Beyond Carr–Madan: A Projection Approach to Risk-Neutral Moment Estimation

Tjeerd De Vries*

January 28, 2026

Abstract

We propose a projection method to estimate risk-neutral moments from option prices. We derive a finite-sample bound implying that the projection estimator attains (up to a constant) the smallest pricing error within the span of traded option payoffs. No analogous guarantee is available for the widely used Carr–Madan approximation. We then extend the framework to multiple underlyings, deriving necessary and sufficient conditions under which simple options complete the market in higher dimensions, and providing estimators for joint moments. Simulations demonstrate that the method remains accurate in sparse-strike settings and in higher dimensions. In the empirical application, we recover risk-neutral correlations and joint tail risk from FX options alone, addressing a longstanding measurement problem raised by Ross (1976). Our joint tail-risk measure predicts future joint currency crashes and identifies periods in which currency portfolios are particularly useful for hedging.

1 Introduction

Option prices provide real-time, forward-looking information about the state of the economy. Their tractability and informational content have made them central to a wide range of empirical applications, including forecasting the equity premium,

*Department of Finance, HEC Paris. Email: de-vries@hec.fr. I thank Nikolay Kudrin, Evgenii Vladimirov, participants of the HEC Brownbag seminar, and especially Irina Zviadadze for helpful feedback.

predicting volatility, and measuring skewness and higher-order risk-neutral moments.[1] A widely used approach for extracting such quantities is the method of Carr and Madan (2001) (henceforth, CM), which expresses the risk-neutral expectation of a twice-differentiable payoff as a weighted integral over put and call prices. Because option prices are observed across a range of strikes on any given day, the integral can be approximated numerically, enabling the practical estimation of objects such as the VIX and other risk-neutral measures.

Given the substantial notional amounts traded in derivatives such as VIX options, accurate measurement of risk-neutral quantities is essential. Measurement error in these quantities can also distort inference about the informational content of option prices and their predictive power for future market outcomes. This paper proposes a new method for estimating risk-neutral quantities that improves significantly on the standard approach. Rather than approximating payoffs using a second-order Taylor expansion around the forward price, as in CM, we project the target payoff function onto the linear span of payoffs from traded instruments—specifically, puts, calls, and the underlying.

The approach generalizes the classical put-call parity identity, which arises from an exact replication of a constant payoff using a portfolio of the underlying, a put, and a call. In our framework, the constant function is just one element of a broader class of payoffs that can be projected onto this same payoff space. For any such projection, the risk-neutral expectation can be computed directly from observed option prices, yielding a tractable, model-free estimator.

This *projection-based approach* offers several advantages over the widely used method of CM. First, it allows for extrapolation beyond the range of observed strike prices, which is particularly important when option quotes do not extend sufficiently into the tails. This allows the researcher to incorporate prior beliefs about the relevant support of the risk-neutral distribution even when strikes are sparse in the tails. Effectively, the observed option payoffs are used to form the best approximation to the target payoff over the chosen domain. Moreover, the resulting estimate corresponds directly to an investable portfolio constructed from traded options, whereas common extensions of the CM formula rely on curve fitting and extrapolation to impute unobserved option prices (e.g., Jiang and Tian (2005)).

[1] See, for example, Bates (1991), Andersen et al. (2017), Martin (2017), Kremens and Martin (2019), and Schneider and Trojani (2019) for predicting the equity premium; Britten-Jones and Neuberger (2000), Carr and Madan (2001), Jiang and Tian (2005), Bollerslev et al. (2009), and Carr and Wu (2009) for volatility forecasting; and Bakshi et al. (2003), Kozhan et al. (2013), and Chabi-Yo and Loudis (2020) for higher-order moment estimation.

Second, the projection approach enjoys good finite-sample properties. In particular, we derive a bound which implies that the projection-based pricing error is, up to a constant, the smallest attainable among portfolios spanned by the traded option payoffs. An analogous guarantee is not available for the CM approach, even though it uses the same set of observed options. This finite-sample optimality complements our asymptotic results. In an idealized framework, we show that projection and CM converge at the same rate to the true risk-neutral expectation, and under strong assumptions they asymptotically assign the same portfolio weights. These equivalence results break down in realistic settings with irregular strike spacing and limited tail coverage. Simulations illustrate the resulting finite-sample gains, showing that projection yields substantially more accurate estimates of key quantities such as VIX and SVIX. This improvement is particularly relevant in our FX application, where only five strikes are available and quotes do not extend far into the tails.

Third, unlike the CM approach, the projection method can be used to estimate the full risk-neutral distribution. This is central to a large literature on recovering measures of risk aversion and pricing kernels.[2] Our estimator satisfies a key internal consistency condition: it exactly reproduces the observed option prices. This is not guaranteed by most existing approaches. Furthermore, unlike the classical method of Breeden and Litzenberger (1978), our approach does not require numerical differentiation of the option price surface. This is an important advantage, as estimating second derivatives is often unstable in practice due to the irregular spacing of strike prices.

Fourth, projection generalizes to higher dimensions. Prior work shows that options on individual stocks cannot pin down joint risk-neutral expectations (Martin, 2018, 2025). We formalize this in Proposition 8, which proves the impossibility of identifying correlation from single-name options alone. To overcome this, we incorporate information from index options, which embed constraints on the joint distribution of the constituents' returns.

In this more complicated setting, we derive necessary and sufficient conditions under which simple options complete the market for the payoff class we study. The key step is an equivalence: market completeness obtains precisely when ridge

[2]See, for example, Aït-Sahalia and Lo (2000), Jackwerth (2000), Bliss and Panigirtzoglou (2004), and Almeida and Freire (2022) for estimates of risk aversion; and Ross (1976), Breeden and Litzenberger (1978), Jackwerth and Rubinstein (1996), Aït-Sahalia and Lo (1998), Rosenberg and Engle (2002), Bondarenko (2003), Figlewski (2010), Filipović et al. (2013), Ross (2015), Beare and Schmidt (2016), Linn et al. (2017), and Figlewski (2018) for estimates of the pricing kernel or risk-neutral density.

functions $x \to g(w'x)$ are dense in the uniform topology, and the latter question is well studied in approximation theory (e.g., Pinkus, 2015). Ridge representations are also familiar in econometrics through projection pursuit (Friedman and Stuetzle, 1981): the difference here is that the directions w are fixed by portfolio weights, whereas projection pursuit also optimizes over w.

The density result for ridge functions requires observations on infinitely many distinct portfolio options, or equivalently, an unbounded set of portfolio weights w. In practice only a finite collection is observed. For example, options on the SPDR ETF together with its 11 sector funds yield 12 distinct weights $\{w_j\}_{j=1}^{12}$. Estimating correlations or other measures of joint dependence therefore becomes an inverse problem: we seek to recover those quantities from the finite set of portfolio returns, i.e. from line projections in $\mathbb{R}^d$. Closely related problems arise in tomography and compressed sensing, where functionals of a distribution are reconstructed from line integrals (e.g., Candès et al., 2006). Despite the finite menu of portfolios, projection can yield informative estimates of joint risk-neutral moments.

We also consider joint dependence estimation in FX returns, focusing on EUR/USD and GBP/USD. This setting is particularly clean because triangular parity introduces a traded cross rate, EUR/GBP, satisfying $S_{\text{EUR/GBP}} = S_{\text{EUR/USD}}/S_{\text{GBP/USD}}$. Options on the cross therefore contain information about the joint risk-neutral distribution of the two leg returns. While we show that vanilla options do not complete the market for the two legs, our projection approach nevertheless recovers option-implied correlations with very high accuracy in simulations and allows accurate estimation of joint probabilities, addressing a longstanding measurement problem for return dependence.[3] These estimates can be used, for instance, to infer the option-implied variance of currency portfolios and to calibrate empirical models of joint currency risk (e.g., Chernov et al. (2018)).

Particular care is required when constructing portfolios that replicate joint-dependence measures because options on the cross rate are quoted in GBP, whereas options on the two dollar rates are quoted in USD. Valuing all payoffs under a common (USD) numéraire introduces a state-dependent conversion term, namely the pricing kernel that converts GBP-denominated payoffs into USD units. Our projection approach incorporates this numéraire-change term directly, yielding a portfolio that is fully implementable for a U.S. investor. This contrasts with existing approaches in the FX literature which effectively treat the conversion

[3]See, for example, Ross (1976), Martin (2018), Bondarenko and Bernard (2024), and Martin (2025) on estimating joint risk-neutral probabilities.

kernel as constant (e.g., Mueller et al. (2017)).

We estimate the forward-looking (risk-neutral) correlation between EUR/USD and GBP/USD to average about 0.7 over the sample, with pronounced time variation. The correlation reaches a local minimum around the June 2016 Brexit vote, near 0.2. A variance decomposition indicates that this decline is largely accounted for by a spike in the volatility of GBP/USD, with little contemporaneous change in EUR/USD volatility. We also estimate the risk-neutral probability that both monthly returns fall by at least 3%. This measure forecasts subsequent downside outcomes: in a predictive regression, its coefficient is statistically significant in-sample. Reduced-form evidence points to state dependence in risk compensation. In tranquil periods, the joint crash probability under the risk-neutral measure is below its physical counterpart, consistent with option portfolios providing hedge-like payoffs. During stress episodes (e.g., the 2008 financial crisis), the ordering reverses, implying higher compensation required for exposure to joint crash risk.

The rest of this paper is structured as follows. Section 2 reviews the CM approach and introduces the projection method. Section 3 derives the convergence properties of the projection approach and establishes an equivalence with risk-neutral density estimation. Section 4 extends the projection method to higher dimensions and shows how joint risk-neutral moments can be estimated. Section 5 presents evidence on the finite-sample performance using Monte Carlo simulation, and Section 6 presents the main empirical findings. Finally, Section 7 concludes.

2 Estimating nonlinear payoffs using projection

In this section, we introduce the projection method to estimate risk-neutral moments. We first review Carr and Madan (2001) to benchmark our approach.

2.1 Carr-Madan approach

Let $g(S_T)$ denote a payoff at maturity T as a function of the realized stock price S_T. Our object of interest is the conditional risk-neutral expectation $\mathbf{E}_t^Q[g(S_T)]$. The CM approach constructs a portfolio of puts and calls that replicates $g(S_T)$ state by state. By the law of one price, $\mathbf{E}_t^Q[g(S_T)]$ equals the time-t value of this replicating portfolio, which can be computed from observed option prices.

To implement this idea, CM start from a second-order Taylor expansion with

integral remainder,

$$g(S_T) = g(F_{t\to T}) + g'(F_{t\to T})(S_T - F_{t\to T}) + \int_0^{F_{t\to T}} g''(K)\,(K - S_T)^+\,\mathrm{d}K + \int_{F_{t\to T}}^{\infty} g''(K)\,(S_T - K)^+\,\mathrm{d}K, \tag{1}$$

where $F_{t\to T}$ is the time-t forward price for maturity T. Using risk-neutral valuation, we obtain

$$\mathbf{E}_t^Q g(S_T) = g(F_{t\to T}) + R_{f,t\to T}\int_0^{F_{t\to T}} g''(K)P_{t\to T}(K)\,\mathrm{d}K + R_{f,t\to T}\int_{F_{t\to T}}^{\infty} g''(K)C_{t\to T}(K)\,\mathrm{d}K, \tag{2}$$

where $R_{f,t\to T}$ is the gross risk-free rate from t to T, and $P_{t\to T}(K)$ and $C_{t\to T}(K)$ denote European put and call option prices with strike K and maturity T.

In practice, option prices are observed only at a discrete set of strikes, so the integrals in (2) are approximated by a trapezoidal rule. For example, for observed put strikes $K_0 < \cdots < K_J \le F_{t\to T}$,

$$\int_0^{F_{t\to T}} g''(K)P_{t\to T}(K)\,\mathrm{d}K \approx \sum_{j=0}^{J} g''(K_j)P_{t\to T}(K_j)\,\Delta K_j, \tag{3}$$

$$\Delta K_0 \coloneqq K_1 - K_0, \quad \Delta K_J \coloneqq K_J - K_{J-1}, \quad \Delta K_j \coloneqq \frac{K_{j+1} - K_{j-1}}{2}\ (1 \le j \le J-1).$$

This is the trapezoidal discretization used in the CBOE's VIX methodology and in related model-free moment estimators. We refer to (3) as the CM *approximation or discretization*, to distinguish it from the exact CM formula in (2). Before introducing our projection-based alternative, we illustrate how (3) is used in two canonical applications.

Example 1 (Risk-neutral variance (SVIX))**.** Martin (2017) derives a bound on the conditional expected market return using the risk-neutral variance:

$$\mathbf{E}_t R_{t\to T} - R_{f,t\to T} \ge \frac{1}{R_{f,t\to T}}\mathbf{Var}_t^Q R_{t\to T},$$

where $R_{t\to T} = S_T/S_t$ is the return on the stock. To compute this bound from the data, it is necessary to calculate $\mathbf{E}_t^Q S_T^2$. The CM approximation can then be used with $g(S_T) = S_T^2$ and $g''(S_T) = 2$.

Example 2 (Risk-neutral entropy (VIX))**.** The VIX is a popular measure of market uncertainty and is defined by the risk-neutral entropy of returns (Martin,

2017):

$$\text{VIX}^2_{t\to T} = \frac{2}{T-t}\left(\log R_{f,t\to T} - \mathbf{E}^Q_t \log R_{t\to T}\right). \tag{4}$$

Entropy, just like variance, is a measure of variability of a random variable. In this case it is necessary to calculate the expectation of a log-return, which can be accomplished with the CM approximation using $g(S_T) = \log(S_T)$ and $g''(S_T) = -1/S_T^2$. Britten-Jones and Neuberger (2000) further show that the VIX measures the risk-neutral expected volatility from time t to $t+T$.

In addition to these examples, there are important settings in which the CM formula does not directly apply. The next two examples illustrate cases that are central for empirical work.

Example 3 (Risk-neutral distribution)**.** The estimation of the risk-neutral density is not covered by the CM formula because the payoff function necessary to calculate the PDF corresponds to a "discontinuous function". However, Breeden and Litzenberger (1978) show that the risk-neutral CDF and PDF can be derived from

$$\begin{aligned} F^Q_{t\to T}(K) &= \mathbf{E}^Q_t \mathbb{1}\left(\{S_T \le K\}\right) = 1 + R_{f,t\to T}\frac{\partial}{\partial K} C_{t\to T}(K) \\ f^Q_{t\to T}(K) &= \frac{\partial}{\partial K} F^Q_{t\to T}(K) = R_{f,t\to T}\frac{\partial^2}{\partial K^2} C_{t\to T}(K). \end{aligned}$$

These formulas are widely used to estimate risk-neutral densities and, when combined with additional information on physical probabilities, to infer pricing kernels and risk aversion. We will show that projection can also be used to estimate the risk-neutral distribution, thereby treating Examples 1–3 in a unified manner.

Example 4 (Risk-neutral covariance and correlation)**.** For hedging purposes, it is often useful to estimate the risk-neutral covariance between two stock returns (see, e.g., Lustig et al. (2014)). In a different direction, the risk-neutral covariance between the market return and an individual stock also allows us to infer that stock's equity premium when the representative investor has log utility (Martin (2025)):

$$\mathbf{E}_t R_{i,t\to T} - R_{f,t\to T} = \frac{1}{R_{f,t\to T}} \mathbf{Cov}^Q_t\left(R_{i,t\to T}, R_{t\to T}\right).$$

In this case, the CM formula neither applies because it is inherently univariate. Generally, estimating a covariance from options remains an open problem.[4] Sec-

[4] In certain settings the covariance is identifiable from option prices, e.g., for quanto options (Kremens and Martin, 2019), or one can estimate it by imposing additional constraints, such as maximizing entropy (see Bondarenko and Bernard (2024)).

tion 4 shows how the projection approach extends to the multivariate setting, allowing one to estimate these correlations.

It can also be of interest to estimate the joint risk-neutral distribution. However, there is no higher-dimensional analogue of Breeden and Litzenberger (1978). We derive necessary and sufficient conditions on the option market that guarantee a unique multivariate risk-neutral measure. Although these conditions are typically not met in practice, the projection approach can nonetheless yield accurate approximations.

2.2 A simple illustration of the projection method

To illustrate the projection approach to estimating risk-neutral expectations of non-linear payoffs, consider the following simple example.

Example 5 (Projection approach)**.** Suppose the stock price at time T can take four possible values: $S_T = [10, 11, 12, 13]'$. We aim to replicate the payoff of the squared stock value, S_T^2. Assume we can trade a risk-free asset with return $R_{f,t\to T}$, the stock itself, and a call option on the stock with strike $K = 12$. The squared stock value and the payoffs of the tradable assets, denoted by the matrix X, are given by

$$S_T^2 = \begin{pmatrix} 100 \\ 121 \\ 144 \\ 169 \end{pmatrix}, \qquad X = \begin{pmatrix} 1 & 10 & 0 \\ 1 & 11 & 0 \\ 1 & 12 & 0 \\ 1 & 13 & 1 \end{pmatrix}.$$

Clearly the market in this example is not complete because the value of S_T^2 cannot be replicated perfectly by a portfolio of tradable assets. To find a portfolio that comes closest to replicating S_T^2, a natural idea is to project S_T^2 onto the space spanned by X:

$$S_T^2 \approx X\hat{\beta}, \quad \text{where } \hat{\beta} = (X'X)^{-1} X'S_T^2.$$

Because the prices of the tradable assets are observable, we can estimate the risk-neutral expectation of S_T^2 via

$$\mathbf{E}_t^Q S_T^2 \approx [1, F_{t\to T}, R_{f,t\to T} C_{t\to T}(12)]\hat{\beta}.$$

This approximation follows from risk-neutral pricing because $F_{t\to T} = \mathbf{E}_t^Q[S_T]$ and $C_{t\to T}(12) = (1/R_{f,t\to T})\mathbf{E}_t^Q[\max(S_T - 12, 0)]$. In general, the projection estimate will differ from the CM estimate, because in this example the CM approach always assigns a portfolio weight of 2 to the option, regardless of the strike price.

The projection approach also generalizes the familiar put–call parity. For example, if we replace S_T^2 with the payoff of a put option, $\max(12 - S_T, 0)$, the projection on X yields zero error, thereby recovering the classical parity relation. By contrast, put–call parity is not covered by the CM formula because the payoff functions are not twice differentiable.

2.3 General projection approach

This section generalizes the example above and introduces notation. Let the observed (ordered) out-of-the-money put and call strikes be

$$\mathbf{K}^P \coloneqq [K_1^P, \ldots, K_{n_k^P}^P]', \qquad \mathbf{K}^C \coloneqq [K_1^C, \ldots, K_{n_k^C}^C]',$$

with $K_{n_k^P}^P \leq F_{t\to T}$ and $K_1^C > F_{t\to T}$, and define the total number of strikes by $n_k \coloneqq n_k^P + n_k^C$. Let

$$\mathbf{s} \coloneqq [s_1, \ldots, s_{n_s}]'$$

denote a researcher-chosen grid of stock prices at maturity T. The choice of the endpoints (s_1, s_{n_s}) amounts to a stance on the relevant support of the risk-neutral distribution; we discuss a data-driven choice in Appendix C. Importantly, this allows us to estimate risk-neutral expectations *even outside the range of observed strikes*.

Define the payoff design matrices for puts and calls on the grid $\mathbf{s}$ by

$$X_{ij}^P \coloneqq (K_j^P - s_i)_+, \qquad X_{ij}^C \coloneqq (s_i - K_j^C)_+, \qquad i = 1, \ldots, n_s.$$

When it creates no confusion, we drop the superscripts P and C on strikes. Let $\mathbf{1}_{n_s}$ denote an n_s-vector of ones and define the state-by-state payoff matrix

$$X \coloneqq \begin{bmatrix} \mathbf{1}_{n_s} & \mathbf{s} & X^P & X^C \end{bmatrix} \in \mathbb{R}^{n_s \times (2+n_k)}.$$

If a put and a call share the same strike, including both is redundant given put–call parity and the presence of the bond and stock columns. Let $Y \in \mathbb{R}^{n_s}$ be the payoff evaluated on the grid, $Y_i \coloneqq g(s_i)$. We compute the projection of Y onto the column span of X:

$$Y = X\widehat{\beta} + \widehat{\varepsilon}, \qquad \widehat{\beta} \coloneqq (X'X)^{-1}X'Y.$$

Equivalently, this yields the approximation

$$g(S_T) \approx \hat{\beta}_1 + \hat{\beta}_2 S_T + \sum_{j=1}^{n_k^P} \hat{\beta}_j^P (K_j - S_T)_+ + \sum_{j=1}^{n_k^C} \hat{\beta}_j^C (S_T - K_j)_+ =: \hat{g}(S_T). \tag{5}$$

Taking risk-neutral expectations on both sides, we obtain a projection estimate of the risk-neutral expectation.

Definition 1 (Projection estimator)**.** Let X collect terminal payoffs at T (cash, the underlying, and options) evaluated on a state grid, and let $\hat{\beta}$ be the OLS coefficient vector from projecting the target payoff Y on X. Then the projection estimator is defined by

$$\mathbf{E}_t^Q \hat{g}(S_T) \coloneqq \hat{\beta}_1 + \hat{\beta}_2 F_{t\to T} + R_{f,t\to T} \left(\sum_{j=1}^{n_k^P} \hat{\beta}_j^P P_{t\to T}(K_j) + \sum_{j=1}^{n_k^C} \hat{\beta}_j^C C_{t\to T}(K_j) \right). \tag{6}$$

Remark 1 (Constrained least squares). In some applications—such as estimating risk-neutral variance—it is natural to impose that the estimate be nonnegative. With very few options, the least-squares replicating portfolio implied by $\hat{\beta}$ can produce a payoff that is negative over parts of the state space, which in turn can yield a negative variance estimate. In such cases, it is natural to require the replicating payoff to be nonnegative pointwise. This is achieved by solving the constrained least-squares problem

$$\min_{\beta} \ \|Y - X\beta\|_2^2 \quad \text{subject to} \quad X\beta \geq 0,$$

where the inequality is interpreted componentwise on the chosen state grid. This convex quadratic program enforces a nonnegative replication in every state and, hence, a nonnegative variance estimate. Similarly, one may impose direct restrictions on the portfolio weights, for example, the componentwise bound $\beta \geq -c$ for some $c > 0$ to reflect borrowing constraints.

Remark 2 (Weighted least squares). The replicating portfolio in (5) penalizes deviations equally across states (stock prices). In applications it can be preferable to penalize errors more heavily near the forward price—where the risk-neutral measure places more mass—and less heavily in the tails. This can be implemented via weighted least squares:

$$\hat{\beta}_{\text{wls}} = (X'WX)^{-1} X'Wy,$$

where $W = \text{diag}(w_1, \ldots, w_{n_s})$ collects state weights. The (infeasible) theoretically optimal choice sets weights proportional to the risk-neutral density, $w_i \propto f^Q_{t\to T}(s_i)$. A practical alternative is to use, for example, a Gamma weighting density calibrated so that its mean matches the forward price and its standard deviation matches the ATM (forward) implied volatility. As we show below, however, the choice of weighting affects the portfolio weights only at second order, and as the strike grid becomes sufficiently dense the resulting replicating portfolio coincides with the one obtained from the unweighted regression.

Remark 3 (Redundancy of option-implied regressors). Because the projection estimator is an OLS linear projection of the target payoff onto the span of the option basis functions, the Frisch–Waugh–Lovell theorem implies that adding any payoff that already lies in this span does not change the fitted values. For example, the CBOE VIX (Example 2) corresponds to a log contract that is replicated from options. Hence adding $\log(S_T)$ as an additional basis element and using the VIX price does not improve the estimation of a general payoff. By contrast, if there were a genuinely tradable claim delivering the log payoff (or a variance claim) whose price were not implied by the options in the basis, then adding $\log(S_T)$ would enlarge the span and improve estimation. Notice that the CM formula does not provide a generic way to exploit information from non-option payoffs.

To illustrate the benefits of the replicating portfolio obtained by projection in (5) relative to the CM discretization in (3), Figure 1 plots both replicating portfolios for a nonlinear payoff. The projection-based portfolio is nearly indistinguishable from the true payoff across the entire domain, including outside the range of observed strikes. In contrast, the CM approximation replicates the payoff much less accurately, especially in the tails. The discrepancy arises because the CM formula relies on a Taylor expansion around the forward price (see (1)), and strike prices do not go far enough in the tails to yield an accurate approximation. As a result, the risk-neutral expectation can be estimated with substantial error.

2.4 Continuous-state limit

To implement the projection method, the researcher needs to choose a grid of possible future stock values, $\mathbf{s}$. This is analogous to specifying the up and down states in the binomial option pricing model. Since the grid can be made arbitrarily fine, a natural question is what the discrete projection converges to as the mesh size tends to zero.

Throughout, we denote the set of basis functions used for portfolio replication

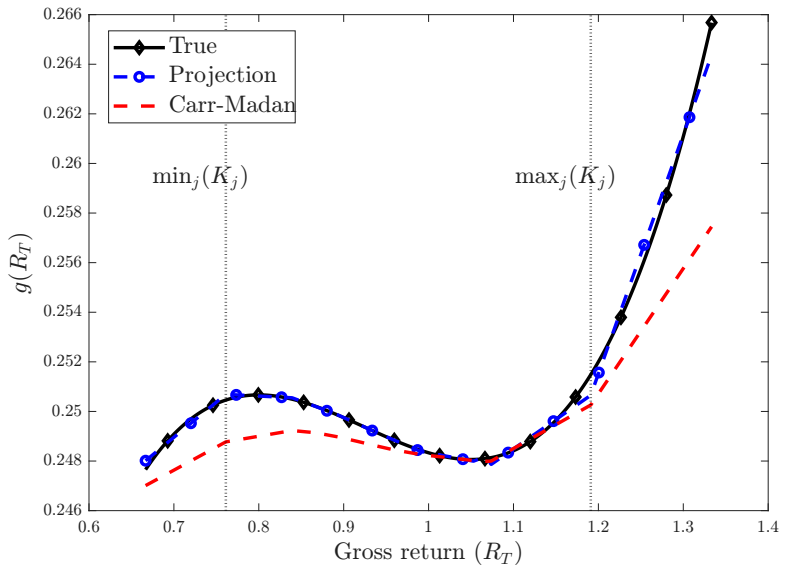


Figure 1: **Replication of cubic payoff**. The figure shows the function $g(R_{t\to T}) = (2/3)R_{t\to T}^3 - (37/40)R_{t\to T}^2 + (21/25)R_{t\to T}$ (black), together with the projection-based portfolio (blue) and CM portfolio (red). The approximations are based on 15 strike prices drawn from a uniform distribution. Dashed vertical lines indicate the minimum and maximum strike values used.

by

$$\mathcal{F}_{2+n_k} = \left\{1, S_T, (K_1 - S_T)^+, \ldots, \left(K_{n_k^P} - S_T\right)^+, (S_T - K_1)^+, \ldots, \left(S_T - K_{n_k^C}\right)^+\right\}.$$

When convenient, we index the basis as $\phi_i \in \mathcal{F}_{2+n_k}$ for $i = 1, \ldots, 2+n_k$. To derive the limiting value as $\max_i |s_{i+1} - s_i| \to 0$, we make the following assumption.

Assumption 1. *Let $A = [a_{\min}, a_{\max}]$ be a compact interval in $\mathbb{R}_{++}$ such that $a_{\min} < K_1^P$ and $a_{\max} > K_{n_k^C}^C$, and all strike prices are unique. Moreover, $g \in L^2(A)$: $\int_A g(S)^2 \, \mathrm{d}S < \infty$.*

Assumption 1 guarantees that the projection estimator is well defined when n_s is sufficiently large. In particular, because the strike prices are assumed to be unique, all basis functions are linearly independent over $L^2(A)$. The next result establishes the continuous-grid limit. By slight abuse of notation, let $\hat{\beta}_{n_s}$ denote the projection coefficients obtained from a grid of size n_s.

Proposition 1. *Let Assumption 1 hold and define an inner product on $L^2(A)$ by*

$$\langle \phi_i, \phi_j \rangle = \int_A \phi_i(S_T)\phi_j(S_T) \, \mathrm{d}S_T.$$

If $\max_i |s_{i+1} - s_i| \to 0$ *as* $n_s \to \infty$*, then* $\hat{\beta}_{n_s} \to \hat{\beta}$*, where*

$$\hat{\beta}_{n_s} \to \begin{bmatrix} \langle \phi_1, \phi_1 \rangle & \dots & \langle \phi_1, \phi_{2+n_k} \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi_{2+n_k}, \phi_1 \rangle & \dots & \langle \phi_{2+n_k}, \phi_{2+n_k} \rangle \end{bmatrix}^{-1} \begin{bmatrix} \langle \phi_1, g \rangle \\ \vdots \\ \langle \phi_{2+n_k}, g \rangle \end{bmatrix} =: \hat{\beta}. \tag{7}$$

Moreover, $\hat{\beta}$ *solves the minimization problem*

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^{2+n_k}}{\arg\min} \int_A \left(g(S_T) - \sum_{j=1}^{2+n_k} \beta_j \phi_j(S_T) \right)^2 \mathrm{d}S_T. \tag{8}$$

Longer proofs are delegated to Appendix A. The minimization property in (8) states that $\hat{\beta}$ minimizes the L^2-distance between $g(\cdot)$ and the basis functions. In this sense, the basis functions optimally replicate $g(\cdot)$ over the entire domain. This property is attractive because A is allowed to be much wider than the range of available strike prices, which is beneficial if we believe the strikes only cover a limited range of the stock price's support. The approach of Carr and Madan (2001) does not have this property. The continuous-state limit is also a convenient tool in some of the proofs. However, for practical computations we will mostly rely on the discrete approximation, as it is faster and numerically more stable.

3 Completeness, convergence, and distribution estimation

This section establishes conditions under which options complete the market and the risk-neutral measure is uniquely determined. We then derive the convergence rate of the projection estimator for risk-neutral expectations. Finally, we show how the same projection framework can be used to estimate the risk-neutral distribution.

3.1 Market completeness

Market completeness implies that every contingent claim can be hedged and, equivalently, that the risk-neutral measure is unique. As is well known, options complete the market for a single underlying security. For example, the CM portfolio in (3) converges to the true risk-neutral moment under certain assumptions on the strike prices. We now establish the analogous result for projection. Specif-

ically, if there is a portfolio of options, the risk-free asset, and the underlying stock that perfectly replicates the payoff $g(S_T)$, then projection will find it, as the following proposition shows.

Proposition 2. *Let $A \subset \mathbb{R}_+$ be compact and let $C(A)$ denote the space of continuous functions on A equipped with the sup norm $\|g\| = \sup_{x \in A} |g(x)|$. If the strikes $\{K_j\}_{j=1}^{n_k}$ satisfy*

$$\min_{j=1,\ldots,n_k} |x - K_j| \to 0 \quad \text{for every } x \in A \qquad \text{as } n_k \to \infty,$$

then $\operatorname{span}(\mathcal{F}_{2+n_k})$ *is dense in $C(A)$. Equivalently, for every $g \in C(A)$ there exists $f_{n_k} \in \operatorname{span}(\mathcal{F}_{2+n_k})$ such that $\|g - f_{n_k}\|_\infty \to 0$.*

Intuitively, the condition above means that strikes become dense in A, which is necessary to replicate g well in the tails. Proposition 2 is a restatement of the classical fact that piecewise linear splines are dense in $C(A)$ (see, e.g., Lebesgue (1898)). It is also more general than the CM approximation, which requires additional smoothness (e.g., g twice differentiable a.e.).

The replication property in Proposition 2 connects to market completeness, which means that the risk-neutral measure is unique (Back, 2017). When the prices of options are given and each contingent claim can be replicated, the risk-neutral measure is indeed uniquely pinned down.

Corollary 1 (Market completeness)**.** *Let A and the strikes be as in Proposition 2, and suppose absence of arbitrage. If two risk-neutral measures agree on the prices of all traded payoffs in* $\operatorname{span}(\mathcal{F}_{2+n_k})$ *for all n_k, then they coincide on $C(A)$ in the limit, and therefore induce the same risk-neutral distribution on A.*

This result is closely related to the Breeden and Litzenberger (1978) formula from Example 3. While that formula is theoretically elegant, its practical implementation can be challenging because recovering densities requires numerical differentiation of option prices, which is often unstable. For this reason, researchers and practitioners commonly use the CM approximation to compute risk-neutral expectations. However, the CM approximation is not designed for discontinuous payoffs such as indicator functions and therefore does not directly deliver estimates of the full risk-neutral distribution. In finite samples, this can lead to substantial differences between the risk-neutral expectation implied by Breeden and Litzenberger (1978) and that implied by the CM approximation, which is undesirable. As shown in Proposition 7 below, the projection method provides a unified approach that closes this gap.

3.2 Convergence rate

In this section, we establish the rate at which the estimated risk-neutral expectation converges as a function of the number of strikes. From approximation theory, we expect the convergence rate to depend on the smoothness of the underlying function (see, e.g., Canuto et al. (2006, Chapter 5) or Trefethen (2018, Chapter 10)). To facilitate the comparison with the CM formula, we assume that the underlying function is twice continuously differentiable. The following proposition derives the convergence rate of the projection approach under this assumption.

Proposition 3. *Suppose* $g \in C^2[a_{\min}, a_{\max}]$ *and that the risk-neutral density is square-integrable on* A*:* $\int_{a_{\min}}^{a_{\max}} f^Q_{t\to T}(x)^2 \,\mathrm{d}x < \infty$*. Let* $\Delta = \max_j K_{j+1} - K_j$*, where the strikes are ordered* $a_{\min} < K_1 < K_2 < \ldots, K_{n_k} < a_{\max}$*, and assume that* $\Delta = O(1/n_k)$*,* $K_1 - a_{\min} = O(1/n_k^{4/5})$*, and* $a_{\max} - K_{n_k} = O(1/n_k^{4/5})$*. Then as* $n_k \to \infty$

$$\mathbf{E}_t^Q\left[g(S_T)\mathbb{1}\left(S_T \in A\right)\right] = \mathbf{E}_t^Q\left[\hat{g}(S_T)\mathbb{1}\left(S_T \in A\right)\right] + O\left(\frac{1}{n_k^2}\right),$$

where $\hat{g}$ *is the function estimated by* (5)*.*

Proposition 3 can be viewed as a quantitative version of the statement that options complete the market (Corollary 1). For the CM formula, the integral representation can be approximated using the composite trapezoidal rule, which is the method employed by the CBOE to calculate the VIX. Under the same assumptions, the CM approximation with the trapezoidal rule attains the same convergence rate.

Proposition 4. *Let everything be as in Proposition 3, and denote the CM replicating portfolio by*

$$\begin{aligned}\hat{g}_{\mathrm{CM}}(S_T) = {} & g(F_{t\to T}) + g'(F_{t\to T})(S_T - F_{t\to T}) \\ & + \sum_{j:K_j \le F_{t\to T}} \Delta K_j \, g''(K_j)\left(K_j - S_T\right)^+ + \sum_{j:K_j > F_{t\to T}} \Delta K_j \, g''(K_j)\left(S_T - K_j\right)^+.\end{aligned}$$

where

$$\Delta K_j = \begin{cases} \frac{K_{j+1} - K_{j-1}}{2}, & j = 2, \ldots, n_k - 1 \\ K_2 - K_1, & j = 1 \\ K_{n_k} - K_{n_k - 1}, & j = n_k. \end{cases}$$

Then, as $n_k \to \infty$

$$\mathbf{E}_t^Q \left[g(S_T) \mathbb{1} \left(S_T \in A \right) \right] = \mathbf{E}_t^Q \left[\hat{g}_{\text{CM}}(S_T) \mathbb{1} \left(S_T \in A \right) \right] + O\left(\frac{1}{n_k^2} \right).$$

Because projection and the CM approximation attain the same convergence rate, it seems plausible that the coefficients are similar when there are lots of options in the market. In fact, under certain assumptions one can show that asymptotically the projection approach and the CM approximation attach the same weights to each option in the portfolio.

Proposition 5. *Let* $A = [a_{\min}, a_{\max}]$ *and let* $a_{\min} < K_1 < \cdots < K_{n_k} < a_{\max}$ *be uniformly spaced with*

$$h \coloneqq K_i - K_{i-1} \quad (i = 1, \ldots, n_k), \qquad K_0 \coloneqq a_{\min}, \quad K_{n_k+1} \coloneqq a_{\max}.$$

Assume $g \in C^4(A)$ *and let* $\hat{g}$ *be the* $L^2(A)$*-projection of* g *onto* $\operatorname{span}(\mathcal{F}_{2+n_k})$*,*

$$\hat{g}(x) = \hat{\beta}_1 + \hat{\beta}_2 x + \sum_{i=1}^{n_k} \hat{\gamma}_i (x - K_i)_+.$$

Then for interior indices $i = 2, \ldots, n_k - 1$*,*

$$\hat{\gamma}_i = \underbrace{h\, g''(K_i)}_{\textit{CM weight}} + O(h^3) \qquad \textit{as } h \to 0,$$

where the $O(h^3)$ *term is uniform in* $i = 2, \ldots, n_k - 1$*. Moreover, at the boundary one has*

$$\hat{\gamma}_1 = \underbrace{h\, g''(K_1)}_{\textit{CM weight}} + O(h^2), \qquad \hat{\gamma}_{n_k} = \underbrace{h\, g''(K_{n_k})}_{\textit{CM weight}} + O(h^2), \qquad \textit{as } h \to 0.$$

This result may appear surprising at first because the projection method seems global, in the sense that each coefficient estimate depends on the full set of strikes. However, results from the series regression literature suggest that it depends on the number of basis functions: when the number of strikes is small the estimator is effectively global, whereas as the strike grid becomes dense the projection behaves increasingly like a local method (see, e.g., Hansen (2022, Section 20.7)).

Why, then, prefer the projection method? First, the results above are asymptotic and may not accurately describe the finite-sample behavior that is relevant in practice. Second, Proposition 5 relies on idealized assumptions, such as

a uniformly spaced strike grid and a mesh that becomes dense all the way to the endpoints of A. When either assumption fails, as is typical in option data, the asymptotic approximation in Proposition 5 need not hold, and the implied portfolio weights can differ substantially from those obtained by a CM type approximation.

It is therefore desirable to derive a finite-sample bound that does not rely on these assumptions. The next proposition provides an exact finite-sample bound on the projection error.

Proposition 6. *Let $g \in C(0,\infty)$, $A = [a_{\min}, a_{\max}]$ with $a_{\min} < K_1 < \dots K_{n_k} < a_{\max}$, and let $\hat{g}$ be the $L^2(A)$-projection of g onto $\mathcal{F}_{2+n_k}$. Suppose that $\mathbf{E}_t^Q |g(S_T)| < \infty$, $\mathbf{E}_t^Q S_T < \infty$, and that*

$$\mathbf{E}_t^Q \left[(|g(S_T)| + |\hat{g}(S_T)|) \mathbb{1}\left(S_T \notin A\right) \right] \leq \varepsilon.$$

Then

$$\left| \mathbf{E}_t^Q g(S_T) - \mathbf{E}_t^Q \hat{g}(S_T) \right| \leq \varepsilon + 4 \operatorname{dist}(g, \mathcal{F}_{2+n_k}), \tag{9}$$

where

$$\operatorname{dist}(g, \mathcal{F}_{2+n_k}) = \inf \left\{ \max_{x \in A} |g - f| : f \in \mathcal{F}_{2+n_k} \right\}.$$

If the support of S_T is contained in A, then (9) holds with $\varepsilon = 0$.

The tail assumption effectively says that A, which can be chosen by the researcher, covers most of the support of S_T and that the contribution of the risk-neutral moment outside A is small. The main message of Proposition 6 is that the estimation error is controlled by how well g is spanned by the *given* option basis functions (together with the constant and linear payoffs). For example, suppose the only option payoff observed is a call option with strike K, and let $g(S_T) = (K - S_T)^+$. Using the identity

$$(K - S_T)^+ = (S_T - K)^+ + K - S_T,$$

the put payoff lies in the span of $\mathcal{F}_{2+n_k}$. Hence, $\operatorname{dist}(g, \mathcal{F}_{2+n_k}) = 0$, and the estimation error is zero. This exactly recovers put–call parity and illustrates that the bound is genuinely finite-sample: it only uses the actually traded payoff(s), without any asymptotic market completeness assumption.

A clean substitute for the CM approximation appears unlikely, though we do not have a formal proof. Under the stated assumptions, no analogue of Proposition 6 can hold for CM, since the formula relies on second derivatives of g and

thus defines an unbounded operator with respect to the sup-norm.[5] In sum, the projection error is well controlled in finite samples and leads to a notion of finite sample near-optimality, while a clean substitute for CM is not available. The simulation results in Section 5 also confirm this.

3.3 Estimation of the risk-neutral CDF and PDF

The convergence rate and error bound derived above are valid when the function is twice differentiable, or merely continuous. However, the projection method need not be restricted to such functions. A primary motivation to broaden the class of estimable functions comes from estimating the risk-neutral distribution, which requires approximating indicator functions. Since we are ultimately interested in the risk-neutral expectation of a function, the averaging inherent in the expectation operator suggests that the projection approach can still estimate the risk-neutral distribution reliably, even when g is not smooth.

More precisely, consider $g(S_T) = \mathbb{1}(S_T \leq x)$, which is used to compute the risk-neutral CDF: $F^Q_{t\to T}(x) = \mathbf{E}^Q_t \mathbb{1}(S_T \leq x)$. In this case, the projection estimates obtained in (7) will also depend on x, because

$$\langle \phi_j, \mathbb{1}(\cdot \leq x) \rangle = \int_A \phi_j(S_T) \mathbb{1}(S_T \leq x) \, \mathrm{d}S_T = \int_{a_{\min}}^{x} \phi_j(S_T) \, \mathrm{d}S_T.$$

We will let $\hat{\beta}(x)$ denote the coefficient estimate corresponding to the function $\langle \phi_j, \mathbb{1}(\cdot \leq x) \rangle$. The risk-neutral CDF is then simply estimated by

$$\hat{F}^Q_{t\to T}(x) = \hat{\beta}_1(x) + \hat{\beta}_2(x) F_{t\to T} + R_{f,t\to T} \left(\sum_{j=1}^{n^P_k} \hat{\beta}^P_j(x) P_{t\to T}(K_j) + \sum_{j=1}^{n^C_k} \hat{\beta}^C_j(x) C_{t\to T}(K_j) \right). \tag{10}$$

The following proposition shows that $\hat{F}^Q_{t\to T}(x)$ obtained in this way satisfies many of the natural CDF requirements.

Proposition 7 (Risk-neutral distribution)**.** *Let Assumption 1 hold. Then:*

(i) The estimated CDF satisfies the natural boundary limits

$$\lim_{x \to a^+_{\min}} \hat{F}^Q_{t\to T}(x) = 0, \quad \text{and} \quad \lim_{x \to a^-_{\max}} \hat{F}^Q_{t\to T}(x) = 1.$$

[5] Even if the assumptions were strengthened to, say, $g \in C^2(0,\infty)$, a finite-sample bound in the spirit of Proposition 6 still appears unlikely, because the operator that sends g to $\sum_i g''(K_i)(K_{i+1} - K_i)$ is unbounded as a linear functional on $(C(A), \|\cdot\|_\infty)$; one can construct sequences of bump functions with $\|g\|_\infty$ bounded but $\sum_i g''(K_i)(K_{i+1} - K_i)$ diverging.

(ii) $\hat{F}^Q_{t\to T}(x)$ *is continuously differentiable on the interior of* A*, with density estimate* $\hat{f}^Q_{t\to T} = (\hat{F}^Q_{t\to T})'$*; moreover,* $\hat{f}^Q_{t\to T}$ *is piecewise linear.*

(iii) (Moment consistency) The estimated value of a nonlinear contract in (6) *equals the moment implied by the estimated distribution:*

$$\mathbf{E}^Q_t\left[\hat{g}(S_T)\mathbb{1}\left(S_T \in A\right)\right] = \int_A g(x)\,\mathrm{d}\hat{F}^Q_{t\to T}(x).$$

Property (iii) is the most important: for any finite set of strikes, the estimated risk-neutral distribution produces the same moment as obtained by directly approximating g.[6] This moment-consistency is typically not guaranteed by existing risk-neutral density estimators. In particular, the value of a nonlinear contract computed from a density estimate will almost never coincide with the estimate given by the CM method. The CM approach is often used in applications where the full risk-neutral distribution is not of primary interest, as it is empirically more robust (see, e.g., Martin (2017)). This discrepancy between density-implied moments and CM-implied moments calls into question the accuracy of the density estimate. By construction, the projection approach avoids this issue and yields a density that is consistent with any moment obtained by direct projection. Furthermore, our density estimator requires only mild assumptions on the underlying distribution: it is sufficient for the first moment of the stock price to exist.

Despite these desirable properties, the projection-based CDF estimate need not be monotone. In simulations, violations of monotonicity occur mainly in the extreme tails, where sparse strike coverage makes the distribution hard to estimate. A remedy is to apply the rearrangement approach of Chernozhukov et al. (2013), which amounts to sorting the estimated CDF values on the grid to enforce monotonicity. In fact, Chernozhukov et al. (2009) show that, unless the original estimate is already monotone, the rearranged CDF has better finite-sample properties.

4 Completeness in multiple asset markets and joint dependence

It is of great interest to generalize the projection approach to higher dimensions. For example, the risk-premium of an individual return can often be related to its

[6] That is, using the estimate in (6).

risk-neutral covariance with the market return (see Example 4). The key challenge is that the claim paying $S_{1,T}S_{2,T}$ is not traded; hence $\mathbf{E}_t^Q(S_{1,T}S_{2,T})$ needs to be identified from tradable options.

A naive extension of the univariate approach is to consider a projection of $g(S_{1,T}, S_{2,T}) = S_{1,T}S_{2,T}$ onto

$$\begin{aligned}\hat{g}\left(S_{1,T}, S_{2,T}\right) = \hat{\beta}_0 + \hat{\beta}_1 S_{1,T} + \sum_{j=1}^{n_k^P} \hat{\beta}_{1,j}^P \left(K_j - S_{1,T}\right)^+ + \sum_{j=1}^{n_k^C} \hat{\beta}_{1,j}^C \left(S_{1,T} - K_j\right)^+ \\ + \hat{\beta}_2 S_{2,T} + \sum_{j=1}^{n_k^P} \hat{\beta}_{2,j}^P \left(K_j - S_{2,T}\right)^+ + \sum_{j=1}^{n_k^C} \hat{\beta}_{2,j}^C \left(S_{2,T} - K_j\right)^+ . \quad (11)\end{aligned}$$

Notice that the strike prices can be different across assets and basis functions, but we suppress this dependence for notational clarity. The risk-neutral expectation of each of the basis functions is known, and thus provides a way to estimate $\mathbf{E}_t^Q S_{1,T}S_{2,T}$. However, the next proposition shows that this separable specification cannot capture dependence: the implied correlation is always zero.

Proposition 8 (Zero correlation)**.** *Assume that the support of $S_{1,T}$ and $S_{2,T}$ be defined on compact intervals with midpoints equal to $\mathbf{E}_t^Q S_{1,T} = F_{1,t\to T}$ and $\mathbf{E}_t^Q S_{2,T} = F_{2,t\to T}$ respectively. Let the projection of $S_{1,T}S_{2,T}$ be defined by $\hat{g}$ in (11), then*

$$\mathbf{E}_t^Q \left[\hat{g}(S_{1,T}, S_{2,T})\right] = \left(\mathbf{E}_t^Q S_{1,T}\right)\left(\mathbf{E}_t^Q S_{2,T}\right).$$

Intuitively, options on the individual stocks are sufficient to identify the marginal distributions, but not the joint distribution. To estimate a nonzero correlation, the basis must include nonlinear terms that depend on both assets or incorporate multi-asset instruments such as basket options.

To incorporate additional information that depends on the joint distribution of returns, options on the S&P500 can be used. As Kelly et al. (2016) noted, there are eleven sector ETFs that also have options available, and whose weighted returns sum to the S&P500 return:

$$\sum_{i=1}^{11} w_{i,t} R_{i,t\to T} = R_{t\to T},$$

where $w_{i,t}$ and $R_{i,t\to T}$ denote the weight and realized return on sector ETF i, and $R_{t\to T}$ represents the return on the market portfolio. Thus, options on the S&P500 reveal information about the joint distribution of returns. In combination with options on the individual sectors, they allow more precise inference about

correlations. Nevertheless, the information conveyed by options on the market index and on the sectors is limited: with three or more sectors, correlations cannot be identified from these derivatives alone. We establish this non-identification result below.

4.1 Identifying joint dependence from options on multiple portfolios

We are looking for an extension of Proposition 2 that is valid in higher dimensions. In particular, we would like to understand when option payoffs are rich enough to approximate multivariate contingent claims, and how the set of available portfolios governs what can be learned about joint dependence. Suppose, as in practice, that there are d sectors (or stocks) that span the market return:[7]

$$\sum_{i=1}^{d} w_{i,t} R_{i,t\to T} = R_{t\to T}.$$

Assume now that for each sector, as well as for the market return, the assumptions of Proposition 2 hold, so that any continuous function of the sector return (or market return), can be uniformly approximated by options. By combining the options on each of the sectors and on the market return in a portfolio, we thus conclude that the set of option payoff functions span the space

$$\mathcal{M}(\Omega) := \operatorname{span}\Big\{ x \mapsto f(a'x) \;:\; a \in \Omega,\; f \in C(\mathbb{R}) \Big\},$$

where $\Omega \subset \mathbb{R}^d$ is the set of available portfolio directions. In our baseline setting,

$$\Omega = \{e_1, \ldots, e_d,\; w_t\}, \qquad w_t = (w_{1,t}, \ldots, w_{d,t})', \tag{12}$$

where e_i corresponds to the ith basis vector in $\mathbb{R}^d$ (i.e. it gives full weight to sector i). Functions of the form $f(a'x)$ are known as *ridge functions* in the approximation theory literature (Pinkus, 2015). Thus, the question of multivariate spanning by simple options can be phrased as a question about when ridge functions with directions in Ω are dense (in the uniform topology on compact sets). The following result by Vostrecov and Kreines (1961) provides necessary and sufficient conditions (see also Lin and Pinkus (1993)):

[7] When dealing with sectors, there are thus $d = 11$ sectors spanning the S&P500 return. When dealing with individual returns, there are $d = 500$ returns spanning the S&P500 return.

Theorem 2. *$\mathcal{M}(\Omega)$ is dense in $C(\mathbb{R}^d)$ in the topology of uniform convergence on compacta if and only if no non-trivial homogeneous polynomial vanishes on Ω.*[8]

In the special case $d = 2$, for the set of option payoffs to be dense Theorem 2 requires Ω to contain an infinite number of pairwise linearly independent vectors. This result is related to Ross (1976) and Martin (2018, Result 2), but is stronger, because the condition is necessary and sufficient. Furthermore, Theorem 2 applies to any $d \geq 1$, not just to the case $d = 2$. In applications, we therefore cannot hope to approximate the price of *every* multivariate contingent claim arbitrarily well, since we only observe the finite set of twelve direction vectors in (12) associated with the $d = 11$ sector portfolios. Nevertheless, it is still possible to approximate the payoff of an arbitrary claim using projection on the sector and market option payoff functions. Furthermore, Theorem 2 suggests that better approximations can be obtained if we also consider options on a portfolio of sectors, where the weights are different from the market portfolio. Recently, options were introduced on an equally weighted sector portfolio (called "EQL"). This additional variation can allow us to obtain better estimates of the sector correlations.

Remark 4. It was noted by Ross (1976) and Nachman (1988) that payoffs formed by products of call options on multiple assets are sufficient to complete the market. In particular, derivatives with payoffs of the form

$$(S_{1,T} - K_1)^+ (S_{2,T} - K_2)^+ ,$$

for arbitrary strikes K_1 and K_2, together with standard call options on $S_{1,T}$ and $S_{2,T}$, span all contingent claims in dimension two.[9] In our framework, this market completeness result follows directly from Proposition 2 and the fact that tensor products of spline basis functions are dense in the space of continuous functions on compact sets. This argument immediately generalizes to arbitrary dimension d by considering tensor products of spline basis functions in $\mathbb{R}^d$. For example, in $\mathbb{R}^3$, market completeness would require observing prices of derivatives with payoffs of the form

$$(S_{i,T} - K_i)^+ , \; (S_{i,T} - K_i)^+ (S_{j,T} - K_j)^+ , \; (S_{i,T} - K_i)^+ (S_{j,T} - K_j)^+ (S_{p,T} - K_p)^+ ,$$

for all choices of indices i, j, p and strikes K_i, K_j, K_p. In practice, however, derivatives involving products of options on more than one underlying are not traded

[8] A polynomial in several variables is homogeneous if all monomials have the same total degree.
[9] Bakshi and Madan (2000) refer to such securities as correlation options.

on major exchanges and are typically only available over the counter. For this reason, we focus on market completeness achieved using simple options, which are widely traded and liquid.

4.2 Identification of risk-neutral covariances and correlations

Theorem 2 suggests that it is impossible to identify the price of an *arbitrary* claim using options, unless we observe an infinite number of different portfolio options.

However, in specific cases, such as the covariance in two dimensions, it is possible to obtain positive results. Furthermore, in higher dimensions, one can still approximate the covariance well even if it is not strictly identified. Focusing on two dimensions first, and letting $R_{t\to T} = w_{1,t}R_{1,t\to T} + w_{2,t}R_{2,t\to T}$, the following identity obtains:

$$R_{1,t\to T}R_{t\to T} = \frac{1}{2w_{1,t}}R_{t\to T}^2 + \frac{w_{1,t}}{2}R_{1,t\to T}^2 - \frac{w_{2,t}^2}{2w_{1,t}}R_{2,t\to T}^2.$$

The prices of each of the payoffs on the right-hand side can be inferred from options on the market index, sector 1, and sector 2, respectively. Hence, in this case, the covariance between any of the returns can be identified from option prices.[10]

Generally, the question of identifying the price of a payoff thus depends on whether there is an exact algebraic identity linking the payoff function and a linear combination of ridge functions. It is useful to have a simple algebraic condition that determines whether such a separable identity holds. Diaconis and Shahshahani (1984) derived the following necessary and sufficient condition for a function $g(x, y)$ to admit a representation of the form

$$g(x, y) = \sum_{i=1}^{r} g_i(a_i x + b_i y)$$

In this case, the following differential identity is necessary and sufficient:

$$\prod_{i=1}^{r}\left(b_i\frac{\partial}{\partial x} - a_i\frac{\partial}{\partial y}\right)[g] = 0.$$

When $d \geq 3$, the situation becomes more involved. Necessary and suffi-

[10] This is unsurprising, since $\mathbf{Var}_t^Q R_{t\to T} = w_{1,t}^2\mathbf{Var}_t^Q R_{1,t\to T} + w_{2,t}^2\mathbf{Var}_t^Q R_{2,t\to T} + 2w_{1,t}w_{2,t}\,\mathbf{Cov}_t^Q\left(R_{1,t\to T}, R_{2,t\to T}\right)$ and because each individual variance is identified from option prices, the covariance must also be identifiable.

cient conditions were derived by Lin and Pinkus (1993), although they are not straightforward to verify in practice. For completeness, we state their result in Appendix A.10 and provide a more elementary argument showing why correlations in dimensions $d \geq 3$ cannot be identified solely from options on the individual sectors and the market portfolio. The following Proposition summarizes this result.

Proposition 9 (Non-replication of covariance payoff)**.** *Let $d \geq 3$. Fix $i \in \{1, \ldots, d\}$ and a weight vector $w \in \mathbb{R}^d$ such that there exist two distinct indices $j, k \neq i$ with $w_j \neq 0$ and $w_k \neq 0$. Consider the function class*

$$\mathcal{F} = \left\{ \sum_{z=1}^{d} g_z(x_z) + h(w \cdot x) : g_z, h \in C(\mathbb{R}) \right\}.$$

Then the polynomial $g(x) = x_i(w \cdot x)$ is not in $\mathcal{F}$. Consequently, no static portfolio formed from European options on each single return x_z and on the market return $w \cdot x$ can replicate the payoff $x_i(w \cdot x)$.

Remark 5. Despite this non-identification result, Appendix B proposes a correlation estimator that can nonetheless *approximate* the risk-neutral correlation accurately. There we also show how the projection approach generalizes the CBOE equicorrelation estimator.

4.3 Completeness in FX markets

We now extend the above results to foreign-exchange options. Let $S_{1,T}$ denote the EUR/USD exchange rate, $S_{2,T}$ the GBP/USD rate, and $S_{3,T}$ the EUR/GBP rate at maturity T. By triangular no-arbitrage, $S_{3,T} = S_{1,T}/S_{2,T}$. Hence, options on EUR/GBP reveal joint information not captured by options on EUR/USD and GBP/USD, which only reveal the marginal distribution. Incorporating this additional source of variation is thus expected to yield a better estimate of the covariance and correlation. Throughout we use the convention that S_1 and S_2 are quoted in USD, while S_3 is in GBP units.[11]

With $R_{f,t\to T}$ and $R^{£}_{f,t\to T}$ denoting the US and UK gross risk-free rates, the European call prices are

$$C^{\$}_{i,t\to T}(K) = \frac{1}{R_{f,t\to T}} \mathbf{E}_t^{Q^{\$}} \left(S_{i,T} - K\right)^+, \quad i = 1, 2,$$
$$C^{£}_{t\to T}(K) = \frac{1}{R^{£}_{f,t\to T}} \mathbf{E}_t^{Q^{£}} \left(S_{3,T} - K\right)^+,$$

[11] This convention is the same as for the Bloomberg options data that we use in Section 6.

where $Q^{\$}$ and $Q^{£}$ are the risk-neutral measures using the US and UK money-market accounts as numéraires, respectively. This distinction is needed because EUR/GBP options are GBP-quoted.

Using the change of numéraire result (Shreve, 2004, Chapter 9), it follows that the Radon-Nikodym derivative between the two risk-neutral measures is given by

$$\left.\frac{\mathrm{d}Q^{\$}}{\mathrm{d}Q^{£}}\right|_{\mathcal{F}_T} \Big/ \left.\frac{\mathrm{d}Q^{\$}}{\mathrm{d}Q^{£}}\right|_{\mathcal{F}_t} = \frac{R_{f,t\to T}}{R^{£}_{f,t\to T}}\frac{S_{2,t}}{S_{2,T}},$$

where $\mathcal{F}_t$ denotes the information set up to time t. Using this result, we obtain the following expression for a judicious choice of payoff function under $Q^{\$}$

$$\begin{aligned}\mathbf{E}_t^{Q^{\$}}\left[S_{2,T}\left(\frac{S_{1,T}}{S_{2,T}}-K\right)^{+}\right] &= \frac{R_{f,t\to T}}{R^{£}_{f,t\to T}}S_{2,t}\mathbf{E}_t^{Q^{£}}\left[(S_{3,T}-K)^{+}\right]\\ &= R_{f,t\to T}S_{2,t}C^{£}_{t\to T}(K).\end{aligned}$$

Hence, the reason we consider this specific type of payoff is that the right-hand side involves quantities that are all observed in the market. Notice how the change of numéraire ensures that the quantity on the right is in dollar units, because $S_{2,t}$ converts GBP prices to USD. A key advantage of projection is that it can incorporate the state-dependent change of numéraire kernel when combining options quoted in different currencies, yielding a theoretically consistent USD-denominated replicating portfolio. By contrast, much of the existing FX literature effectively ignores this state dependence (or treats the conversion kernel as approximately constant) when extracting dependence measures from option prices (e.g., Mueller et al. (2017)). Further, it is possible to obtain the expected value of EUR/USD and GBP/USD under the USD risk-neutral measure because

$$\mathbf{E}_t^{Q^{\$}}S_{1,T} = \frac{R_{f,t\to T}}{R^{€}_{f,t\to T}}S_{1,t} = F_{1,t\to T}, \quad \mathbf{E}_t^{Q^{\$}}S_{2,T} = \frac{R_{f,t\to T}}{R^{£}_{f,t\to T}}S_{2,t} = F_{2,t\to T},$$

where $F_{i,t\to T}$ denotes the T-maturity forward FX rate for pair $i = 1, 2$.

The foregoing discussion suggests a way to obtain the covariance and correlation between EUR/USD and GBP/USD. Namely, project the function

$$\left(S_{1,T}-F_{1,t\to T}\right)\left(S_{2,T}-F_{2,t\to T}\right)$$

on basis functions of the form

$$1,\ S_{1,T}, (S_{1,T}-K)^+, S_{2,T},\ (S_{2,T}-K)^+,\ S_{2,T}\left(\frac{S_{1,T}}{S_{2,T}}-K\right)^+.$$

Upon taking risk-neutral expectations using the US money market as numéraire, all expectations of the basis functions reduce to market observables: constant, forward levels, USD call prices multiplied by a known discount factor, and EUR/GBP call prices multiplied by known discount and FX conversion factors. In particular,

$$\begin{aligned}\mathbf{Cov}_t^{Q^\$}(S_{1,T}, S_{2,T}) &= \mathbf{E}_t^{Q^\$}\left(S_{1,T}-F_{1,t\to T}\right)\left(S_{2,T}-F_{2,t\to T}\right)\\ &\approx \hat\beta_0 + \hat\beta_{1,1}F_{1,t\to T} + R_{f,t\to T}\sum_{j=1}^{n_k}\hat\beta_{1,j+1}C^\$_{1,t\to T}(K_j)\\ &+ \hat\beta_{2,1}F_{2,t\to T} + R_{f,t\to T}\sum_{j=1}^{n_k}\hat\beta_{2,j+1}C^\$_{2,t\to T}(K_j)\\ &+ F_{2,t\to T}R^{\pounds}_{f,t\to T}\sum_{j=1}^{n_k}\hat\beta_{3,j}C^{\pounds}_{t\to T}(K_j).\end{aligned}$$

The number of options and the strike grids generally differ across currencies; we omit this from the notation to avoid clutter.

If options on all three bilateral rates are available and, for each rate, the assumptions of Proposition 2 hold, then static portfolios in these options can uniformly approximate any payoff of the form

$$g(S_{1,T}, S_{2,T}) \;=\; g_1(S_{1,T}) \;+\; g_2(S_{2,T}) \;+\; S_{2,T}\cdot g_3\left(\frac{S_{1,T}}{S_{2,T}}\right), \tag{13}$$

with g_i continuous. This function class, however, is not universal on $C(A)$ for a compact $A \subset \mathbb{R}^2_{++}$ with nonempty interior. In particular, the function $g(x,y)=xy$ cannot be represented by the display above. Thus, the covariance of exchange rates is not strictly identified from vanillas on the three bilateral rates alone. Nevertheless, we find in simulations that projecting $S_{1,T}S_{2,T}$ onto the class (13) yields highly accurate approximations of the covariance and correlation. In our empirical application, we exploit this observation to estimate conditional risk-neutral correlations between exchange rates.

5 Simulation

5.1 Univariate projection

To illustrate the benefits of the projection based approach, we consider the problem of approximating the value of the SVIX and VIX discussed in Examples 1–2. The Monte-Carlo experiment randomly draws strike prices from a uniform grid with cardinality $\{10, 20, \ldots, 130\}$. We also consider the case where the strike grid is equally spaced.[12] This allows us to study the approximation error as a function of the number of strikes available in the market. In addition, we also consider a design where the number of strikes is fixed, but the range of the strike prices is increasing to cover a bigger part of the distribution's support.

Based on the strikes, we obtain the corresponding call and put option prices from either the Black and Scholes (1973) model or the stochastic volatility and jump (SVCJ) model of Eraker et al. (2003). The latter model incorporates jumps in both the return and volatility dynamics which makes estimation more challenging relative to Black-Scholes. More details on the simulation and calibration of these models are given in Appendix E. The accuracy of the approximation for each number of strikes is measured by the relative error,

$$\text{Relative error} = \frac{\left|\widehat{\text{SVIX}} - \text{SVIX}\right|}{\text{SVIX}},$$

where $\widehat{\text{SVIX}}$ is the SVIX estimate obtained by either CM or the projection method. The relative error for VIX is defined analogously.

Figure 2 illustrates the results. Panels 2a–2d show convergence as the number of strikes increases, while the strike range remains fixed at 90% of the support. When the strike grid is equally spaced, the relative errors of both methods are roughly half as large as when the strikes are drawn uniformly at random, but both designs convey the same message. The convergence of the CM method is gradual and levels off at a relative error of about 10%. By contrast, the projection approach stabilizes already around 20 strikes, at which point its relative error is roughly an order of magnitude smaller. At 130 strikes, the relative error remains close to 2% in all cases. Moreover, for nearly all strike counts, the projection estimate is pointwise closer to SVIX/VIX than the CM estimate. Because both methods underestimate

[12] In our implementation, A covers 99.8% of the distribution's support, while observed strikes extend only into the 5% tail. Thus Proposition 5 does not apply, and CM and projection weights can differ substantially.

SVIX/VIX due to the limited strike range, the projection estimate—being closer to the truth—is almost always larger than the corresponding CM estimate. In the empirical application in Appendix C, we find the same behavior in actual data.

The strike range appears more important for the convergence rate of the projection approach, as shown in Panels 2e and 2f. In this case, convergence is much faster as the strike range increases while the number of strikes is held fixed at $n_k = 30$. This result can be understood via the proof of Proposition 3, which shows that the error arising from the tails converges to zero faster than the error induced by strike spacing. When the strike range covers almost the entire support, the relative projection error is close to zero and roughly 63 times smaller than for CM.

By contrast, the CM approach shows little improvement when the strike range increases. As the range widens while the number of strikes remains fixed, the average strike spacing becomes larger, which offsets the benefit of better tail coverage because the accuracy of the integral approximation deteriorates as the spacing increases.[13]

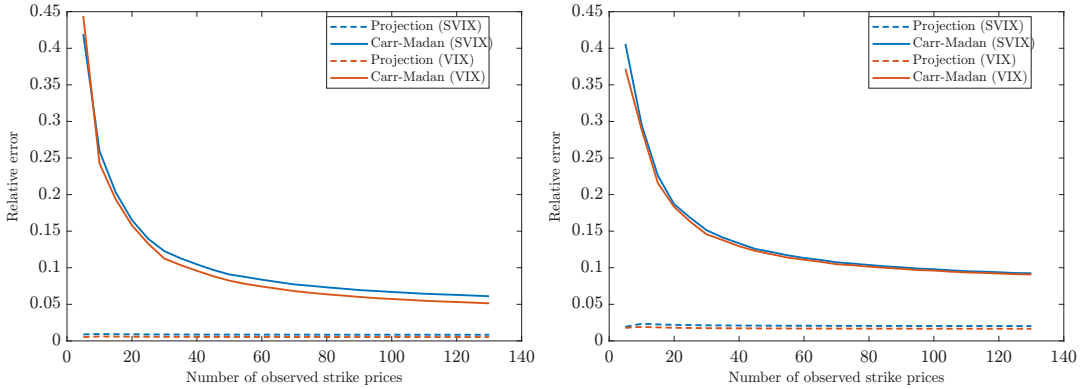
5.2 Multivariate projection for exchange rates

We simulate exchange-rate outcomes under the risk-neutral measure from a bivariate normal distribution:

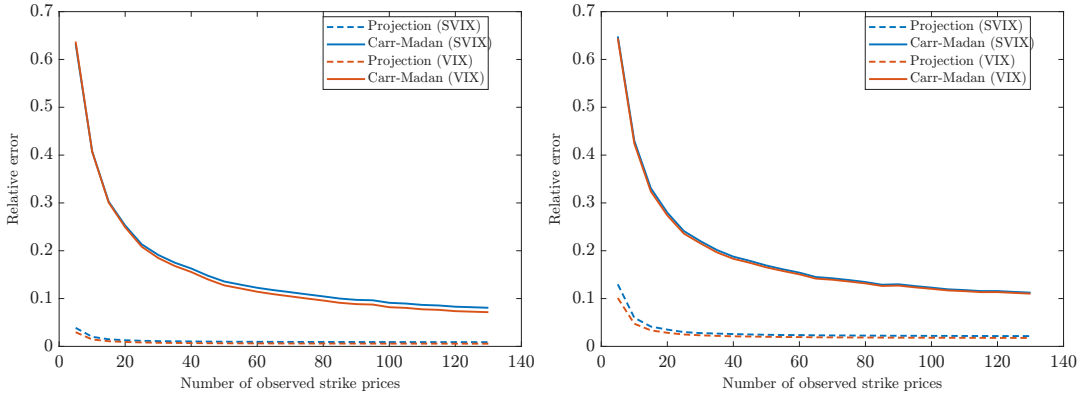
$$\begin{bmatrix} S_{1,T} \\ S_{2,T} \end{bmatrix} \sim \mathsf{N}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0.1^2 & 0.1 \cdot 0.05 \cdot \rho \\ 0.1 \cdot 0.05 \cdot \rho & 0.05^2 \end{bmatrix} \right).$$

In each Monte Carlo iteration, we draw the correlation independently as $\rho \sim \mathsf{Unif}(-1, 1)$. For the option inputs, we take five strikes each on $S_{1,T}$, $S_{2,T}$, and $S_{1,T}/S_{2,T}$. The strikes are evenly spaced between the 5th and 95th percentiles of the respective marginal distributions. This choice mirrors OTC FX practice: quotes out to the 5-delta call and 95-delta put (under forward-delta conventions) roughly correspond to the 5th and 95th percentiles for 1-month tenors. The approximation grid is taken to be equally spaced between the 2nd and 98th percentiles of each variable; for two-dimensional quantities we use the tensor product of the univariate grids.

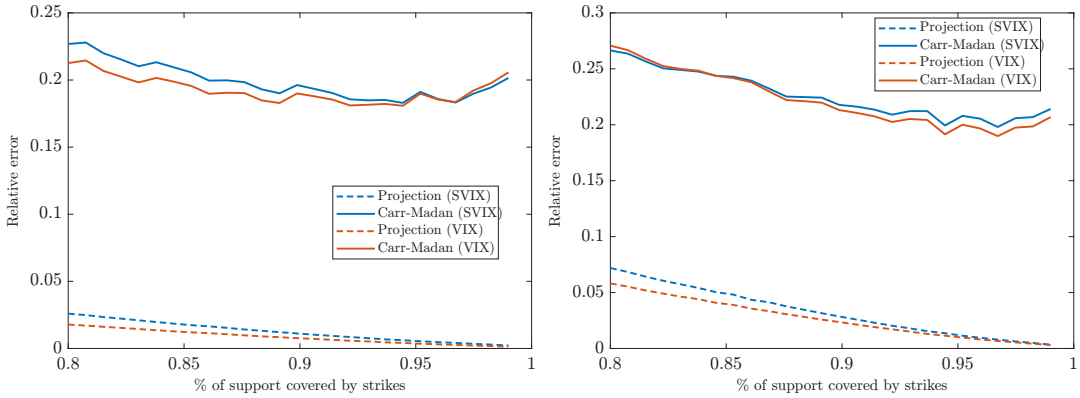
[13] In unreported simulations, we replace the trapezoidal rule in the CM approximation by Simpson's rule. The numerical results are very similar in all cases and the projection method continues to dominate.



(a) Black-Scholes model (equal) (b) SVCJ model (equal)



(c) Black-Scholes model (d) SVCJ model



(e) Black-Scholes model (f) SVCJ model

Figure 2: **MSE of approximation**. The figure shows the convergence rate as a function of the number of strikes (upper and middle panels) and as a function of the strike range (bottom panels). In the top panels, the strike grid is equally spaced, while in the middle panels the strikes are uniformly distributed.

We then project the payoff $(S_{1,T}-1)(S_{2,T}-1)$ onto the span of the payoffs

$$1, \quad S_{1,T}, \quad (S_{1,T}-K_1)^+, \quad S_{2,T}, \quad (S_{2,T}-K_2)^+, \quad S_{2,T}\,(S_{1,T}/S_{2,T}-K_3)^+,$$

with strikes $\{K_1, K_2, K_3\}$ generated as above. To recover the correlation, we also estimate the standard deviations by projecting $(S_{1,T} - 1)^2$ onto the constant function, $S_{1,T}$, and options on $S_{1,T}$ (and analogously for $S_{2,T}$).

In addition, we consider a setting where $S_{2,T}$ is generated as above and then perturbed to $\tilde{S}_{2,T} = S_{2,T} + 0.1S_{1,T}^3$. $\tilde{S}_{2,T}$ is further normalized so that the mean is 1. We estimate the correlation between $S_{1,T}$ and $\tilde{S}_{2,T}$ to introduce nonlinear dependence and verify that our results are not driven by the normality assumption.

The upper panels in Figure 3 report results from 1,000 Monte Carlo simulations. In both panels, the projection approach recovers the true correlation with high accuracy: the scatter points lie nearly on the 45° line. This is encouraging because the correlation is not exactly identifiable within the restricted function class (see Section 4.3). We conclude that projection delivers an excellent approximation to the true correlation in the FX setting, irrespective of the underlying distribution of the data.

In the bottom panels, we use the same generated data to estimate the joint probability that both returns are below a certain threshold, which can be interpreted as a measure of joint tail risk. Specifically, we estimate $\mathbf{P}(S_{1,T} \leq 0.95, S_{2,T} \leq 0.95)$, by projecting the payoff

$$\mathbb{1}\left(S_{1,T} \leq 0.95\right) \mathbb{1}\left(S_{2,T} \leq 0.95\right)$$

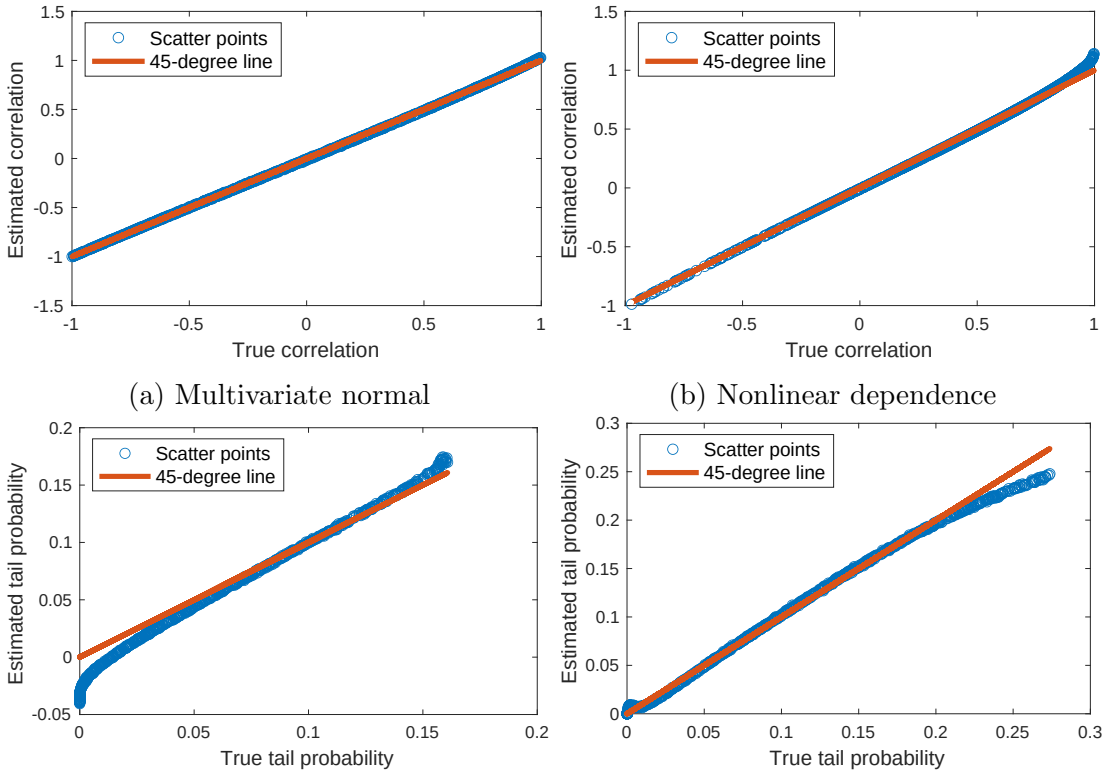
onto the basis functions. The bottom panels of Figure 3 report fitted versus true probabilities. The estimates line up closely with the 45° line—albeit slightly less tightly than for the correlation results—indicating that the projection method recovers joint tail probabilities with high accuracy.

6 Empirical application

This section estimates risk-neutral correlations and tail risk using the method of Section 4.3 in the FX setting.

6.1 Data collection

From Bloomberg we obtain daily end-of-day composite (OTC) quotes for money-market deposit rates at 1 month EUR, USD and GBP. We also retrieve daily spot FX rates and construct 1 month forwards for EUR/USD, GBP/USD and EUR/GBP via covered interest parity.



(a) Multivariate normal (b) Nonlinear dependence

(c) Multivariate normal (d) Nonlinear dependence

Figure 3: **Estimated correlation and joint tail risk in exchange-rate markets.** Each point is one of 1,000 Monte Carlo simulations. Top: true correlation versus its projection-based estimate. Bottom: true joint left-tail probability $\mathbf{P}(S_{1,T} \leq 0.95, S_{2,T} \leq 0.95)$ versus its projection-based estimate.

For FX options, we use Bloomberg's OTC constant-maturity implied volatilities at 1M. Each day we observe the standard smile pillars: the ATM delta-neutral volatility and the 10- and 25-delta risk reversals (RR) and butterflies (BF), quoted under the spot-delta, premium-included convention. When fixed-delta call/put vols are not directly provided, we recover them from ATM, RR and BF via the standard identities. We then map quotes to strikes and compute option prices using the Garman–Kohlhagen model, the reference model with respect to which the implied volatilities are quoted.[14] For each currency pair, this yields only five strikes per day, yet the simulations indicate that projection remains accurate under such sparse strike coverage. Our sample spans July 2008 to April 2023 and contains 3,721 trading days. Finally, returns on each currency are defined relative to the forward price: $R_{i,t\to T} = S_{i,T}/F_{i,t\to T}$. Thus, by construction, $\mathbf{E}_t^Q R_{i,t\to T} = 1$.

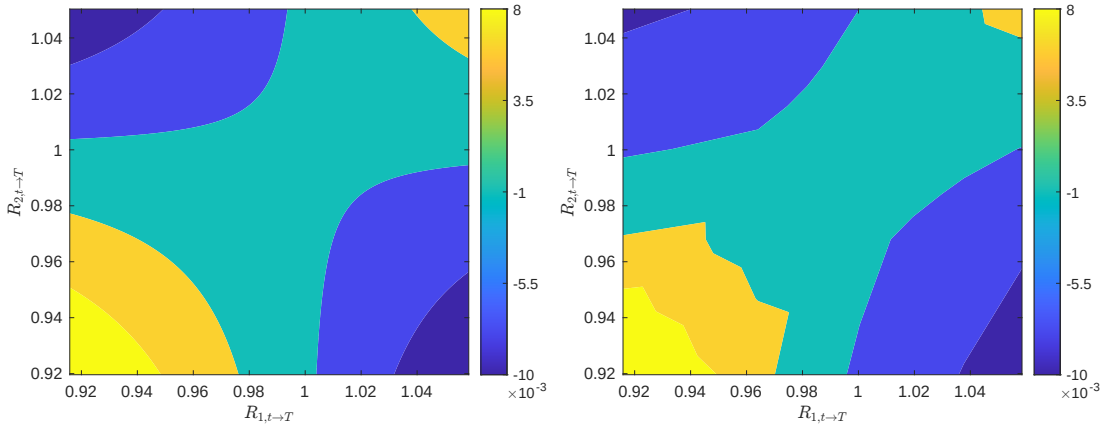
[14]Using the Garman–Kohlhagen formula in this step simply converts implied volatilities into option prices and does not impose Garman–Kohlhagen as the true pricing model.

6.2 Correlation estimates

Before presenting the estimates, it is instructive to assess the quality of the projection approximation in our sample. We take option quotes on the three exchange rates on July 1, 2008 (the first day of our sample) and project the covariance payoff

$$(R_{1,t\to T} - 1)\,(R_{2,t\to T} - 1)$$

onto the span of the 15 observed option payoffs, together with the cash position and the two underlying exchange rates. The left panel of Figure 4 plots the true covariance payoff, while the right panel plots the payoff of the replicating portfolio obtained by projection. The contour shapes are highly similar overall, and the projection captures the key tail regions, corroborating our simulation evidence that the risk-neutral covariance is well approximated from the available options in the exchange rate setting.



(a) Covariance payoff (b) Projection replication

Figure 4: **Covariance replicating portfolio.** The figure shows the covariance payoff (left) and its projection-based replicating portfolio payoff (right) for EUR/USD and GBP/USD gross returns. Option strikes and forward prices are from July 1, 2008.

Panel 5a reports the 1-month forward-looking risk-neutral correlation between the EUR/USD and GBP/USD exchange rates over time. As expected, almost all estimates lie below one; the few instances slightly above one are consistent with small measurement noise, as in our simulations. The sample-average correlation is about 0.7, in line with the view that major exchange rates co-move due to a handful of common risk factors. The lowest estimate—about 0.2—occurs just before the Brexit referendum, on June 9–10, 2016. Panel 5b indicates that the decline in correlation is driven primarily by a sharp increase in GBP/USD volatility.

The high frequency of option quotes also lets us zoom in on short-lived episodes. One stands out: a sharp decline in the 1-month risk-neutral correlation between December 12, 2012 and February 14, 2013, from nearly one to roughly 0.4. As the right panel of Figure 5 shows, this drop was not accompanied by a spike in the (annualized) volatilities, pointing to a genuine change in dependence rather than a level-volatility effect. Several contemporaneous developments are consistent with this interpretation: unexpectedly weak UK Q4-2012 GDP (weighing on GBP) alongside improving euro-area conditions such as the tightening peripheral spreads and the first LTRO repayments, which would have supported EUR. We therefore view this episode as a period in which currency-specific risks dominated shared USD drivers, temporarily depressing the implied correlation.

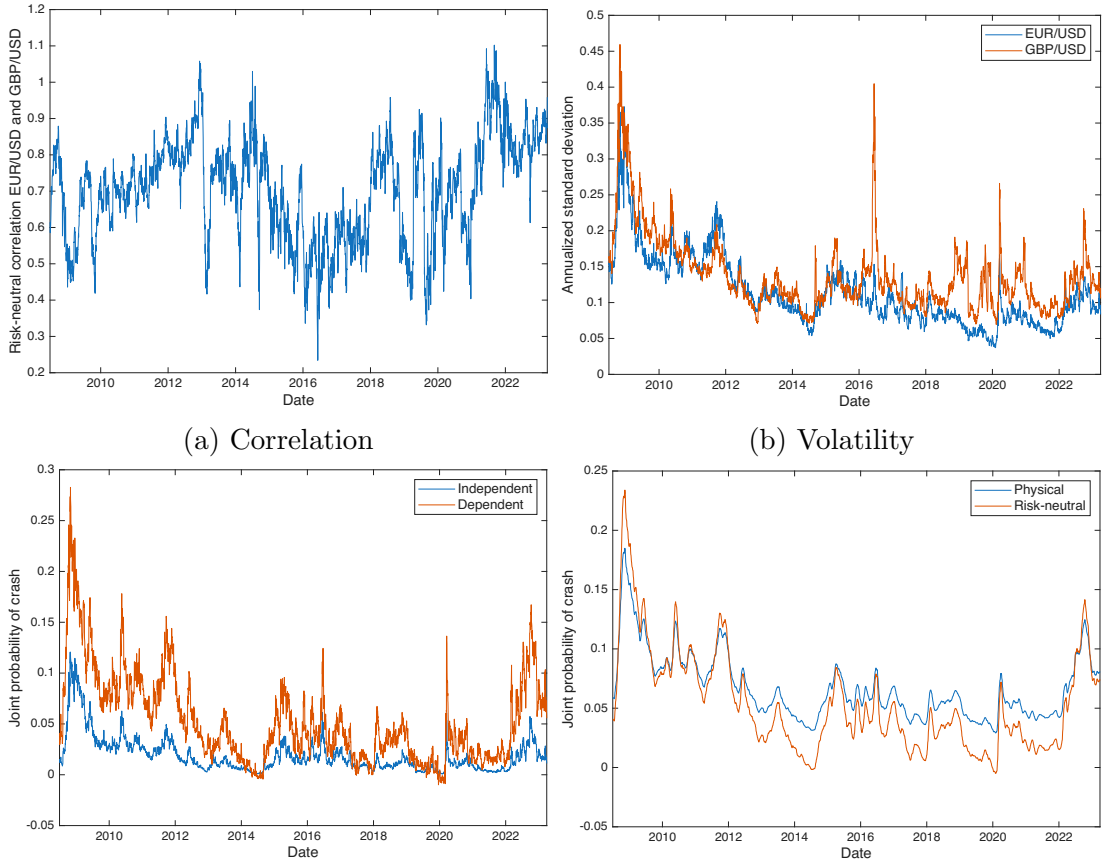
6.3 Tail probability estimates

Second, we examine the joint risk-neutral crash probability, defined as the probability that both EUR/USD and GBP/USD monthly returns are less than 3%. The estimate from our projection approach is shown in red in Panel 5c (labeled "dependent"). For comparison, we also plot the independence benchmark (labeled "independent"), obtained by multiplying the estimated marginal crash probabilities. The figure shows that accounting for dependence is crucial: the joint (dependent) probability is typically well above the independence benchmark, especially during periods of market stress.

To evaluate the informativeness of the joint risk-neutral crash probability, we estimate the forecasting model

$$\texttt{Crash}_{t\to T} = \beta_0 + \beta_1 \texttt{RiskNeutralProb}_{t\to T} + \varepsilon_{t\to T}, \tag{14}$$

where $\texttt{Crash}_T = \mathbb{1}\left(R_{1,t\to T} \leq 0.97\right)\mathbb{1}\left(R_{2,t\to T} \leq 0.97\right)$. The regressor is either the dependent (option-implied) joint crash probability or the independence benchmark (product of marginal crash probabilities). Results appear in Table 1. The dependent joint probability is a significant predictor, and the associated R^2 is substantially larger than for the independence benchmark. If physical and risk-neutral crash probabilities coincided at each date, the restriction $[\beta_0, \beta_1] = [0, 1]$ would hold; the bottom row reports the p-value of this Wald test, which is not rejected only for the dependent regressor. We also report an out-of-sample R^2,



(a) Correlation

(b) Volatility

(c) Joint crash probability

(d) Joint crash probability (smoothed)

Figure 5: **Daily risk-neutral correlation, volatility, and crash risk (30-day horizon).** Panels (a)–(d) plot: (a) the 30-day risk-neutral correlation between EUR/USD and GBP/USD; (b) the corresponding annualized 30-day risk-neutral standard deviations for each exchange rate; (c) the 30-day joint crash probability under independence ($p_{\text{EUR}}\, p_{\text{GBP}}$) and under the option-implied dependence structure; (d) the option-implied (dependent) crash probability alongside the physical crash probability estimated from OLS. The estimates in this last panel are smoothed using a 30-day moving average.

R^2_{OOS}, defined as

$$R^2_{OOS} = 1 - \frac{\sum_T (\texttt{Crash}_{t\to T} - \widehat{\texttt{Crash}}_{t\to T})^2}{\sum_T (\texttt{Crash}_{t\to T} - \overline{\texttt{Crash}}_{t\to T})^2},$$

where forecasts are $\widehat{\texttt{Crash}}_{t\to T} = \texttt{RiskNeutralProb}_{t\to T}$, and $\overline{\texttt{Crash}}_{t\to T}$ is the historical prevailing crash probability computed using an expanding window that begins after 1,000 historical observations are available. This design avoids any in-sample bias and yields a strict out-of-sample evaluation. In both specifications R^2_{OOS} is positive, with larger values when using the dependent covariate, indicating that

risk-neutral probabilities outperform the prevailing-mean benchmark.

The last column of Table 1 includes both predictors; the incremental R^2 gain is modest, and the coefficient on the independence benchmark enters with the opposite sign. We conclude that the option-implied (dependent) joint crash probability performs markedly better, providing evidence that it helps forecast joint physical tail risk.

Panel 5d plots the inferred physical joint crash probability based on the regression with the dependent (option-implied) covariate, alongside the risk-neutral series; both are smoothed for readability. The figure illustrates a time-varying premium for joint crash risk. During turbulent periods (e.g., the Global Financial Crisis), the risk-neutral probability exceeds the physical estimate, consistent with a positive compensation for bearing joint tail risk. In contrast, in calmer markets the ordering often reverses—the physical probability exceeds the risk-neutral one—suggesting that currency exposures may provide a hedging benefit and earn a negative tail-risk premium. Overall, the evidence points to currencies serving as tail-risk hedges in normal times, but commanding compensation during stress episodes.

This evidence is consistent with the structural explanation of Lustig et al. (2014). They argue that, in times of stress when the marginal utility of wealth is high, U.S. investors who are long foreign currencies are exposed to the risk that the dollar appreciates. Consequently, the conditional expected return on such a strategy should be high. In contrast, during normal times the strategy behaves more like a hedge: investors bear the risk of a dollar depreciation following a positive shock to the U.S. pricing kernel, so the conditional expected return is low or even negative.

7 Conclusion

This paper introduces a new approach to estimating risk-neutral expectations from option prices. The core idea is to project the target payoff function onto the space spanned by observed option payoffs and the underlying asset. Like the method of Carr and Madan (2001), the resulting estimate is a linear combination of option prices and the underlying. However, the projection approach makes optimal use of the available strike prices to minimize the approximation error. We show that this method much better finite sample properties. Simulation results confirm this advantage: the projection method delivers approximation errors that are orders of magnitude smaller.

	(1)	(2)	(3)
Constant (β_0)	0.042 (0.0171)	0.032 (0.0182)	0.031 (0.0178)
Independent	1.350 (0.8585)		−1.380 (1.8274)
Dependent		0.652 (0.3497)	1.176 (0.7555)
$R^2(\%)$	0.94	1.33	1.45
$R^2_{OOS}(\%)$	2.70	3.75	
p-value (const= 0, slope= 1)	0.00	0.16	

Table 1: **OLS estimates of** (14). Each column reports a different forecasting model. Newey–West standard errors (20 trading-day lag) are shown beneath the coefficients. The bottom row reports the p-value of the Wald test on the joint restriction $[\beta_0, \beta_1] = [0, 1]$.

We extend the projection approach to higher dimensions and, using approximation-theoretic tools (ridge functions), derive necessary and sufficient conditions under which simple options complete multiple asset markets. Although these conditions are rarely satisfied exactly in practice, we show that projection still estimates joint risk-neutral expectations robustly—most notably for risk-neutral covariances/correlations in the FX setting. Thus, projection provides a unified framework for estimating risk-neutral quantities not only for a single asset but also in the multi-asset case.

The empirical application, FX, provides a clean setting for multivariate estimation. In simulations, the projection estimator recovers the true correlation with near-zero error. In the data, we estimate the conditional 1-month risk-neutral correlation between EUR/USD and GBP/USD returns, which averages around 0.7. Thanks to the high frequency and forward-looking nature of option quotes, we detect a notable shift in this correlation at the end of 2012. We interpret this as a genuine change in dependence: bearish U.K. news contrasted with more bullish euro-area developments that temporarily weakened the co-movement between the two USD majors.

Relatedly, we also estimate the joint risk-neutral crash probability and find that it forecasts future realized crashes. Furthermore, when comparing the risk-neutral crash probability to the physical probability inferred from an OLS regression, we find that the risk-neutral probability is higher during crises but generally lower outside these periods. We interpret this as data-driven evidence that U.S. investors in portfolios of foreign currencies demand crash compensation in bad times, but value these positions as a hedge in normal market conditions.

References

Aït-Sahalia, Yacine and Andrew W. Lo (1998). "Nonparametric estimation of state-price densities implicit in financial asset prices". *Journal of Finance* 53.2, 499–547. DOI: 10.1111/0022-1082.215228.

Aït-Sahalia, Yacine and Andrew W. Lo (2000). "Nonparametric risk management and implied risk aversion". *Journal of Econometrics* 94.1-2, 9–51. DOI: 10.1016/S0304-4076(99)00016-0.

Almeida, Caio and Gustavo Freire (2022). "Pricing of index options in incomplete markets". *Journal of Financial Economics* 144.1, 174–205. DOI: 10.1016/j.jfineco.2021.05.041.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov (2017). "Short-Term Market Risks Implied by Weekly Options". *Journal of Finance* 72.3, 1335–1386. DOI: 10.1111/jofi.12486.

Back, Kerry (2017). *Asset Pricing and Portfolio Choice Theory*. second. Oxford University Press.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan (2003). "Stock return characteristics, skew laws, and the differential pricing of individual equity options". *Review of Financial Studies* 16.1, 101–143. DOI: 10.1093/rfs/16.1.0101.

Bakshi, Gurdip and Dilip Madan (2000). "Spanning and derivative-security valuation". *Journal of Financial Economics* 55.2, 205–238. DOI: 10.1016/S0304-405X(99)00050-1.

Bates, David S. (1991). "The crash of '87: Was it expected? Evidence from options markets". *Journal of Finance* 46.3, 1009–1044. DOI: 10.1111/j.1540-6261.1991.tb03775.x.

Beare, Brendan K. and Lawrence D. W. Schmidt (2016). "An empirical test of pricing kernel monotonicity". *Journal of Applied Econometrics* 31.2, 338–356.

Billingsley, Patrick (1999). *Convergence of Probability Measures*. second. John Wiley & Sons.

Black, Fischer and Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities". *Journal of Political Economy* 81.3, 637–654. DOI: 10.1086/260062.

Bliss, Robert R. and Nikolaos Panigirtzoglou (2004). "Option-Implied Risk Aversion Estimates". *Journal of Finance* 59.1, 407–446. DOI: 10.1111/j.1540-6261.2004.00637.x.

Bollerslev, Tim, George Tauchen, and Hao Zhou (2009). "Expected stock returns and variance risk premia". *Review of Financial Studies* 22.11, 4463–4492. DOI: 10.1093/rfs/hhp008.

Bondarenko, Oleg (2003). "Estimation of risk-neutral densities using positive convolution approximation". *Journal of Econometrics* 116.1, 85–112. DOI: 10.1016/S0304-4076(03)00104-0.

Bondarenko, Oleg and Carole Bernard (2024). "Option-Implied Dependence and Correlation Risk Premium". *Journal of Financial and Quantitative Analysis* 59.7, 3139–3189. DOI: 10.1017/S0022109023000960.

Breeden, Douglas T. and Robert H. Litzenberger (1978). "Prices of state-contingent claims implicit in option prices". *Journal of Business*, 621–651.

Britten-Jones, Mark and Anthony Neuberger (2000). "Option Prices, Implied Price Processes, and Stochastic Volatility". *Journal of Finance* 55.2, 839–866. DOI: 10.1111/0022-1082.00228.

Broadie, Mark, Mikhail Chernov, and Michael Johannes (2007). "Model specification and risk premia: Evidence from futures options". *Journal of Finance* 62.3, 1453–1490. DOI: 10.1111/j.1540-6261.2007.01241.x.

Candès, Emmanuel J., Justin Romberg, and Terence Tao (2006). "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information". *IEEE Transactions on Information Theory* 52.2, 489–509.

Canuto, Claudio, M. Youssuff Hussaini, Alfio Quarteroni, and Thomas A. Zang (2006). *Spectral Methods: Fundamentals in Single Domains*. Springer.

Carr, Peter and Dilip Madan (2001). "Towards a Theory of Volatility Trading". In: *Handbooks in Mathematical Finance: Option Pricing, Interest Rates and Risk Management*. Cambridge University Press, 458–476.

Carr, Peter and Liuren Wu (2009). "Variance risk premiums". *Review of Financial Studies* 22.3, 1311–1341. DOI: 10.1093/rfs/hhn038.

Chabi-Yo, Fousseni and Johnathan Loudis (2020). "The conditional expected market return". *Journal of Financial Economics* 137.3, 752–786. DOI: 10.1016/j.jfineco.2020.03.009.

Chernov, Mikhail, Jeremy Graveline, and Irina Zviadadze (2018). "Crash Risk in Currency Returns". *Journal of Financial and Quantitative Analysis* 53.1, 137–170.

Chernozhukov, Victor, Iván Fernández-Val, and Alfred Galichon (2009). "Improving point and interval estimators of monotone functions by rearrangement". *Biometrika* 96.3, 559–575. DOI: 10.1093/biomet/asp030.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly (2013). "Inference on counterfactual distributions". *Econometrica* 81.6, 2205–2268.

de Boor, Carl (2001). *A Practical Guide to Splines*. Revised. Vol. 27. Applied Mathematical Sciences. Springer-Verlag.

Diaconis, Persi and Mehrdad Shahshahani (1984). "On nonlinear functions of linear combinations". *SIAM Journal on Scientific and Statistical Computing* 5.1, 175–191.

Embree, Mark (2010). *Numerical Analysis I*. Lecture notes, Rice University. Pages 1–207.

Engle, Robert F. and Bryan Kelly (2012). "Dynamic equicorrelation". *Journal of Business & Economic Statistics* 30.2, 212–228.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson (2003). "The impact of jumps in volatility and returns". *Journal of Finance* 58.3, 1269–1300. DOI: 10.1111/1540-6261.00566.

Figlewski, Stephen (2010). "Estimating the Implied Risk-Neutral Density for the US Market Portfolio". In: *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*. Oxford University Press. DOI: 10.1093/acprof:oso/9780199549498.003.0015.

Figlewski, Stephen (2018). "Risk-Neutral Densities: A Review". *Annual Review of Financial Economics* 10.1, 329–359. DOI: 10.1146/annurev-financial-110217-022944.

Filipović, Damir, Eberhard Mayerhofer, and Paul Schneider (2013). "Density approximations for multivariate affine jump-diffusion processes". *Journal of Econometrics* 176.2, 93–111. DOI: j.jeconom.2012.12.003.

Friedman, Jerome H. and Werner Stuetzle (1981). "Projection pursuit regression". *Journal of the American Statistical Association* 76.376, 817–823.

Hansen, Bruce E. (2022). *Econometrics*. Princeton University Press.

Jackwerth, Jens Carsten (2000). "Recovering risk aversion from option prices and realized returns". *Review of Financial Studies* 13.2, 433–451.

Jackwerth, Jens Carsten and Mark Rubinstein (1996). "Recovering Probability Distributions from Option Prices". *Journal of Finance* 51.5, 1611–1631.

Jiang, George J. and Yisong S. Tian (2005). "The model-free implied volatility and its information content". *Review of Financial Studies* 18.4, 1305–1342.

Kelly, Bryan, Hanno Lustig, and Stijn Van Nieuwerburgh (2016). "Too-systemic-to-fail: What option markets imply about sector-wide government guarantees". *American Economic Review* 106.6, 1278–1319. DOI: 10.1257/aer.20120389.

Kozhan, Roman, Anthony Neuberger, and Paul Schneider (2013). "The skew risk premium in the equity index market". *Review of Financial Studies* 26.9, 2174–2203. DOI: 10.1093/rfs/hht039.

Kremens, Lukas and Ian Martin (2019). "The quanto theory of exchange rates". *American Economic Review* 109.3, 810–843. DOI: 10.1257/aer.20180019.

Lebesgue, Henri (1898). "Sur l'approximation des fonctions". *Bulletin des Sciences Mathématiques* 22.10, 278–287.

Lin, V.Y. and A. Pinkus (1993). "Fundamentality of Ridge Functions". *Journal of Approximation Theory* 75.3, 295–311. DOI: 10.1006/jath.1993.1104.

Linn, Matthew, Sophie Shive, and Tyler Shumway (2017). "Pricing Kernel Monotonicity and Conditional Information". *Review of Financial Studies* 31.2, 493–531. DOI: 10.1093/rfs/hhx095.

Lustig, Hanno, Nikolai Roussanov, and Adrien Verdelhan (2014). "Countercyclical currency risk premia". *Journal of Financial Economics* 111.3, 527–553.

Martin, Ian (2017). "What is the Expected Return on the Market?" *Quarterly Journal of Economics* 132.1, 367–433.

Martin, Ian (2018). "Options and the gamma knife". *Journal of Portfolio Management* 44.6, 47–55. DOI: 10.3905/jpm.2018.44.6.047.

Martin, Ian (2025). "Information in Derivatives Markets: Forecasting Prices with Prices". *Annual Review of Financial Economics*. ISSN: 1941-1367. DOI: 10.1146/annurev-financial-082123-105811.

Mueller, Philippe, Andreas Stathopoulos, and Andrea Vedolin (2017). "International correlation risk". *Journal of Financial Economics* 126.2, 270–299. DOI: 10.1016/j.jfineco.2016.09.012.

Nachman, David C. (1988). "Spanning and completeness with options". *Review of Financial Studies* 1.3, 311–328. DOI: 10.1093/rfs/1.3.311.

Pinkus, Allan (2015). *Ridge Functions*. Vol. 205. Cambridge University Press.

Rosenberg, Joshua V. and Robert F. Engle (2002). "Empirical pricing kernels". *Journal of Financial Economics* 64.3, 341–372. DOI: 10.1016/S0304-405X(02)00128-9.

Ross, Stephen A. (1976). "Options and Efficiency". *Quarterly Journal of Economics* 90.1, 75–89.

Ross, Stephen A. (2015). "The recovery theorem". *Journal of Finance* 70.2, 615–648. DOI: 10.1111/jofi.12092.

Schneider, Paul and Fabio Trojani (2019). "(Almost) model-free recovery". *Journal of Finance* 74.1, 323–370. DOI: 10.1111/jofi.12737.

Shreve, Steven E. (2004). *Stochastic Calculus for Finance II: Continuous-Time Models*. Vol. 11. Springer.

Trefethen, Lloyd N. (2018). *Approximation Theory and Approximation Practice*. Extended. Philadelphia, PA: Society for Industrial and Applied Mathematics. DOI: 10.1137/1.9781611975949.

Vostrecov, B. A. and M. A. Kreines (1961). "Approximation of continuous functions by superpositions of plane waves". In: *Doklady Akademii Nauk SSSR*. Vol. 140, 1237–1240.

A Proofs

A.1 Proof of Proposition 1

Proof. The normal equations yield $X'X\hat{\beta}_{n_s} = X'Y$. The (i,j)-element of $X'X$ and the ith element of $X'Y$ are given by

$$(X'X)_{ij} = \sum_{z=1}^{n_s} \phi_i(s_z)\phi_j(s_z), \quad (X'Y)_i = \sum_{z=1}^{n_s} \phi_i(s_z)g(s_z).$$

Assuming that the grid is equally spaced with length $m(n_s) = (a_{\max} - a_{\min})/n_s$, it follows by the Riemann sum approximation that as $n_s \to \infty$

$$m(n_s)(X'X)_{ij} \to \int_A \phi_i(S_T)\phi_j(S_T)\,\mathrm{d}S_T, \quad m(n_s)(X'Y)_i \to \int_A \phi_i(S_T)g(S_T)\,\mathrm{d}S_T.$$

The proof continues to hold if the grid in not equally spaced but the mesh goes to zero. The associated Gram matrix is invertible because the basis functions are linearly independent in $L^2(A)$, so the solution to the normal equations exists and is unique if n_s is sufficiently large.

The proof that $\hat{\beta}$ also solves the minimization problem follows immediately from the first order conditions

$$\left(\int_A \phi(S_T)\phi(S_T)'\,\mathrm{d}S_T\right)\hat{\beta} = \int_A g(S_T)\phi(S_T)\,\mathrm{d}S_T,$$

where $\phi(S_T) = [\phi_1(S_T), \ldots, \phi_{2+n_k}(S_T)]'$. □

A.2 Proof of Proposition 2

The following proof is well known (see Lebesgue (1898)), but we include it for completeness and because the assumption on the strikes results in some slight modifications of the original proof. The proof below is presented for call options, but applies verbatim to put options as well.

Proof. Let $g \in C(A)$. Because g is continuous on a compact set it is uniformly continuous: for every $\varepsilon > 0$ there exists a $\delta > 0$ (independent of x), such that $\sup_{|x-y|<\delta} |g(x) - g(y)| < \varepsilon$. Let $a_{\min} = x_1 < x_2 < \cdots < x_n = a_{\max}$ be a partition of A such that $x_{j+1} - x_j < \delta \ \forall j$, where $a_{\min} = \min(A)$ and $a_{\max} = \max(A)$. On each interval $[x_j, x_{j+1}]$ construct a linear function $\tilde{g}_j(x) = a_j x + b_j$ such that $\tilde{g}_j(x_j) = g(x_j)$ and $\tilde{g}_j(x_{j+1}) = g(x_{j+1})$. For every $x_c \in (x_j, x_{j+1})$ it follows that

$$|g(x_c) - \tilde{g}_j(x_c)| \leq |g(x_c) - g(x_j)| + |g(x_j) - \tilde{g}_j(x_c)| < 2\varepsilon,$$

because

$$|g(x_j) - \tilde{g}_j(x_c)| = \left| \frac{x_c - x_j}{x_{j+1} - x_j} \right| |g(x_{j+1}) - g(x_j)| .$$

Since x_c is arbitrary, it follows that $\sup_{x\in[x_j,x_{j+1}]} |g(x) - \tilde{g}_j(x)| < 2\varepsilon$. Now, define the polygonal function

$$\tilde{g}(x) = \sum_{j=1}^{n-2} \tilde{g}_j(x) \mathbb{1}\left(x \in [x_j, x_{j+1})\right) + \tilde{g}_{n-1}(x) \mathbb{1}\left(x \in [x_{n-1}, x_n]\right). \tag{15}$$

From the construction above it follows that $\tilde{g}$ is continuous and $\sup_{x\in A} |g(x) - \tilde{g}(x)| < 2\varepsilon$. We claim that the polygonal function constructed in this way can be written as

$$\tilde{g}(x) = \beta_1 + \sum_{j=1}^{n-1} \beta_{j+1} (x - x_j)^+ . \tag{16}$$

To see this, proceed inductively. On $[x_1, x_2]$, (15) can be written as

$$\tilde{g}(x) = a_1 x + b_1 = a_1 (x - x_1)^+ + \tilde{b}_1,$$

where $\tilde{b}_1 = b_1 + a_1 x_1$. On $[x_1, x_3]$, we can write

$$\tilde{g}(x) = a_1 (x - x_1)^+ + \tilde{b}_1 + \tilde{a}_2 (x - x_2)^+ + \tilde{b}_2,$$

where

$$a_1 + \tilde{a}_2 = a_2 \quad \text{and} \quad \tilde{b}_2 = b_2 + a_1 x_1,$$

which can be solved for to obtain $\tilde{a}_2, \tilde{b}_2$. Continuing inductively, we obtain (16). It remains to show that $\tilde{g}$ can be uniformly approximated by a function of the form

$$\tilde{g}_{n_k}(x) = \beta_1 + \sum_{j=1}^{n_k-1} \beta_{j+1} (x - K_j)^+ ,$$

where K_j is among the observed call option strike prices. But this can be achieved if n_k is large enough. Specifically, let n_k be large enough such that $\max_{j=1,\ldots,n-1} |x_j - K_j| < \varepsilon$. By assumption such n_k can always be found since $\{K_j\}_{j=1}^{n_k}$ is dense in A as $n_k \to \infty$. Considering that

$$\sup_{x \in A} \left| (x - x_j)^+ - (x - K_j)^+ \right| < \varepsilon,$$

it follows by another application of the triangle inequality that

$$\sup_{x \in A} |g(x) - \tilde{g}_{n_k}(x)| < 3\varepsilon.$$

□

A.3 Proof of Corollary 1

Proof. According to Billingsley (1999, Theorem 1.2), a probability measure $\mathbf{P}$ on a metric space is completely determined by the expected values $\mathbf{E}f(X)$, for all bounded, uniformly continuous functions f, where $X \sim \mathbf{P}$. Proposition 2 shows there is a sequence of functions $f_{n_k} \in \text{span}(\mathcal{F}_{2+n_k})$ converging uniformly to f. Because A is compact and f, f_{n_k} are continuous (hence bounded), the dominated convergence theorem shows that $\mathbf{E}f(X)$ is pinned down uniquely for every bounded, uniformly continuous f. □

A.4 Proof of Proposition 3

Proof. Without loss of generality, we assume that all strike prices correspond to call options. We start by deriving an error bound on the piecewise linear polynomial, denoted by $\tilde{g}$, that interpolates the points

$$\left\{ (a_{\min}, g(a_{\min})), (K_j, g(K_j))_{j=1}^{n_k}, (a_{\max}, g(a_{\max})) \right\}.$$

Letting $\tilde{g}_j$ denote the interpolating polynomial on $[K_j, K_{j+1}]$, it follows from standard results in approximation theory (e.g., Embree (2010, Lecture 11)) that

$$\begin{aligned}\max_{x\in[K_j,K_{j+1}]} |g(x)-\tilde{g}_j(x)| &\le \left(\max_{\xi\in[K_j,K_{j+1}]} \frac{|g''(\xi)|}{2}\right)\left(\max_{x\in[K_j,K_{j+1}]} (x-K_j)(K_{j+1}-x)\right)\\ &\le \|g'\|_\infty \frac{1}{8}\left(K_{j+1}-K_j\right)^2, \end{aligned} \tag{17}$$

where $\|g'\|_\infty = \max_{\xi\in[a_{\max},a_{\min}]} |g''(\xi)|$. Hence,

$$\int_{K_j}^{K_{j+1}} \left(g(x)-\tilde{g}_j(x)\right)^2 \mathrm{d}x \le \frac{1}{64}\|g'\|_\infty^2 \left(K_{j+1}-K_j\right)^5.$$

Since $\tilde{g}$ equals $\tilde{g}_j(x)$ on $[K_j, K_{j+1})$, it follows that

$$\begin{aligned}\int_{K_1}^{K_{n_k}} \left(g(x)-\tilde{g}(x)\right)^2 \mathrm{d}x &= \sum_{j=1}^{n_k-1} \int_{K_j}^{K_{j+1}} \left(g(x)-\tilde{g}_j(x)\right)^2 \mathrm{d}x\\ &\le \frac{n_k}{64}\|g'\|_\infty^2 \Delta^5\\ &= O(1/n_k^4), \end{aligned} \tag{18}$$

where in the last line we used that $\Delta = O(1/n_k)$. Applying (17) again on $[K_{n_k}, a_{\max}]$ renders the estimate

$$\max_{x\in[K_{n_k},a_{\max}]} |g(x)-\tilde{g}(x)| \le \|g'\|_\infty \frac{1}{8}\left(a_{\max}-K_{n_k}\right)^2. \tag{19}$$

A similar bound can be derived on $[a_{\min}, K_1]$. From the proof of Proposition 2 we know that $\tilde{g}(x)$ can be written in the form

$$\tilde{g}(x) = \beta_1 + \beta_2 x + \sum_{j=1}^{n_k} \beta_{2+j}\left(x-K_j\right)^+.$$

Then we can bound the estimation error as follows

$$
\begin{aligned}
&\left|\mathbf{E}_t^Q\left[g(S_T)\mathbb{1}\left(S_T\in A\right)\right]-\mathbf{E}_t^Q\left[\hat{g}(S_T)\mathbb{1}\left(S_T\in A\right)\right]\right| \\
&=\left|\int_{a_{\min}}^{a_{\max}}\left(g(x)-\hat{g}(x)\right)f_{t\to T}^Q(x)\,\mathrm{d}x\right| \\
&\leq\int_{a_{\min}}^{a_{\max}}\left|g(x)-\hat{g}(x)\right|f_{t\to T}^Q(x)\,\mathrm{d}x \\
&\leq\left(\int_{a_{\min}}^{a_{\max}}\left(g(x)-\hat{g}(x)\right)^2\mathrm{d}x\right)^{1/2}\left(\int_{a_{\min}}^{a_{\max}}f_{t\to T}^Q(x)^2\,\mathrm{d}x\right)^{1/2} \\
&\leq\left(\int_{a_{\min}}^{a_{\max}}\left(g(x)-\tilde{g}(x)\right)^2\mathrm{d}x\right)^{1/2}\left(\int_{a_{\min}}^{a_{\max}}f_{t\to T}^Q(x)^2\,\mathrm{d}x\right)^{1/2} \\
&=\left(\int_{a_{\min}}^{a_{\max}}f_{t\to T}^Q(x)^2\,\mathrm{d}x\right)^{1/2}\left(\int_{a_{\min}}^{K_1}\left(g(x)-\tilde{g}(x)\right)^2\mathrm{d}x\right. \\
&\left.+\int_{K_1}^{K_{n_k}}\left(g(x)-\tilde{g}(x)\right)^2\mathrm{d}x+\int_{K_{n_k}}^{a_{\max}}\left(g(x)-\tilde{g}(x)\right)^2\mathrm{d}x\right)^{1/2} \\
&=:\left(\int_{a_{\min}}^{a_{\max}}f_{t\to T}^Q(x)^2\,\mathrm{d}x\right)^{1/2}\left(B_1+B_2+B_3\right)^{1/2},
\end{aligned}
$$

where we successively used the Cauchy-Schwarz inequality combined with the square-integrability of $f_{t\to T}^Q$, and the minimization property of $\hat{g}$. From (18), we know that $B_2=O(\Delta^4)=O(1/n_k^4)$. Moreover, by (19) and the assumption that $a_{\max}-K_{n_k}=O(1/n_k^{4/5})$, B_3 is of order $O(a_{\max}-K_{n_k})^5=O(1/n_k^4)$. Analogous reasoning yields $B_1=O(K_1-a_{\min})^5=O(1/n_k^4)$.

□

A.5 Proof of Proposition 4

Proof. Over A, the CM Taylor expansion in (1) is given by

$$
\begin{aligned}
g(x)&=g(F_{t\to T})+g'(F_{t\to T})\left(x-F_{t\to T}\right) \\
&+\int_{a_{\min}}^{F_{t\to T}}g''(K)\left(K-x\right)^+\mathrm{d}K+\int_{K_{n_k}}^{a_{\max}}g''(K)\left(x-K\right)^+\mathrm{d}K.
\end{aligned}
$$

We will focus on the case $x\leq F_{t\to T}$ (the case $x>F_{t\to T}$ is identical). The integral is discretized using the trapezoidal rule, which is known to satisfy

$$
\sum_{j:K_j\leq F_{t\to T}}\Delta K_j\,g''(K_j)\left(K_j-S_T\right)^+=\int_{K_1}^{F_{t\to T}}g''(K)\left(K-x\right)^+\mathrm{d}K+O\left(\frac{1}{n_k^2}\right),
$$

uniformly in x. Hence, for $x \in [K_1, F_{t\to T}]$, we obtain

$$\max_{x\in[K_1,F_{t\to T}]} |g(x) - \hat{g}_{\mathrm{CM}}(x)| = O\left(\frac{1}{n_k^2}\right).$$

For $x \in [a_{\min}, K_1]$, we get

$$\begin{aligned}|g(x) - \hat{g}_{\mathrm{CM}}(x)| &= \left|\int_x^{K_1} g''(K)(K-x)\,\mathrm{d}K\right| \\ &\le \|g''\|_\infty \frac{1}{2}(K_1 - x)^2.\end{aligned}$$

Analogous reasoning yields a similar bound for $x > F_{t\to T}$. The same reasoning at the end of Proposition 3 then finally gives

$$\begin{aligned}&\left|\int_{a_{\min}}^{a_{\max}} (g(x) - \hat{g}_{\mathrm{CM}}(x))\, f^Q_{t\to T}(x)\,\mathrm{d}x\right| \\ &\le \left(\int_{a_{\min}}^{a_{\max}} f^Q_{t\to T}(x)^2\,\mathrm{d}x\right)^{1/2} \left(\int_{a_{\min}}^{K_1} (g(x) - \hat{g}_{\mathrm{CM}}(x))^2\,\mathrm{d}x\right. \\ &\left. + \int_{K_1}^{K_{n_k}} (g(x) - \hat{g}_{\mathrm{CM}}(x))^2\,\mathrm{d}x + \int_{K_{n_k}}^{a_{\max}} (g(x) - \hat{g}_{\mathrm{CM}}(x))^2\,\mathrm{d}x\right)^{1/2} \\ &= \left(O\left(K_1 - a_{\min}\right)^5 + O\left(\frac{1}{n_k^4}\right) + O\left(a_{\max} - K_{n_k}\right)^5\right)^{1/2} \\ &= O\left(\frac{1}{n_k^2}\right).\end{aligned}$$

□

A.6 Proof of Proposition 5

Proof. Let PL denote the space of continuous piecewise linear functions on this knot sequence. It is standard that (e.g. using the proof of Proposition 2)

$$PL = \mathrm{span}\big(\mathcal{F}_{2+n_k}\big) = \mathrm{span}\{1, x, (x-K_1)_+, \ldots, (x-K_{n_k})_+\}.$$

Equivalently, PL is spanned by the nodal tent functions $\{\varphi_i\}_{i=0}^{n_k+1}$ defined by

$$\varphi_i(K_j) = \delta_{ij}, \qquad \mathrm{supp}(\varphi_i) = [K_{i-1}, K_{i+1}],$$

(where $\varphi_0, \varphi_{n_k+1}$ are the boundary hats). In particular, any $s \in PL$ can be written uniquely as

$$s(x) = \sum_{i=0}^{n_k+1} \alpha_i \varphi_i(x), \qquad \alpha_i = s(K_i).$$

Let $g \in C^4(A)$ and let $\hat{g}$ be its $L^2(A)$-projection onto PL. Write $\hat{g}(x) = \sum_{i=0}^{n_k+1} \alpha_i \varphi_i(x)$ and define

$$b_i \coloneqq \int_A \varphi_i(x) g(x)\,\mathrm{d}x, \qquad M_{ij} \coloneqq \int_A \varphi_i(x)\varphi_j(x)\,\mathrm{d}x.$$

Then the normal equations are $M\alpha = b$. For interior indices $i = 1, \dots, n_k$ (away from the boundary), the matrix entries on a uniform grid are

$$M_{ii} = \frac{2h}{3}, \qquad M_{i,i\pm1} = \frac{h}{6}, \qquad M_{ij} = 0 \ \text{ if } |i-j| > 1,$$

so the interior normal equations read (see also de Boor (2001, p.34))

$$\frac{h}{6}\alpha_{i-1} + \frac{2h}{3}\alpha_i + \frac{h}{6}\alpha_{i+1} = b_i, \qquad i = 1, \dots, n_k. \tag{20}$$

Define $y_i \coloneqq b_i/h$ and the discrete operator $\mathcal{T}$ by

$$(\mathcal{T}\alpha)_i \coloneqq \frac{1}{6}\alpha_{i-1} + \frac{2}{3}\alpha_i + \frac{1}{6}\alpha_{i+1}.$$

Then (20) is equivalently

$$(\mathcal{T}\alpha)_i = y_i, \qquad i = 1, \dots, n_k. \tag{21}$$

Step 1 (expansion of y_i). For interior i, the hat function satisfies $\varphi_i(K_i + u) = 1 - |u|/h$ for $u \in [-h, h]$, hence

$$b_i = \int_{K_{i-1}}^{K_{i+1}} g(x)\varphi_i(x)\,\mathrm{d}x = \int_{-h}^{h} g(K_i + u)\Big(1 - \frac{|u|}{h}\Big)\,\mathrm{d}u.$$

Expanding $g(K_i + u)$ around $u = 0$ and using symmetry (odd moments vanish), we obtain

$$\frac{b_i}{h} = y_i = g(K_i) + \frac{h^2}{12} g''(K_i) + O(h^4), \qquad i = 1, \dots, n_k,$$

where the $O(h^4)$ term is uniform in i.

Step 2 (candidate solution). Define the candidate sequence

$$\tilde{\alpha}_i \coloneqq g(K_i) - \frac{h^2}{12} g''(K_i). \tag{22}$$

A Taylor expansion yields, for interior i,

$$(\mathcal{T}\tilde{\alpha})_i = \tilde{\alpha}_i + \frac{h^2}{6}\tilde{\alpha}''(K_i) + O(h^4).$$

Since $\tilde{\alpha}''(K) = g''(K) - \frac{h^2}{12} g^{(4)}(K)$, this implies

$$(\mathcal{T}\tilde{\alpha})_i = g(K_i) + \frac{h^2}{12} g''(K_i) + O(h^4).$$

Combining with Step 1 gives the residual

$$r_i \coloneqq (\mathcal{T}\tilde{\alpha})_i - y_i = O(h^4), \qquad i = 1, \ldots, n_k.$$

The operator $\mathcal{T}$ corresponds to a tridiagonal Toeplitz matrix on interior indices, and is strictly diagonally dominant. Hence $\mathcal{T}$ is uniformly invertible on interior indices and $\|\mathcal{T}^{-1}\| \le C$ for a constant C independent of h. Therefore, solving (21) and using $\mathcal{T}\alpha = y$,

$$\alpha - \tilde{\alpha} = \mathcal{T}^{-1}(y - \mathcal{T}\tilde{\alpha}) = -\mathcal{T}^{-1} r,$$

so $\alpha_i - \tilde{\alpha}_i = O(h^4)$ for interior i. In particular,

$$\alpha_i = g(K_i) - \frac{h^2}{12} g''(K_i) + O(h^4), \qquad i = 1, \ldots, n_k. \tag{23}$$

Step 3 (translate to option basis). Write the same projected spline in the option payoff basis,

$$\hat{g}(x) = \hat{\beta}_1 + \hat{\beta}_2 x + \sum_{i=1}^{n_k} \hat{\gamma}_i \left(x - K_i\right)^+ .$$

For $x \neq K_i$, differentiating gives

$$\hat{g}'(x) = \hat{\beta}_2 + \sum_{j : K_j < x} \hat{\gamma}_j,$$

hence $\hat{\gamma}_i$ is the jump in slope at K_i. Let

$$p_i := \frac{\alpha_{i+1} - \alpha_i}{h} \quad \text{(the slope of } \hat{g} \text{ on } [K_i, K_{i+1}]).$$

Then the jump in slope at K_i is

$$\hat{\gamma}_i = p_i - p_{i-1} = \frac{\alpha_{i+1} - 2\alpha_i + \alpha_{i-1}}{h}.$$

A Taylor expansion yields $\alpha_{i+1} - 2\alpha_i + \alpha_{i-1} = h^2\alpha''(K_i) + O(h^4)$, hence

$$\hat{\gamma}_i = h\,\alpha''(K_i) + O(h^3).$$

Using (23), we have $\alpha''(K_i) = g''(K_i) + O(h^2)$, and therefore

$$\hat{\gamma}_i = h\,g''(K_i) + O(h^3), \qquad i = 2, \ldots, n_k - 1,$$

i.e. for interior strikes the leading-order term of the projection coefficient in the truncated power basis is $h\,g''(K_i)$.

The slower convergence rate at the boundary coefficient $\hat{\gamma}_1$ follows because the kernel function ϕ_0 is one sided, so odd moments under the kernel function no longer vanish. The same observation applies to $\hat{\gamma}_{n_k}$. □

A.7 Proof of Proposition 6

Proof. The space spanned by $\mathcal{F}_{2+n_k}$ is equal to the span of the B-spline basis functions of order 2 with knots at $a_{\min} < K_1 < \cdots < K_{n_k} < a_{\max}$. In particular, this implies that the $L^2(A)$-projections concur. de Boor (2001, Theorem 12 in Chapter 2) then shows that

$$\max_{x \in A} |g(x) - \hat{g}(x)| \leq 4\,\mathrm{dist}(g, \mathcal{F}_{2+n_k}).$$

Consequently,

$$\begin{aligned}\left|\mathbf{E}_t^Q g(S_T) - \mathbf{E}_t^Q \hat{g}(S_T)\right| &\leq \int_0^\infty |g(x) - \hat{g}(x)|\, f_{t\to T}^Q(x)\,\mathrm{d}x \\ &= \int_0^{a_{\min}} |g(x) - \hat{g}(x)|\, f_{t\to T}^Q(x)\,\mathrm{d}x + \int_{a_{\min}}^{a_{\max}} |g(x) - \hat{g}(x)|\, f_{t\to T}^Q(x)\,\mathrm{d}x \\ &+ \int_{a_{\max}}^\infty |g(x) - \hat{g}(x)|\, f_{t\to T}^Q(x)\,\mathrm{d}x \\ &\leq \mathbf{E}_t^Q \left[(|g(S_T)| + |\hat{g}(S_T)|)\,\mathbb{1}\left(S_T \notin A\right)\right] + 4\,\mathrm{dist}(g, \mathcal{F}_{2+n_k}) \\ &\leq \varepsilon + 4\,\mathrm{dist}(g, \mathcal{F}_{2+n_k}).\end{aligned}$$

Notice that $\mathbf{E}_t^Q S_T < \infty$ implies that $\mathbf{E}_t^Q |\hat{g}(S_T)| < \infty$, since $\hat{g}$ is a piecewise linear function of S_T, and therefore has at most linear growth. □

A.8 Proof of Proposition 7

Proof. Part (i) follows immediately from the continuous-state problem (8), as $\mathbb{1}\left(S_T \leq a_{\min}\right) \equiv 0$ and $\mathbb{1}\left(S_T \leq a_{\max}\right) \equiv 1$. Since the approximating function class contains the constant function, it follows that the solution to (8) in both cases is $\beta = 0$ and $[\beta_1, \beta_2, \ldots, \beta_{2+n_k}] = [1, 0, \ldots, 0]$ respectively.

Part (ii): We need to establish differentiability of $\hat{\beta}(x)$. The risk-neutral distribution can easily be derived from (7) and (10). In particular, from (7) we deduce that

$$\frac{\partial}{\partial x}\hat{\beta}(x) = \begin{bmatrix} \langle\phi_1, \phi_1\rangle & \cdots & \langle\phi_1, \phi_{2+n_k}\rangle \\ \vdots & \ddots & \vdots \\ \langle\phi_{2+n_k}, \phi_1\rangle & \cdots & \langle\phi_{2+n_k}, \phi_{2+n_k}\rangle \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ x \\ \vdots \\ \phi_j(x) \\ \vdots \\ \phi_{2+n_k}(x) \end{bmatrix}.$$

Each component of $\frac{\partial}{\partial x}\hat{\beta}(x)$ is therefore a piecewise linear function due to the structure of the basis functions. The final claim then follows because a linear combination of piecewise linear functions is piecewise linear.

Part (iii): By the Gram-Schmidt process, we can assume that $\{\phi_i\}_{i=1}^{2+n_k}$ is an orthonormal basis w.r.t. the inner product $\langle\phi_i, \phi_j\rangle = \int_A \phi_i(x)\phi_j(x)\,\mathrm{d}x$. This integral is finite because all basis functions are continuous and A is compact.

Hence, for $x \in A$ the risk-neutral CDF and PDF can be expressed as

$$\begin{aligned}\hat{F}^Q_{t\to T}(x) &= \sum_{j=1}^{2+n_k} \langle \mathbb{1}\left(S_T \le x\right), \phi_j(S_T)\rangle \, \mathbf{E}^Q_t \phi_j(S_T) \\ \hat{f}^Q_{t\to T}(x) &= \frac{\partial}{\partial x}\hat{F}^Q_{t\to T}(x) = \sum_{j=1}^{2+n_k} \phi_j(x)\mathbf{E}^Q_t \phi_j(S_T). \end{aligned} \tag{24}$$

Notice that $\mathbf{E}^Q_t \phi_j(S_T)$ is now a linear combination of put and call option prices due to the Gram-Schmidt process. It follows from (24) that

$$\begin{aligned}\int_A g(x)\, \mathrm{d}\hat{F}^Q_{t\to T}(x) &= \sum_{j=1}^{2+n_k} \mathbf{E}^Q_t\left[\phi_j(S_T)\right] \int_A g(x)\phi_j(x)\,\mathrm{d}x \\ &= \sum_{j=1}^{2+n_k} \mathbf{E}^Q_t\left[\phi_j(S_T)\right] \langle g, \phi_j\rangle \\ &= \mathbf{E}^Q_t \hat{g}(S_T).\end{aligned}$$

The last line follows because, under the Gram-Schmidt process, $\hat{\beta}_j$ from (7) equals $\langle g, \phi_j\rangle$ since $\langle \phi_i, \phi_j\rangle = \delta_{ij}$ by orthonormality. □

A.9 Proof of Proposition 8

Proof. To simplify notation in the proof, we let x denote stock 1 ($S_{1,T}$) and y denotes stock 2 ($S_{2,T}$). Similarly, the support of both stock will be denoted by the intervals $[x_1, x_n]$ and $[y_1, y_n]$. By a straightforward extension of Equation (8), $\hat{g}$ solves the approximation problem

$$\int_{x_1}^{x_n}\int_{y_1}^{y_n} \left(xy - \hat{g}(x,y)\right)^2 \mathrm{d}y\,\mathrm{d}x. \tag{25}$$

We first solve a simpler problem where the function xy is projected on

$$\hat{g}(x,y) = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 y.$$

The first order conditions for the (simplified) approximation problem (25) imply

$$\int_{x_1}^{x_n}\int_{y_1}^{y_n} xy - \hat{\beta}_0 - \hat{\beta}_1 x - \hat{\beta}_2 y \,\mathrm{d}y\,\mathrm{d}x = 0 \tag{26a}$$

$$\int_{x_1}^{x_n}\int_{y_1}^{y_n} x\left(xy - \hat{\beta}_0 - \hat{\beta}_1 x - \hat{\beta}_2 y\right)\mathrm{d}y\,\mathrm{d}x = 0 \tag{26b}$$

$$\int_{x_1}^{x_n}\int_{y_1}^{y_n} y\left(xy - \hat{\beta}_0 - \hat{\beta}_1 x - \hat{\beta}_2 y\right)\mathrm{d}y\,\mathrm{d}x = 0. \tag{26c}$$

Now define the constants

$$\begin{aligned}
\bar{x} &= \frac{1}{x_n - x_1}\int_{x_1}^{x_n} x\,\mathrm{d}x = (x_n + x_1)/2 = \mathbf{E}_t^Q S_{1,T} \\
\bar{y} &= \frac{1}{y_n - y_1}\int_{y_1}^{y_n} y\,\mathrm{d}y = (y_n + y_1)/2 = \mathbf{E}_t^Q S_{2,T} \\
\bar{x}\bar{y} &= \frac{1}{x_n - x_1}\frac{1}{y_n - y_1}\int_{x_1}^{x_n}\int_{y_1}^{y_n} xy\,\mathrm{d}y\,\mathrm{d}x
\end{aligned}$$

The fact that $\bar{x}$ and $\bar{y}$ are equal to the risk-neutral expectations of the first and second stock follows from the assumption. The first constraint in (26) forces

$$\hat{\beta}_0 = \bar{x}\bar{y} - \hat{\beta}_1\bar{x} - \hat{\beta}_2\bar{y}.$$

The second and third constraints can thus be expressed as

$$\begin{aligned}
\int_{x_1}^{x_n}\int_{y_1}^{y_n} (x - \bar{x})\left[xy - \bar{x}\bar{y} - \hat{\beta}_1(x - \bar{x}) - \hat{\beta}_2(y - \bar{y})\right]\mathrm{d}y\,\mathrm{d}x = 0 \\
\int_{x_1}^{x_n}\int_{y_1}^{y_n} (y - \bar{y})\left[xy - \bar{x}\bar{y} - \hat{\beta}_1(x - \bar{x}) - \hat{\beta}_2(y - \bar{y})\right]\mathrm{d}y\,\mathrm{d}x = 0.
\end{aligned}$$

From here, we readily obtain the solution

$$\hat{\beta}_1 = \frac{\int_{x_1}^{x_n}\int_{y_1}^{y_n}(x - \bar{x})(xy - \bar{x}\bar{y})\;\mathrm{d}y\,\mathrm{d}x}{\int_{x_1}^{x_n}\int_{y_1}^{y_n}(x - \bar{x})^2\;\mathrm{d}y\,\mathrm{d}x} = \bar{y} \tag{27a}$$

$$\hat{\beta}_2 = \frac{\int_{x_1}^{x_n}\int_{y_1}^{y_n}(y - \bar{y})(xy - \bar{x}\bar{y})\;\mathrm{d}y\,\mathrm{d}x}{\int_{x_1}^{x_n}\int_{y_1}^{y_n}(y - \bar{y})^2\;\mathrm{d}y\,\mathrm{d}x} = \bar{x} \tag{27b}$$

$$\hat{\beta}_0 = -\bar{x}\bar{y} \tag{27c}$$

Finally we verify that adding a put or call option basis function yields a coefficient of zero. To see this, without loss of generality, we focus on a basis function of the

form $(x-K)^+$. Using the first order conditions, it is sufficient to show that

$$\int_{x_1}^{x_n}\int_{y_1}^{y_n}(x-K)^+\left(xy-\hat{\beta}_0-\hat{\beta}_1x-\hat{\beta}_2y\right)\mathrm{d}y\,\mathrm{d}x=0,$$

where $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\beta}_2$ are given by (27a) – (27c). Notice that $xy-\hat{\beta}_0-\hat{\beta}_1x-\hat{\beta}_2y = (x-\bar{x})(y-\bar{y})$. So the integral can be written as

$$\int_{x_1}^{x_n}(x-K)^+(x-\bar{x})\,\mathrm{d}x\int_{y_1}^{y_n}y-\bar{y}\,\mathrm{d}y=0.$$

□

A.10 Sufficient conditions for ridge representation and the proof of Proposition 9

For completeness, we state the result of Lin and Pinkus (1993), giving necessary and sufficient conditions for ridge representation to hold. To state the result, some additional terminology is necessary. A polynomial $p(x_1,\ldots,x_d)$ can be associated to the differential operator $p(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_d})$. Let $P(a^1,\ldots,a^r)$ be the set of polynomials which vanish on all lines $\{\lambda a^i, \lambda\in\mathbb{R}\}$. Let Q be the set of polynomials $q(x_1,\ldots,x_d)$ such that $p(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_d})q=0$ for all $p\in P(a^1,\ldots,a^r)$.

Proposition 10 (Lin and Pinkus (1993))**.** *Let $a^1,\ldots,a^r$ be pairwise linearly independent vectors in $\mathbb{R}^d$. A function $g\in C(\mathbb{R}^d)$ can be expressed in the form*

$$g(x)=\sum_{i=1}^{r}g_i(a^i\cdot x)$$

if and only if g belongs to the closure of the linear span of Q.

In many practical situations, a more elementary argument suffices to show that a function cannot be written as a ridge combination with given directions a^i. For example, in the case $d=3$, the following reasoning shows that $g(x)=x_1(w'x)$ cannot be expressed as

$$g(x)=g_1(x_1)+g_2(x_2)+g_3(x_3)+g_4(w'x).$$

Suppose, by contradiction, that such a representation exists. Then, by differentiating twice, we have $\frac{\partial^2 g}{\partial x_2\partial x_3}=0$. However, $\frac{\partial^2 g_i}{\partial x_2\partial x_3}$ for $i=1,\ldots,3$, while $\frac{\partial^2 g_4}{\partial x_2\partial x_3}=w_2w_3g_4''(w'x)$. This implies that g_4 must be affine, but this cannot possi-

bly hold since $g(x)$ contains the cross terms x_1x_2 and x_1x_3. This proves Proposition 9 in case $d = 3$. Notice that we tacitly assume the most favorable scenario where options complete the market for each asset (e.g. using the same assumptions as in Proposition 2), so that each g_i can be estimated with arbitrary accuracy.

The argument generalizes directly to $d \geq 3$, thus showing that in higher dimensions it is not possible to perfectly estimate the risk-neutral covariance or correlation of sector i with the market portfolio.

A.11 Proof of Proposition 11

Proof. Let $\hat{\beta}$ denote the projection coefficients obtained from the quadratic projection. We need to show that

$$\int_A \left(x_i x_j - \hat{\beta}_0 - \sum_{r=1}^{d} \hat{\beta}_r x_r^2 - \hat{\beta}_M x_M^2 \right) x_k^n \, \mathrm{d}x = 0,$$

for odd $n \in \mathbb{N}$, and $k \in \{1, \ldots, d, M\}$. Because each x_r^n and x_M^n are symmetric around $R_{f,t\to T}$, it follows that $\hat{\beta}_0 \int_A x_k^n \, \mathrm{d}x = 0$ for $k = \{1, \ldots, d, M\}$.

Next suppose that $k \in \{1, \ldots, d\}$. Then

$$\int_A x_i x_j x_k^n \, \mathrm{d}x = 0. \tag{28}$$

This holds because the integral factors and it always contains an odd moment which vanishes. Using the same reasoning, it follows that

$$\int_A \sum_{r=1}^{d} \hat{\beta}_r x_r^2 x_k^n \, \mathrm{d}x = 0.$$

Now we handle the excess market return. Note that because $\sum_{r=1}^{d} w_r = 1$, it follows that

$$x_M^2 = \sum_{r=1}^{d} w_r^2 x_r^2 + 2 \sum_{1 \leq j_1 < j_2 \leq d} w_{j_1} w_{j_2} x_{j_1} x_{j_2}.$$

Then, using identical reasoning as before we get

$$\int_A x_M^2 x_k^n \, \mathrm{d}x = 0.$$

Suppose now that $k = M$ (the market return). Ordering the indices $i_1, \ldots, i_n$ as $j_1 < \cdots < j_m$ for some $1 \leq m \leq n$ with each j_r occurring with multiplicity a_r, we

then obtain that for $n \in \mathbb{N}$

$$\left(\sum_{i=1}^{d} w_i x_i\right)^n = \sum_{1 \le j_1 < \cdots < j_m \le d} c_{n,a_1,\ldots a_m} w_{j_1}^{a_1} x_{j_1}^{a_1} \ldots w_{j_m}^{a_m} x_{j_m}^{a_m} \tag{29}$$

where $1 \le m \le n$, $a_1, \ldots, a_m$ are positive integers adding up to n, and $c_{n,a_1,\ldots a_m}$ is the multinomial coefficient

$$c_{n,a_1,\ldots a_m} = \frac{n!}{a_1! \ldots a_m!}.$$

From the identity (29), it follows that for odd $n \ge 3$

$$\int_A x_i x_j x_M^n \, \mathrm{d}x = 0.$$

The identity holds by splitting cases. The only way for the integral to be non-zero is if the summand in (29) contains even powers of x_i and x_j. But if that is the case, then there must be at least one odd power of x_k for some $k \neq i, j$. As shown at the beginning of the proof, the integral of an odd power of x_k is zero.

Similar reasoning shows that

$$\int_A x_i^2 x_M^n \, \mathrm{d}x = 0, \tag{30}$$

because the only reason the integral cannot vanish is when (29) contains even powers of x_i. But then by implication there must be at least one odd moment of x_k in the product, whose integral vanishes. Because the overall integral factors as a product we conclude (30).

Finally, the fact that $\int_A x_M^2 x_M^n \, \mathrm{d}x = 0$ follows again because x_M^{n+2} is an odd function.

□

A.12 Proof of Proposition 12

Proof. We start from the identity

$$x_M^2 = \sum_{k=1}^{d} w_k^2 x_k^2 + 2 \sum_{1 \le i < j \le d} w_i w_j x_i x_j.$$

Because $\mathcal{F}$ contains the quadratic monomials, and because the projection operator $\widehat{\Pi}_{\mathcal{F}}$ is linear and idempotent, it follows that

$$x_M^2 = \sum_{k=1}^{d} w_k^2 x_k^2 + 2 \sum_{1 \le i < j \le d} w_i w_j \widehat{\Pi}_{\mathcal{F}}[x_i x_j].$$

Taking risk-neutral expectations on both sides then completes the proof. □

B Projection and equicorrelation

Given that vanilla options on the individual sectors and the market portfolio do not, in general, identify the full matrix of pairwise correlations, one must introduce additional structure. A common approach is to impose equicorrelation. We show that this equicorrelation estimator can be interpreted as a replicating portfolio, and then use projection to generalize it: the projection step chooses portfolio weights that are optimal (in an L^2 sense) for estimating heterogeneous covariances and correlations. In this section we assume that no dividends are paid, although it is straightforward to incorporate them at the cost of slightly heavier notation.[15]

The equicorrelation estimator of Engle and Kelly (2012) assumes that the correlation between any two assets is the same. In that case, the correlation estimate can be written as

$$\hat{\rho}_t = \frac{\mathbf{Var}_t^Q(R_{t\to T}) - \sum_{j=1}^{d} w_{j,t}^2 \mathbf{Var}_t^Q(R_{j,t\to T})}{2\sum_{1\le i<j\le d} w_{i,t} w_{j,t} \sqrt{\mathbf{Var}_t^Q(R_{i,t\to T})\mathbf{Var}_t^Q(R_{j,t\to T})}}.$$

This formula is also used by the CBOE to construct its implied correlation index. It is useful to reinterpret this as a portfolio replication problem. The target payoff is

$$\frac{(R_{i,t\to T} - R_{f,t\to T})(R_{j,t\to T} - R_{f,t\to T})}{\sqrt{\mathbf{Var}_t^Q(R_{i,t\to T})\mathbf{Var}_t^Q(R_{j,t\to T})}},$$

and the basis functions are the quadratic payoffs

$$(R_{t\to T} - R_{f,t\to T})^2 \quad \text{and} \quad (R_{j,t\to T} - R_{f,t\to T})^2, \quad \text{for } j = 1, \dots, d$$

Viewed this way, the replicating portfolio is the same for all $i \neq j$, with weights proportional to a weighted average of sector-specific standard deviations.

[15] Under this assumption, $\mathbf{E}_t^Q R_{i,t\to T} = R_{f,t\to T}$. If we include dividends, then $\mathbf{E}_t^Q R_{i,t\to T} = F_{i,t\to T}/S_t$.

The projection approach allows us to optimize and generalize these features. For shorthand, let $x_k := R_{k,t\to T} - R_{f,t\to T}$ and $x_M := R_{t\to T} - R_{f,t\to T}$ denote the excess returns on asset k and on the market, respectively. Let $x = [x_1, \ldots, x_d]'$, so that $x_M = w \cdot x$, where w is the vector of market weights. To generalize the equicorrelation estimator, we seek the *optimal* replicating portfolio for $x_i x_j$, which directly targets the risk-neutral covariance between returns i and j.[16]

First, consider the continuous-state analogue. Let $A = A_1 \times \cdots \times A_d \subset \mathbb{R}^d$ be compact. We seek univariate functions $g_1, \ldots, g_d, g_M \in C(\mathbb{R})$ that minimize

$$\int_A \Big(x_i x_j - \sum_{k=1}^{d} g_k(x_k) - g_M(x_M) \Big)^2 \, \mathrm{d}x,$$

where $x_M = w'x$. Rather than solving this infinite-dimensional problem directly, we approximate it by restricting attention to low-degree polynomial payoffs. This is motivated by two considerations: (i) polynomials are dense in $C(A)$ (Stone–Weierstrass); and (ii) higher-order risk-neutral moments are empirically difficult to estimate. The following result implies that we can restrict attention to quadratic and quartic terms, because the coefficients on odd moments are zero.

Proposition 11 (Odd-moment orthogonality)**.** *Fix $i \neq j$. Let $\mathcal{F} = \{1, x_1^2, \ldots, x_d^2, x_M^2\}$, and let $\widehat{\Pi}_{\mathcal{F}}[x_i x_j]$ be the L^2-projection onto $\mathcal{F}$ under the inner product $\langle f, g \rangle = \int_A f(x) g(x) \, \mathrm{d}x$, where $A = A_1 \times \ldots A_d$, and $A_i = [a^i_{\min}, a^i_{\max}]$ is symmetric around 0. Define the residual function by*

$$\hat{\varepsilon}_{ij} = x_i x_j - \widehat{\Pi}_{\mathcal{F}}[x_i x_j].$$

Then for every odd integer $n \geq 1$,

$$\langle \hat{\varepsilon}_{ij}, x_k^n \rangle = 0 \quad \textit{for all } k \in \{1, \ldots, d, M\}.$$

Remark 6. In practice, the interval for each excess return will typically not have 0 as midpoint, because options data are skewed and there tends to be more information going further in the left-tail. Nevertheless, the midpoint of each interval will be close to 0, and we find in simulation that the projection coefficients of odd moments still tend to be negligible in that case.

In contrast to odd-moments, the projection coefficients of even degree will generally not vanish, and including these monomials will generally decrease the

[16] Working with covariance instead of correlation involves no loss of generality, since the equicorrelation estimator maps directly to a replicating portfolio for $x_i x_j$.

approximation error. Compared to the equicorrelation estimator, we thus gain generality in that we incorporate not only variance but also the 4th moment (a measure of tail-thickness), and the portfolio weights are allowed to differ for each pair of assets, thus allowing to estimate the correlation between an arbitrary pair of assets, instead of assuming all correlations are the same.

Projecting $x_i x_j$ on the subspace

$$\mathcal{F} = \left\{1, x_1^2, \ldots, x_d^2, x_1^4, \ldots, x_d^4, x_M^2, x_M^4\right\} \tag{31}$$

also circumvents the computational burden of projecting $x_i x_j$ directly onto the full set of option payoffs. The latter would require minimizing an objective that depends on an 11-dimensional integral and a large number of parameters. A discretized OLS approach is likewise infeasible: with 1000 grid points per return, the state grid would contain 1000^{11} rows.

Instead, we first project $x_i x_j$ onto $\mathcal{F}$ and then project each resulting power payoff onto the corresponding univariate option basis. This two-step procedure yields the same result as projecting directly onto the smallest subspace, because for orthogonal projections onto nested subspaces one has $\Pi_F\, g = \Pi_F\, \Pi_G\, g$ whenever $F \subseteq G$ (with respect to the same inner product).

Moreover, the projection of $x_i x_j$ onto $\mathcal{F}$ can be derived in closed form, and the subsequent projection of a monomial such as x_k^2 onto option payoffs that depend only on asset k is a one-dimensional problem, which can be solved using the method in Section 2.3. Based on the projection coefficient on the subspace in (31), we define an estimator of the covariance by

$$\begin{aligned}\widehat{\mathbf{Cov}}^{Q}_{ij,t} \coloneqq \mathbf{E}_t^Q \widehat{\Pi}_{\mathcal{F}}[x_i x_j] = \hat{\beta}_{0,ij} + \sum_{k=1}^{d} \Big[\hat{\beta}_{k,ij} \mathbf{Var}_t^Q(R_{k,t\to T}) + \hat{\gamma}_{k,ij} \mathbf{E}_t^Q \left(R_{k,t\to T} - R_{f,t\to T}\right)^4\Big] \\ + \hat{\beta}_{M,ij} \mathbf{Var}_t^Q R_{t\to T} + \hat{\gamma}_{M,ij} \mathbf{E}_t^Q \left(R_{t\to T} - R_{f,t\to T}\right)^4 .\end{aligned} \tag{32}$$

Because we can identify the risk-neutral variance, for consistency, it is desirable that the covariance estimator satisfies

$$\mathbf{Var}_t^Q R_{t\to T} = \sum_{i=1}^{d} w_i^2 \mathbf{Var}_t^Q R_{i,t\to T} + 2 \sum_{1\le i<j\le d} w_i w_j \widehat{\mathbf{Cov}}^{Q}_{ij,t}. \tag{33}$$

The next proposition shows that the addition formula holds whenever the projection space contains all univariate quadratic terms.

Proposition 12. *Let $\mathcal{F}$ be a function space such that $\{x_1^2, \ldots, x_d^2, x_M^2\} \subset \mathcal{F}$.*

Define the covariance estimator based on $\mathcal{F}$ by

$$\widehat{\mathbf{Cov}}^{Q}_{ij,t} = \mathbf{E}^{Q}_{t}\widehat{\Pi}_{\mathcal{F}}[x_i x_j].$$

Then, (33) *holds.*

Remark 7. Motivated by the empirical setting, the results above extend to the case with multiple index portfolios. Suppose there are two index returns $x_{M,1} = w_1 \cdot x$ and $x_{M,2} = w_2 \cdot x$ with corresponding options. Then Proposition 11 holds verbatim for each index. Likewise, Proposition 12 holds simultaneously provided the projection space contains all univariate quadratic terms, including $x^2_{M,1}$ and $x^2_{M,2}$; under this condition, (33) holds for each weight vector w_ℓ ($\ell = 1, 2$), with the variance on the left-hand side taken for the corresponding portfolio.

B.1 Simulation evidence for sector ETFs

We now evaluate the efficacy of the covariance estimator in (32) for the eleven sector ETFs using a simple factor structure under the risk-neutral measure. Let $X \in \mathbb{R}^{11}$ denote log-returns and $R = \exp(X)$ the corresponding gross returns. We simulate

$$X = Bf + \varepsilon, \qquad f \sim \mathsf{N}\big(0, \operatorname{diag}(\sigma^2_{1,f}, \sigma^2_{2,f})\big), \quad \varepsilon \sim \mathsf{N}\big(0, \operatorname{diag}(\sigma^2_1, \ldots, \sigma^2_{11})\big), \quad f \perp \varepsilon,$$

with $B \in \mathbb{R}^{11\times 2}$. Hence

$$\mathbf{Var}^{Q}_{t}(X) \;=\; B\,\operatorname{diag}(\sigma^2_1, \sigma^2_2)\,B' \;+\; \operatorname{diag}(\sigma^2_1, \ldots, \sigma^2_{11}).$$

The factor structure captures systematic risk and cross-sectional correlation. We set the gross-return means to one and winsorize R at $[0.4,\ 1.5]$ componentwise. Entries of B are drawn IID from $\mathsf{Unif}[-0.4, 1]$.

We run 1,000 Monte Carlo simulations. In each run we compute the mean squared error (MSE) between the vector of true pairwise correlations and the estimated correlations. As a benchmark, we include the equicorrelation estimator. Table 2 reports summary statistics: the projection-based estimator attains lower MSE across the distribution. We also report the correlation between the true correlation vector and the projection-based estimate within each run; the average is about 20%, indicating that the projection approach captures meaningful cross-sectional heterogeneity. By construction, the equicorrelation estimator does not capture such heterogeneity, as it imposes a common correlation across all pairs.

	Min	Median	Max	Mean	Std. dev.
Equicorrelation	0.0231	0.1361	0.3860	0.1436	0.0472
Projection correlation	0.0253	0.1259	0.3312	0.1284	0.0408

Table 2: **Summary statistics of MSE**. This table summarizes, across 1,000 Monte Carlo simulations, the distribution of the MSE for the equicorrelation estimator and the projection-based correlation estimator in (32).

C Empirical estimates of SVIX and VIX

According to the simulation results, the projection approach compares favorably to the CM formula especially when the number of observed option prices is small. When the number of observed options is large it is a priori not so clear whether a more refined approximation yields economically different results. To investigate the benefits of the projection approach in the latter case, we estimate the SVIX and VIX from Examples 1–2 using both methods. The calculation of both indexes requires options on the S&P500, which is one of the most liquid option markets worldwide. The SVIX and VIX thus stand a natural test case.

The options data on the SP500 are coming from OptionMetrics and span the period January 4, 1996 until July 20, 2023. Several data cleaning procedures are applied before each volatility index is calculated. The procedure is almost identical to CBOE's method when it calculates the VIX. A detailed description of our procedure is included in Appendix D.

First, consider the SVIX defined by

$$\text{SVIX}^2_{t\to T} = \frac{1}{T-t}\mathbf{Var}^Q_t\left(\frac{R_{t\to T}}{R_{f,t\to T}}\right). \tag{34}$$

Martin (2017) derives conditions under which the conditional equity premium satisfies

$$\frac{1}{T-t}\left(\mathbf{E}_t R_{t\to T} - R_{f,t\to T}\right) \geq R_{f,t\to T}\text{SVIX}^2_{t\to T}.$$

In fact, when running the regression

$$\frac{1}{T-t}\left(\mathbf{E}_t R_{t\to T} - R_{f,t\to T}\right) = \beta_0 + \beta_1 R_{f,t\to T}\text{SVIX}^2_{t\to T} + \varepsilon_T, \tag{35}$$

Martin (2017, 2025) cannot reject the null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$, thus suggesting that the lower bound is tight. This conclusion is particularly interesting as it gives a model-free way to measure the equity premium in real time. Given

its importance, we reassess this claim by using our projection method to measure $\text{SVIX}^2_{t\to T}$. Table 3 shows the results. For each prediction horizon, the difference between the CM and projection method are very small, suggesting that in very liquid option markets it is immaterial which method is used.

	30 days		90 days		180 days	
	Projection	CM	Projection	CM	Projection	CM
β_0	0.002	0.005	−0.002	−0.005	−0.046	−0.052
	(0.0407)	(0.0400)	(0.0512)	(0.0504)	(0.0361)	(0.0365)
β_1	1.434	1.493	1.395	1.589	2.455	2.865
	(1.0160)	(1.0816)	(1.2693)	(1.3602)	(0.7914)	(0.8371)
R^2 (%)	1.12	1.08	2.09	2.35	6.91	7.94
# obs	6932	6932	6865	6865	6745	6745

Table 3: **Equity premium regression.** This table reports estimates from regression (35) for return horizons of 30, 90, and 180 days. Newey–West standard errors, using a bandwidth equal to the number of trading days in the horizon, are reported in parentheses below the coefficients.

In addition to SVIX, we also estimate the VIX. Figure 6 plots the time series of the difference between the two VIX estimates; the solid orange line is its 60-day moving average, which remains positive throughout, consistent with the simulation. The largest gaps occur early in the sample when option coverage is sparser. We mark the 20 largest differences with blue dots, which can reach close to 8 percentage points. Such a gap is economically significant: portfolios with hundreds of VIX futures contracts can experience multi-million-dollar P&L swings. The single largest peak occurs on March 2, 2009, at the height of the global financial crisis. On that day, the projection-implied VIX is 52%, while the CM approximation yields 44%. During periods of heightened uncertainty, risk-neutral mass shifts to the left tail, which amplifies entropy because $\log(x)$ decays steeply near zero (see (4)). In such episodes the CM method—linearized around the risk-free rate—can be inaccurate, whereas the projection method remains reliable because it approximates $\log(x)$ well over the entire domain. In line with this intuition, the largest measurement differences cluster around the dot-com bust (2000), the global financial crisis (2008), and COVID-19 (2020).

D Option data preprocessing

We use SP500 option data from OptionMetrics, covering the period January 4, 1996 to July 20, 2023. Following the CBOE procedure, we discard all in-the-money

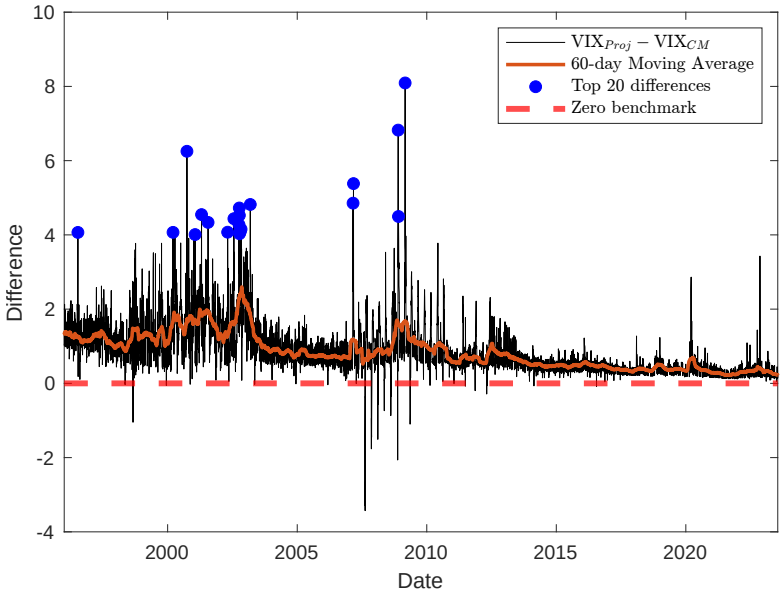


Figure 6: **VIX estimate**. This figure shows the projection VIX estimate minus the VIX estimate obtained by CM. The solid orange line denotes the 60-day moving average of this difference. The blue dots indicate the 20 largest observed differences.

put and call options, as well as any option with a bid price of zero. When there are two consecutive strikes with a bid price equal to zero, all options with higher strikes (for calls) or lower strikes (for puts) are discarded. For each remaining option, the price is defined as the average of the bid and ask prices. In total, this filtering yields 11.738 million option prices. The risk-free rate for each return horizon is obtained from the zero-coupon yield curve dataset provided by OptionMetrics.

D.1 ETF options and conversion of American option price

Options on SPY, XLK, and SPXT are recorded as American in OptionMetrics. To estimate the risk-neutral volatility, we first convert these quotes to European-equivalent prices. For each option we compute the Black–Scholes price using the forward price and implied volatility reported by OptionMetrics; this conversion accounts for dividends via the forward.[17]

After this conversion, our preprocessing for SPY is identical to Section D. For XLK and SPXT, by contrast, in-the-money options are often liquid, so we retain both in- and out-of-the-money quotes. Furthermore, we discard only options with zero bid prices, rather than also truncating the strike range after two consecutively

[17] As in Martin (2017) and Kremens and Martin (2019), we assume dividends are known in advance and paid at time T.

observed zero-bid options.

E Details on simulation

In the Monte-Carlo simulation, we use two different models to generate option prices. In both cases the time to maturity is 1 year. The first model is the standard Black and Scholes (1973) model with a risk-free rate of 5% and volatility of 20%. The simulation of the stochastic volatility jump (SVCJ) model is based on Eraker et al. (2003). In their setup, the log asset price follows

$$\begin{pmatrix} \mathrm{d}\log S_t \\ \mathrm{d}V_t \end{pmatrix} = \begin{pmatrix} \mu \\ \kappa\left(\theta - V_{t-}\right) \end{pmatrix} \mathrm{d}t + \sqrt{V_{t-}} \begin{pmatrix} 1 & 0 \\ \rho\sigma_v & \sqrt{1-\rho^2}\sigma_v \end{pmatrix} \mathrm{d}W_t + \begin{pmatrix} \xi^y \\ \xi^v \end{pmatrix} \mathrm{d}N_t,$$

where $V_{t-} = \lim_{s\uparrow t} V_s$ denotes the left limit, W_t is a standard two-dimensional Brownian motion, N_t is a Poisson process with intensity λ, and ξ^y, ξ^v are the jump sizes in returns and volatility. These jump sizes are correlated and have distributions $\xi^v \sim \exp(\mu_v)$ and $\xi^y|\xi^v \sim \mathsf{N}\left(\mu_y + \rho_J\xi^v, \sigma_y^2\right)$. For simulation, we only need to calibrate the model under the risk-neutral measure. The risk-neutral parameters are taken from Broadie et al. (2007) and are summarized in Table 4.

Parameter	Value
κ	0.0570
θ	0.0062
ρ	−0.4838
σ_v	0.0800
μ_v	0.2213
μ_y	−0.0539
ρ_J	0.0000
σ_y	0.0578
λ	1.5120
r	0.0500

Table 4: **SVCJ model calibration**