	−0.06
Moneyness: OTM												
1-2M	−0.04	0.110	−0.01	−0.292	0.009	−0.07	0.42	0.181	0.14	1.376	0.016	−0.16
2-3M	−0.25	0.032	−0.07	1.434	0.032	−0.15	0.61	0.191	0.21	1.728	0.015	−0.18
3-6M	−0.45	0.011	−0.14	−1.774	0.011	0.00	0.71	0.197	0.24	3.879	0.014	−0.29
All	−0.25	0.015	−0.07	−0.236	0.003	−0.07	0.58	0.192	0.19	2.334	0.012	−0.21
Moneyness: DOTM												
1-2M	0.00	0.040	0.00	2.507	0.058	−0.23	−0.12	0.127	−0.03	2.136	0.013	−0.25
2-3M	−0.20	0.053	−0.06	1.575	0.045	−0.11	0.54	0.161	0.19	3.526	0.016	−0.38
3-6M	−0.50	0.071	−0.16	−2.708	0.024	0.07	1.03	0.196	0.34	8.082	0.028	−0.63
All	−0.24	0.049	−0.07	0.428	0.002	−0.09	0.50	0.180	0.17	4.688	0.017	−0.43

Table 8: Alternative variance proxies

This table summarizes the consequences of replacing VAR with alternative variance factor proxies. In Panel A we report the summary statistics of the alternative variance proxies. In Panel B we report the summary statistics of the premia on the MKT factor, and on each variance proxy. The standard deviation is reported on an annualized basis and can be understood as a volatility-of-volatility. ρ_{VAR} denotes the correlation of the proxy with the VAR factor, and ρ_{SPOT} denotes the correlation of a proxy with the SPOT proxy. In Panel C we report aggregate information on option sensitivities with respect to the alternative variance proxies, and the proxies' explanatory power, measured as the Shapley-Owen value of R^2. In Panel D we report the coefficients of aggregate Mincer-Zarnowitz regressions of realized option returns on expected option returns obtained with each variance factor proxy. The proxies are: SPOT, i.e. the daily increment in the square of the spot volatility of Todorov (2019) obtained from www.tailindex.com, GARCH, i.e. the daily increment in the annualized daily filtered variance from a GJR-GARCH(1,1) (Glosten et al., 1993) model with Student-t innovations estimated on our underlying SPX excess return dataset, RV and IQV, i.e., realized variance and integrated quadratic variation (estimated using the MedRV measure of Andersen, Dobrev, and Schaumburg (2012)) of intraday S&P 500 index returns, obtained from realized.oxford-man.ox.ac.uk.

	Panel A: Summary statistics of alternative variance factor proxies				
	(1) VAR	(2) SPOT	(3) GARCH	(4) RV	(5) IQV
Std. dev.	0.19	0.19	0.19	0.68	0.36
Skewness	0.35	−0.34	5.89	13.09	6.13
Kurtosis	94.2	70.8	104.3	452.2	217.7
ρ_{VAR}		0.67	0.76	0.22	0.32
ρ_{SPOT}			0.61	0.17	0.17

	Panel B: Premia with alternative variance factor proxies									
	(1) Baseline		(2) SPOT		(3) GARCH		(4) RV		(5) IQV	
	MKT	VAR	MKT	SPOT	MKT	GARCH	MKT	RV	MKT	IQV
Average (%)	9.1	−22.1	9.5	−51.6	8.0	−36.8	1.3	−664.0	5.3	−62.8
Std. dev. (%)	0.6	1.5	0.6	4.6	0.5	4.2	1.5	116.9	0.8	16.4
Skewness	5.0	−4.9	5.0	−4.7	5.9	−5.9	−5.1	−7.6	−6.1	−11.1
Kurtosis	40.8	40.5	39.1	33.4	53.3	53.3	59.3	83.8	110.4	186.8
Exp. sign (%)	100.0	100.0	99.7	100.0	100.0	86.4	68.1	68.2	85.6	49.1
Sample	1996–2019		2008–2019		1996–2019		2000–2019		2000–2019	

	Panel C: Explanatory power of variance proxies									
	(1) Baseline		(2) SPOT		(3) GARCH		(4) RV		(5) IQV	
	Puts	Calls	Puts	Calls	Puts	Calls	Puts	Calls	Puts	Calls
Sens.	0.09	0.05	0.07	0.04	1.05	2.72	2.54	2.56	2.24	2.48
R^2	0.36	0.23	0.24	0.10	0.02	0.03	0.03	0.03	0.03	0.03

	Panel D: Mincer-Zarnowitz regressions for variance proxies									
	(1) Baseline		(2) SPOT		(3) GARCH		(4) RV		(5) IQV	
	Puts	Calls	Puts	Calls	Puts	Calls	Puts	Calls	Puts	Calls
$\widehat{\alpha}$	0.05	0.42	1.64	1.34	−1.45	0.69	−2.18	0.81	−1.53	−0.50
$\widehat{\beta}$	1.00	1.13	1.32	0.08	0.36	0.62	0.94	0.47	0.98	1.10