

Updating Density Estimates using Conditional Information Projection: Stock Index Returns and Stochastic Dominance Relations

Stelios Arvanitis^{}, Richard McGee[†], and Thierry Post[‡]*

Abstract

We propose, analyze, and apply a conditional density estimator that projects a prior time-series estimator onto the set of distributions satisfying given conditional moment conditions with a functional nuisance parameter. Theoretically, we show that the estimator reduces the limiting divergence from the latent population density when the prior is inconsistent and the moment conditions are correctly specified. A Monte Carlo simulation and an empirical application to stock index returns illustrate the method, using moment conditions derived from option pricing restrictions and treating the stochastic discount factor as a nuisance parameter. The simulations demonstrate substantial improvements in density estimation under realistic yet controlled conditions. Empirically, the estimator enhances monthly S&P 500 index return density forecasts in likelihood ratio tests and improves the out-of-sample investment performance of option-enhanced portfolios by refining the timing and composition of one-month S&P 500 index option combinations.

Keywords: Conditional density estimation, information projection, Stochastic Dominance, stock index returns, stock index options

1 Introduction

Financial decision making in risk management, derivative pricing, and portfolio optimization often requires conditional density estimates of asset returns. Approaches

^{*}Queen's University, Canada, and Athens University of Economics and Business, Greece. Email: stelios@aueb.gr

[†]University College Dublin, Ireland. Email: richard.mcgee@ucd.ie

[‡]Graduate School of Business of Nazarbayev University, Kazakhstan. E-mail: thierry-post@hotmail.com

based on Expected Utility, Stochastic Dominance (SD), or Mean-Lower Partial Moment orders rely on a full probabilistic view that accounts for asymmetry and tail risk beyond point forecasts and variances. In addition, the return distribution should be updated to reflect prevailing market conditions. Common sources of conditioning information include market prices and investment yields of securities, and theoretical pricing restrictions.

A key challenge is incorporating new information to refine an existing density estimate. A large forecasting literature estimates conditional return distributions using GARCH-type models, quantile regressions, and nonparametric methods. These approaches rely primarily on historical data and often do not explicitly integrate real-time market conditions or theoretical pricing restrictions. How can an existing density estimate be updated to incorporate the available conditioning information?

We propose a new estimator for this task: the Conditional Information Projection Density Estimator (CIPDE). It starts with a potentially inconsistent prior estimate of the conditional return distribution and updates it to satisfy conditional moment conditions via information projection (see for example ([Csiszár, 1975](#); [Schennach, 2005](#); [Komunjer and Ragusa, 2016](#))) to minimize the divergence from the prior. This results in a new density estimate that integrates information from both the prior estimate and the moment conditions. To acknowledge that side information is generally incomplete, the conditional moment conditions may take the form of inequalities and include functional nuisance parameters such as the ubiquitous Stochastic Discount Factor (SDF) in Finance.

While the CIPDE estimator is valid for multivariate densities, our primary focus is on forecasting univariate conditional densities, where the curse of dimensionality is more manageable and the projection-based methods show tangible improvements in finite samples.

Although the CIPDE estimator is not derived from a likelihood, it admits a Bayesian interpretation: the prior density estimate is updated via an information projection that incorporates moment conditions as constraints. This can be seen as a generalized Bayesian update that replaces the likelihood with a set of economically motivated admissible distributions. The resulting posterior minimizes Kullback–Leibler divergence from the prior while satisfying the conditioning information, echoing principles of robust and entropic Bayesian inference under partial specification—see [Jewson et al. \(2018\)](#).

In our empirical analysis, we anticipate that the initial density estimator may be statistically inconsistent (and thus may asymptotically violate moment conditions) because it is constructed under assumptions of debatable distributional shapes or erroneous conditioning information.

We establish a statistical theory for the CIPDE, including pseudo-consistency, convergence rates, and the limiting distribution, using a combination of information-theoretic and variational techniques. Among other things, the theory shows that CIPDE asymptotically reduces KL divergence to the latent true conditional distribution under general regularity conditions.

A Monte Carlo simulation experiment and an empirical analysis apply CIPDE to stock index returns using conditional moment restrictions based on market prices and pricing restrictions for index options. The experiment foreshadows the empirical analysis by density forecasting based on repeated random samples drawn from a simplified but known simulation process. The empirical analysis focuses on density forecasting for monthly S&P 500 index (SPX) returns and optimization of combinations of one-month Chicago Board Options Exchange (CBOE) SPX options given the posterior density estimate.

For the pricing and trading of options, it is essential to account for the conditional and non-Gaussian nature of the return distribution. Most liquid index options have a short maturity of several days or weeks and their valuation and risk assessment critically depends on up-to-date market information. Since option payoffs are inherently asymmetric, mean-variance analysis is not appropriate and higher-order risk needs to be taken into account. Encouragingly, payoffs at expiry of concurrent option series are driven solely by the value of the index at the option expiration date, so that the estimation can focus the one-dimensional density of SPX returns.

The conditional moment conditions in the simulation experiment and empirical analysis are based on observed market price quotes for index options and general option pricing restrictions that exclude SD relations in the spirit of [Constantinides et al. \(2009\)](#). In the simulation experiment, these conditions are known to be true for a subset of all options; in the empirical analysis, they are assumed to be true for the options that are selected based on a number of option characteristics (maturity, moneyness, delta and premium).

The optimization of option combinations is inspired by [Constantinides et al. \(2020\)](#); [Post and Longarela \(2021\)](#); [Beare et al. \(2025\)](#), but seeks to improve the density forecast of index returns that is plugged into the solver. The earlier studies employed conditional density estimates that are statistically inconsistent because they fix or restrict the variation of the shape of the distribution and its parameters and use a limited set of conditioning information (such as the risk-free rate and a market volatility index). In the present study, these estimates serve as the initial estimator, and we attempt to correct their biases using the conditional moment conditions.

The simulations demonstrate significant potential improvements in density estimates through information projection, particularly when the prior is biased or

imprecise. The reduction in relative entropy increases with the number of qualified options and decreases in the bid-ask spread, highlighting the crucial role of conditioning information.

In the empirical analysis of SPX returns and options, CIPDE significantly improves the density forecasts out-of-sample (OOS) for a range of prior density forecasts. In optimizing index option combinations, the use of CIPDE yields significant OOS performance improvements compared with the use of the prior or a risk-neutral posterior density forecast, accounting for quoted bid-ask spreads using a simple buy-and-hold strategy.

Apart from the link with the aforementioned studies on the pricing and trading of index options using SD, our study is related to several earlier applications of EL and Bayesian econometrics in asset pricing and portfolio optimization.

Most notably, [Stutzer \(1996\)](#) applied Bayesian updating with moment conditions to option prices for option valuation. However, that study focuses on estimating the risk-neutral distribution rather than the physical (or real-world) distribution. The risk-neutral distribution is a biased estimator of the physical distribution because it conflates the physical distribution with the SDF. Although the risk-neutral distribution can be used for valuing options, it is less suitable for real-world decision making and utility-based portfolio optimization due to its biased nature. Furthermore, [Stutzer \(1996\)](#) does not provide a formal statistical theory for the estimated risk neutral distribution, instead focusing on practical application rather than developing a comprehensive statistical framework for the distribution itself.

In a related study, [Stutzer \(1995\)](#) pioneers the application of Bayesian econometrics for model selection and validation in empirical asset pricing. Similarly, [Almeida and Garcia \(2012\)](#) and [Post and Potì \(2017\)](#) develop and apply asset pricing tests based on general Minimum Discrepancy estimation and a combination of EL and SD, respectively. These studies however do not focus on density estimation and its application in forecasting and optimization.

[Post et al. \(2018\)](#) developed statistical theory and numerical algorithms for portfolio optimization based on EL and applied them to an equity industry allocation problem. The present study focuses on information projection instead of moment projection, considers conditional moment conditions, and it account for potentially inconsistent priors and the role of functional nuisance parameters. In addition, the index option combination problem seems more suitable for a model-free approach than the multi-asset allocation problem, due to its lower dimensionality (one underlying stock index vs. a multitude of industries).

The remainder of the paper is structured as follows. Section 2 introduces the formal framework of the CIPDE, presents its variational representation, and establishes

the statistical properties of the estimator. Section 3 introduces the principles of option pricing and option trading that will be used in our simulation study and empirical study. Section 4 conducts Monte Carlo experiments, and Section 5 applies the methodology to the estimation of the conditional distribution of monthly S&P 500 index returns and optimization of SPX option combinations. Section 5 concludes.

2 Statistical Theory

2.1 Preliminaries

We aim to estimate a latent conditional distribution F , using (i) an initial estimator $\hat{F}$ (the prior) and (ii) a set of J valid conditional moment inequalities. The proposed estimator belongs to the class of information projection-based methods, relying on the minimization of the KL divergence subject to the moment constraints. Unlike classical Bayesian updating which uses a likelihood function, this procedure integrates information through conditional moment restrictions.

The framework admits the possibility that the distribution depends on conditioning information summarized by observed and/or latent variables (e.g., covariates). While those are not explicitly modeled here, all results are to be understood as applying conditionally, for almost every fixed value of the conditioning variables.

The support $\mathcal{S} \subseteq \mathbb{R}^p$ is assumed independent of the conditioning variables. In our simulation experiment and empirical analysis, a univariate distribution ($p = 1$) is used, but the method also applies to multivariate cases ($p \geq 2$), although the curse of dimensionality presumably makes direct applicability in high-dimensional applications elusive; see the Conclusions.

This curse arises because in high dimensions, the linear growth of the available data is not usually able to match the exponentially growing complexity of the feasible set, making it increasingly difficult to accurately estimate the high-dimensional conditional density or ensure feasibility of the moment inequalities without substantial regularization or structural assumptions.

The prior $\hat{F}$, which may be a kernel density estimator or a parametric model-based estimate, is assumed to be absolutely continuous with respect to the Lebesgue measure, almost surely. This means that $\hat{F}$ admits a density function $\hat{f}$ on its support. This assumption is standard in nonparametric and semiparametric density estimation and holds for most practical estimators used in empirical work. It ensures that the KL divergence between $\hat{F}$ and other distributions with densities is well-defined. $\hat{F}$ is allowed to be an inconsistent estimator of F .

The conditional moment conditions include a functional parameter $m \in \mathcal{M}$, where

$\mathcal{M}$ is a convex and closed subset of a Sobolev space of smooth functions. More precisely, $\mathcal{M} \subset W^{2,2}(\mathcal{S})$, the space of square-integrable functions on $\mathcal{S}$ whose first and second weak derivatives are also square-integrable; see [Adams and Fournier \(2003\)](#) for a formal definition. In our simulation experiment and empirical analysis, the functional parameter is an SDF. The moment conditions take the form of inequalities:

$$\int_{\mathcal{S}} \mathbf{H}(S)m(S)g(S)dS \geq \mathbf{0},$$

where g is the density of a candidate conditional distribution G , and $\mathbf{H} : \mathcal{S} \rightarrow \mathbb{R}^J$ is a known vector-valued function.

The convex set of admissible densities that satisfy the moment inequalities at the functional parameter m is denoted by

$$\mathbb{M}(\mathbf{H}, \hat{F}; m) := \left\{ G \ll \hat{F} : \int_{\mathcal{S}} \mathbf{H}(S)m(S)g(S)dS \geq \mathbf{0} \right\},$$

where $\ll$ denotes absolute continuity.

The CIPDE estimator $\hat{G}$ is introduced in the next subsection. Its asymptotic properties as a sample size $T \rightarrow \infty$ are investigated in the subsequent subsections. The analysis builds upon variational representations, convexity properties, and information-theoretic tools to derive results concerning pseudo-consistency, convergence rates, robustness to prior inconsistency, and inference validity for SD comparisons.

In the following, convergence of densities is mostly assumed to be taking place in $\ell^1(S)$; the space of (equivalence classes of) integrable real functions on S equipped with the usual integral of absolute value norm. The symbol $\rightsquigarrow$ denotes weak convergence, and $\overset{\text{epi}}{\rightsquigarrow}$ denotes epi-convergence in distribution; see [Knight \(1999\)](#).

2.2 The CIPDE Estimator

To define the CIPDE estimator we use a variational problem formulation of the traditional constrained minimization of the relative entropy:

$$\hat{G} \in \arg \min_G \min_{m \in \mathcal{M}} \left\{ \text{KL}(G \| \hat{F}) + \chi_{\mathbb{M}(\mathbf{H}, \hat{F}; m)}(G) \right\}, \quad (1)$$

where $\text{KL}(G \| \hat{F})$ is the Kullback-Leibler divergence from G to $\hat{F}$ and $\chi_{\mathbb{M}}$ is the characteristic function of the admissible set, taking the value 0 for admissible and $+\infty$ for inadmissible. The estimation involves joint optimization over the density G and the functional parameter m .

Using duality and convex optimization arguments (e.g., [Feydy et al. \(2019\)](#), [Yilmaz \(2021\)](#)), the variational problem can be equivalently written in a saddle-point form:

$$\inf_{m \in \mathcal{M}} \inf_G \sup_{\lambda \leq 0} \left\{ \text{KL}(G \| \hat{F}) + \boldsymbol{\lambda}^\top \int_S \mathbf{H}(S) m(S) g(S) dS - \mu \left(\int_S g(S) dS - 1 \right) \right\}, \quad (2)$$

with Lagrange multipliers $\boldsymbol{\lambda} \in \mathbb{R}_-^J$ and $\mu \in \mathbb{R}$ enforcing the moment and normalization constraints, respectively. Solving the inner problem yields a closed-form expression for the optimal density $\hat{g}_m$ (for a fixed m) in the form of a Gibbs posterior:

$$\hat{g}_m(S) = \frac{\hat{f}(S) \exp(\boldsymbol{\lambda}_T^\top \mathbf{H}(S) m(S))}{\int_S \hat{f}(S) \exp(\boldsymbol{\lambda}_T^\top \mathbf{H}(S) m(S)) dS}, \quad (3)$$

where $\boldsymbol{\lambda}_T$ is the empirical Lagrange multiplier vector satisfying feasibility and complementary slackness. The estimator $\hat{G}$ is then given by:

$$\hat{G}(s) := \int_S 1_{\{\mathbf{x} \leq s\}}(\mathbf{x}) \hat{g}_{m_T}(\mathbf{x}) d\mathbf{x},$$

where m_T is the minimizer over $\mathcal{M}$.

The CIPDE can be viewed as a generalized Bayesian estimator: Moment inequalities replace the likelihood in Bayesian updating. The KL divergence plays the role of a regularizer that ensures proximity to the prior $\hat{F}$, given that the solution respects the moment conditions. The additional optimization over the functional parameter m creates further deviation from the classical Bayesian setting; it provides leeway in the projection procedure to choose optimally between the projection sets indexed by the functional parameter.

If an additional prior distribution were available over the functional space $\mathcal{M}$ —see, for example, [Li and Zhao \(2002\)](#), then a density estimator could also be definable by integrating $\hat{g}_m$ w.r.t. the m -values for which $\mathbb{M}(\mathbf{H}, \hat{F}; m) \neq \emptyset$, via the aforementioned prior that must be supported on them. Such a procedure could then be characterized as Bayesian information projection averaging, contrasting the current one where the information projection additionally involves optimal selection of the parameter.

2.3 Pseudo-Consistency

We begin the investigation of the asymptotic properties, by establishing the pseudo-consistency of the CIPDE estimator; high-level assumptions about the convergence of the initial estimator $\hat{F}$ and the asymptotic identification of the optimal functional parameter are introduced.

Assumption 1 (Prior Convergence-Exponential Moments). *There exists a density f_∞ , such that $\hat{f} \rightsquigarrow f_\infty$ in $\ell^1(S)$. Furthermore, $\int_S \exp(\mathbf{t}^T \mathbf{H}(s)m(s)f_\infty(s)ds < +\infty$ for some $\mathbf{t} > \mathbf{0}$, $m \in \mathcal{M}$.*

This first part of the assumption ensures that the initial prior density estimator converges in probability to a deterministic limit f_∞ inside $\ell^1(S)$. The limit may differ from the true DGP density f . $F_\infty(s) := \int_S 1_{\{\mathbf{x} \leq s\}}(\mathbf{x})f_\infty(\mathbf{x})d\mathbf{x}$ need not equal the DGP cdf F . In many applications, it can be expected that $F_\infty \notin \mathbb{M}(\mathbf{H}, F_\infty; m)$ for all $m \in \mathcal{M}$ for a set of values for the conditioning variables of positive probability, even though $F \in \mathbb{M}(\mathbf{H}, F_\infty; m)$ for some SDF and that the DGP satisfies strictly the moment inequalities, because the estimator may use a counterfactual distribution shape or does not correctly take into account the conditioning information or it is incorrectly smoothed, etc.

In parametric models, the ℓ^1 convergence follows if the likelihood has a bounded derivative w.r.t. the parameter, and the parametric estimator has a unique pseudo-true value at which it converges weakly; see Blasques et al. (2018) for examples in the context of GARCH-type models. In nonparametric kernel-based density estimation, such convergence can be established under standard bandwidth and smoothness assumptions; see, for example, Tsybakov (2009). In semiparametric settings, such as sieve maximum likelihood estimation or partially linear models, uniform convergence results for $\hat{f}$ are also available when the parametric component is estimated at $\sqrt{n}$ -rate and the nonparametric correction satisfies a Donsker-type condition; see, for example, Blundell et al. (2012). Thus, the first part of Assumption 1 encompasses a wide range of estimation frameworks relevant to practitioners.

The second part of Assumption 1 requires the existence of some f_∞ -exponential moments for the product $\mathbf{H}m$, for almost every value of the conditioning variables. It is a strong condition that is, however, weaker than support boundedness. Along with the first part of the assumption, it implies the weak epi-continuity-see Knight (1999)-of the functionals $g \rightarrow \mathbb{M}(\mathbf{H}, g; m)$ for any g that converges to f_∞ in $\ell^1(S)$.

The second assumption is useful for asymptotic identification.

Assumption 2 (Functional Parameter Identification). *If $f_\infty = f$, then there exists some $m_0 \in \mathcal{M}$, such that $\int_S \mathbf{H}(S)m_0(S)f(S)dS > \mathbf{0}$. Otherwise, the following hold:*

1. *The set $M^U := \cup_{m \in \mathcal{M}} \mathbb{M}(\mathbf{H}, F_\infty, m)$ is convex.*
2. *The mapping $m \mapsto G_m$, where $G_m := \arg \min_{G \in \mathbb{M}(\mathbf{H}, F_\infty; m)} \text{KL}(G \| F_\infty)$, is injective on M^U .*

The first part of Assumption 2 requires the existence of a functional parameter value w.r.t. which the moment conditions are non-binding for the limiting f_∞

when this is actually the DGP. This is a mild restriction given that $\mathbf{H}$ can be always conveniently modified. The remaining part of the assumption ensures the identifiability of the limiting functional parameter m_∞ by leveraging the strict convexity of the KL divergence in its first argument and the convex structure of the feasible sets $\mathbb{M}(\mathbf{H}, F_\infty; m)$. The injectivity condition ensures that different m 's lead to distinct projections, thereby preventing flat regions in the objective function.

The assumption implies that the composite map

$$m \rightarrow \inf_{G \in \mathbb{M}(\mathbf{H}, F_\infty; m)} \text{KL}(G \| F_\infty)$$

is strictly convex on M^U , and thus admits a unique minimizer m_∞ . A sufficient condition for Assumption 2.1 it that there exists a $G \ll F_\infty$ which is a member of any non empty $\mathbb{M}(\mathbf{H}, F_\infty, m)$. A sufficient condition for the injectivity of the mapping $m \mapsto G_m := \arg \min_{G \in \mathbb{M}(\mathbf{H}, F_\infty; m)} \text{KL}(G \| F_\infty)$ over the set M^U is that the moment function $\mathbf{H}(S)m(S)$ is injective in m , for every S in a set of positive F_∞ probability.

The following result establishes then (pseudo) consistency for $\hat{G}$ if for some m , the set $\mathbb{M}(\mathbf{H}, F_\infty; m)$ is non empty. It derives the existence of a unique limiting and sample independent m_∞ at which the empirical parameter m_T weakly converges. The limiting density will then be an element of $\mathbb{M}(\mathbf{H}, F_\infty; m_\infty)$. If on the other hand emptiness is the case for any $\mathbb{M}(\mathbf{H}, F_\infty; m)$ then the optimization problem is asymptotically ill-posed; weak convergence to a limiting criterion that is identically equal to $+\infty$ is obtained:

Theorem 1. *Under Assumptions 1 and 2, as $T \rightarrow \infty$, then: a) If for some m , $\mathbb{M}(\mathbf{H}, F_\infty; m) \neq \emptyset$, then there exists a unique $m_\infty \in \mathcal{M}$, such that $\hat{g}_{m_T} \rightsquigarrow g_{m_\infty}$ in $\ell^1(S)$, for a unique $G_{m_\infty} \in \mathbb{M}(\mathbf{H}, F_\infty; m_\infty)$, with $G_{m_\infty}(s) := \int_S 1_{\{\mathbf{x} \leq s\}}(\mathbf{x}) g_{m_\infty}(\mathbf{x}) d\mathbf{x}$. b) If $\mathbb{M}(\mathbf{H}, F_\infty; m) = \emptyset$ for all $m \in \mathcal{M}$, then the optimization problem in (1) is ill-posed.*

Proof. First, notice that $\text{KL}(\cdot \| \cdot)$ is lower semicontinuous due to Proposition 8 in [Feydy et al. \(2019\)](#) and the fact that ℓ^1 convergence of densities imply weak convergence of the underlying measures. Also due to Assumption 1, and via Skorokhod representations applicable due to Theorem 1 of [Cortissoz \(2007\)](#), we have $\mathbb{M}(\mathbf{H}, \hat{F}, m) \xrightarrow{\text{epi}} \mathbb{M}(\mathbf{H}, F_\infty, m)$, continuously w.r.t. m , which is then directly translated to the weak epiconvergence of the indicators. Using again the Skorokhod representations it is obtained that:

$$\text{KL}(\cdot \| \hat{F}) + \chi_{\mathbb{M}(\mathbf{H}, \hat{F}, m)}(\cdot) \xrightarrow{\text{epi}} \text{KL}(\cdot \| F_\infty) + \chi_{\mathbb{M}(\mathbf{H}, F_\infty, m)}(\cdot),$$

continuously w.r.t. m . If $\mathbb{M}(\mathbf{H}, F_\infty, m) \neq \emptyset$ for some $m \in \mathcal{M}$, then existence and uniqueness of m_∞ and subsequently of G_{m_∞} follows from the strict convexity of

$\text{KL}(\cdot\|F)$ on the convex $\mathbb{M}(\mathbf{H}, F_\infty, m)$ for each m for which this is non-empty as well as from the existence and the identifiability of the limiting parameter implied by Assumption 2. $\square$

Thus, the CIPDE procedure guarantees the weak approximation of a unique limiting posterior distribution G_{m_∞} for which the moment conditions are satisfied, as long as there exists an m such that $\mathbb{M}(\mathbf{H}, F_\infty, m)$ is non-empty. When moreover F_∞ obeys the moment conditions, hence $F_\infty \in \mathbb{M}(\mathbf{H}, F_\infty, m)$, then the CIPDE procedure asymptotically recovers it. G_{m_∞} need not equal the DGP distribution F ; this is the case only if $F \ll F_\infty$, and the functional parameter space $\mathcal{M}$ is well-specified.

2.4 Rate and Limiting Distribution

The results in case a) of Theorem 1 are refined to obtain standard convergence rates and the limiting distribution for the scaled discrepancy between the CIPDE density estimator and its limit.

The derivations are based on the convenient representation of the CIPDE estimator as the solution of the inner Kuhn-Tucker problem in (2) evaluated at the optimal SDF. This representation along with Theorem 1.a) already entails that $\boldsymbol{\lambda}_T(m_T)$ weakly converges to a non-sample dependent pointwise negative Lagrange multiplier $\boldsymbol{\lambda}_\infty(m_\infty)$; its j^{th} component is zero if $\int_{\mathcal{S}} \mathbf{H}_j(S)m(S)f_\infty(S)dS > 0$. Using an argument that is based on Skorokhod representations-see Knight (1999), if $\int_{\mathcal{S}} \mathbf{H}_j(S)m(S)f_\infty(S)dS > 0$, then $\boldsymbol{\lambda}_{j,T}(m)$ is eventually zero almost surely.

The following high-level assumption enables the refinement of the aforementioned limiting properties of the optimal empirical Lagrange multipliers:

Assumption 3 (Prior Rate and Weak Convergence). *For some $r_T \rightarrow \infty$, $r_T(\hat{f} - f_\infty) \rightsquigarrow \mathcal{G}$ pointwisely over S , where $\mathcal{G}$ is a process with almost surely continuous sample paths. If $J(\mathbf{H}, F_\infty; m) \neq \emptyset$, there exists a neighborhood of m_∞ , such that for any m inside this, $M(\mathbf{H}, F_\infty, m)$ and $J(\mathbf{H}, F_\infty; m)$ are non-empty, $J(\mathbf{H}, F_\infty; m)$ is independent of m , and the matrix*

$$V_{J_m} := \int_{\mathbb{R}} \mathbf{H}_{J_m}(S) \mathbf{H}_{J_m}^T(S) \exp(\boldsymbol{\lambda}_\infty^T(m)m(S)\mathbf{H}(S)) f_\infty(S) dS,$$

is well-defined, and has a minimum eigenvalue bounded below a positive constant that is independent of m ; there $J_m := J(\mathbf{H}, F_\infty; m)$.

For the Empirical Cumulative Distribution Function (ECDF), the assumption can be verified via results like the ones in Doukhan et al. (1994) for stationary and strong

mixing processes. For parametric models with differentiable likelihoods, it can be verified in cases where the derivative is bounded, the parameter estimator has a unique pseudo-true value and it is asymptotically Gaussian with standard rates; see again Blasques et al. for GARCH-type models. Other semi-non-parametric estimators, like the kernel-based estimators or the non-parametric MLEs may have more complicated rates that could depend on bandwidths, and/or non-Gaussian limiting distributions; see for example Ch. 24 of [Van der Vaart \(1998\)](#). The second part of the assumption holds whenever the J_{m_∞} components of $\mathbf{H}$ are linearly independent, and the number of binding constraints depends continuously on m locally around m_∞ at the limiting F_∞ . This among others enables the invocation of implicit function arguments that ensure asymptotic smoothness of the associated Lagrange multiplier vectors w.r.t. the integral constraints.

Assumptions 1-3 and standard expansions then suffice for the derivation of the limiting behavior of the random element (ξ_T, ζ_t) comprised of the translated and rescaled by r_T components $\xi_T := r_T(\lambda_T(m_T) - \lambda_\infty(m_\infty))$, and $\zeta_T := r_T(m_T - m_\infty)$. This is summarized in the following auxiliary lemma, which is in turn useful for the derivation of the rates and the limiting distribution of the CIPDE estimator:

Lemma 1. *Suppose that Assumptions 1-3 hold. Then, on $\ell_\infty(\mathbb{R}^{J_{m_\infty}} \times \mathcal{S})$,*

$$(\xi_T, \zeta_T) \rightsquigarrow (\xi_\infty, \zeta_\infty), \quad (4)$$

for

$$\xi_\infty := \begin{cases} \arg \min_{\xi \in \mathcal{H}} \|\xi + V_{J_{m_\infty}}^{-1} z(\mathcal{G}, \lambda_\infty, m_\infty)\|_{V_{J_{m_\infty}}}^2, & \text{on } J_{m_\infty} \\ \mathbf{0}, & \text{on } \{1, 2, \dots, J\} - J_{m_\infty}, \end{cases}$$

$$z(\mathcal{G}, \lambda_\infty, m_\infty) := \int_{\mathbb{R}} \mathbf{H}_{J_{m_\infty}}(S) \mathcal{G}^*(S) dS,$$

$$\mathcal{G}^*(S) := \exp(\lambda_\infty^T(m_\infty) m_\infty(S) \mathbf{H}(S)) \mathcal{G}(S), \quad \mathcal{H} := \prod_{j \in J_{m_\infty}} \mathcal{H}_j, \quad \text{with}$$

$$\mathcal{H}_j := \begin{cases} \mathbb{R}_-, & \lambda_{j,\infty}(m_\infty) = 0 \\ \mathbb{R}, & \lambda_{j,\infty}(m_\infty) < 0 \end{cases}, \quad \|\mathbf{u}\|_A^2 := \mathbf{u}^T A \mathbf{u}, \quad \text{while}$$

$$\zeta_\infty := \arg \min_{\zeta \in \mathcal{H}_{m_\infty}} Z(\zeta)$$

where,

$$Z(\zeta) := (\lambda_\infty(m_\infty) + D_{\lambda_\infty} [\int_{\mathcal{S}} \mathbf{H}(S) \zeta(S) f_\infty(S) dS])^T \int_{\mathcal{S}} \mathbf{H}(S) \zeta(S) \mathcal{G}^*(S) dS + \frac{1}{2} \int_{\mathcal{S}} \kappa(S)^2 g_{m_\infty}(S) dS,$$

for $\kappa(S) := \lambda_\infty^\top(m_\infty) \mathbf{H}(S) \zeta(S) + (D_{\lambda_\infty} [\int_S \mathbf{H}(S) \zeta(S) f_\infty(S) dS])^\top \mathbf{H}(S) m_\infty(S) + \xi_\infty^\top \mathbf{H}(S) m_\infty(S)$, with D_{λ_∞} denoting the derivative of the limiting vector of Lagrange multipliers w.r.t. the integral constraints, and $\mathcal{H}_{m_\infty}$ denotes the convex cone obtained as the Painleve-Kuratowski limit of $r_T(\mathcal{M} - m^*)$.

Proof. Duality implies that for each m , the optimal Lagrange multiplier solves the optimization problem $\min_{\lambda \leq 0} \int_{\mathbb{R}} \exp(\lambda^\top \mathbf{H}(S) m_T(S)) d\hat{F}$. Assumption 1 implies that the objective converges locally uniformly over the multiplier, and continuously w.r.t. m_∞ in probability to the limiting criterion $\int_{\mathbb{R}} \exp(\lambda^\top \mathbf{H}(S) m_\infty(S)) d\hat{F}_\infty$, while Assumption 3 implies that the latter is strictly convex. The previous imply the existence of the limiting multiplier $\lambda_\infty(m_\infty)$ as the unique solution to the asymptotic problem $\min_{\lambda \leq 0} \int_{\mathbb{R}} \exp(\lambda^\top \mathbf{H}(S) m_\infty(S)) dF_\infty$. Consider now the rescaled and translated problem:

$$r_T^2 \left(\int_S \exp(\lambda_T^\top \mathbf{H}(S) m_T(S)) \hat{f}(S) dS - \int_S \exp(\lambda_\infty^\top(m_\infty) \mathbf{H}(S) m_\infty(S)) \hat{f}(S) dS \right),$$

which can be expanded as

$$\begin{aligned} &= r_T^2 \int_S \exp(\lambda_\infty^\top(m_\infty) \mathbf{H}(S) m_T(S)) \left(1 + \frac{1}{r_T} \xi^\top \mathbf{H}(S) m_T(S) + \frac{1}{r_T^2} (\xi^\top \mathbf{H}(S) m_T(S))^2 \right) \hat{f}(S) dS \\ &\quad - r_T^2 \int_S \exp(\lambda_\infty^\top(m_\infty) \mathbf{H}(S) m_\infty(S)) \hat{f}(S) dS, \end{aligned}$$

which weakly converges locally uniformly in ξ to the strictly convex $\xi^\top \int_S \mathbf{H}(S) m_\infty(S) \mathcal{G}^*(S) dS + \frac{1}{2} \xi^\top \int_S \mathbf{H}(S) \mathbf{H}^\top(S) m_\infty^2(S) g_{m_\infty}(S) dS \xi$ establishing the first result. An analogous expansion w.r.t. ζ produces

$$\begin{aligned} &r_T \lambda_\infty^\top(m_\infty) \int_S \exp(\lambda_\infty^\top(m_\infty) \mathbf{H}(S) m_\infty(S)) \mathbf{H}(S) \zeta(S) \hat{f}(S) dS + \\ &r_T \int_S \exp(\lambda_\infty^\top(m_\infty) \mathbf{H}(S) m_\infty(S)) D_{\lambda_T} \int_S \mathbf{H}(S) \zeta(S) \hat{f}(S) dS \mathbf{H}(S) m_\infty(S) \hat{f}(S) dS + \\ &\int_S \exp(\lambda_\infty^\top(m_\infty) \mathbf{H}(S) m_\infty(S)) D_{\lambda_T} \int_S \mathbf{H}(S) \zeta(S) \hat{f}(S) dS \mathbf{H}(S) \zeta(S) \hat{f}(S) dS + \\ &\int_S \exp(\lambda_\infty^\top(m_\infty) \mathbf{H}(S) m_\infty(S)) \xi_T^\top \mathbf{H}(S) \zeta(S) \hat{f}(S) dS + \\ &\int_S \exp(\lambda_\infty^\top(m_\infty) \mathbf{H}(S) m_\infty(S)) \frac{1}{2} (k_T(S))^2 \hat{f}(S) dS, \end{aligned}$$

where

$$\begin{aligned} k_T(S) &:= \lambda_\infty^\top(m_\infty) \mathbf{H}(S) \zeta(S) \\ &+ D_{\lambda_T} \int_S \mathbf{H}(S) \zeta(S) \hat{f}(S) dS \mathbf{H}(S) m_\infty(S) + a \xi_T^\top \mathbf{H}(S) m_\infty(S), \end{aligned}$$

while differentiability of the optimal limiting Lagrange multiplier follows from Assumptions 1-3, and the Implicit Function Theorem in Banach spaces, see for example

Rockafellar and Wets (1998). The previous expansion can be similarly seen to converge, locally uniformly w.r.t. ζ , to

$$\begin{aligned} & (\lambda_\infty(m_\infty) + K)^T \int_{\mathcal{S}} \mathbf{H}(S) \zeta(S) \mathcal{G}^*(S) dS + \\ & \frac{1}{2} \int_{\mathcal{S}} [\lambda_\infty^T(m_\infty) \mathbf{H}(S) \zeta(S) + K^T \mathbf{H}(S) m_\infty(S) + \xi_\infty^T \mathbf{H}(S) m_\infty(S)]^2 g_{m_\infty}(S) dS, \end{aligned}$$

where $K := D_{\lambda_\infty} [\int_{\mathcal{S}} \mathbf{H}(S) \zeta(S) f_\infty(S) dS]$, due to that under Assumptions 1 and 3, the following term converges to zero in probability as $T \rightarrow \infty$:

$$\lambda_\infty^T(m_\infty) \int_{\mathcal{S}} \mathbf{H}(S) \zeta(S) \exp(\lambda_\infty^T \mathbf{H}(S) m_\infty(S)) r_T \hat{f}(S) dS.$$

This last result holds because $\lambda_\infty(m_\infty)$ satisfies the first-order condition of the limiting dual problem:

$$\int_{\mathcal{S}} \mathbf{H}(S) m_\infty(S) \exp(\lambda_\infty^T \mathbf{H}(S) m_\infty(S)) f_\infty(S) dS = \mathbf{0},$$

implying that first-order deviations in the moment function are asymptotically orthogonal to the gradient direction. The limiting criterion is then strictly convex due to that the integrand of the second term is a strictly convex quadratic form in ζ , provided that $g_{m_\infty}(S) > 0$ almost everywhere and $\mathbf{H}(S)$ has full rank on the support $\mathcal{S}$. The first term is linear in ζ . Therefore, $\mathbf{Z}(\zeta)$ is a strictly convex functional.

Also, the feasible set $\mathcal{H}_{m_\infty}$, defined as the Painlevé–Kuratowski limit of $r_T(\mathcal{M} - m_\infty)$, is a convex cone. Hence, the strict convexity of the objective over a convex domain implies the existence and uniqueness of the minimizer. The previous establish the second result. $\square$

Then the required limit theory of the CIPDE density estimator is readily obtained from (3) and the Delta method; the following theorem describes it:

Theorem 2. *Suppose that Assumptions 1-3 hold, and that $\mathbb{M}(\mathbf{H}, F_\infty; m) \neq \emptyset$ for some $m \in \mathcal{M}$. a. If $J_{m_\infty} = \emptyset$, then in $\ell^\infty(\mathcal{S})$,*

$$r_T(\hat{g} - f_\infty) \rightsquigarrow \mathcal{G}. \quad (5)$$

b. If $J_{m_\infty} \neq \emptyset$, then in $\ell^\infty(\mathcal{S})$,

$$r_T(\hat{g} - g_{m_\infty}) \rightsquigarrow \mathcal{G}_\infty, \quad (6)$$

where

$$\mathcal{G}_\infty(S) := \frac{\Lambda(S) + g_{m_\infty}(S) \int_{\mathcal{S}} \Lambda(S) dS}{\int_{\mathcal{S}} \exp(\lambda_\infty^T(m_\infty) \mathbf{H}(S) m_\infty(S)) f_\infty(S) dS},$$

$$\Lambda(S) := \mathcal{G}^*(S) + (\xi_\infty^T(S) m_\infty(S) + \lambda_\infty^T(m_\infty) \zeta_\infty(S)) \mathbf{H}(S) g_{m_\infty}(S),$$

and ξ_∞ , ζ_∞ , $\mathcal{G}^*$, as in Lemma 1.

Proof. The first case follows from the eventual nullification of λ_T w.h.p. and Assumption 3. The second case follows from the fact that Assumptions 1-3 along with Lemma 1, imply that $\begin{pmatrix} r_T(f_T - f_\infty) \\ \xi_T \\ \zeta_T \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathcal{G} \\ \xi_\infty \\ \zeta_\infty \end{pmatrix}$, the fact that $\hat{g}_{m_T}(S) = \frac{\hat{f}(S) \exp(\lambda_T^\top \mathbf{H}(S) m_T(S))}{\int_{\mathcal{S}} \hat{f}(S) \exp(\lambda_T^\top \mathbf{H}(S) m_T(S)) dS}$, $\hat{g}_{m_\infty}(S) = \frac{f(S) \exp(\lambda_\infty^\top \mathbf{H}(S) m_\infty(S))}{\int_{\mathcal{S}} f(S) \exp(\lambda_\infty^\top \mathbf{H}(S) m_\infty(S)) dS}$, and two successive applications of the Delta method. $\square$

Theorem 2 specifies a rate of convergence, inherited by the limit theory of the prior, and a limiting distribution for the translated conditional density estimator of F . In all the considered cases the CIPDE estimator is not asymptotically independent of the limiting prior; this result is not surprising because the KL inconsistency reduction obtained by the CIPDE is asymptotically defined by restrictions that depend on the limiting prior F_∞ .

Whenever $J_{m_\infty} = \emptyset$, and thus F_∞ strictly satisfies the moment conditions, the eventual nullification of the multiplier implies that the limit theory of the initial estimator is recovered. Due to the first part of Assumption 2 this encompasses the case where the prior is consistent, i.e. $f_\infty = f$.

When the limiting Lagrange multipliers are non-zero, i.e., $J_{m_\infty} \neq \emptyset$, the CIPDE estimator incorporates binding moment inequality constraints. These constraints act as a source of side information, effectively restricting the estimator to a lower-dimensional feasible set. The resulting limiting distribution $\mathcal{G}_\infty$ is a projection of the unconstrained limit process $\mathcal{G}$ onto a constrained tangent cone, modulated by the Lagrange multipliers and the local geometry of the feasible set (via the influence functions ξ_∞ and ζ_∞). As is well known from the theory of constrained M -estimation and semiparametric efficiency bounds, such projections reduce asymptotic variance in directions aligned with the active constraints. Hence, the CIPDE estimator may achieve variance reduction-i.e., efficiency gains-relative to the unconstrained estimator $\hat{f}$, provided that the constraints are informative. This interpretation aligns with results in Andrews (1999), van der Vaart (1998), and Bickel et al. (1998), and formalizes the intuition that valid economic restrictions, when binding, yield more precise inference.

In the special case where $r_T = \sqrt{T}$, $\mathcal{G}$ is zero mean Gaussian, $J_{m_\infty} \neq \emptyset$, $\lambda_{j,\infty} < 0$, $\forall j \in J_{m_\infty}$, and for any $S \in \mathcal{S}$, $\cup_{\zeta \in \mathcal{H}_{m_\infty}} \{\zeta(S)\} = \mathbb{R}$, then $\mathcal{G}_\infty$ is a zero mean Gaussian process, and the likelihood ratio tests performed in Section 5.4, can be proven to be asymptotically valid.

2.5 Inconsistency Correction

Theorem 1 provides limited information regarding the relation between the limiting G and the true F . If $\mathbb{M}(\mathbf{H}, F_\infty; m) = \{F\}$ for some $m \in \mathcal{M}$, then $G = F$, and $\hat{G}$ is then a weakly consistent estimator of the DGP distribution. However, in many applications, F cannot be expected to be the single element of $\mathbb{M}(\mathbf{H}, F_\infty; m)$ for any premissible SDF. For instance, when F satisfies moment inequalities with strict inequality, there will typically be other elements nearby it.

The result below, utilizing the Pythagorean Theorem within information geometry (refer to Nielsen (2020) for instance), asserts that $\hat{G}$ effectively diminishes the asymptotic divergence to F in scenarios where F_∞ fails to meet the moment inequalities:

Theorem 3. *Under Assumption 1, and if $F \in \mathbb{M}(\mathbf{H}, F_\infty; m)$ for some $m \in \mathcal{M}$, then:*

$$\text{KL}(F\|G) = \text{KL}(F\|F_\infty) - \text{KL}(G\|F_\infty). \quad (7)$$

Proof. If $F = F_\infty$ or $G = F_\infty$ the result follows trivially by Theorem 1. If $F \neq F_\infty$, and $G \neq F_\infty$, then Theorem 1 along with the CIPDE estimator representation in (3) implies that there exists a limiting non-sample dependent, pointwise non-positive and non identically zero multiplier $\lambda_\infty(m_\infty)$, such that $g_{m_\infty}(S) = \frac{f_\infty(S) \exp(\lambda_\infty^\top(m_\infty)m_\infty(S)\mathbf{H}(S))}{\int_{\mathcal{S}} f_\infty(S) \exp(\lambda_\infty^\top(m_\infty)m_\infty(S)\mathbf{H}(S))dS}$. The second part of Theorem 3.1 in Csiszár (1975) then implies the result. $\square$

Equation (7) has a nontrivial information-theoretic interpretation when $F_\infty \notin \mathbb{M}(\mathbf{H}, F_\infty; m_\infty)$: the information lost by approximating the DGP F by the limiting posterior G , compared to the information lost by approximating F by the limiting prior F_∞ is reduced by the non-negative quantity $\text{KL}(G\|F_\infty) \geq 0$, that is, the information gained by using moment inequalities to update the inference to the limiting posterior G , instead of using the moment conditions ignorant limiting prior. Hence, whenever F_∞ violates the moment conditions, the CIPDE density estimator performs asymptotically a (partial) KL inconsistency correction.

2.6 Inconsistency Correction under Approximate Feasibility

In practice, F may not exactly satisfy the pricing-implied moment inequalities required by the feasible set; $F \notin \mathbb{M}(\mathbf{H}, F_\infty, m)$, $\forall m \in \mathcal{M}$. Notably, asset pricing restrictions may be violated if the stock market or option market is out of equilibrium or if the observed prices are not fully synchronized. This motivates the study of cases of approximate feasibility where F violates the constraints only by a small amount at

least for some m . The question then becomes whether CIPDE still asymptotically yields a meaningful Kullback-Liebler correction towards F .

Suppose that the limiting prior F_∞ is both different from F and does not satisfy the moment inequalities, while additionally F is only approximately feasible:

$$\inf_{j \in J_{m_\infty}} \int_{\mathcal{S}} \mathbf{H}_j(S) m_\infty(S) f(S) dS = -\delta, \quad \text{for some small } \delta > 0,$$

thus the Pythagorean decomposition of KL divergence does not necessarily hold exactly. However, it remains plausible to ask whether the limiting CIPDE still satisfies:

$$\text{KL}(F \| G) < \text{KL}(F \| F_\infty),$$

thus constituting an asymptotic partial KL-correction on $\hat{F}$. This is formalized in the following result:

Proposition 1 (Partial KL Correction under Approximate Feasibility). *Under Assumption 1, suppose also that:*

- (a) $F \notin \mathbb{M}(\mathbf{H}, F_\infty, m)$ for all $m \in \mathcal{M}$,
- (b) the following exponential moment condition holds:

$$\int_{\mathcal{S}} \exp(\boldsymbol{\lambda}_\infty^\top \mathbf{H}(S) m_\infty(S)) f_\infty(S) dS > \exp\left(|J_{m_\infty}| \cdot \sup_{j \in J^*} |\boldsymbol{\lambda}_{\infty, j}| \cdot \delta\right),$$

where

$$\begin{aligned} \delta &:= - \inf_{j \in J_{m_\infty}} \int_{\mathcal{S}} \mathbf{H}_j(S) m_\infty(S) f(S) dS, \\ J^* &:= \arg \min_{j \in J_{m_\infty}} \int_{\mathcal{S}} \mathbf{H}_j(S) m_\infty(S) f(S) dS. \end{aligned}$$

Then,

$$\text{KL}(F \| G) < \text{KL}(F \| F_\infty).$$

Proof. Using elementary calculus, we obtain the KL decomposition:

$$\text{KL}(F \| F_\infty) = \text{KL}(F \| G) + \text{KL}(G \| F_\infty) + \int_{\mathcal{S}} (\log g(S) - \log f_\infty(S)) (f(S) - g(S)) dS.$$

Rearranging the expression with respect to $\text{KL}(F\|G)$, we get:

$$\begin{aligned} \text{KL}(F\|G) &= \text{KL}(F\|F_\infty) - \text{KL}(G\|F_\infty) \\ &\quad - \int_{\mathcal{S}} (\log g(S) - \log f_\infty(S)) (f(S) - g(S)) dS. \end{aligned}$$

By the first-order condition of the limiting dual problem, the cross term equals:

$$\begin{aligned} \int_{\mathcal{S}} (\log g(S) - \log f_\infty(S)) (f(S) - g(S)) dS &= \log \left(\int_{\mathcal{S}} \exp(\boldsymbol{\lambda}_\infty^\top \mathbf{H}(S) m_\infty(S)) f_\infty(S) dS \right) \\ &\quad - \int_{\mathcal{S}} \boldsymbol{\lambda}_\infty^\top \mathbf{H}(S) m_\infty(S) f(S) dS. \end{aligned}$$

The second term above can be bounded using the worst-case component in the set J_{m_∞} :

$$\int_{\mathcal{S}} \boldsymbol{\lambda}_\infty^\top \mathbf{H}(S) m_\infty(S) f(S) dS \leq |J_{m_\infty}| \cdot \sup_{j \in J^*} |\boldsymbol{\lambda}_{\infty,j}| \cdot \delta.$$

Therefore, we obtain the inequality:

$$\text{KL}(F\|G) \leq \text{KL}(F\|F_\infty) - \text{KL}(G\|F_\infty) + |J_{m_\infty}| \cdot \sup_{j \in J^*} |\boldsymbol{\lambda}_{\infty,j}| \cdot \delta.$$

Hence, partial correction is ensured whenever:

$$\text{KL}(G\|F_\infty) > |J_{m_\infty}| \cdot \sup_{j \in J^*} |\boldsymbol{\lambda}_{\infty,j}| \cdot \delta.$$

Finally, by the dual representation of the projection G , we have:

$$\text{KL}(G\|F_\infty) = \log \int_{\mathcal{S}} \exp(\boldsymbol{\lambda}_\infty^\top \mathbf{H}(S) m_\infty(S)) f_\infty(S) dS,$$

which completes the proof. $\square$

The above result shows that even when the moment conditions fail to hold exactly, the CIPDE projection may still produce a KL-partially corrected update, provided the dual potential (expressed through the exponential moment) outweighs the severity of the moment violations. Intuitively, this implies that the model's “pull” through the dual weighting structure can still dominate minor economic inconsistencies — a robustness property particularly valuable in contexts where asset pricing restrictions rely on data from disequilibrium or unsynchronized sources. The proposition formalizes a sufficiency threshold: if the prior is bad enough and feasibility violations are small enough, the update is justified.

The bound involving $|J_{m_\infty}| \cdot \sup_{j \in J^*} |\boldsymbol{\lambda}_{\infty,j}| \cdot \delta$ is conservative. A refined inequality could use Jensen's inequality or concentration bounds to exploit curvature in the exponential term, especially if the vector $\boldsymbol{\lambda}_\infty$ is sparse or structured.

3 Option Pricing and Option Combinations

CIPDE is applied in this study to stock index returns using moment conditions based on an option pricing system in the spirit of [Constantinides et al. \(2009\)](#), and optimization of combinations of index options in the spirit of [Constantinides et al. \(2020\)](#); [Post and Longarela \(2021\)](#); [Beare et al. \(2025\)](#). This approach is applied both to randomly generated data sets from a known DGP in a simulation experiment (Section 4) and to historical monthly SPX returns and price quotes of CBOE SPX options in an empirical application (Section 5).

[Constantinides et al. \(2009\)](#) develop a system of model-free pricing restrictions for multiple concurrent option series with different option types (put or call) and different strike prices. These restrictions are quite general because they don't assume that the option market is complete or perfect, and they furthermore don't assume a specific functional form for the SDF. Still, the truth of these restrictions may depend on the characteristics of the selected option series and the prevailing market conditions.

Violations of the equilibrium system, or pricing errors, imply that the market index is inefficient in the sense of being dominated by second-degree SD (SSD) by portfolios that are enhanced using properly constructed option combinations. [Constantinides et al. \(2020\)](#); [Post and Longarela \(2021\)](#); [Beare et al. \(2025\)](#) develop and apply optimization problems to construct such enhanced portfolios. These studies do not find robust evidence for mispricing in the form of significant out-of-sample out-performance for one-month options, consistent with the notion that the price quotes correctly reflect the latent conditional return distribution.

A limitation of these studies is that they do not account for time-variation of the shape of the distribution and risk premiums which generally results in statistical inconsistency of the conditional distribution estimator. For instance, the empirical analysis by [Constantinides et al. \(2009\)](#) is based on a scaled version of the *unconditional* ECDF, which uses a fixed shape, fixed Market Risk Premium and fixed Variance Risk Premium. Estimation error for the conditional distribution in turn introduces the risk of false rejection of options market efficiency and poor OOS performance of optimized option combinations.

The present section discusses the use of CIPDE in this application area. We analyze the statistical significance of the improvements in OOS forecasting ability using a Likelihood Ratio Test (LRT). We analyze the economic significance using the OOS investment performance of option combinations that are optimized under the posterior density ($\hat{G}$).

We deliberately deviate from the aforementioned studies in two subtle ways. First, the equilibrium system of [Constantinides et al. \(2009\)](#) is tightened by requiring the

SDF to be log-convex, which refines SSD into Decreasing Absolute Risk Aversion (DARA) SD (henceforth DSD) (Vickson, 1975; Bawa, 1975). This approach, also used in Post and Potì (2017), sharpens the analysis by excluding pathological shapes of the SDF and facilitates numerical implementation by allowing for a convenient log-linearization of the CIPDE problem.

Second, rather than accepting or rejecting the pricing efficiency of all options, we identify a subset of options for which it is plausible that the pricing system is valid. Specifically, we consider a superset of N tradable options and a subset of $N_0 \leq N$ options that pass a stricter set of filters for assumed pricing efficiency. This approach parallels Stutzer (1996), who constructed a nonparametric risk-neutral distribution from one set of option prices and applied it to value other options.

To illustrate the importance of estimating the physical distribution and modeling the SDF, we also estimate the risk-neutral distribution by applying CIPDE under a constant SDF, which resembles the methodology of Stutzer (1996) even more closely. While the risk-neutral distribution is, by construction, well-suited for valuing options, it is less suitable for forecasting and optimization because of its inherent pessimistic bias.

3.1 Preliminaries

The focus is on European-type stock index options. A total of N distinct option series are considered with different option types (put or call) and different strike prices K_i , $i = 1, \dots, N$. All options have the same time to expiry (denoted here by one time unit), and their expiration date equals the forecast horizon for the underlying index return.

The annualized risk-free rate on a maturity-matched Treasury bill is R_B . The current index value and the index value at expiry date are S_0 and $S_1 = S_0(1 + R_S)$, respectively. The (annualized and maturity-matched) dividend yield to the index is R_D .

The price return R_S is treated as a random variable with latent distribution F that is estimated with priors $\hat{F}$ and posteriors $\hat{G}$. The specification of the priors will be discussed in more detail in Section 4 (Monte Carlo simulation) and Section 5 (empirical study); the specification of the moment conditions is discussed in Subsection 3.2 below.

The option payoffs are $X_{i,1} = P_{i,1} := \max(K_i - S_1, 0)$ for puts and $X_{i,1} = C_{i,1} := \max(S_1 - K_i, 0)$ for calls, $i = 1, \dots, N$. The payoffs of all options are driven exclusively by the index price return R_S , so that a perfect single-factor structure arises for the payoffs of all options.

The present values are denoted by $X_{i,0} = P_{i,0}$ for puts and $X_{i,0} = C_{i,0}$ for calls, respectively. Whereas the present values are latent, the quoted ask price a_i and quoted bid price $b_i \geq a_i$ are observable. In addition, if prices are efficient, then the ranking $b_i \geq X_{i,0} \geq a_i$ is obtained.

Dynamic replication or hedging of options using rebalanced combinations of bills and stocks is not considered because it is challenging to estimate the entire dynamic process and the relevant transaction costs of portfolio rebalancing, which depend on the rebalancing frequency and portfolio turnover. Instead, the investment universe consists of combinations of bills, stocks, and options that are held until the option expiry date.

Buying or selling bills and stocks is assumed to involve proportional one-way transaction costs of 0.1%, which has a negligible effect on the reported results based on our buy-and-hold strategies (in contrast to dynamic replication or hedging). Options can be bought at the quoted ask price a_i and written at the quoted bid price b_i . Relative to the option premium, the bid-ask spread ($a_i - b_i$) is much larger and more consequential than the proportional transaction costs for trading bills and stocks, especially for illiquid option series and during market turbulence. The profit or loss at the expiration date is given by $(X_{i,1} - a_i(1 + R_B))$ for a bought option and $(b_i(1 + R_B) - X_{i,1})$ for a written option.

The focus is on one-month options. For this maturity segment, a relatively long time series, broad cross-section of liquid options, and rich research literature are available for SPX options. In addition, this maturity aligns with the common use of one-month returns in financial econometrics, portfolio optimization, and empirical asset pricing.

3.2 Option pricing conditions

To update the prior $\hat{F}$ to a posterior $\hat{G}$, CIPDE uses conditioning information in the form of conditional moment conditions. It is important that the moment conditions are not controversial because they are not directly tested but used to incorporate side information, in this study. For this reason, we eschew parametric option pricing formulas, fully-specified asset pricing theories or models that assume a complete or perfect option market.

Instead, we employ and extend a system of model-free pricing restrictions based on [Constantinides et al. \(2009\)](#). The restrictions are imposed for a subset of $N_0 \leq N$ options that pass a set of data filters. In arbitrage-free equilibrium, an SDF for pricing cash flows at the option expiration date, $m : S_1 \rightarrow \mathbb{R}$, exists that is consistent with the prevailing market prices of the securities (index, bill and options):

$$0.999 \leq \mathbb{E}_F[m(1 + R_S)] \leq 1.001; \quad (8)$$

$$0.999 \leq \mathbb{E}_F[m(1 + R_B)] \leq 1.001; \quad (9)$$

$$b_i \leq \mathbb{E}_F[mX_{i,1}] \leq a_i; \quad i = 1, \dots, N_0. \quad (10)$$

Attractively, these pricing restrictions do not assume that options can be replicated without costs using dynamic combinations of bills and stocks, and they furthermore accounts for the relatively large bid-ask spread for index options. In addition, the restrictions are imposed in the empirical application only for options that pass our tight filters based on maturity, moneyness, option delta and option premium (see Section 5.1).

The SDF is generally not unique under these conditions, and the estimation benefits from imposing structure on the SDF to prevent overfitting it to a misspecified prior $\hat{F}$. The SDF is assumed to represent the Intertemporal Marginal Rate of Substitution (IMRS) of index investors and it is partially identified by a set of functions that obey standard regularity conditions for the IMRS:

$$\{m \in \mathcal{C}^2 : m(S_T) \geq 0; m'(S_T) \leq 0; \ln(m(S_T))'' \geq 0\} =: \mathcal{M} \ni m. \quad (11)$$

These SDFs are positive, increasing and log-convex, as required for the IMRS of standard utility functions. Although [Constantinides et al. \(2009\)](#) do not require this property, convexity directly follows from the generally accepted property of DARA for standard utility functions ([Kimball \(1990\)](#)). Violations of the equilibrium system (8)-(11), or pricing errors, imply that the market index is dominated by DSD by portfolios that enhance the index using certain option combinations.

If the log-convexity condition is not imposed, as in [Constantinides et al. \(2009\)](#), the analysis allows for SDFs with a pathological shape such an implausible reverse S-shape with increasing risk aversion followed by decreasing risk aversion. As a result, pricing errors occur less frequently under the prior estimate; the posterior density will diverge less from the prior density if pricing errors do occur; and the evidence for OOS forecasting success and investment outperformance weakens. Similarly, [Basso and Pianca \(1975\)](#) show that general n -th degree SD allow for financial option prices that are inconsistent with DARA.

The risk-neutral distribution (Q) can be estimated using the same pricing conditions (((8))-((10))) by restricting the SDF to be constant: $m(S) = (1 + R_B)^{-1}$. This approach naturally introduces a pessimistic bias for estimating the physical distribution which may weaken forecasting success and investment performance.

The pricing restrictions (8)-(11) are conditional moment conditions because the probability distribution and option prices are updated every month. The truth of the

moment conditions is not directly tested in this study. In the simulation study, the moment conditions are known to be true. In the empirical study, we select a subset of options for which the moment conditions appear most plausible using option filters. In addition, the truth of the moment conditions (for the selected options) is tested indirectly based on the OOS forecasting and investment success.

The CIPDE relies on the empirical counterparts of the moment conditions (8)-(11) based on $\mathbb{E}_{\hat{G}}[m(1 + R_S)]$, $\mathbb{E}_{\hat{G}}[m(1 + R_B)]$, and $\mathbb{E}_{\hat{G}}[mX_{i,1}]$, $i = 1, \dots, N_0$. These empirical conditions are expected to be effective because specification error and estimation error is likely to lead to violations of these conditions for at least some of the option series.

If the density estimator is discretized using a finite set of atoms $\{S_{j,T}\}$, the condition $m \in \mathcal{M}$ can be discretized and linearized to a finite set of linear restrictions on the SDF values $\{m(S_{j,T})\}$. In this study, we discretize the density estimators using an equally spaced grid in the range $[-0.6, 0.6]$ with 25 bps (0.0025) grid size, resulting in 481 atoms. After discretization and linearization, CIPDE becomes a Bi-convex Programming problem that is convex in the SDF values $m(S_{j,T})$ and convex in the probability weights $\hat{g}(S_{j,T})$. The present study solves the problem using the two-step optimization procedure of [Post and Potì \(2017\)](#).

3.3 Likelihood Ratio Test

To evaluate the OOS forecasting ability, we apply the Likelihood Ratio (LR) test of [Vuong \(1989\)](#), as per [Amisano and Giacomini \(2007\)](#).

$$\text{WLR}_{t+1}(\hat{F}, \hat{G}) := \left(\log \hat{f}_t(S_{t+1}) - \log \hat{g}_t(S_{t+1}) \right); \quad (12)$$

$$\text{LRT}(\hat{F}, \hat{G}) := \frac{n^{-1} \sum_{t=1}^n \text{WLR}_{t+1}(\hat{F}, \hat{G})}{\sqrt{n^{-1} \sum_{t=1}^n \text{WLR}_{t+1}(\hat{F}, \hat{G})^2} / \sqrt{n}}. \quad (13)$$

where $t = 1$ refers to the month that the first one-step-ahead OOS forecast is produced and n is the number of forecasts evaluated.

As per [Amisano and Giacomini \(2007\)](#), the numerator does not include a correction for serial dependence because serial dependence is relatively weak for monthly stock index returns. Ljung-Box tests and autocorrelation function analysis of the log likelihood ratios, $\text{WLR}_{t+1}(\hat{F}, \hat{G})$, do not detect significant serial autocorrelation over our sample of 321 months. Short truncation lags in correcting for serial dependence in the denominator of Eq. (13) have been shown to improve the finite-sample properties

of tests of equal predictive ability (see [Diebold and Mariano, 2002](#)), this is used to motivate excluding the correction term in [Amisano and Giacomini \(2007\)](#).

3.4 Optimized option combinations

To analyze the economic significance of the improvements in forecasting ability in the empirical application, we analyze the investment performance of option combinations that are optimized based on prior or posterior estimates. The aim is to identify risk arbitrage opportunities that occur if an option-enhanced portfolio dominates the index by SD.

The analysis differs from [Constantinides et al. \(2020\)](#), [Post and Longarela \(2021\)](#), and [Beare et al. \(2025\)](#) because it assumes pricing efficiency for the N_0 options included in the moment conditions used to construct the posterior. Consequently, performance improvement under the posterior can stem from (i) avoiding unprofitable positions in the options included in the moment conditions and (ii) exploiting pricing errors in options not included in the moment conditions.

The optimization problem for a given density forecast $H = \hat{F}, \hat{G}, \hat{Q}$ is specified as follows:

$$\max \sum_{i=1}^N \beta_i b_i - \sum_{i=1}^N \alpha_i a_i \quad (14)$$

$$\text{s.t.} \quad \left(S_1 + \sum_{i=1}^N \alpha_i X_{i,1} - \sum_{i=1}^N \beta_i X_{i,1} \right) \succeq_H S_1, \quad (15)$$

$$\sum_{i=1}^N \alpha_i + \sum_{i=1}^N \beta_i \leq 1, \quad (16)$$

$$\alpha_i, \beta_i \geq 0, \quad i = 1, \dots, N. \quad (17)$$

The model variables $\{\alpha_i\}$ and $\{\beta_i\}$ represent the open long positions entered at the best ask price and short positions entered at the best bid price, respectively, in the individual options.

The objective is to maximize the net option premium, following [Post \(2003\)](#) and [Beare et al. \(2025\)](#). A key advantage of this objective is that the premium is independent of distributional assumptions or estimation error. Maximizing this deterministic return component is expected to improve the robustness of the dominance relation.

The operator $\succeq_H$ denotes the partial order induced by the SSD relation. The optimization is based on the second-order SD rule instead of DSD, as in the aforementioned benchmark studies. Tightening the dominance rule in this way is expected

to enhance robustness to distributional assumptions and estimation error ($H \neq F$). Second, exact and tractable problem formulations are available for optimization with SSD constraints, while optimization under higher-order SD constraints relies on approximations and is computationally more demanding (see, e.g., [Post and Kopa \(2017\)](#)).

The SSD condition is generally infinite-dimensional and non-smooth, and thus numerically intractable. For discretized density estimates, however, it reduces to a finite system of Mixed Integer Linear Programming constraints on the Expected Shortfall levels of the option-enhanced portfolio and the index. We adopt the computational strategy of [Post and Longarela \(2021\)](#), which builds on the linearization of Expected Shortfall by [Rockafellar and Uryasev \(2000\)](#).

Finally, the position limit $\sum_{i=1}^N \alpha_i + \sum_{i=1}^N \beta_i \leq 1$ is imposed, following [Constantinides et al. \(2020\)](#). This limit reduces concerns about market liquidity, since at most one open option position is taken per unit of index. In addition, it acts as a vector norm restriction on the model variables, which is expected to enhance robustness and realized performance; see [DeMiguel et al. \(2009\)](#). As the objective function, the position limit is not affected by distributional assumptions or estimation error.

Optimal positions that are expected if the density forecasts are accurate include writing overpriced covered calls (if any) and buying underpriced protective puts (if any), because such positions simultaneously lower downside risk (by construction) and improve risk-adjusted expected returns (in case of mispricing), as required by the dominance condition.

At every formation date, we solve the linearized problem using Linear Programming for the various discretized prior and posterior density forecasts ($H = \hat{F}, \hat{G}, \hat{Q}$). For every forecast, we evaluate the consequences for index investors of buying the chosen options at their ask prices or writing at bid prices, and holding the options until cash settlement at expiration. We evaluate the OOS investment performance of the option-enhanced portfolio. In addition to the mean and the standard deviation of return, we evaluate the Certainty Equivalent Return (CER) for a power utility function with $RRA = 1, 4, 10$, computed across the expiration dates.

4 Monte Carlo Experiment

A Monte Carlo simulation experiment is used to analyze the pre-asymptotic behavior of the updated density estimator. We simulate the divergence $KL(\cdot \| F)$ for $\hat{F}$, $\hat{G}$, and $\hat{Q}$, and the test statistic $LRT(\cdot, \hat{F})$ for F , $\hat{G}$, and $\hat{Q}$. The experiment assumes that all options are correctly priced ($N_0 = N$), so risk arbitrage opportunities do not exist. In this case, option-enhanced portfolios based on $\hat{F}$ necessarily underperform those

based on $\hat{G}$, making such a comparison tautological. We therefore focus exclusively on statistical performance and do not report investment performance in this section.

4.1 Design

Suppose that $S_0 = 100$ and R_S obeys a log-normal distribution (F). The expected return is based on the classic mean-variance approximation $\mu_S = \mu(\gamma) := R_B + \gamma\sigma_S^2$, where γ is a latent aggregate RRA coefficient.

The variance is based on a second-order Taylor series approximation by [Kang and Yoon \(2010\)](#): $\sigma_S^2 = \sigma^2(\gamma) := \sigma_Q^2 + \gamma\sigma_Q^3\theta_Q + 0.5\gamma^2\sigma_Q^4(\kappa_Q - 3)$, where $(\sigma_Q^2, \theta_Q, \kappa_Q)$ are the risk-neutral variance, skewness and kurtosis.

By the probability calculus of the lognormal, the location parameter is $M(\gamma) = \ln(1 + \mu(\gamma)) - \frac{1}{2}V(\gamma)$ and the scale parameter is $V(\gamma) = \ln(1 + \sigma^2(\gamma)/(1 + \mu(\gamma))^2)$. The risk-neutral distribution Q arises here as the special case with $\gamma = 0$.

The prior $\hat{F}$ takes the correct, lognormal form but its parameters are based on the RRA estimator $\hat{\gamma}_T$: $M(\hat{\gamma}_T)$ and $V(\hat{\gamma}_T)$. It could be based on time-series estimates for the physical variance σ_S^2 , combined with the extracted Q moments, in a GMM estimation procedure, as in [Bakshi and Madan \(2006\)](#) or [Kang and Yoon \(2010\)](#). Instead of modeling the precise volatility dynamics and estimation method, we assume that $\hat{\gamma}_T \sim U(\gamma + \delta_0 - \delta_1, \gamma + \delta_0 + \delta_1)$, where δ_0 is the bias and $\delta_1 \geq 0$ is the semi-range.

One may think of δ_0 as capturing the non-vanishing effect of specification error for the volatility dynamics and δ_1 as capturing the vanishing effect of sampling variation reflecting the sample size (T). The limiting case $\delta_1 \rightarrow 0$ then represents $T \rightarrow \infty$.

The conditional information for constructing the posterior $\hat{G}$ includes the quotes b_i and a_i for N call options with strikes $K_i = S_0 + 0.01S_0(i - 0.5(N + 1))$, $i = 1, \dots, N$, as well as the equilibrium system (8)-(11).

The Present Value of the call options amounts to $X_{i,0} := \mathbb{E}_F[mX_{i,1}]$, where the latent SDF is assumed to take a power shape: $m \propto (1 + R_S)^{-\gamma}$. The bid-ask spread has width $W_i := (a_i - b_i) = wX_{i,0}$, where $w \geq 0$ is the fractional spread; it is randomly located around the Present Value: $a_i \sim U(X_{i,0}, X_{i,0}(1 + w))$, so that it includes the latent true Present Value: $a_i \leq X_{i,0} \leq b_i$, $i = 1, \dots, N$.

We set $\gamma = 5$ and $(R_B, \sigma_Q^2, \theta_Q, \kappa_Q) = (0, 0.04/12, -1, 4.5)$, all representative values for monthly S&P500 returns and one-month SPX options. The RRA level $\gamma = 5$ is relatively high compared to estimates from risky choice experiments, to generate realistic values for the Market Risk Premium and Variance Risk Premium, as is also the case for typical option-implied RRA estimates in [Bakshi and Madan \(2006\)](#) or [Kang and Yoon \(2010\)](#). We abstract from dividends in the simulations: $R_D = 0$.

We vary the bias $\delta_0 = 0, 5$ and semi-range $\delta_1 = 1, 5$ to analyze the effect of the

goodness of the prior, and we vary the cross-sectional breadth $N = 5, 25$ and the spread $w = 0.010, 0.002$ to analyze the effect of the information contents of the moment conditions. The assumed breadth is smaller than the number of qualified SPX options in our empirical analysis; the assumed spreads are representative for the quoted spreads under normal market conditions in recent years.

The improvements in density estimation are quantified by repeatedly drawing a random RRA estimate $\hat{\gamma}_t \sim U(\gamma + \delta_0 - \delta_1, \gamma + \delta_0 + \delta_1)$, $t = 1, \dots, 300$, each representing a random underlying time series sample, and computing the parameters of $\hat{F}$ accordingly $(M(\hat{\gamma}_t), V(\hat{\gamma}_t))$. Next, we construct $\hat{G}$ using CIPDE based on system (8)-(11) and the simulated bid-ask prices $\{a_i, b_i\}$. We also construct the risk-neutral $\hat{Q}$ using the constant SDF $m(S_T) = (1 + R_B)^{-1}$.

Having generated $\hat{F}$, $\hat{G}$ and $\hat{Q}$ for a given $\hat{\gamma}_t$, we proceed to compute the divergence measures $\text{KL}(\hat{F}||F)$, $\text{KL}(\hat{G}||F)$, and $\text{KL}(\hat{Q}||F)$. In addition, we apply forecasting ability tests based on $\text{LRT}(F, \hat{F})$, $\text{LRT}(\hat{G}, \hat{F})$ or $\text{LRT}(\hat{Q}, \hat{F})$ to 10,000 random samples of 300 return realizations from the true F . Finally, we construct the optimal active portfolio with open option positions and evaluate its economic goodness relative to the true F using various investment performance measures.

The simulation design aligns with the statistical theory outlined in Section 2. The form of the SDF and the moment functions, together with the assumption of log-normal index returns, imply the existence of exponential moments. The choice of the bias and semi-range parameters is consistent with RRA estimates that converge, as $\delta_1 \rightarrow 0$ (or $T \rightarrow \infty$ in case of time-series estimation), to a deterministic value at standard rates and with asymptotic Gaussianity.

4.2 Divergence

Table 1 summarizes the results for the divergence measure $\text{KL}(H||F)$, $H = \hat{F}, \hat{G}, \hat{Q}$. Reported are the quartile boundaries (p25, p50 and p75) across the 300 random draws of $\hat{\gamma}_T$.

Not surprisingly, if $\hat{F}$ is unbiased and precise $((\delta_0, \delta_1) = (0, 1))$, improvements are difficult to achieve, even if many tight moment inequalities are available $((N, w) = (25, 0.002))$.

However, the divergence $\text{KL}(\hat{F}||F)$ quickly increases as the bias (δ_0) or semi-range (δ_1) is increased. Whereas the bias (δ_0) mostly affects the median value of the divergence (p50), the semi-range mostly affects the inter-quartile range (p75-p25). Even a small number of loose moment inequalities $((N, w) = (5, 0.010))$ now suffices for achieving material improvements.

The strongest results naturally are obtained if $\hat{F}$ is biased and imprecise, and

many tight moment inequalities are available $((\delta_0, \delta_1, N, w) = (5, 5, 25, 0.002))$. In this case, $\hat{G}$ generates more than 50% reduction of the $\text{KL}(\cdot|F)$ levels.

The estimated risk-neutral $\hat{Q}$ strongly diverges from the true F , due to its false assumption that $\gamma = 0$, even in the benign case with $(\delta_0, \delta_1) = (0, 1)$. The bias is already severe if only a few loose moment conditions are imposed. If the prior is consistent ($\delta_0 = 0$) or precise ($\delta_1 = 1$), then $\hat{Q}$ is clearly inferior to the physical $\hat{G}$. Even in the least favorable case $(\delta_0, \delta_1) = (5, 5)$, $\hat{Q}$ shows greater divergence than $\hat{G}$ in most cases.

[Insert Table 1 about here.]

4.3 Predictive Ability

An alternative way to evaluate the effect of the information projection is to analyze the rejection rates for the forecasting ability test. For this purpose, $\text{LRT}(F, \hat{F})$, $\text{LRT}(\hat{G}, \hat{F})$, or $\text{LRT}(\hat{Q}, \hat{F})$ are computed based on 10,000 random samples of 300 return realizations from the true F . Table 2 reports the rejection rates for nominal significance levels of 2.5%, 5%, and 10%. The rejection rate measures the relative frequency of samples in which F , $\hat{G}$, or $\hat{Q}$ is classified as significantly better than $\hat{F}$ at the assumed significance level.

The rejection rates for $\text{LRT}(F, \hat{F})$ are slightly above the nominal significance level if $\hat{F}$ is unbiased and precise $((\delta_0, \delta_1) = (0, 1))$. This finding is reassuring about the statistical test size, given that F_∞ lies in the null for $\delta_0 = 0$.

The rejection rate increases significantly if $\hat{F}$ is estimated with lower accuracy. The high rejection rate for $\delta_0 = 5$ is reassuring about the statistical test power, given that F_∞ lies deep in the alternative in this case. The rejection rate for $\text{LRT}(\hat{G}, \hat{F})$ is even higher than for $\text{LRT}(F, \hat{F})$, as a result of the relatively high correlation between the likelihood scores of $\hat{F}$ and $\hat{G}$.

Consistent with the high $\text{KL}(\hat{Q} || F)$ in Table 1, rejection rates are very low for $\text{LRT}(\hat{Q}, \hat{F})$ across all specifications, reflecting the inherent pessimistic bias of $\hat{Q}$.

[Insert Table 2 about here.]

5 Empirical Analysis

The analysis now shifts to real-life stock index returns and option price quotes. The focus is on density forecasting for monthly SPX returns and optimization of combinations of monthly CBOE SPX options using various prior and posterior density forecasts.

5.1 Data

The primary source for SPX option price data is intra-day quotes captured at 15:45 ET from the iVolatility database, available from January 2004 through January 2023. We backfill these intra-day quotes with closing price data from OptionMetrics IvyDB, covering the period from January 1996 to December 2003. We include options with 29 calendar days to expiry.

Before January 2016, SPX option expiry dates were recorded as the Saturday following the third Friday of the month, although the Friday index level was used for settlement. We use the time to the settlement value in our time-to-expiry estimation. In four months, the sample trade dates fell on Thanksgiving Day, when the options exchange was closed (22 November 2001, 2007, 2012, and 2018). The absence of quotes on these days leaves $M = 321$ trade dates with valid option data for analysis.

We apply a series of filters based on moneyness, option delta, and option premium to select options for inclusion in the analysis. A first, more permissive set of filters defines the investable universe of N tradable options, while a second, more restrictive set of filters specifies the subset of N_0 options for which the moment conditions are imposed in the posterior density estimation.

Following standard conventions in empirical research, we exclude in-the-money (ITM) options, which tend to be less liquid than out-of-the-money (OTM) options. We also impose a minimum delta of 0.15 for calls and a maximum delta of -0.15 for puts, in order to exclude deep OTM options. **Additionally, we exclude options with bid prices below 15 cents.**

Table 3 presents summary statistics for the options that are filtered in. Separate results are provided for puts and calls. The table reports the quartile boundaries (p25, p50, and p75), computed across all $M = 321$ monthly strips, for the cross-sectional breadth (N_P for puts and N_C for calls), as well as the average and standard deviation (computed across all N_P or N_C qualified options) of moneyness, IV, and the bid-ask spread.

Consistent with the maturation of the options market, the cross-sectional breadth increases and quoted spreads decline over time. In the most recent five years, the number of qualified option series ($N_P + N_C$) exceeds 100, and the quoted spread falls below 1% of the midpoint premium for both puts and calls. These patterns suggest increasing market efficiency and declining profitability of active option trading over time. They also support the calibration of option breadth and bid-ask spreads used in the Monte Carlo simulations in Section 4.

[Insert Table 3 about here.]

To characterize general market conditions, our analysis also uses data on the

risk-free rate, dividends paid on the index constituent stocks, and the overall level of market volatility. The one-month T-bill rate and dividend yield are obtained from OptionMetrics IvyDB, while S&P 500 Index and VIX values are sourced from Thomson Reuters Eikon.

5.2 Prior Density Estimators

To generate prior density estimates ($\hat{F}$), we apply three common methods for estimating the conditional distribution of stock index returns:

$\hat{F}_{\text{LN}}$: a lognormal distribution with conditional mean and volatility,

$\hat{F}_{\text{ECDF}}$: a transformation of the historical unconditional ECDF, and

$\hat{F}_{\text{FHS}}$: a forecast density generated using a GJR-GARCH model with returns simulated via Filtered Historical Simulation (FHS).

The conditional means for all three forecast distributions, and the conditional standard deviation for $\hat{F}_{\text{LN}}$ and $\hat{F}_{\text{ECDF}}$, are estimated using arguments from [Constantinides et al. \(2009\)](#). Specifically, the conditional mean is based on the prevailing one-month Treasury bill rate (R_B), the estimated market risk premium ($\hat{\text{MRP}}$), and the prevailing dividend yield (R_D), such that $\mathbb{E}_{\hat{F}}[R_S] = R_B + \hat{\text{MRP}} - R_D$.

We set the annualized $\hat{\text{MRP}}$ to 4%, and verify that that our results and conclusions are robust to plausible variations in this parameter—presumably because option prices are more sensitive to the scale and shape of the return distribution than to its location.

Conditional volatility is proxied by the implied volatility (IV) of the nearest-to-the-money call option. Our results and conclusions about information projection are robust to reasonable variations in the estimator of conditional standard deviation. However, it should be noted that the CIPDE problem may fail to admit a solution if an excessively poor prior is used.

The conditional ECDF $\hat{F}_{\text{ECDF}}$ is derived by transforming the unconditional ECDF computed over a rolling window of 800 monthly returns before each trade date. This broad window is chosen to capture sufficiently many historical tail events. The ECDF is build on a regular grid of return values in the range $[-0.6, 0.6]$, spaced at 25 basis points. Translation and scale transformations are applied to the atoms of the unconditional ECDF to match the target mean $\mathbb{E}_{\hat{F}}[R_S]$ and volatility (IV). Using this approach, the conditional distribution preserves the shape of the unconditional distribution.

The GJR-GARCH FHS model is implemented following [Barone-Adesi et al. \(2020\)](#).¹ On each trade date, the model is fitted using daily return data in an 800-month historical moving window up to the trade date. A set of 10,000 return paths is then simulated, with drift adjusted to the target mean, and the conditional ECDF is estimated from the simulated series using a Gaussian kernel smoothing function with bandwidth set using Silverman’s rule.

To facilitate OOS evaluation, the discrete forecast distributions—namely, the transformed unconditional ECDF and the conditional ECDF from FHS—are smoothed using kernel density estimation to yield continuous densities: $p_h^*(S) := \frac{1}{h} \sum_{i=1}^N \pi_i^* k\left(\frac{S_i - S}{h}\right)$, where k is a Gaussian kernel and h is the bandwidth chosen via Silverman’s rule. Smoothing is essential because the OOS realized returns almost surely do not coincide with atoms in the discrete forecast distributions. In contrast, $\hat{F}_{\text{LN}}$ already defines a smooth density and thus requires no additional smoothing.

All three priors satisfy our statistical theory. Under standard regularity conditions—including stationarity, ergodicity, and mixing—their estimated densities converge uniformly in probability to pseudo-true limits. For $\hat{F}_{\text{LN}}$, convergence follows if the estimated mean and volatility converge uniformly. For $\hat{F}_{\text{ECDF}}$, convergence derives from the Glivenko–Cantelli theorem applied to the rolling-window empirical distribution, with affine transformations preserving uniform convergence. For $\hat{F}_{\text{FHS}}$, provided the GARCH model is pseudo-consistently estimated and the empirical residuals are ergodic, the smoothed ECDF also converges uniformly to a pseudo-true limit. Hence, Assumption 1 is satisfied in all three cases.

Regarding Assumption 3, and under regularity conditions, $\hat{F}_{\text{LN}}$ exhibits $\sqrt{T}$ -rate weak convergence via the delta method, since its parameters (mean and variance) converge at parametric rates. $\hat{F}_{\text{ECDF}}$ satisfies a Donsker-type weak convergence result under standard assumptions (e.g., strong mixing and a Lipschitz-continuous kernel), as discussed in [Mojirsheibani \(2006\)](#); the use of a fixed bandwidth preserves this rate. For $\hat{F}_{\text{FHS}}$, the GARCH volatility model is estimated via quasi-maximum likelihood, which guarantees $\sqrt{T}$ -consistency under suitable conditions, and the simulation of 10,000 paths regularizes the empirical forecast distribution. Uniform-type empirical process convergence results, under stationarity and mixing, justify the assumption; see, for example, [Rao and Krishnaiah \(1988\)](#).

Table 4 compares the three prior density estimators based on quartile boundaries (p_{25} , p_{50} , and p_{75}), computed over $M = 321$ months, for the first four central moments, 99% Value-at-Risk (VaR) and 99% Value-at-Gain (VaG), and pairwise KL divergences.

$\hat{F}_{\text{ECDF}}$ and $\hat{F}_{\text{FHS}}$ tend to exhibit pronounced negative skewness and fat tails, consistent with empirical stylized facts. These features are most pronounced in $\hat{F}_{\text{FHS}}$,

¹ We thank Carlo Sala for kindly providing the code to implement this method.

which helps explain why the largest divergence is observed between $\hat{F}_{\text{LN}}$ and $\hat{F}_{\text{FHS}}$. Despite their negative skewness and excess kurtosis, $\hat{F}_{\text{ECDF}}$ and $\hat{F}_{\text{FHS}}$ tend to have less downside risk and upside potential than $\hat{F}_{\text{LN}}$ due to lower standard deviation or higher peakedness.

The bottom row shows the number of months in which prior distribution violates the conditional moment conditions (M_{Inf}). In these months, updating is required to obey the moment conditions. Violations are most frequent for $\hat{F}_{\text{LN}}$ (96% of the months) and least frequent for $\hat{F}_{\text{ECDF}}$ (61%).

The prior density estimates fix or constrain the shape and parameters of the return distribution using limited conditioning information (R_B , R_D , and IV). These simplifying features make the priors potentially biased and inefficient, allowing for improvements via additional conditional moment conditions based on index option market prices.

[Insert Table 4 about here.]

5.3 Posterior Density Estimation

The prior distributions are updated to posterior distributions using the moment conditions (8)–(11), applied to all options that pass the data filters for pricing efficiency. For the risk-neutral posterior, the SDF is restricted to $m(S_T) = (1 + R_B)^{-1}$. Table 5 summarizes the properties of the posterior physical distributions ($\hat{G}$) and the posterior risk-neutral distributions ($\hat{Q}$), using the same format as Table 4.

The top panel reports results for the three posterior physical distributions ($\hat{G}_{\text{LN}}$, $\hat{G}_{\text{ECDF}}$, $\hat{G}_{\text{FHS}}$). Relative to their priors ($\hat{F}_{\text{LN}}$, $\hat{F}_{\text{ECDF}}$, $\hat{F}_{\text{FHS}}$), they tend to display more negative skewness and greater excess kurtosis. Given the typical smiles and smirks of the IV curve for SPX options, it is unsurprising that the skewness and kurtosis of the lognormal $\hat{F}_{\text{LN}}$ are substantially revised. Even for the ECDF and FHS specifications, the prior skewness and kurtosis tend to be too moderate to align with option prices. For $\hat{G}_{\text{LN}}$, the increase in negative skewness and excess kurtosis is accompanied by both lower downside risk and reduced upside potential, reflecting a combination of lower standard deviation and increased peakedness. Given these large revisions in central moments and tail risk, it is unsurprising that the divergence $\text{KL}(\hat{G}, \hat{F})$ is largest for the lognormal specification.

The last row reports the number of months in which a feasible posterior distribution could not be found (M_{Inf}), in which case the prior was retained as the posterior. This occurred in only one out of $M = 321$ months, indicating that the moment conditions are generally consistent.

The bottom panel presents results for the posterior risk-neutral distributions ($\hat{Q}$).

As expected, they are more pessimistic than the corresponding physical posteriors, exhibiting lower means, higher standard deviations, more negative skewness, greater excess kurtosis, higher downside risk, and lower upside potential. The risk-neutral posteriors also diverge more strongly from their priors than the physical posteriors do across all three specifications. Although these adjustments are consistent with the prices of the filtered options by construction, they are considerably larger than would be required if a strictly risk-averse SDF were assumed.

[Insert Table 5 about here.]

5.4 Predictive Ability

To evaluate the improvements in forecasting ability achieved by the posteriors, the LR test discussed in Section ?? is applied based on the realized index return on the $M = 321$ option expiry dates. In addition to comparing posteriors with their corresponding priors (e.g., $\text{LRT}(\hat{G}_{\text{LN}}, \hat{F}_{\text{LN}})$ and $\text{LRT}(\hat{Q}_{\text{LN}}, \hat{F}_{\text{LN}})$), comparisons are also made across priors (e.g., $\text{LRT}(\hat{F}_{\text{ECDF}}, \hat{F}_{\text{LN}})$, $\text{LRT}(\hat{G}_{\text{ECDF}}, \hat{F}_{\text{LN}})$, and $\text{LRT}(\hat{Q}_{\text{ECDF}}, \hat{F}_{\text{LN}})$). Table 6 summarizes the test results.

Among the three priors, $\hat{F}_{\text{LN}}$ exhibits the lowest predictive ability—significantly worse than both $\hat{F}_{\text{ECDF}}$ and $\hat{F}_{\text{FHS}}$. This finding is consistent with the counterfactual moments and tail risk of the lognormal reported in Table 4, as well as with the large divergence between $\hat{F}_{\text{LN}}$ and its posterior counterpart $\hat{G}_{\text{LN}}$ shown in Table 5, reflecting its low consistency with the prices of filtered-in options.

The posterior physical distribution $\hat{G}$ provides a significantly better density forecast than the corresponding prior $\hat{F}$ for all three specifications (LN, ECDF, FHS), as indicated by the low p -values for $\text{LRT}(\hat{G}_{\text{LN}}, \hat{F}_{\text{LN}})$, $\text{LRT}(\hat{G}_{\text{ECDF}}, \hat{F}_{\text{ECDF}})$, and $\text{LRT}(\hat{G}_{\text{FHS}}, \hat{F}_{\text{FHS}})$. This pattern aligns with the theoretical benefits of information projection and the assumption that the option pricing system holds for the filtered options.

The most pronounced improvements are observed for the lognormal prior $\hat{F}_{\text{LN}}$, reflecting its relatively poor forecasting ability. The improvements remain statistically significant for $\hat{F}_{\text{FHS}}$, which exhibits the highest predictive performance among the three priors, as evidenced by the p -value for $\text{LRT}(\hat{G}_{\text{FHS}}, \hat{F}_{\text{FHS}})$.

The posterior risk-neutral distribution $\hat{Q}$ also improves upon the prior $\hat{F}$ for all three specifications. However, these improvements are less pronounced than those for the posterior physical distribution $\hat{G}$, reflecting the intrinsic pessimistic bias of the risk-neutral distribution.

[Insert Table 6 about here.]

5.5 Investment Performance

Table 8 summarizes the composition and performance of the optimized option combinations for the various priors and posteriors. As discussed in Section 3.4, the option combinations are obtained by maximizing net premium income subject to the dominance constraints and the position limit.

The top panel reports the number of formation dates on which arbitrage opportunities were detected and active (non-zero) option positions were taken. The next three panels summarize the composition and characteristics of the open option positions, using the median (across all non-zero solutions) of the number of options and their weighted-average moneyness and implied volatility for bought calls, written calls, bought puts, and written puts.

The fifth and sixth panels evaluate performance in-sample and out-of-sample. The fifth panel reports medians of descriptive statistics (mean, standard deviation, skewness, and 99% CVaR) of the projected distributions of the index and the option-enhanced portfolio, including the median net premium income (the objective function). The sixth panel summarizes realized investment performance using the average return and the CER for $RRA = 1, 4$, and 10 .

Finally, the bottom panel reports p -values for SSD tests. Given the modest length of the time series of realized monthly returns ($M = 321$), we employ the Empirical Likelihood Ratio (ELR) test proposed by [Arvanitis and Post \(2024, Subsection 4.2\)](#). The null hypothesis is that the option-enhanced portfolio dominates the index by SSD out of sample, since the enhanced portfolio was explicitly engineered to achieve this outcome.²

Under the prior distributions $(\hat{F}_{LN}, \hat{F}_{ECDF}, \hat{F}_{FHS})$, option positions are opened in the large majority of months. Contrary to standard risk-reduction strategies, the optimal combinations involve substantial purchases of calls and sales of puts. Moreover, realized performance falls significantly short of projected performance and trails the index. A similar pattern was also observed by [Beare et al. \(2025\)](#) for optimization based on the lognormal model.

The composition, divergence, and underperformance likely reflect the combined effects of inaccuracies in the prior density forecasts, relatively high pricing efficiency in the index option market, and the impact of bid–ask spreads. The performance deterioration is largest for the lognormal prior forecast, $\hat{F}_{LN}$, in which case the SSD hypothesis is decisively rejected, and smallest for the GARCH prior forecast, $\hat{F}_{FHS}$,

² An alternative approach to test the same hypothesis, based on resampling Kolmogorov–Smirnov or Cramér–von Mises-type statistics, provides consistent inference under very general conditions (see, e.g., [Linton et al. \(2005, 2010\)](#)). However, such resampling-based methods tend to be less powerful in settings with relatively short time series, as is typical for a few decades of index option data.

consistent with the notion that investment performance is driven by forecasting ability, as shown in Table 6.

Under the posterior physical distributions ($\hat{G}_{\text{LN}}$, $\hat{G}_{\text{ECDF}}$, $\hat{G}_{\text{FHS}}$), it is substantially more difficult to detect arbitrage opportunities, and the number of non-zero solutions drops to less than one-third of the months. In addition, covered calls become more frequent due to the reduction in projected upside potential. Realized investment performance improves markedly relative to the priors, and outperformance of the index is achieved with a perfect dominance relation out of sample (p -value = 1) in each of the three cases. A discernible pattern among the three posteriors is that the use of covered calls becomes more common, the strike prices of the covered calls are higher, and the investment performance is stronger for the posteriors with the highest forecasting ability, especially $\hat{G}_{\text{FHS}}$.

The performance improvements highlight the economic significance of the forecast gains achieved through conditional information projection. The results also validate the dichotomy introduced by the option filters: the prices of the options that were filtered were sufficiently informative about the conditional distribution of index returns to identify pricing inefficiencies among the remaining, filtered-out options.

A very different picture emerges for the posterior risk-neutral distributions ($\hat{Q}_{\text{LN}}$, $\hat{Q}_{\text{ECDF}}$, $\hat{Q}_{\text{FHS}}$). In the majority of months, active positions are again opened. Due to the strong pessimistic bias of these forecasts, protective puts are bought and covered calls are written in most months, leading to a substantial reduction in downside risk. The option-enhanced portfolio, however, strongly underperforms the index by average realized return and CER_1 , owing to reduced equity exposure and the associated decline in the risk premium, as well as transaction costs from bid–ask spreads. As expected, SSD is strongly rejected. These findings highlight that a pessimistic forecast bias can lead to economically significant decision errors.

[Insert Table 8 about here.]

5.6 Robustness Analysis

To evaluate the robustness of our results and conclusions, we analyze the effects of excluding months with extreme economic circumstances from the time series and excluding illiquid deep OTM options from the monthly option strips.

We consider three alternative time-series subsamples: (i) a subsample excluding the 34 months where the US economy was classified as being in recession by the NBER; (ii) a subsample excluding the 32 months where the financial uncertainty index of [Jurado et al. \(2015\)](#) was above its 90th percentile level; (iii) a subsample excluding the 32 months where the VIX was above its 90th percentile level for the

sample.

In addition, we consider a more restrictive option filter for the conditional moment conditions used to estimate the posterior forecasts. Specifically, the OTM delta filter is tightened to exclude OTM options with absolute delta in the range $[0, 0.15]$ instead of $[0, 0.05]$. **Is the filter for optimization also changed based on $[x, 1-x]$? How do these changes affect the median and IQR of the number of options used for estimation and for optimization?**

Table 9 summarizes the LRT results for the alternative samples. The improvements in forecasting accuracy from conditional information projection are highly robust to the exclusion of extreme months and deep OTM options. Across all four sub-samples and three prior forecast distributions, the posterior forecast distribution consistently enhances the OOS goodness of fit.

[Insert Table 9 about here.]

Table 10 summarizes the OOS investment performance of the option-enhanced strategies after the exclusions. The performance improvements are robust for $\hat{F}_{LN}$ and $\hat{F}_{FHS}$. However, they are less robust for $\hat{F}_{ECDF}$, which is unsurprising given the low frequency of option position openings and low LWS10 p-value under this prior.

[Insert Table 10 about here.]

6 Conclusions

This study has introduced CIPDE to estimate conditional densities by integrating prior estimates with conditional moment conditions with functional nuisance parameters. The prior density estimate is updated through information projection onto the set of distributions that satisfy the moment conditions. Theoretical analysis and Monte Carlo simulations demonstrate that CIPDE achieves lower relative entropy to the latent conditional distribution when the prior is inconsistent and the moment conditions hold.

In option market research, conditional density estimates are essential due to the dynamic and non-Gaussian nature of the index return distribution and the short maturity of the most liquid series. Our empirical analysis shows that conditional information projection enhances the forecasting performance of standard conditional density estimators used in this domain.

The conditional moment conditions incorporate observed SPX option prices and general option pricing restrictions that rule out SD relations for filtered-in options. CIPDE, based on these moment conditions, significantly improves OOS forecasting accuracy. It also enhances investment performance by improving the timing and composition of option combinations. The improvements critically depend

on incorporating the SDF as a parameter in the moment conditions to avoid the pessimistic bias of the risk-neutral density. Robustness checks confirm the stability of the improvements across alternative subsamples that exclude extreme economic conditions and illiquid options.

The documented forecast improvements support the truth of the moment conditions based on the CJP pricing system. At the same time, our objective is not to establish the overall pricing efficiency of the index option market, and on this broader issue our results remain inconclusive. Two distinctions are important here: first, between the stricter estimation filters and the weaker investment filters; and second, between avoiding underperformance and achieving outperformance.

The forecast improvements support pricing efficiency for the options that pass the estimation filters. However, they are also consistent with the equilibrium system being only approximately—rather than exactly—correct, provided that pricing errors are rare or small for the filtered options. Moreover, while the optimized option positions based on the posterior physical distribution clearly improve on those based on the prior or the risk-neutral distribution—thereby avoiding the underperformance of the latter—the outperformance relative to the market index remains limited to tens of basis points per annum. Finally, for the options that pass the investment filters but not the estimation filters, tradability remains uncertain once slippage risk and market impact are taken into account.

Our statistical theory is designed to work with conditioning variable ranges in arbitrary Euclidean spaces. In practice, the curse of dimensionality is expected to kick in for dimensions greater than two. Recent advances in non-parametric statistics have extended projection methods to high-dimensional settings (potentially diverging with the sample size) by leveraging structured regularization techniques based on Hilbert norms (see, for example, [Wainwright \(2014\)](#) and the references therein). The consideration of such technologies in our framework seems like an interesting path for further research.

References

- ADAMS, R. A. AND J. J. FOURNIER (2003): *Sobolev Spaces*, vol. 140 of *Pure and Applied Mathematics*, Academic Press, 2nd ed.
- ALMEIDA, C. AND R. GARCIA (2012): “Assessing Misspecified Asset Pricing Models with Empirical Likelihood Estimators,” *Journal of Econometrics*, 519–537.
- AMISANO, G. AND R. GIACOMINI (2007): “Comparing density forecasts via weighted likelihood ratio tests,” *Journal of Business & Economic Statistics*, 25, 177–190.

- ANDREWS, D. W. (1999): “Estimation when a parameter is on a boundary,” *Econometrica*, 67, 1341–1383.
- ARVANITIS, S. AND T. POST (2024): “Generalized stochastic arbitrage opportunities,” *Management Science*, 70, 4629–4648.
- BAKSHI, G. AND D. MADAN (2006): “A theory of volatility spreads,” *Management Science*, 52, 1945–1956.
- BARONE-ADESI, G., N. FUSARI, A. MIRA, AND C. SALA (2020): “Option market trading activity and the estimation of the pricing kernel: A Bayesian approach,” *Journal of Econometrics*, 216, 430–449.
- BASSO, A. AND P. PIANCA (1975): “Decreasing Absolute Risk Aversion And Option Pricing,” *Management Science*, 43, 206—216.
- BAWA, V. (1975): “Optimal Rules for Ordering Uncertain Prospects,” *Journal of Financial Economics*, 2, 95–121.
- BEARE, B. K., J. SEO, AND Z. ZHENG (2025): “Stochastic arbitrage with market index options,” *Journal of Banking & Finance*, 173.
- BICKEL, P. J., C. A. KLAASSEN, Y. RITOV, AND J. A. WELLNER (1998): *Efficient and Adaptive Estimation for Semiparametric Models*, Springer.
- BLASQUES, F., S. J. KOOPMAN, AND A. LUCAS (2018): “Information-Theoretic Optimality of Observation-Driven Time Series Models for Likelihood Inference,” *Journal of Econometrics*, 202, 112–131.
- BLUNDELL, R., J. L. HOROWITZ, AND M. PAREY (2012): “Nonparametric Engel curves and revealed preference,” *Econometrica*, 80, 197–229.
- CONSTANTINIDES, G. M., M. CZERWONKO, AND S. PERRAKIS (2020): “Mispriced index option portfolios,” *Financial Management*, 49, 297–330.
- CONSTANTINIDES, G. M., J. C. JACKWERTH, AND S. PERRAKIS (2009): “Mispricing of S&P 500 Index Options,” *Review of Financial Studies*, 22, 1247–1277.
- CORTISSOZ, J. (2007): “On the Skorokhod Representation Theorem,” *Proceedings of the American Mathematical Society*, 135, 3995–4007.
- CSISZÁR, I. (1975): “I-divergence Geometry of Probability Distributions and Minimization Problems,” *The Annals of Probability*, 3, 146–158.

- DEMIGUEL, V., L. GARLAPPI, F. J. NOGALES, AND R. UPPAL (2009): “A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms,” *Management Science*, 55, 798–812.
- DIEBOLD, F. X. AND R. S. MARIANO (2002): “Comparing predictive accuracy,” *Journal of Business & economic statistics*, 20, 134–144.
- FEYDY, J., T. SÉJOURNÉ, F.-X. VIALARD, S.-I. AMARI, A. TROUVÉ, AND G. PEYRÉ (2019): “Interpolating between optimal transport and mmd using sinkhorn divergences,” in *The 22nd international conference on artificial intelligence and statistics*, PMLR, 2681–2690.
- JEWSON, J., J. Q. SMITH, AND C. HOLMES (2018): “Principled Bayesian minimum divergence inference,” *Entropy*, 20, 442.
- JURADO, K., S. C. LUDVIGSON, AND S. NG (2015): “Measuring uncertainty,” *American Economic Review*, 105, 1177–1216.
- KANG, B.J., T. K. AND S.-J. YOON (2010): “Information content of volatility spreads,” *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 30, 533–558.
- KIMBALL, M. S. (1990): “Precautionary Saving in the Small and in the Large,” *Econometrica*, 58, 53–73.
- KNIGHT, K. (1999): “Epi-convergence in distribution and stochastic equi-semicontinuity,” *Unpublished manuscript*.
- KOMUNJER, I. AND G. RAGUSA (2016): “Existence and characterization of conditional density projections,” *Econometric Theory*, 32, 947–987.
- LI, X. AND L. H. ZHAO (2002): “Bayesian nonparametric point estimation under a conjugate prior,” *Statistics & Probability Letters*, 58, 23–30.
- LINTON, O., E. MAASOUMI, AND Y.-J. WHANG (2005): “Consistent testing for stochastic dominance under general sampling schemes,” *Review of Economic Studies*, 72, 735–765.
- LINTON, O., K. SONG, AND Y.-J. WHANG (2010): “An improved bootstrap test of stochastic dominance,” *Journal of Econometrics*, 154, 186–202.

- MOJIRSHEIBANI, M. (2006): “A Note on the Strong Approximation of the Smoothed Empirical Process of α -mixing Sequences,” *Statistical inference for stochastic processes*, 9, 97–107.
- NIELSEN, F. (2020): “An Elementary Introduction to Information Geometry,” *Entropy*, 22.
- POST, T. (2003): “Empirical tests for stochastic dominance efficiency,” *The Journal of Finance*, 58, 1905–1931.
- POST, T., S. KARABATI, AND S. ARVANITIS (2018): “Portfolio optimization based on stochastic dominance and empirical likelihood,” *Journal of Econometrics*, 206, 167–186.
- POST, T. AND M. KOPA (2017): “Portfolio choice based on third-degree stochastic dominance,” *Management Science*, 63, 3381–3392.
- POST, T. AND I. R. LONGARELA (2021): “Risk arbitrage opportunities for stock index options,” *Operations Research*, 69, 100–113.
- POST, T. AND V. POTÌ (2017): “Portfolio analysis using stochastic dominance, relative entropy, and empirical likelihood,” *Management Science*, 63, 153–165.
- RAO, M. S. AND Y. R. KRISHNAIAH (1988): “Weak convergence of empirical processes of strong mixing sequences under estimation of parameters,” *Sankhyā: The Indian Journal of Statistics, Series A*, 26–43.
- ROCKAFELLAR, R. T. AND S. URYASEV (2000): “Optimization of conditional value-at-risk,” *Journal of Risk*, 2, 21–42.
- ROCKAFELLAR, R. T. AND R. J.-B. WETS (1998): *Variational Analysis*, vol. 317 of *Grundlehren der Mathematischen Wissenschaften*, Berlin, Heidelberg: Springer.
- SCHENNACH, S. M. (2005): “Bayesian Exponentially Tilted Empirical Likelihood,” *Biometrika*, 92, 31–46.
- STUTZER, M. (1995): “A Bayesian approach to diagnosis of asset pricing models,” *Journal of Econometrics*, 68, 367–397.
- (1996): “A simple nonparametric approach to derivative security valuation,” *The Journal of Finance*, 51, 1633–1652.
- TSYBAKOV, A. B. (2009): *Introduction to Nonparametric Estimation*, Springer.

- VAN DER VAART, A. W. (1998): *Asymptotic Statistics*, Cambridge University Press.
- VAN DER VAART, A. W. (1998): *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.
- VICKSON, R. (1975): “Stochastic Dominance Tests for Decreasing Absolute Risk Aversion. I. Discrete Random Variables.” *Management Science*, 21, 1438—1446.
- VUONG, Q. H. (1989): “Likelihood ratio tests for model selection and non-nested hypotheses,” *Econometrica: journal of the Econometric Society*, 307–333.
- WAINWRIGHT, M. J. (2014): “Structured regularizers for high-dimensional problems: Statistical and computational issues,” *Annual Review of Statistics and Its Application*, 1, 233–253.
- YILMAZ, H. (2021): “A generalization of multiplier rules for infinite-dimensional optimization problems,” *Optimization*, 70, 1825–1835.

δ_0	δ_1	N	w	KL($\hat{F} F$)			KL($\hat{G} F$)			KL($\hat{Q} F$)		
				p25	p50	p75	p25	p50	p75	p25	p50	p75
0	1	5	0.010	0.000	0.000	0.001	0.001	0.001	0.001	0.027	0.027	0.028
			0.002	0.000	0.000	0.001	0.001	0.001	0.002	0.029	0.029	0.029
		25	0.010	0.000	0.000	0.001	0.001	0.001	0.002	0.029	0.030	0.031
			0.002	0.000	0.000	0.001	0.001	0.002	0.003	0.032	0.032	0.033
	5	5	0.010	0.003	0.012	0.023	0.003	0.011	0.020	0.028	0.032	0.035
			0.002	0.003	0.012	0.023	0.003	0.009	0.016	0.029	0.030	0.034
		25	0.010	0.003	0.012	0.023	0.003	0.008	0.016	0.030	0.031	0.033
			0.002	0.003	0.012	0.023	0.002	0.007	0.014	0.032	0.033	0.034
5	1	5	0.010	0.031	0.038	0.045	0.025	0.031	0.036	0.036	0.037	0.039
			0.002	0.031	0.038	0.045	0.016	0.020	0.023	0.030	0.030	0.030
		25	0.010	0.031	0.038	0.045	0.016	0.019	0.023	0.030	0.031	0.032
			0.002	0.031	0.038	0.045	0.013	0.016	0.020	0.033	0.033	0.033
	5	5	0.010	0.010	0.040	0.081	0.008	0.032	0.063	0.030	0.038	0.043
			0.002	0.010	0.040	0.081	0.005	0.021	0.042	0.030	0.030	0.030
		25	0.010	0.010	0.040	0.081	0.005	0.020	0.042	0.030	0.031	0.032
			0.002	0.010	0.040	0.081	0.004	0.017	0.037	0.033	0.033	0.033

Tab. 1: Simulated Divergence. For each combination of the bias ($\delta_0 = 0, 5$), semi-range ($\delta_1 = 1, 5$), number of option series ($N = 5, 25$) and fractional bid-ask spread ($w = 0.010, 0.002$), 300 prior forecasts are generated based on 300 simulated RRA estimates $\hat{\gamma}_T$. Given the simulated prior and simulated option prices, the posterior is estimated using the CIPDE procedure described in Section 2.2. Quartile boundary values (p25, p50, p75) are reported for the KL divergence from the prior, $\hat{F}$, and the posterior, $\hat{G}$, to the true distribution F . The posterior forecast can be seen to move significantly closer (in terms of KL divergence) than the prior forecast to the true distribution in each case, with the exception of $(\delta_0, \delta_1) = (0, 1)$, where the prior is already very close to the true distribution. The KL divergence is shown to reduce with increased number of option series (25 vs. 5) and reduced spread (20 bps vs. 100 bps) and, for given semi-range and bias.

δ_0	δ_1	N	w	LRT($F, \hat{F}$)			LRT($\hat{G}, \hat{F}$)			LRT($\hat{Q}, \hat{F}$)		
				2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
0	1	5	0.010	0.040	0.060	0.140	0.020	0.040	0.070	0.000	0.000	0.000
			0.002	0.040	0.060	0.140	0.020	0.020	0.050	0.000	0.000	0.000
		25	0.010	0.040	0.060	0.140	0.020	0.030	0.060	0.000	0.000	0.000
			0.002	0.040	0.060	0.140	0.000	0.020	0.040	0.000	0.000	0.000
	5	5	0.010	0.290	0.410	0.590	0.100	0.200	0.280	0.000	0.000	0.010
			0.002	0.290	0.410	0.590	0.180	0.240	0.370	0.000	0.000	0.000
		25	0.010	0.290	0.410	0.590	0.180	0.290	0.370	0.000	0.000	0.000
			0.002	0.290	0.410	0.590	0.190	0.250	0.390	0.000	0.000	0.000
5	1	5	0.010	0.720	0.850	0.890	0.720	0.830	0.890	0.020	0.040	0.130
			0.002	0.720	0.850	0.890	0.880	0.910	0.960	0.050	0.080	0.170
		25	0.010	0.720	0.850	0.890	0.890	0.910	0.970	0.060	0.080	0.190
			0.002	0.720	0.850	0.890	0.890	0.930	0.980	0.050	0.080	0.150
	5	5	0.010	0.750	0.810	0.890	0.810	0.870	0.940	0.040	0.120	0.200
			0.002	0.750	0.810	0.890	0.940	0.950	0.970	0.070	0.160	0.280
		25	0.010	0.750	0.810	0.890	0.910	0.960	0.970	0.070	0.150	0.270
			0.002	0.750	0.810	0.890	0.940	0.960	0.970	0.070	0.110	0.250

Tab. 2: Simulated LRT Rejection Rates. For each combination of the bias ($\delta_0 = 0, 5$), semi-range ($\delta_1 = 1, 5$), number of option series ($N = 5, 25$) and fractional bid-ask spread ($w = 0.010, 0.002$), 300 random RRA estimates $\hat{\gamma}_T$ are generated from the uniform distribution. For each RRA estimate, the density estimators $\hat{F}$, $\hat{G}$ and $\hat{Q}$ constructed and their predictive ability is evaluated. LRT($F, \hat{F}$), LRT($\hat{G}, \hat{F}$) and LRT($\hat{Q}, \hat{F}$) are computed based on random samples of 10,000 return realizations from the true F . The Table reports the rejection rates for nominal significance levels of 2.5%, 5%, or 10%.

	Calls			Puts		
	p25	p50	p75	p25	p50	p75
# options	7	10	14	10	15	22
IV (Mean)	10.98	14.76	19.22	14.04	18.69	23.24
St. Dev.	0.53	0.73	1.03	1.19	1.43	1.92
Mnes (Mean)	101.50	101.99	102.83	96.66	97.39	97.85
St. Dev.	0.90	1.19	1.55	1.36	1.72	2.16
Spread (Mean)	7.03	10.10	12.28	6.21	8.88	10.79
St. Dev.	2.45	3.77	5.29	1.65	2.75	3.85

Tab. 3: Option strip statistics, estimated over 251 monthly call and put option strips (Jan. 1996 - Jan 2023). The sample includes OTM options with $\text{abs}(\text{delta}) \geq 0.15$ and bid price $> \$0.15$. The mean and standard deviation of the annualised implied volatility of the options in each monthly strip are reported at the quartile boundaries in the sample; MNES reports the average and st. dev. of the ratio of the option strip strikes to the underlying spot price; Spread captures the average and standard deviation of the option spread in percentage terms (of the midpoint price) across the option strips. All values are displayed as percentages.

	$\hat{F}_{\text{LN}}$			$\hat{F}_{\text{ECDF}}$			$\hat{F}_{\text{FHS}}$		
	p25	p50	p75	p25	p50	p75	p25	p50	p75
Mean (ann.)	0.022	0.043	0.084	0.022	0.044	0.084	0.023	0.044	0.086
St. Dev.(ann.)	0.120	0.161	0.211	0.124	0.166	0.216	0.113	0.147	0.198
Skewness	-0.000	-0.000	0.000	-0.617	-0.593	-0.506	-0.917	-0.810	-0.726
Kurtosis	3.000	3.000	3.000	5.397	5.677	5.960	5.589	6.334	7.321
VaR_{90}	0.042	0.055	0.070	0.040	0.052	0.068	0.037	0.048	0.064
VaG_{90}	0.047	0.062	0.082	0.042	0.054	0.071	0.040	0.053	0.069
$SemiDev_-$ (ann.)	0.080	0.107	0.136	0.089	0.118	0.151	0.082	0.107	0.144
$SemiDev_+$ (ann.)	0.094	0.121	0.160	0.090	0.116	0.154	0.081	0.106	0.136
$KL(\cdot \hat{F}_{\text{LN}})$	0	0	0	0.0208	0.0233	0.0285	0.0216	0.0375	0.0771
$KL(\cdot \hat{F}_{\text{ECDF}})$	0.0361	0.0394	0.0470	0	0	0	0.0190	0.0411	0.1132
$KL(\cdot \hat{F}_{\text{FHS}})$	0.0356	0.0474	0.0786	0.0152	0.0223	0.0462	0	0	0
M_{Inf}	307			194			238		

Tab. 4: Forecast mean, standard deviation, skewness, kurtosis and pairwise KL divergence (reported as $KL(\text{col}, \text{row})$), between the three prior density estimates. Quartile values are based on 321 forecast dates with valid option data (Jan. 1996–Jan. 2023).

	$\hat{G}_{\text{LN}}$			$\hat{G}_{\text{ECDF}}$			$\hat{G}_{\text{FHS}}$		
	p25	p50	p75	p25	p50	p75	p25	p50	p75
Mean (ann.)	0.052	0.078	0.098	0.031	0.053	0.074	0.024	0.043	0.071
St. Dev. (ann.)	0.102	0.140	0.187	0.119	0.159	0.206	0.116	0.148	0.195
Skewness	-0.590	-0.414	-0.268	-1.333	-0.976	-0.728	-1.417	-1.007	-0.784
Kurtosis	3.354	3.519	3.767	6.263	7.267	9.997	5.747	6.795	9.364
VaR_{90}	0.035	0.048	0.066	0.039	0.052	0.067	0.039	0.051	0.066
VaG_{90}	0.039	0.053	0.070	0.037	0.049	0.066	0.037	0.050	0.066
$SemiDev_-$ (ann.)	0.068	0.094	0.127	0.087	0.116	0.151	0.085	0.111	0.146
$SemiDev_+$ (ann.)	0.081	0.106	0.140	0.080	0.109	0.142	0.076	0.100	0.134
$KL(\cdot \hat{F}_{\text{LN}})$	0.0233	0.0414	0.0672	0.0163	0.0246	0.0380	0.0149	0.0326	0.0584
$KL(\cdot \hat{F}_{\text{ECDF}})$	0.0551	0.0751	0.1123	0.0035	0.0114	0.0237	0.0249	0.0507	0.1015
$KL(\cdot \hat{F}_{\text{FHS}})$	0.0473	0.0628	0.0970	0.0179	0.0277	0.0405	0.0035	0.0175	0.0404
M_{Inf}	1			1			1		
	$\hat{Q}_{\text{LN}}$			$\hat{Q}_{\text{ECDF}}$			$\hat{Q}_{\text{FHS}}$		
	p25	p50	p75	p25	p50	p75	p25	p50	p75
Mean (ann.)	-0.038	-0.022	0.007	-0.039	-0.022	0.007	-0.038	-0.020	0.009
St. Dev. (ann.)	0.127	0.168	0.214	0.136	0.179	0.231	0.140	0.183	0.229
Skewness	-1.100	-0.849	-0.641	-1.974	-1.629	-1.234	-2.476	-1.900	-1.444
Kurtosis	3.514	3.862	4.299	7.488	9.350	10.763	8.468	11.322	15.174
VaR_{90}	0.055	0.072	0.092	0.047	0.061	0.077	0.046	0.060	0.076
VaG_{90}	0.036	0.050	0.065	0.035	0.048	0.063	0.036	0.049	0.065
$SemiDev_-$ (ann.)	0.103	0.133	0.169	0.112	0.149	0.191	0.119	0.151	0.189
$SemiDev_+$ (ann.)	0.072	0.100	0.130	0.073	0.102	0.131	0.072	0.098	0.128
$KL(\cdot \hat{F}_{\text{LN}})$	0.0601	0.0931	0.1505	0.0354	0.0538	0.0774	0.0409	0.0679	0.1034
$KL(\cdot \hat{F}_{\text{ECDF}})$	0.1042	0.1441	0.1800	0.0131	0.0234	0.0404	0.0432	0.0724	0.1327
$KL(\cdot \hat{F}_{\text{FHS}})$	0.0810	0.1105	0.1518	0.0297	0.0431	0.0644	0.0197	0.0393	0.0657
M_{Inf}	1			1			1		

Tab. 5: Forecast mean, standard deviation, skewness, kurtosis and pairwise KL divergence between the three posterior density estimates (reported as $KL(\text{col}, \text{row})$), for the TSD ($\hat{G}$) and risk-neutral ($\hat{Q}$) kernels. Quartile values are based on 321 forecast dates with valid option data (Jan. 1996–Jan. 2023).

	$\hat{F}_{\text{LN}}$	$\hat{G}_{\text{LN}}$	$\hat{Q}_{\text{LN}}$	$\hat{F}_{\text{ECDF}}$	$\hat{G}_{\text{ECDF}}$	$\hat{Q}_{\text{ECDF}}$	$\hat{F}_{\text{FHS}}$	$\hat{G}_{\text{FHS}}$	$\hat{Q}_{\text{FHS}}$
LRT($\cdot, \hat{F}_{\text{LN}}$)	-	5.06	3.68	4.67	5.00	4.24	4.70	5.21	4.18
	(1.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LRT($\cdot, \hat{F}_{\text{ECDF}}$)	-4.67	0.88	0.98	-	3.27	2.32	1.23	2.90	2.29
	(1.00)	(0.19)	(0.16)	(0.00)	(0.01)	(0.11)	(0.00)	(0.01)	(0.01)
LRT($\cdot, \hat{F}_{\text{FHS}}$)	-4.70	-0.21	-0.18	-1.23	0.69	0.37	-	1.82	1.67
	(1.00)	(0.58)	(0.57)	(0.89)	(0.24)	(0.36)	(0.03)	(0.05)	(0.05)

Tab. 6: LRT statistics: test statistics are reported for $\text{LRT}(\text{col}, \text{row})$, with p-values in brackets. For example, for the lognormal prior, the LRT of the posterior, $\hat{G}_{\text{LN}}$, to the prior, $\hat{F}_{\text{LN}}$, is 5.06.

	Index	$\tilde{F}_{LN}$			$\tilde{F}_{ECDF}$			$\tilde{F}_{FHS}$		
		$\hat{F}$	$\hat{G}$	$\hat{Q}$	$\hat{F}$	$\hat{G}$	$\hat{Q}$	$\hat{F}$	$\hat{G}$	$\hat{Q}$
Prot. Puts		0.11	0.01	0.92	0.05	0.00	0.93	0.14	0.02	0.90
Avg Mnes		100.25	94.56	99.45	100.16	100.30	99.22	99.85	93.83	98.58
St.Dev. Mnes		0.66	8.12	1.15	0.54	0.00	1.66	0.89	7.43	2.23
Avg IV		0.24	0.46	0.18	0.19	0.15	0.18	0.19	0.40	0.19
St.Dev. IV		0.13	0.44	0.09	0.07	0.00	0.09	0.11	0.26	0.09
Cov. Calls		0.00	0.07	0.55	0.02	0.11	0.55	0.36	0.37	0.63
Avg Mnes		0.00	101.97	100.77	101.40	101.64	101.00	101.02	101.18	100.82
St.Dev. Mnes		0.00	2.03	1.01	1.37	1.77	1.34	1.02	1.26	1.02
Avg IV		0.00	0.26	0.17	0.30	0.23	0.18	0.19	0.20	0.18
St.Dev. IV		0.00	0.16	0.08	0.16	0.13	0.08	0.06	0.09	0.08
Ann. Mean (%)	7.88	6.19	7.13	1.38	7.60	7.80	2.56	7.69	7.79	3.06
Ann. St. Dev. (%)	17.58	16.52	16.80	7.37	17.04	16.81	7.65	14.82	15.66	7.54
t-test p-value		0.04	0.28	0.02	0.65	0.90	0.05	0.88	0.93	0.08
bootstrap p-value		0.04	0.28	0.02	0.64	0.90	0.05	0.87	0.92	0.08
Ann. CER ₁ (%)	6.24	4.73	5.61	1.11	6.05	6.28	2.26	6.51	6.46	2.77
t-test p-value		0.06	0.33	0.08	0.74	0.95	0.17	0.85	0.83	0.23
bootstrap p-value		0.06	0.32	0.08	0.74	0.94	0.16	0.84	0.83	0.23
Ann. CER ₂ (%)	4.52	3.17	4.00	0.84	4.41	4.67	1.98	5.27	5.05	2.49
t-test p-value		0.04	0.28	0.02	0.65	0.90	0.05	0.88	0.93	0.08
bootstrap p-value		0.04	0.28	0.02	0.64	0.90	0.05	0.87	0.92	0.08
Ann. CER ₄ (%)	0.58	-0.45	0.25	0.30	0.62	0.91	1.39	2.36	1.72	1.93
t-test p-value		0.04	0.28	0.02	0.65	0.90	0.05	0.88	0.93	0.08
bootstrap p-value		0.21	0.52	0.94	0.94	0.49	0.83	0.28	0.22	0.73
ELRT(Optstrat SDs Index,)		0.0697	0.2924	0.0394	0.7855	1.00	0.0818	0.9532	0.9998	0.1119
ELRT(Index SDs Optstrat)		0.0545	0.1046	0	0.0136	0	0	0	0	0

Tab. 7: Option Combinations: Composition and Performance. Returns to a portfolio consisting of the index and max one protective put and max one covered call, where the included options (if any) are selected through expected utility maximization, under the relevant forecast. The Prot. Puts row records the fraction of months that a protective put is included. The Cov. Calls row records the fraction of months that a covered call is sold. CIPDEs are estimated using OTM Puts and Calls (absolute delta range 0.15-0.5). Also reported is the ELR p-value for the null that the option-enhanced index dominates the passive index.

Count	Non-Zero	Index	$\hat{F}_{LN}$			$\hat{F}_{ECDF}$			$\hat{F}_{FHS}$		
			$\hat{F}$	$\hat{G}$	$\hat{Q}$	$\hat{F}$	$\hat{G}$	$\hat{Q}$	$\hat{F}$	$\hat{G}$	$\hat{Q}$
			303	62	240	235	53	226	256	107	218
Composition	Buy Call		0.34	0.22	0.00	0.30	0.18	0.00	0.29	0.00	0.00
	Sell Call		0.16	0.26	0.39	0.16	0.31	0.40	0.23	0.42	0.46
	Buy Put		0.21	0.21	0.56	0.24	0.22	0.56	0.20	0.24	0.50
	Sell Put		0.29	0.31	0.00	0.30	0.21	0.00	0.20	0.28	0.00
Moneyness	Buy Call		103.00	101.04	102.03	102.58	100.98	101.95	101.74	95.68	101.75
	Sell Call		95.22	98.63	95.49	96.19	99.16	96.11	98.23	99.73	97.30
	Buy Put		102.54	102.94	102.53	101.88	102.09	102.30	102.07	103.27	102.76
	Sell Put		95.34	96.87	95.62	96.02	97.58	97.12	97.05	97.08	98.04
Implied Vol.	Buy Call		13.75	19.34	11.78	13.68	19.37	14.43	16.27	19.98	13.87
	Sell Call		20.65	20.49	19.39	19.63	20.12	19.39	18.12	18.50	18.49
	Buy Put		13.89	16.20	13.30	13.96	16.93	13.67	13.56	15.73	13.63
	Sell Put		20.24	21.61	19.02	19.67	21.62	19.59	19.28	21.87	16.37
Projected Distr.	Passive Mean		7.78	10.59	1.14	7.89	8.43	1.42	7.85	7.38	1.57
	Passive St.Dev.		16.24	13.94	16.53	16.49	15.61	17.26	14.62	14.55	17.31
	Passive Skew		0.14	-0.25	-0.68	-0.28	-0.64	-1.21	-0.53	-0.71	-1.43
	Passive VaR		11.17	11.08	14.26	16.29	16.82	21.67	14.49	15.33	22.44
	Prem. Income		1.66	0.00	0.47	0.39	0.00	0.31	0.94	0.00	0.18
	Active Mean		9.19	10.68	1.96	8.51	8.43	2.03	8.83	7.63	2.04
	Active St.Dev.		14.16	13.46	6.95	15.02	15.37	8.75	12.67	13.80	10.30
	Active Skew		0.59	-0.25	-0.41	-0.07	-0.64	-1.28	-0.55	-0.94	-1.86
	Active VaR		9.83	11.04	6.59	15.67	16.71	12.30	13.43	15.18	14.80
OOS Performance	Avg Ret.	7.88	5.57	7.85	4.66	6.90	8.11	5.61	7.44	8.04	5.36
	CER ₁	6.24	4.36	6.25	4.00	5.55	6.51	4.95	6.16	6.52	4.64
	CER ₄	0.58	-0.04	0.71	1.81	0.77	0.95	2.83	1.73	1.21	2.30
	CER ₁₀	-18.89	-17.33	-18.56	-4.46	-17.03	-18.38	-2.71	-14.32	-17.64	-3.81
SSD Test	p-value		0	1.00	0.15	0.14	1.00	0.30	0.67	1.00	0.22

Tab. 8: Option Combinations: Composition and Performance. SSD, $[abs(delta) > x, abs(delta) < 1 - x]$, $x = 0.15$, bid price > 0.15 .

	$\hat{F}_{LN}$	$\hat{G}_{LN}$	$\hat{Q}_{LN}$	$\hat{F}_{ECDF}$	$\hat{G}_{ECDF}$	$\hat{Q}_{ECDF}$	$\hat{F}_{FHS}$	$\hat{G}_{FHS}$	$\hat{Q}_{FHS}$
Excl. High VIX									
LRT($\cdot, \hat{F}_{LN}$)	-	3.89***	2.92***	3.84***	4.06***	3.46***	3.94***	4.34***	3.54***
LRT($\cdot, \hat{F}_{ECDF}$)	-3.84	0.29	0.61	-	2.53**	1.85**	1.05	2.37**	2.02**
LRT($\cdot, \hat{F}_{FHS}$)	-3.94	-0.53	-0.31	-1.05	0.45	0.24	-	1.49*	1.54*
Excl. Recessions									
LRT($\cdot, \hat{F}_{LN}$)	-	4.72***	3.33***	4.33***	4.60***	3.79***	4.76***	5.03***	4.11***
LRT($\cdot, \hat{F}_{ECDF}$)	-4.33	0.96	0.69	-	3.01***	1.80**	1.34*	3.24***	2.19**
LRT($\cdot, \hat{F}_{FHS}$)	-4.76	-0.02	-0.29	-1.34	0.86	0.26	-	1.46*	0.65
Excl. High Fin. Unc.									
LRT($\cdot, \hat{F}_{LN}$)	-	4.62***	3.45***	4.31***	4.54***	3.86***	4.60***	4.91***	4.14***
LRT($\cdot, \hat{F}_{ECDF}$)	-4.31	0.63	0.96	-	2.81***	2.04**	0.88	2.69***	2.44**
LRT($\cdot, \hat{F}_{FHS}$)	-4.60	-0.01	0.21	-0.88	1.04	0.70	-	1.50*	1.15
Alt. Filters									
LRT($\cdot, \hat{F}_{LN}$)	-	4.49***	3.49***	4.67***	4.76***	4.06***	4.70***	4.96***	4.04***
LRT($\cdot, \hat{F}_{ECDF}$)	-4.67	0.34	0.32	-	2.78***	1.78**	1.23	2.43**	2.06**
LRT($\cdot, \hat{F}_{FHS}$)	-4.70	-0.61	-0.66	-1.23	0.42	0.05	-	1.38*	1.43*

Tab. 9: LRT Robustness Analysis: LR tests are repeated in three alternative samples, excluding (i) the 10% of months with the highest VIX on the trade date (ii) all months the US economy is classified as being in recession by the NBER and (iii) the 10% of months with the highest financial uncertainty on the trade date. The Alt. Filters results are generated using a more restrictive filter on OTM options where $\text{abs}(\text{delta})$ of included options must be in the range $[0.2, 0.5]$.

		$\hat{F}_{LN}$			$\hat{F}_{ECDF}$			$\hat{F}_{FHS}$		
	Index	$\hat{F}$	$\hat{G}$	$\hat{Q}$	$\hat{F}$	$\hat{G}$	$\hat{Q}$	$\hat{F}$	$\hat{G}$	$\hat{Q}$
Excl. High VIX										
Ann. Mean (%)	4.16	3.03	4.34	2.84	3.76	4.60	3.68	4.06	4.76	3.55
Ann. St. Dev. (%)	15.99	13.99	15.78	10.59	14.69	15.80	10.60	14.48	15.35	11.23
CER ₁ (%)	2.80	4.36	6.25	4.00	5.55	6.51	4.95	6.16	6.52	2.89
CER ₄ (%)	-2.08	-0.04	0.71	1.81	0.77	0.95	2.83	1.73	1.21	0.71
CER ₁₀ (%)	-19.84	-17.33	-18.56	-4.46	-17.03	-18.38	-2.71	-14.32	-17.64	-5.13
ELRT p-value		0.00	1.00	0.15	0.14	1.00	0.30	0.67	1.00	0.22
Excl. Recessions										
Ann. Mean (%)	10.40	8.03	10.33	6.46	9.18	10.67	7.33	9.76	10.74	7.67
Ann. St. Dev. (%)	14.07	11.75	13.79	8.96	12.52	13.80	9.22	12.22	13.24	9.76
CER ₁ (%)	9.36	4.36	6.25	4.00	5.55	6.51	4.95	6.16	6.52	7.17
CER ₄ (%)	6.17	-0.04	0.71	1.81	0.77	0.95	2.83	1.73	1.21	5.71
CER ₁₀ (%)	-1.47	-17.33	-18.56	-4.46	-17.03	-18.38	-2.71	-14.32	-17.64	2.44
ELRT p-value		0.00	1.00	0.15	0.14	1.00	0.30	0.67	1.00	0.22
Excl. High Fin. Unc.										
Ann. Mean (%)	9.02	6.80	8.97	5.35	7.78	9.25	6.12	8.23	9.11	6.14
Ann. St. Dev. (%)	14.72	12.46	14.45	9.91	13.23	14.48	10.16	13.07	13.99	10.80
CER ₁ (%)	7.88	4.36	6.25	4.00	5.55	6.51	4.95	6.16	6.52	5.54
CER ₄ (%)	4.33	-0.04	0.71	1.81	0.77	0.95	2.83	1.73	1.21	3.67
CER ₁₀ (%)	-4.31	-17.33	-18.56	-4.46	-17.03	-18.38	-2.71	-14.32	-17.64	-0.82
ELRT p-value		0.00	1.00	0.15	0.14	1.00	0.30	0.67	1.00	0.22
Alt. Filters										
Ann. Mean (%)	7.88	5.57	7.72	4.06	6.90	7.81	5.01	7.44	8.09	5.56
Ann. St. Dev. (%)	17.58	15.07	17.02	12.05	15.94	17.12	12.11	15.45	16.87	11.30
CER ₁ (%)	6.24	4.36	6.18	3.24	5.55	6.26	4.21	6.16	6.58	4.90
CER ₄ (%)	0.58	-0.04	0.84	-0.18	0.77	0.87	1.16	1.73	1.29	2.83
CER ₁₀ (%)	-18.89	-17.33	-17.85	-21.99	-17.03	-17.84	-13.20	-14.32	-17.38	-2.18
ELRT p-value		0	0.91	0.08	0.14	1.00	0.18	0.67	1.00	0.30

Tab. 10: Investment Performance: Robustness Analysis. Option trading performance is evaluated in three alternative samples, excluding (i) the 10% of months with the highest VIX on the trade date (ii) all months the US economy is classified as being in recession by the NBER and (iii) the 10% of months with the highest financial uncertainty on the trade date. The Alt. Filters results are generated using a more restrictive filter on OTM options where abs(delta) of included options must be in the range [0.2,0.5]. **[What about the delta filters for optimization – are they updated to [0.2, 0.8]?]** Also reported is the ELR p-value for the null that the option-enhanced index dominates the passive index.