Empirical implications of the pricing kernel puzzle for the return on contingent claims

Brendan K. Beare and Lawrence D. W. Schmidt*

Department of Economics, University of California, San Diego

July 6, 2014

Abstract

A recent literature in empirical finance documents the nonmonotone shape of pricing kernel estimates for several major market indices. This unexpected phenomenon implies that the return on the market portfolio may be stochastically dominated by the return on a contingent claim written on the market portfolio. We investigate by using a multiobjective evolutionary algorithm to search for a portfolio of European put and call options written on the S&P 500 index whose return stochastically dominates the market return. Monthly out-of-sample returns of our selected portfolios over a twenty year period are consistent with stochastic dominance over the market return.

*We thank Eric Becker, Geert Bekaert, and seminar participants at Humboldt University of Berlin, Macquarie University, the University of California – Los Angeles, the University of New South Wales, the University of Nottingham, and the University of Southern California for helpful comments. Jong-Myun Moon provided valuable research assistance at an early stage.
1 Introduction

Since the early 2000s, empirical researchers have documented mounting evidence of non-monotone pricing kernels for a variety of major market indices. This apparent nonmonotonicity has become known as the \textit{pricing kernel puzzle}. In this paper we define the pricing kernel as the ratio of the option implied risk neutral density of the price of the market portfolio in one period, and the objective (or physical) density of the price of the market portfolio in one period conditional on current information.\footnote{If $M_t$ and $S_t$ are the stochastic discount factor and the price of the market portfolio at time $t$, and $M^*_t = E_t(M_{t+1}|S_{t+1})$ is the projection of the stochastic discount factor on the price of the market portfolio next period, then the pricing kernel $\pi_t$ satisfies $\pi_t(S_{t+1}) = (1+r_t)M^*_t$, where $r_t$ is the one-period risk-free interest rate at time $t$. See Beare and Schmidt (2014) for further details.} Nonmonotonicity of pricing kernels is puzzling because it appears to contradict the basic principle, ubiquitous in economic and financial theory, that output in good states of the world should be cheaper than output in bad states of the world. If the pricing kernel is not monotonically decreasing, this principle is violated, and in fact there must be two states of the world in which (a) the market index is lower in the first state than in the second state, and (b) the price-per-probability-unit of a security paying one unit in the second state exceeds the price-per-probability-unit of a security paying one unit in the first state.

The first empirical evidence of nonmonotone pricing kernels was reported in a trio of papers by Jackwerth (2000), A"ıt-Sahalia and Lo (2000) and Rosenberg and Engle (2002). Nonparametric estimates of the pricing kernel for the S&P 500 index appearing in each of these papers were monotonically decreasing at low and high return levels, but increasing over an intermediate range of returns. Subsequent empirical assessments of pricing kernel monotonicity include Jackwerth (2004), Barone-Adesi et al. (2008), Chabi-Yo et al. (2008), Bakshi et al. (2010), Barone-Adesi and Dall’O (2012), Christoffersen et al. (2013), Chaudhuri and Schroder (2013), Golubev et al. (2014), H"ardle et al. (2014) and Beare and Schmidt (2014). In the last of these papers, a version of a statistical test due to Carolan and Tebbs (2005) and Beare and Moon (2014) was applied to 180 cross sections of European put and call options on the S&P 500 index between 1997 and 2013. Pricing kernel monotonicity was rejected around 40\% of the time at the 5\% significance level. Many additional references on the pricing kernel puzzle may be found in a recent paper by Hens and Reichlin (2012), who provided a review of the literature in this area, and
considered a range of theoretical explanations for pricing kernel nonmonotonicity.

An important early contribution of Dybvig (1988) identified a striking implication of pricing kernel nonmonotonicity under complete markets: if the pricing kernel is not a nonincreasing function of aggregate output, then there exists a portfolio of Arrow securities delivering the payoff distribution of the market portfolio, but at a lower price. The rate of return on such a portfolio necessarily stochastically dominates the rate of return on the market portfolio. Beare (2011) pursued some extensions of this result, obtaining an explicit formula for the cheapest contingent claim delivering the payoff distribution of the market portfolio. This claim is referred to as the *optimal measure preserving derivative*, or OMPD. The OMPD corresponds to the market portfolio if and only if the pricing kernel is nonincreasing. If the pricing kernel is not nonincreasing, then the OMPD delivers a rate of return that stochastically dominates that of the market portfolio. Related results are given in a forthcoming paper by Bernard et al. (2014).

In this paper we use the results just described as the basis for an indirect assessment of pricing kernel nonmonotonicity. Specifically, we seek to identify portfolios of actively traded contingent claims with returns stochastically dominant over those of the market portfolio. To this end, we design a selection algorithm that exploits information about the shape of the pricing kernel in order to identify such portfolios. We confine ourselves to portfolios of European put and call options written on the S&P 500 index. Our selection algorithm proceeds roughly as follows: (1) estimate the conditional distribution of the market payoff next period using the method of filtered historical simulation (Barone-Adesi et al., 1998), employing an estimate of current volatility that allows for an asymmetric response to past positive and negative return innovations (leverage); (2) compute a Pareto frontier of option portfolios that each minimize price subject to having an estimated conditional payoff distribution that differs from the estimated conditional market payoff distribution by no more than some fixed quantity; (3) select a portfolio on the Pareto frontier with a rate of return that conditionally stochastically dominates the market rate of return, if possible. Step (2) can be implemented using a multiobjective evolutionary algorithm (see e.g. Fontes and Gaspar-Cunha, 2010), which computes an entire Pareto frontier of portfolios in a single run, commencing from an initial population of randomly generated portfolios.
We evaluate the out-of-sample performance of our option portfolio selection algorithm over a sample of 219 months running from January 1990 to December 2009. Depending on the selection of a tuning parameter, the empirical distribution of our out-of-sample portfolio return stochastically dominates, or nearly stochastically dominates, that of the market portfolio. The average excess return of our option portfolios exceeds the average excess return of the market portfolio by around 3–4 percentage points at the annual level. Formal tests of statistical significance are consistent with stochastic dominance of the market portfolio. Linear factor regressions indicate that our superior returns cannot be explained by covariation with factors popular in the empirical asset pricing literature, including the size and book-to-market factors of Fama and French (1993), the momentum factor of Jegadeesh and Titman (1993), and the option-implied market volatility factor used by Ang et al. (2000).

The superior returns of our option portfolios are generated by small price reductions relative to the market portfolio. The price of our selected portfolio is at least 99.5% of the price of the market portfolio in more than half of the 219 months considered in our empirical analysis, and nearly always above 99%. The shape of the payoff function of the selected portfolio varies from month to month, but is in many cases consistent with the locally increasing pricing kernel shape originally identified by Aït-Sahalia and Lo (2000), Jackwerth (2000) and Rosenberg and Engle (2002). In these cases we achieve price reduction by reallocating the payoff of the market portfolio away from mildly positive market outcomes and toward mildly negative market outcomes. Our finding that option portfolios may be used to generate returns stochastically dominant over the market return is consistent with earlier work by Constantinides et al. (2009, 2011), who found that European and American S&P 500 index options frequently violated the stochastic dominance bounds established by Constantinides and Perrakis (2002, 2007) over the period 1983–2006.

Similar to an earlier contribution by Chabi-Yo et al. (2008), two recent working papers by Linn et al. (2014) and Song and Xiu (2014) argue that, when one properly conditions on relevant information such as the level of market volatility, the pricing kernel is in fact monotone. It is certainly true that conditioning has been a problematic issue in existing assessments of pricing kernel monotonicity. Golubev et al. (2014) and Härdle et al. (2014) assess the monotonicity of pricing kernels by comparing the shape of the risk neutral distribution implied by current option prices to the shape of the distribution of historical
market returns. Since this latter distribution is properly viewed as an estimate of the unconditional distribution of market returns, there is a mismatch between the scale of the risk neutral and physical distributions in times of high or low volatility, potentially leading to spurious rejections of pricing kernel monotonicity. Beare and Schmidt (2014) do allow the scale of the estimated physical distribution to depend on a measure of market volatility, but do not permit variation in the shape of the distribution, which may also be problematic. On the other hand, the estimation techniques employed by Linn et al. (2014) and Song and Xiu (2014) may obscure anomalies in the shape of the pricing kernel that do not persist over time. Specifically, Linn et al. (2014) assume that the pricing kernel is constant over time, so that their results are best interpreted to mean that the average shape of the pricing kernel over time is monotone. This does not rule out nonmonotonicity at any given point in time, and indeed, Beare and Schmidt (2014) identify significant time variation in the shape of the pricing kernel. Song and Xiu (2014) assume that the risk neutral distribution varies over time only through its dependence on the time varying market volatility factor and on the forward price of the market portfolio. Their local linear estimate of the risk neutral distribution, which combines data on option prices from many different time periods, may therefore fail to identify pricing anomalies that “average out” over periods of similar volatility.

The remainder of the paper is structured as follows. In Section 2 we review the results of Dybvig (1988) and Beare (2011) that provide the theoretical motivation for our analysis. Our option portfolio selection algorithm is described in Section 3. Empirical results are presented in Section 4. We conclude in Section 5 with additional discussion of the practical implications of our results.

2 Theoretical motivation

To build intuition, we begin with a simple numerical example, summarized in Table 1 below. Suppose we know that in one month from today the market index can take three possible values: $0.90, $1.00 and $1.10, with equal probability. The three states of the world defined by the market outcome are assigned state prices equal to one third of $1.08, $0.94 and $0.97 respectively. Note that we have chosen these state prices to be
nonmonotone in the market index levels associated with each state. The price of any asset whose payoff after one month can be written as a function of the market index is given by the inner product of its payoff vector and the vector of state prices. The price of the market index itself is therefore given by $(1.08 \times 0.9 + 0.94 \times 1 + 0.97 \times 1)/3 = 0.993$.

<table>
<thead>
<tr>
<th>State</th>
<th>State probability</th>
<th>State price ($\times 3$)</th>
<th>Payoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/3</td>
<td>$1.08$</td>
<td>$0.90$</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>$0.94$</td>
<td>$1.00$</td>
</tr>
<tr>
<td>3</td>
<td>1/3</td>
<td>$0.97$</td>
<td>$1.10$</td>
</tr>
<tr>
<td>Price</td>
<td></td>
<td>$0.993$</td>
<td>$0.992$</td>
</tr>
<tr>
<td>Expected return (annualized)</td>
<td></td>
<td>8.5%</td>
<td>9.7%</td>
</tr>
</tbody>
</table>

**Table 1:** Motivating example.

Consider a derivative contract that pays $0.90 when the market index is $0.90, $1.10 when the market index is $1.00, and $1.00 when the market index is $1.10. Since the three outcomes are equally likely, this derivative contract delivers the same payoff distribution as the market index. Its price is $(1.08 \times 0.9 + 0.94 \times 1.1 + 0.97 \times 1)/3 = 0.992$, which is less that the price of the market index. Therefore, the return distribution of our derivative contract (i.e., the distribution of payoff/price minus one) stochastically dominates that of the market. Intuitively, the derivative contract achieves price reduction without deviating from the market payoff distribution by switching payoffs between states in such a way that the payoff is greatest when the state price is least, and the payoff is least when the state price is greatest. We can achieve a lower price than the market index price because the state prices are not a decreasing function of the market payoff.

Results due to Dybvig (1988), extended subsequently by Beare (2011), show how the intuition conveyed by this simple numerical example carries over to the case where the market payoff is continuously distributed. Suppose the value of the market index after one period is a nonnegative random variable whose distribution conditional on current information is $P$. We refer to $P$ as the *physical distribution*. Agents may invest in derivative contracts written on the market index. The value of such a contract at the end of one period is a real valued function $\theta$ of the value of the market index. We assume that

---

\(^2P, Q,\) and other distributions should be interpreted as cumulative distribution functions.
there exists a complete market for self-collateralized\(^3\) derivative contracts written on the market index. The set of such contracts is denoted by \(\Theta\), and includes all (technically, all Borel measurable) functions from \(\mathbb{R}_+\) to \(\mathbb{R}_+\). The price of a contract \(\theta \in \Theta\) is assumed to be given by its discounted expectation relative to some distribution \(Q\), which we refer to as the risk neutral distribution. That is, \(\text{price}(\theta) = (1 + r)^{-1} \int \theta(x) dQ(x)\), where \(r\) is the risk-free rate.

The distributions \(P\) and \(Q\) are assumed to be absolutely continuous, with probability density functions \(p\) and \(q\). In the absence of arbitrage opportunities, there exists a well defined density ratio \(\pi = q/p\), which we refer to as the pricing kernel. Dybvig (1988) and Beare (2011) study the way in which the shape of \(\pi\) — in particular, its monotonicity, or lack thereof — affects the relative prices of the derivative contracts in \(\Theta\). Of particular interest is a subset of \(\Theta\), which we denote \(\Theta_P\), comprised of all the functions in \(\Theta\) that are measure preserving with respect to the physical distribution \(P\). A derivative contract is a member of \(\Theta_P\) if it delivers a payoff distribution that is identical to \(P\), the market payoff distribution. Let \(\theta_M \in \Theta_P\) denote the identity function \(\theta_M(x) = x\). The contract \(\theta_M\) represents a direct investment in the market index. We seek to answer the following question: when is \(\theta_M\) the cheapest function in \(\Theta_P\)? That is, we wish to know when \(\theta_M\) solves the following optimization problem:

\[
\minimize \quad \frac{1}{1+r} \int \theta(x) dQ(x) \quad \text{subject to} \quad \theta \in \Theta_P.
\] (2.1)

The solution is provided by Dybvig (1988, Theorem 1); see also Beare (2011, Theorem 3.1(iv)). We summarize it in the following Proposition.

**Proposition 2.1.** \(\theta_M\) solves (2.1) if and only if \(\pi\) is nonincreasing.

The import of Proposition 2.1 is that, in the presence of a nonincreasing pricing kernel, the cheapest way to obtain the payoff distribution of the market index is to buy the market index. Conversely, if the pricing kernel is not nonincreasing, one may obtain the payoff distribution of the market index by purchasing a derivative contract with a lower price. The exact form of the derivative contract that solves (2.1) when \(\pi\) is not nonincreasing was obtained by Beare (2011, Theorem 3.1(ii)). Before stating it, we require some additional

---

\(^3\)We say that a derivative contract is self-collateralized if its payoff function is nonnegative.
notation. Let $P_\pi$ denote the distribution of $\pi(X)$ when $X$ is a random variable with distribution $P$, and let $P^{-1}$ denote the quantile function for $P$. Let $\vartheta \in \Theta$ be the function given by

$$\vartheta(x) = P^{-1}(1 - P_\pi(\pi(x))), \quad x \in \mathbb{R}_+.$$  \hfill (2.2)

**Proposition 2.2.** $\vartheta$ solves (2.1).

Beare (2011) refers to $\vartheta$ as the *optimal measure preserving derivative* (OMPD). When $\pi$ is decreasing, we have $\vartheta = \theta_M$, and the OMPD is a direct market investment. But when $\pi$ is not nonincreasing, $\vartheta$ differs from $\theta_M$, and achieves price reduction by reallocating payoffs between states in much the same fashion as our derivative in the numerical example given in Table 1. In this case, an agent may achieve a return distribution that stochastically dominates the market return distribution by investing their wealth in some scalar multiple of the OMPD.

It is perhaps worth noting that no arbitrage opportunity is implied by the existence of derivative contracts whose payoffs stochastically dominate a direct market investment. Indeed, the nonexistence of arbitrage opportunities is a key feature of our model. We might instead say that, in the presence of a nonmonotone pricing kernel, there is a possibility of *statistical* arbitrage, relative to a direct market investment.

In view of the growing empirical literature documenting the existence of pricing kernel nonmonotonicity in derivatives markets for a variety of major market indices – see Beare and Schmidt (2014), and other papers cited in Section 1 – the results just discussed present an intriguing possibility. Is it indeed possible to construct derivative contracts yielding payoff distributions that stochastically dominate a direct market investment out-of-sample? We address this question in the remainder of the paper, focussing on self-collateralized portfolios of actively traded European put and call options.

### 3 Construction of portfolio weights

In this Section we propose a data-based approach to choosing a portfolio of options that sensibly adapts to the presence of pricing kernel nonmonotonicity. We augment our
discussion with graphical examples from November 19, 2009, the penultimate date in our sample. We select this date because it is representative in many ways of the types of portfolios which are selected when in-sample estimates suggest that the pricing kernel is not monotone. Details about the data used to implement our portfolio selection procedure are postponed until Section 4.1.

The first step in our procedure is to obtain an empirical estimate \( \hat{P} \) of the physical distribution \( P \). This is discussed in Section 3.1. In Section 3.2 we define the set \( \Theta \) of option portfolios that we consider investing in. The optimization problem used to select our option portfolio from \( \Theta \) is described in Section 3.3. Numerical solution of this optimization problem is discussed in Section 3.4. Some guidelines for the selection of a tuning parameter used in our procedure are provided in Section 3.5.

### 3.1 Physical distribution estimation

Since \( P \) is unknown, the first step in our procedure is to obtain an empirical estimate \( \hat{P} \) of \( P \). Many approaches are possible here. We follow Rosenberg and Engle (2002) and Barone-Adesi et al. (2008) and model the evolution of market returns using the so-called GJR-GARCH model of Glosten et al. (1993). In this model, the GARCH(1,1) model of Bollerslev (1986) is augmented with an additional term that allows volatility to respond asymmetrically to positive and negative returns; this is the so-called leverage effect. The market index values \( X_t \) are assumed to satisfy

\[
\ln \left( \frac{X_t}{X_{t-1}} \right) = \mu + \sigma_t \varepsilon_t,
\]

where \( \varepsilon_t \) is an iid sequence of random variables with \( E\varepsilon_t = 0 \) and \( E\varepsilon_t^2 = 1 \), and the volatility \( \sigma_t^2 \) evolves according to

\[
\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 + \delta \max\{0, -\varepsilon_{t-1}\}^2.
\]

We estimate the parameter vector \((\mu, \omega, \alpha, \gamma, \delta)\) by quasi-maximum likelihood (see e.g. Bollerslev and Wooldridge, 1992) using a sample of historical daily market index values \( X_t \), with \( t = 0, -1, \ldots, -T \). The option portfolio choice problem is undertaken at date
When constructing the empirical distribution we normalize by \(N + 1\) rather than \(N\) to ensure that 
\(0 < \hat{P}(X^*_{h,i}) < 1\) for all \(i = 1, \ldots, N\). This is done so that the computation of the Anderson-Darling distance between \(\hat{P}\) and the estimated payoff distribution of candidate option portfolios, as discussed in Section 3.3, does not involve division by zero.
Figure 1: Estimated physical distribution for November 19, 2009.

This figure displays the estimated physical distribution for November 19, 2009, calculated from historical values of the S&P 500 index. The current index value is normalized to one and the return horizon is \( h = 20 \) trading days. We computed the distribution by applying FHS to a fitted GJR-GARCH(1,1) model, with parameters estimated by QMLE using the previous \( T = 3000 \) trading days of index values. For graphical purposes only, we applied a kernel smoother to the estimated distribution. We also plot a Gaussian distribution with the same location and scale as the estimated distribution.

estimated physical distribution is skewed to the left, with a heavier lower tail than the Gaussian distribution. Though our estimates vary significantly over time, the left skew is a standard feature throughout the full twenty year sample.

3.2 Admissible portfolios

In the theoretical framework of Section 2 we assumed the existence of a complete market for derivatives with arbitrary nonnegative payoff functions in \( \Theta \). Our empirical portfolio selection procedure is designed to choose between a smaller collection of tradeable derivatives, with payoff functions in a set \( \bar{\Theta} \subset \Theta \). Each \( \theta \in \bar{\Theta} \) corresponds to a self-collateralized portfolio formed from long or short positions in the market index, European put and call
options written on the market index, and zero-coupon risk-free bonds. Using put-call parity, we can rewrite all put prices as call prices, so we will assume without loss of generality that all options are calls. We choose our portfolio at date \( t = 0 \). Options expire and bonds mature at date \( t = h \). Options may be written at strikes \( s_1 < \cdots < s_n \). The payoff of such a portfolio at date \( t = h \) is given by

\[
\theta(x) = \beta_1 + \beta_2 x + \sum_{i=1}^{n} \beta_{i+2} \max\{x - s_i, 0\}, \quad x \in \mathbb{R}_+, \tag{3.2}
\]

for some vector of portfolio weights \( \beta \in \mathbb{R}^{n+2} \). In the empirical application reported in Section 4, we aim to use \( n = 11 \) and \( (s_1, \ldots, s_n) = (0.95, 0.96, \ldots, 1.05) \), or as close as possible given the data available at any particular date. Further details about the choice of strikes and calculation of option prices are provided in Section 4.1.

In addition to requiring that each admissible payoff function \( \theta \) satisfy (3.2) for some \( \beta \in \mathbb{R}^{n+2} \), we impose the following two additional constraints.

\[
\theta(x) = x \quad \text{for all} \ x \leq s_1 \text{ and all} \ x \geq s_n. \tag{3.3}
\]

\[
|\beta_i| \leq C \quad \text{for all} \ i \in \{1, \ldots, n + 2\}. \tag{3.4}
\]

The admissible portfolio set \( \bar{\Theta} \) is given by

\[
\bar{\Theta} = \{ \theta \in \Theta : \theta \text{ satisfies (3.2) and (3.3) for some} \ \beta \in \mathbb{R}^{n+2} \text{ satisfying (3.4)} \}. \tag{3.5}
\]

Constraint (3.4) renders the admissible set of portfolio weights compact, assisting the numerical implementation of our optimization procedure. With a generous choice of \( C \), (3.4) is only mildly restrictive. We set \( C = 10 \) in the empirical application reported in Section 4. Constraint (3.3) is imposed to limit the sensitivity of our admissible portfolios to extreme market outcomes. In practice, it is difficult to accurately determine the shape of the pricing kernel in the tails of the market payoff distribution. Similarly, it may not be realistic to expect an unrestricted optimization procedure to reliably select sensible positions in options that are very far out-of-the-money. By forcing our admissible payoff functions to match the market payoff beyond the extreme strikes \( s_1 \) and \( s_n \), we focus attention on optimizing our payoff function over a region where we have more information about the shape of the pricing kernel. This aspect of our selection procedure aids statis-
tical comparison between the performances of the selected portfolio and market portfolio, as observed differences between the two cannot be attributed to uncertainty about the probability of far out-of-the-money options being exercised.

3.3 Portfolio selection problem

The OMPD $\vartheta$ is the cheapest payoff function in $\Theta$ with payoff distribution $P$. Our empirical portfolio selection procedure seeks to identify the cheapest payoff function in $\Theta$ whose estimated payoff distribution is close to $\hat{P}$. The estimated payoff distribution of a function $\theta \in \Theta$, denoted $\hat{P}_\theta$, is the distribution of $\theta(X)$ when $X$ has distribution $\hat{P}$. We employ the following notion of closeness between $\hat{P}_\theta$ and $\hat{P}$:

$$\text{dist}(\hat{P}_\theta, \hat{P}) = \int_0^\infty \frac{(\hat{P}_\theta(x) - \hat{P}(x))^2}{\hat{P}(x)(1 - \hat{P}(x))} d\hat{P}(x).$$  

(3.6)

The distance (more accurately, discrepancy) $\text{dist}(\hat{P}_\theta, \hat{P})$ was proposed by Anderson and Darling (1952), and can be considered a weighted version of the Cramér-von Mises distance between $\hat{P}_\theta$ and $\hat{P}$. The inverse weighting $\hat{P}(x)(1 - \hat{P}(x))$ makes the discrepancy between $\hat{P}_\theta$ and $\hat{P}$ more sensitive to differences in the tails of the distributions. In the context of our portfolio choice problem, this helps to penalize portfolios that achieve price reduction in exchange for overexposure to extreme market outcomes, to the extent that this is possible under constraint (3.3).

Given any payoff function $\theta \in \Theta$, we may calculate $\text{price}(\theta)$, the price of $\theta$, from the option prices, market price, and risk-free interest rate prevailing at date $t = 0$. Our estimated portfolio $\hat{\vartheta}$ is defined as the solution to the following optimization problem:

$$\text{minimize } \text{price}(\theta) \text{ subject to } \theta \in \Theta \text{ and } \text{dist}(\hat{P}_\theta, \hat{P}) \leq d.$$  

(3.7)

The tuning parameter $d \geq 0$ governs how closely the estimated portfolio payoff distribution is required to match the estimated market payoff distribution. Though it may be tempting to set $d = 0$, so that perfect distributional replication is imposed as a constraint, there is generally only a single portfolio in $\Theta$ that satisfies $\text{dist}(\hat{P}_\theta, \hat{P}) = 0$. That portfolio is $\theta_M$, the market portfolio. Other portfolios $\theta \in \Theta$ that satisfy $\text{dist}(\hat{P}_\theta, \hat{P}) = 0$ will (with
smooth $P$ and large enough $T$ and $N$) not satisfy the piecewise linear structure imposed on the portfolios in $\bar{\Theta}$. We therefore need to choose $d > 0$ in order to obtain solutions to (3.7) that differ from the market portfolio. But $d$ cannot be chosen too large, or the payoff distribution of $\hat{\vartheta}$ will poorly approximate $P$. Some guidelines for the selection of $d$ are provided in Section 3.5.

In Section 3.1 we proposed to estimate $P$ using FHS. Our estimated physical distribution $\hat{P}$ is given by (3.1), the empirical distribution of a large number $N$ of simulated realizations $X_{h,i}^*, i = 1, \ldots, N$. With this choice of $\hat{P}$, our distributional distance $\text{dist}(\hat{P}_\theta, \hat{P})$ may be calculated directly from the simulated realizations $X_{h,i}^*$. Specifically, we have

$$\text{dist}(\hat{P}_\theta, \hat{P}) = \frac{1}{N+1} \sum_{j=1}^N \left( \frac{1}{N+1} \sum_{i=1}^N \left( \theta(X_{h,i}^*) \leq X_{h,j}^* \right) - 1 \left( X_{h,i}^* \leq X_{h,j}^* \right) \right)^2,$$

which is straightforward to compute.

3.4 Numerical solution via multiobjective optimization

As discussed in Section 3.3, our estimated portfolio $\hat{\vartheta}$ is chosen to solve (3.7). From a computational perspective, this is a difficult problem. The set of portfolios in $\bar{\Theta}$ satisfying $\text{dist}(\hat{P}_\theta, \hat{P}) \leq d$ may be nonconvex, and the dimension of $\bar{\Theta}$ may be large if we allow options written at many different strikes in our portfolio. To overcome these difficulties as best as possible, we propose to compute $\hat{\vartheta}$ using a multiobjective evolutionary algorithm, or MOEA. With an MOEA, we can numerically solve (3.7) for a wide range of values of $d$ simultaneously. A detailed discussion of MOEAs goes well beyond the scope of this paper; for a recent review, see Fontes and Gaspar-Cunha (2010). Here we shall attempt to communicate only the flavor of the method. “Out of the box” MOEAs are readily available in standard statistical software packages, including MATLAB’s Global Optimization Toolbox.

MOEAs are designed to identify a family of points in the choice set which are Pareto efficient with respect to two or more criteria. In our case, the choice set is $\bar{\Theta}$, and a portfolio $\theta \in \bar{\Theta}$ is said to be Pareto efficient if it minimizes $\text{price}(\theta)$ over $\bar{\Theta}$ subject to
\[ \text{dist}(\hat{P}_\theta, \hat{P}) \leq d, \] for some value of \( d \geq 0 \). The collection of all Pareto efficient portfolios in \( \tilde{\Theta} \), which we obtain by identifying price minimizers for all values of \( d \), is called the *Pareto efficient set*, and the set of pairs \( (\text{dist}(\hat{P}_\theta, \hat{P}), \text{price}(\theta)) \in \mathbb{R}^2_+ \) associated with Pareto efficient portfolios \( \theta \in \Theta \) is called the *Pareto frontier*. The shape of this frontier characterizes the tradeoff between the competing goals of price minimization and distributional replication. Our Pareto frontier is reminiscent of the mean-variance tradeoff in the classical portfolio selection model of Markowitz (1952), and indeed, MOEAs have been fruitfully applied to the study of mean-variance efficient portfolio choice. Tapia and Coello Coello (2007) provide a survey of MOEA applications in economics and finance.

MOEAs deliver a numerical approximation to the Pareto efficient set and Pareto frontier by applying some version of the following broadly defined procedure. We begin with an initial population of starting values in the choice set. Some subset of the population are then selected to be “parents”; selection is based on the values of \( \text{dist}(\hat{P}_\theta, \hat{P}) \) and \( \text{price}(\theta) \), and also with a view to preserving “population diversity”. Next, new individuals are generated by “mutating” and/or “recombining” (i.e., mating) different parents. Some or all of the initial population are then replaced with some or all of the new individuals, and a new generation is formed. The procedure is then iterated until the initial population has evolved through some preassigned number of generations. Finally, we eliminate members of the terminal population that are Pareto dominated by other members. A scatterplot of the pairs \( (\text{dist}(\hat{P}_\theta, \hat{P}), \text{price}(\theta)) \) associated with the remaining population traces out the shape of our estimated Pareto frontier. For a more detailed description of the operation of MOEAs, and numerous additional references, see Fontes and Gaspar-Cunha (2010).

Figure 2 depicts the progression of the MOEA through successive generations when applied to our portfolio selection problem at the date November 19, 2009. In each panel, the vertical axis tracks the price of each portfolio \( \theta \) relative to the price of a direct market investment (i.e., the futures price discounted by the risk free rate), and the horizontal axis indicates the distributional distance \( \text{dist}(\hat{P}_\theta, \hat{P}) \times 100 \) associated with each \( \theta \). Note that \( \hat{P} \) here is the physical distribution displayed in Figure 1. The left, center and right panels of Figure 2 display the evolving population of portfolios in generations 1, 10 and 100 respectively. We terminate the algorithm after 100 generations, so the right panel provides our final estimate of the Pareto frontier.\(^5\) Dark circles indicate the Pareto efficient portfolios

\(^5\)Visual inspection of the estimated Pareto frontiers suggests that there is little improvement after 50
This figure displays three generations of portfolios when the MOEA is applied to our portfolio selection problem at the date November 19, 2009. The left, center and right panels correspond to the 1st, 10th and 100th generations of portfolios respectively. Each point in a scatterplot indicates the distributional distance ($\times 100$) and price associated with a given portfolio in that generation, where the price is defined relative to the price of a futures contract written on the market index. The final selected portfolio, indicated with a square, is included in all three panels but belongs only to the 100th generation.

in each generation, while light triangles indicate Pareto inefficient portfolios. Comparing the three panels, we see that the randomly selected initial population of portfolios evolves south-west toward our final estimate of the Pareto frontier as the MOEA iterates through successive generations. The dark square indicates the portfolio ultimately selected by our algorithm, $\hat{\vartheta}$. It is chosen from the Pareto efficient portfolios in the 100th generation using the selection rule described in Section 3.5 below.

Our estimated Pareto frontier exhibits an important feature: it drops sharply as one moves a little to the right of 0.2 on the horizontal axis. This discontinuity indicates that substantial price reduction may be achieved with a very small distributional tradeoff, and

---

6To accelerate convergence of the MOEA, the initial generation of portfolios was generated by randomly perturbing and combining crude preliminary estimates of the OMPD constructed using the pricing kernel estimator of Rosenberg and Engle (2002). Further details are available on request.
is typical of months in which our selected portfolio differs substantially from the market portfolio. Pareto efficient portfolios that lie to the left of the discontinuity tend to closely resemble the market portfolio, while portfolios to the right generally do not. This pattern is consistent with a nonmonotone pricing kernel, or equivalently, with an OMPD that differs from the market portfolio. Since the OMPD will not generally have a piecewise linear payoff function, sufficient error in distributional replication (i.e., sufficiently large \( d \)) must be permitted before a piecewise linear approximation to the OMPD becomes an admissible candidate solution to our optimization problem.

### 3.5 Tuning parameter selection

The output of the MOEA described in Section 3.4 is a family of Pareto efficient portfolios in \( \Theta \). If our numerical optimization has been successful, each of these portfolios solves (3.7) for some value of \( d \geq 0 \). Selection of the tuning parameter \( d \) therefore amounts to selection among the Pareto efficient portfolios identified by the MOEA.

In view of the discussion in Section 2, our hope is to identify portfolios which deliver a return distribution that stochastically dominates the market return. To this end, we eliminate from consideration those Pareto efficient portfolios that fail to satisfy the following screening criterion. First, we multiply each Pareto efficient portfolio \( \theta \) by \( \text{price}(\theta_M)/\text{price}(\theta) \), yielding a rescaled portfolio \( \theta^* \) with price normalized to be equal to that of a direct market investment. We then calculate a one-sided distributional distance between \( \hat{P}_{\theta^*} \), the estimated payoff of the rescaled portfolio, and \( \hat{P} \), the estimated market payoff:

\[
\text{dist}^+(\hat{P}_{\theta^*}, \hat{P}) = \int_0^\infty \frac{\max\{\hat{P}_{\theta^*}(x) - \hat{P}(x), 0\}^2}{\hat{P}(x)(1 - \hat{P}(x))} d\hat{P}(x).
\]

The one-sided distance \( \text{dist}^+(\hat{P}_{\theta^*}, \hat{P}) \) is simply the Anderson-Darling distance between \( \hat{P}_{\theta^*} \) and \( \hat{P} \) as defined in (3.6), but with the domain of integration restricted to the region over which \( \hat{P}_{\theta^*} \) lies above \( \hat{P} \). It is equal to zero when \( \hat{P}_{\theta^*} \) stochastically dominates \( \hat{P} \), and is positive otherwise. We eliminate portfolios for which \( \text{dist}^+(\hat{P}_{\theta^*}, \hat{P}) \) exceeds some small threshold; a threshold of \( 10^{-3} \) was used to obtain the results reported in Section 4.
After eliminating Pareto efficient portfolios that do not satisfy our screening criterion, it remains to choose between the surviving portfolios. Here we employ a simple rule of thumb: we choose the cheapest surviving \( \theta \) for which \( \text{dist}(\hat{P}_\theta, \hat{P}) \times 100 \) is no greater than some arbitrarily specified cutoff \( \lambda > 0 \). This \( \theta \) becomes our selected option portfolio \( \hat{\theta} \). In Section 4 we report results obtained using \( \lambda = 1/5, 1/3 \) and \( 1/2 \). If there are no surviving portfolios with \( \text{dist}(\hat{P}_\theta, \hat{P}) \times 100 \leq \lambda \), we set \( \hat{\theta} = \theta_M \), the market portfolio. Note that \( \theta_M \) is always Pareto efficient, as it minimizes price over \( \Theta \) subject to \( \text{dist}(\hat{P}_\theta, \hat{P}) \leq 0 \), and always satisfies our screening criterion, but in practice the MOEA may sometimes fail to identify \( \theta_M \) as a Pareto efficient portfolio. For this reason, we always add \( \theta_M \) to the Pareto set before selecting a portfolio from along the frontier.

Figure 3 displays the payoff function of the option portfolio selected for November 19, 2009, when \( \lambda = 1/3 \). This portfolio is represented by the dark square in Figure 2. It is priced at \$0.9945, and solves (3.7) for \( d = 0.0033 \). The price of the direct market investment, with payoff function given by the 45°-line, is \$0.9983 in this period, so the relative price

\[ \begin{array}{c}
\text{Market payoff} \\
\text{Payoff} \\
\text{Market} \\
\text{Option portfolio}
\end{array} \]

**Figure 3:** Payoff function of selected option portfolio for November 19, 2009.

This figure displays the payoff function of the option portfolio selected for November 19, 2009, when \( \lambda = 1/3 \). The price of this portfolio is \$0.9945. The price of the market portfolio, with the 45°-line as payoff function, is \$0.9983.
of the option portfolio is $0.9962. The payoff function for our selected portfolio exhibits a pronounced N-shape when market returns are in the region ±5%, while constraint (3.3) forces it to match the market payoff elsewhere. A payoff function with this general shape may be advantageous when the pricing kernel is decreasing at extreme return levels, but increasing at moderate return levels – as reported in the empirical studies of Aït-Sahalia and Lo (2000), Jackwerth (2000) and Rosenberg and Engle (2002), among others. Beare (2011, Fig. 4.3(b)) plots the payoff function for the OMPD implied by the risk neutral and physical density estimates reported by Jackwerth (2000) for April 15, 1992; its shape is similar to that of the function plotted in Figure 3.

Panel A of Figure 4 displays the estimated payoff distributions $\hat{P}_\theta$ and $\hat{P}_{\theta M}$ of the selected

![Figure 4: Estimated payoff distributions for November 19, 2009.](image)

Panel A displays the estimated payoff distribution of the market portfolio and of our selected option portfolio (with $\lambda = 1/3$) for November 19, 2009. The option portfolio has price $0.9945$, while the market portfolio has price $0.9983$. Panel B displays the estimated distributions of payoff/price for the option and market portfolios, which is equivalent to rescaling both portfolios to have price $1$. 
option portfolio \( \hat{\vartheta} \) and direct market investment \( \theta_M \). Note that \( \hat{P}_{\theta_M} = \hat{P} \), the physical distribution whose density (after smoothing) was plotted in Figure 1. We see that \( \hat{P} \) and \( \hat{P}_{\vartheta} \) closely resemble one another; the distance \( d(\hat{P}_{\vartheta}, \hat{P}) \times 100 \) between the two is 0.33, the horizontal ordinate of the dark square in Figure 2. Though the two payoff distributions are very similar, the price of the option portfolio \( \hat{\vartheta} \) is $0.9945, while the market portfolio \( \theta_M \) has price $0.9983. If we rescale both portfolios to have price $1, so that their payoff functions are now given by \( \hat{\vartheta}^* = \hat{\vartheta} / 0.9945 \) and \( \theta_M^* = \theta_M / 0.9983 \), then we achieve the payoff distributions plotted in Panel B. We see that, after rescaling to equalize price, the estimated payoff distribution of our option portfolio stochastically dominates the estimated market payoff distribution. Thus, our option portfolio appears to generate a rate of return superior to that of the market portfolio.

The estimated distributions appearing in Figure 4 are ex-ante descriptions of our beliefs on 11/19/09 (time \( t = 0 \)) regarding the market payoff and option portfolio payoff on 12/16/09 (time \( t = 20 \)). Ex-post, we know that the market payoff on 12/16/09 was $1.0069. From Figure 3, we see that the option portfolio payoff function is very close to the 45°-line at this point; in fact, it is $1.0075, slightly above the market payoff. When we take into consideration the fact that the price of our option portfolio is lower than that of the market portfolio, we find that the rate of return on our option portfolio is 0.0131, while the market rate of return is 0.0086. Thus, for this one month, the rate of return on our option portfolio exceeds the market rate of return by 0.0045. It turns out that, considering all 219 months in our sample period, the rate of return of our selected option portfolios (with \( \lambda = 1/3 \)) exceeds the market rate of return by an average of 0.0032, and this difference is statistically significant at the 0.05 level; see Table 5 below. The monthly return differential of 0.0032 amounts to 0.0391 at the annual level, which would appear to be substantial. Further discussion of the performance of our option portfolio selection procedure over our full sample is provided in Section 4.

4 Empirical results

In this Section we report and discuss the performance of the option portfolio selection algorithm described in Section 3 when applied to each of 219 cross-sections of option prices
observed between January 1990 and December 2009. We begin by describing our data set in Section 4.1. Our main results are presented in Section 4.2, where we compare the out-of-sample return distribution of our option portfolios with the out-of-sample return distribution of the market portfolio, and describe how the shape of our selected payoff function varies over time. Finally, in Section 4.3 we investigate whether the apparently superior returns associated with our option portfolios can be explained by covariation with risk factors of the kind appearing frequently in linear asset pricing models.

4.1 Data

Our primary dataset consists of prices for European put and call options written on the S&P 500 index from January 1990 to December 2009. Daily option prices were obtained from the data provider DeltaNeutral\(^7\), which collects prices for options reported by the Options Price Reporting Authority (OPRA).\(^8\) OPRA compiles information from a number of different exchanges in order to find a national best bid and offer price for each option. Our data set consists of OPRA’s reported bid and ask prices for different options at the close of the market on each trading day, along with the corresponding trading volumes.\(^9\) We obtain the daily record of closing prices for the S&P 500 index, along with daily dividend yields, from CRSP. To compute the futures price, we follow Constantinides et al. (2009) and assume that the market expectations of dividend amount and timing were correct. We discount each daily dividend at the risk-free rate and aggregate over the requisite number of days to match the maturity of a given cross section of options. For the risk-free rate, we use the one month London Interbank Offered Rate (LIBOR), from the British Banker’s Association, obtained from the St. Louis Federal Reserve FRED database.\(^10\) We compute the futures price by subtracting these discounted dividends from the closing index level and multiplying by the risk-free interest rate.

In order to reduce the impact of pricing errors, we apply an extensive series of filters to our

\(^7\)http://www.deltaneutral.com.

\(^8\)http://www.opradata.com.

\(^9\)OptionMetrics is a standard data vendor used in this literature. The OptionMetrics and DeltaNeutral data are based on the same underlying data from OPRA, but the former do not extend as far back in time as the latter.

\(^10\)http://research.stlouisfed.org/fred2.
options data. We exclude all options that do not have between 18 and 22 trading days to maturity, nonzero bid prices, and more than 100 contracts traded on the date of interest. We compute each option price as the bid-ask midpoint. Following Rosenberg and Engle (2002), we eliminate individual option prices which fail to satisfy simple no-arbitrage bounds. When put and call prices are available at the same strike, we use put-call parity to calculate implied dividend yields. Then, we discard pairs where the implied monthly dividend yield is negative or bigger than 50 basis points.

As discussed in Section 3.2, from each filtered cross-section of option prices we hope to extract prices for an evenly spaced grid of strikes at moneyness \(-5\%, -4\%, \ldots, +5\%\). To accomplish this, we use the futures price and risk-free rate to convert all prices to implied volatilities, then use a local linear regression to smooth the implied volatility smile in a neighborhood of each grid point.\(^{11}\) We then convert these smoothed implied volatilities to call prices and run an additional set of filters. We drop cross-sections of option prices which fail to satisfy monotonicity, convexity, and vertical spread restrictions implied by no arbitrage. We also exclude options which expire at the end of each calendar quarter, keeping the contract that expires in the middle of each month. Finally, for each month, we choose the trading day with the most available prices, requiring prices for at least six strikes. We break ties by choosing the time to maturity closest to 20 trading days. This yields 219 cross-sections of option prices with approximately one month to maturity, covering essentially non-overlapping periods.

Table 2 summarizes some key features of the prices which comprise our sample, both overall and over five year subperiods. The first row reports the number of cross sections of option prices in our sample. Due to our screening criteria, there are between 4–7 months in each subperiod for which we observe no option prices, and 21 months overall. Early on, this tends to occur due to smaller trading volumes in the options markets. In the last subperiod, we lose several months during the financial crisis when it was hard to identify cross-sections which satisfied convexity restrictions implied by no arbitrage.

\(^{11}\)The details of the interpolation method are as follows. In each regression we include volatilities at strikes with moneyness within 50 basis points of the strike around which we are smoothing, weighted by the inverse of the difference between bid and ask volatilities. We exclude puts with moneyness greater than 2.5\% and calls with moneyness less than -2.5\%. If we observe fewer than two option prices with strikes within 50 basis points of a target strike, we do not interpolate. Instead, if there is just a single price, we replace the target strike with the strike associated with that price. If there is no observed price with strike within 50 basis points of the target strike, we drop that strike from our grid.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of months</td>
<td>56</td>
<td>53</td>
<td>54</td>
<td>56</td>
<td>219</td>
</tr>
<tr>
<td>Avg. number of raw option prices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>12.6</td>
<td>19.2</td>
<td>19.2</td>
<td>36.0</td>
<td>21.8</td>
</tr>
<tr>
<td>$-5.5% &lt; \text{moneyness} \leq -2.5%$</td>
<td>2.5</td>
<td>3.5</td>
<td>3.4</td>
<td>7.0</td>
<td>4.1</td>
</tr>
<tr>
<td>$-2.5% &lt; \text{moneyness} \leq 0%$</td>
<td>4.0</td>
<td>6.4</td>
<td>6.2</td>
<td>10.9</td>
<td>6.9</td>
</tr>
<tr>
<td>$0% &lt; \text{moneyness} \leq 2.5%$</td>
<td>3.8</td>
<td>6.0</td>
<td>6.1</td>
<td>11.1</td>
<td>6.8</td>
</tr>
<tr>
<td>$2.5% &lt; \text{moneyness} &lt; 5.5%$</td>
<td>2.4</td>
<td>3.3</td>
<td>3.6</td>
<td>7.0</td>
<td>4.1</td>
</tr>
<tr>
<td>Avg. number of smoothed option prices</td>
<td>8.7</td>
<td>9.4</td>
<td>9.0</td>
<td>10.6</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Table 2: Summary statistics for options data.

This table displays count statistics for the option prices used in our empirical analysis. Statistics are provided for the full twenty year sample, and for four five year subperiods. We report the number of cross-sections of option prices used in each period or subperiod, the average number of “raw” option prices in each cross-section (further broken down by moneyness), and the average number of smoothed option prices ultimately used in our portfolio selection procedure.

The middle section of Table 2 shows the average number of raw option prices used as inputs to the local linear smoothing algorithm, while the last row shows the average number of smoothed option prices included in our portfolio selection procedure. More raw options prices are used in the ±2.5% moneyness range, where both put and call prices are used. The number of option contracts increases throughout our sample, particularly over the last subperiod. Initially, there are 1–2 prices per grid point, but the number of prices increases substantially in the last subperiod. To some extent, this is mechanical: the level of the S&P 500 increased over the sample period while option contracts continue to be written with strikes that are $5 apart. Since we focus on the ±5% moneyness range, more strikes fit within this interval. However, the increase in the number of prices is also driven by increases in trading volume in options markets over time. Over our full sample, our portfolios include options written at between 9 and 10 strike prices on average, computed using an average of around 22 raw option prices.
4.2 Main results

The payoffs delivered by our selected option portfolios (scaled to have unit price) over the full sample period from January 1990 to December 2009 are displayed in Figure 5. Panels A, B, C correspond to $\lambda = 1/5, 1/3, 1/2$ respectively. In each panel, the 219 realized option portfolio payoffs are plotted against the corresponding payoffs of the market portfolio (also scaled to have unit price). We make three observations. (1) A significant proportion of the payoffs are concentrated around the 45°-line. At extreme return levels, this is a consequence of constraint (3.3), but at more moderate return levels, it reflects the fact that our selected option portfolio sometimes closely resembles the market portfolio, as we would expect under pricing kernel monotonicity. (2) There is relatively tighter clustering of payoffs around the 45°-line when $\lambda$ is smaller. This is as we expect, since with larger values of $\lambda$ we permit the selection of option portfolios whose payoff distributions differ more substantially from the market payoff distribution. (3) Many payoffs lie above the 45°-line when the market payoff is between 0.96 and 1, and below the 45°-line when the

**Figure 5:** Realized option portfolio payoffs.

Here we plot the realized payoffs of our selected option portfolios (rescaled to have unit price) against the corresponding payoff of the market portfolio (also rescaled to have unit price). The three panels correspond to $\lambda = 1/5, \lambda = 1/3$ and $\lambda = 1/2$. 

24
market payoff is between 1 and 1.05. This pattern is consistent with the selection of N-shaped payoff functions for our option portfolios, which we would expect in the presence of pricing kernels exhibiting nonmonotone behavior of the kind first identified by Aït-Sahalia and Lo (2000), Jackwerth (2000) and Rosenberg and Engle (2002).

Figure 6 provides more information about the kinds of payoff functions that are most often selected in our sample period, focussing on the case $\lambda = 1/3$. The horizontal axis tracks the date at which an option portfolio is selected. At a given point on the horizontal axis, the shade of the line extending upward tells us about the shape of the payoff function of the option portfolio selected on that date. The argument of each payoff function is the market payoff, plotted on the vertical axis between 0.95 and 1.05. The shade of the graph tells us the difference between the payoff functions of the selected option portfolio and market portfolio (both rescaled to have unit price). More specifically, the shade of the

Figure 6: Difference between option portfolio and market payoff functions over time.

This figure displays the difference between the payoff functions for the selected option portfolios (scaled to have unit price) and the market portfolio (also scaled to have unit price). Light shades indicate that the option portfolio payoff exceeds the market portfolio payoff, while dark shades indicate reverse. White vertical bars indicate months for which there was no cross-section of options satisfying our screening criteria. In all cases $\lambda = 1/3$. 

25
graph (expressed as a number in accordance with the legend to the right) at the point $(t, x)$ is equal to \( \hat{\vartheta}^*(x) - \theta^*_M(x) \), where \( \hat{\vartheta}^* \) is the payoff function of the option portfolio selected at date \( t \) (scaled to have unit price), and \( \theta^*_M(x) \) is the payoff function of the market portfolio at date \( t \) (scaled to have unit price). The exception to this rule is where we see white vertical bands. These represent months during which we did not select an option portfolio due to data limitations.

It is clear from a casual glance at Figure 6 that lighter colors tend to dominate in the lower half of the graph, while darker colors dominate toward the top. This pattern emerges because of the presence of many N-shaped payoff functions among our selected portfolios. We also see a large number of grey vertical bands in Figure 6. These bands correspond to option portfolios that are close to the market portfolio. The bulk of our selected payoff functions fall into one of these two categories.

In Figure 7 we plot the empirical distribution of the price of our option portfolio relative to the price of the market portfolio over the 219 months in our sample, for \( \lambda = 1/5, 1/3, 1/2 \).

![Figure 7: Empirical distribution of option portfolio relative prices.](image)

This figure displays the empirical distribution of the relative price of our selected option portfolio over the 219 months in our sample, for \( \lambda = 1/5, 1/3, 1/2 \).
By construction, the distribution shifts left as \( \lambda \) increases, since with larger \( \lambda \) there is more slack in the distributional constraint on price minimization. We see that, even with \( \lambda = 1/2 \), the majority of option portfolios are priced at no less than 99.5% of the market portfolio price, while the vast bulk of relative prices are greater than 99%.

Figure 8 consists of six panels plotting the empirical distributions of the market portfolio payoffs and option portfolio payoffs, computed over the full sample of 219 months. These payoffs are out-of-sample: each payoff corresponds to a portfolio that was selected using prior data only. The upper, middle and lower rows of panels use option portfolios computed with \( \lambda = 1/5 \), 1/3 and 1/2 respectively. In the left column of panels we plot the empirical distributions of the option portfolio and market portfolio payoffs without first scaling those portfolios to have unit price. We see that the empirical distributions are very close when \( \lambda = 1/5 \), a little less close when \( \lambda = 1/3 \), and still less close when \( \lambda = 1/2 \). This is precisely what we would expect, as the parameter \( \lambda \) constrains how closely the estimated payoff distribution of our option portfolio must match the estimated payoff distribution of the market portfolio. It is also no surprise that the empirical distribution of the option portfolio payoff lies mostly to the left of the empirical distribution of the market payoffs when \( \lambda = 1/2 \), and to some extent also when \( \lambda = 1/3 \). Our option portfolios are chosen to minimize price subject to a distributional constraint; as we relax that constraint by increasing \( \lambda \), their payoff distributions naturally shift left in order to reduce price.

In the right column of panels in Figure 8 we see how the empirical distributions of the option portfolio and market portfolio payoffs change when both portfolios are scaled to have unit price. By construction, the option portfolio is always at least as cheap as the market portfolio, which can have price no greater than one. Scaling both portfolios to have unit price therefore shifts both empirical payoff distributions to the right, with a proportionally larger shift for the option portfolio. Is this shift differential enough to bring the payoff distribution of the option portfolio to the right of that of the market portfolio? This is the critical question to ask when assessing whether our option portfolio selection procedure generates returns that stochastically dominate the market return. The evidence in the right column of panels provides grounds for cautious optimism. When \( \lambda = 1/5 \) the empirical distribution of option portfolio payoffs is indeed to the right of the empirical distribution of market portfolio payoffs. When \( \lambda = 1/3 \) the same is mostly true, but
Figure 8: Out-of-sample payoff distributions.

In the left column of panels we display the empirical distributions of the out-of-sample payoffs of our option portfolios and of the market portfolio, with $\lambda = 1/5, 1/3, 1/2$. In the right column of panels we display the same, but with all portfolios scaled to have unit price.
the empirical distribution of market portfolio payoffs creeps slightly above the empirical distribution of option portfolio payoffs in a region between −3% and −2% returns. This region extends rightward to around the −1% return level when we increase \( \lambda \) to 1/2. On the other hand, for both \( \lambda = 1/3 \) and \( \lambda = 1/2 \), the empirical distribution of option portfolio payoffs lies clearly to the right of that for market portfolio payoffs at higher return levels.

Are the differences between empirical distribution functions that we observe in the right column of panels in Figure 8 statistically significant? We report the outcome of relevant significance tests in Table 3. At each of the market return levels −5%, −4%, ..., +5% and for each \( \lambda = 1/5, 1/3, 1/2 \) we report the difference between the empirical distribution functions (not in parentheses) and the p-value for a \( t \)-test of the null hypothesis that the difference is zero, against the alternative that it is positive\(^{12}\) (in parentheses). These p-values were computed using Newey-West standard errors with a six month bandwidth. We see that the differences between the empirical distribution functions are mostly positive, consistent with the option portfolio returns stochastically dominating the market returns.

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>-5%</th>
<th>-4%</th>
<th>-3%</th>
<th>-2%</th>
<th>-1%</th>
<th>0%</th>
<th>1%</th>
<th>2%</th>
<th>3%</th>
<th>4%</th>
<th>5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/5</td>
<td>0.9</td>
<td>0.5</td>
<td>0.9</td>
<td>1.8</td>
<td>3.7</td>
<td>4.1</td>
<td>5.5</td>
<td>0.5</td>
<td>1.8</td>
<td>2.3</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>(7.4)</td>
<td>(33.3)</td>
<td>(28.1)</td>
<td>(20.3)</td>
<td>(9.0)</td>
<td>(6.8)</td>
<td>(2.8)</td>
<td>(43.5)</td>
<td>(18.4)</td>
<td>(6.1)</td>
<td>(15.6)</td>
</tr>
<tr>
<td>1/3</td>
<td>1.8</td>
<td>1.4</td>
<td>0.9</td>
<td>-0.5</td>
<td>1.4</td>
<td>4.6</td>
<td>4.6</td>
<td>2.7</td>
<td>4.1</td>
<td>4.1</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>(1.8)</td>
<td>(9.4)</td>
<td>(28.4)</td>
<td>(57.4)</td>
<td>(32.3)</td>
<td>(8.1)</td>
<td>(7.4)</td>
<td>(19.1)</td>
<td>(3.4)</td>
<td>(2.0)</td>
<td>(3.6)</td>
</tr>
<tr>
<td>1/2</td>
<td>1.8</td>
<td>1.4</td>
<td>-0.9</td>
<td>-0.9</td>
<td>-1.8</td>
<td>1.8</td>
<td>4.1</td>
<td>3.7</td>
<td>5.5</td>
<td>4.1</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>(1.8)</td>
<td>(9.4)</td>
<td>(69.0)</td>
<td>(65.1)</td>
<td>(71.9)</td>
<td>(30.1)</td>
<td>(12.7)</td>
<td>(14.1)</td>
<td>(0.9)</td>
<td>(3.4)</td>
<td>(3.6)</td>
</tr>
</tbody>
</table>

**Table 3:** Pointwise stochastic dominance tests.

Numbers without parentheses are 100 times the difference between the empirical distribution functions plotted in the right column of panels in Figure 8 at the specified return levels. Positive numbers are consistent with the option portfolio returns stochastically dominating the market returns. The corresponding numbers in parentheses are 100 times the p-value from a \( t \)-test (based on Newey-West standard errors) of the null hypothesis that the difference between distribution functions is zero at the stated return level, against the alternative that it is positive.

---

\(^{12}\)For a one-sided test against the alternative that the difference is negative we need merely subtract these p-values from one.
The difference is negative at one return level when $\lambda = 1/3$, or at three levels when $\lambda = 1/2$, but the p-values indicate that these differences are not statistically significant. On the other hand, positive differences are statistically significant (at level 0.05) at one return level when $\lambda = 1/5$, and at four return levels when $\lambda = 1/3$ and when $\lambda = 1/2$. The relative lack of significant p-values when $\lambda = 1/5$ can be explained by the fact that, for this choice of $\lambda$, our selected option portfolio is often very similar to, or equal to, the market portfolio. This can be seen in Figure 5, where the option portfolio payoffs cluster much more tightly around the 45° line when $\lambda = 1/5$ than when $\lambda = 1/3$ or $\lambda = 1/2$. Since with $\lambda = 1/5$ there are fewer periods in which the payoff of our option portfolio differs substantially from that of the market portfolio, it is difficult to statistically distinguish between their distributions. Nevertheless, it is notable that one estimated distribution lies entirely to the right of the other.

The p-values in Table 3 indicate the statistical significance of the difference between empirical payoff distributions at a given return level. We might also consider the joint significance of observed differences across all eleven return levels. To this end we apply a testing procedure proposed by Patton and Timmermann (2010) as an extension of earlier work by White (2000). Our results are presented in Table 4. We consider two null hypotheses. The first null, labeled “Market SD options” in the table, is that the market portfolio returns stochastically dominate the option portfolio returns. The second null, labeled “Options SD market”, is that the option portfolio returns stochastically dominate the market portfolio returns. We test these two nulls using, respectively, the “Up” and

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$H_0 :$ Market SD options</th>
<th>$H_0 :$ Options SD market</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/5</td>
<td>4.7</td>
<td>96.0</td>
</tr>
<tr>
<td>1/3</td>
<td>1.3</td>
<td>97.6</td>
</tr>
<tr>
<td>1/2</td>
<td>3.3</td>
<td>79.5</td>
</tr>
</tbody>
</table>

**Table 4:** Joint stochastic dominance tests.

Each number in the table is 100 times the p-value for testing stochastic dominance between the distributions whose estimates are plotted in the right column of panels in Figure 8. We use studentized and non-studentized versions of the “Up” and “Down” statistics proposed by Patton and Timmermann (2010, p. 609), with p-values computed using a block bootstrap.
“Down” statistics proposed by Patton and Timmermann (2010, p. 609). We consider both studentized and non-studentized versions of these statistics, with studentization achieved using Newey-West standard errors with a six month bandwidth. A block bootstrap with six month block length was used to calculate p-values. We find that the null “Market SD options” is rejected for all values of $\lambda$ using the studentized test statistic at the 0.05 level, or the non-studentized statistic at the 0.1 level. On the other hand, the null “Options SD market” cannot be rejected for any value of $\lambda$ using either statistic at any reasonable significance level. These results suggest that our option portfolio selection procedure may be effective in exploiting pricing kernel nonmonotonicity to achieve superior returns.

### 4.3 Linear factor regressions

Following a large literature in empirical asset pricing, we regress our out-of-sample option portfolio excess returns on a variety of factors which are thought to help explain observed variation in the expected returns of different assets. The coefficient of interest, often referred to as the portfolio alpha, is the constant from the regression. A significantly positive alpha indicates that our portfolio generates superior excess returns that cannot be explained by covariation with the excess market return (hedging value) or with known sources of alpha (the additional factors included as regressors). We assess the significance of alpha in each factor regression using a one-sided $t$-test based on Newey-West standard errors.

Table 5 shows the results of our factor regressions. Column 1 provides the simplest case. We regress our option portfolio excess returns on excess market returns$^{13}$ and a constant, with the slope coefficient constrained to equal one. Thus the estimated alpha is simply the average difference between the option portfolio and market portfolio returns. It is equal to 28 basis points when $\lambda = 1/5$ or $\lambda = 1/2$, or 32 basis points when $\lambda = 1/3$. This translates, respectively, to 3.4% or 3.9% annualized returns, which seems economically large. We can reject the null hypothesis of zero alpha against the alternative of positive alpha at the 0.05 level when $\lambda = 1/5$ or $\lambda = 1/3$, but not when $\lambda = 1/2$ (though we

---

$^{13}$To maintain consistency with the rest of our paper, we use the return on the S&P 500 futures contract to measure the market return, rather than the CRSP value-weighted index used more commonly in the empirical asset pricing literature.
<table>
<thead>
<tr>
<th>Variable</th>
<th>λ</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1/5</td>
<td>0.276</td>
<td>0.363</td>
<td>0.332</td>
<td>0.350</td>
<td>0.323</td>
<td>0.309</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.828)</td>
<td>(2.519)</td>
<td>(2.400)</td>
<td>(2.208)</td>
<td>(2.347)</td>
<td>(2.282)</td>
<td>(2.094)</td>
</tr>
<tr>
<td></td>
<td>1/3</td>
<td>0.318</td>
<td>0.427</td>
<td>0.387</td>
<td>0.407</td>
<td>0.413</td>
<td>0.378</td>
<td>0.395</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.979)</td>
<td>(2.838)</td>
<td>(2.662)</td>
<td>(2.631)</td>
<td>(2.675)</td>
<td>(2.566)</td>
<td>(2.554)</td>
</tr>
<tr>
<td></td>
<td>1/2</td>
<td>0.276</td>
<td>0.393</td>
<td>0.355</td>
<td>0.379</td>
<td>0.382</td>
<td>0.348</td>
<td>0.370</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.596)</td>
<td>(2.438)</td>
<td>(2.233)</td>
<td>(2.260)</td>
<td>(2.315)</td>
<td>(2.171)</td>
<td>(2.206)</td>
</tr>
<tr>
<td>Market</td>
<td>1/3</td>
<td>1.000</td>
<td>0.835</td>
<td>0.852</td>
<td>0.845</td>
<td>0.896</td>
<td>0.904</td>
<td>0.897</td>
</tr>
<tr>
<td>SMB</td>
<td>1/3</td>
<td>-</td>
<td>-</td>
<td>-0.037</td>
<td>-0.040</td>
<td>-</td>
<td>-0.034</td>
<td>-0.037</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.679)</td>
<td>(-0.729)</td>
<td></td>
<td>(-0.635)</td>
<td>(-0.677)</td>
</tr>
<tr>
<td>HML</td>
<td>1/3</td>
<td>-</td>
<td>-</td>
<td>0.085</td>
<td>0.078</td>
<td>-</td>
<td>0.078</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.867)</td>
<td>(1.649)</td>
<td></td>
<td>(1.671)</td>
<td>(1.463)</td>
</tr>
<tr>
<td>MOM</td>
<td>1/3</td>
<td>-</td>
<td>-</td>
<td>-0.016</td>
<td>-</td>
<td>-</td>
<td>-0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.427)</td>
<td>-</td>
<td></td>
<td>(-0.385)</td>
<td></td>
</tr>
<tr>
<td>∆VIX</td>
<td>1/3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.087</td>
<td>0.074</td>
<td>0.073</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.429)</td>
<td>(1.233)</td>
<td>(1.221)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>1/5</td>
<td>0.741</td>
<td>0.758</td>
<td>0.760</td>
<td>0.760</td>
<td>0.761</td>
<td>0.762</td>
<td>0.763</td>
</tr>
<tr>
<td></td>
<td>1/3</td>
<td>0.652</td>
<td>0.678</td>
<td>0.684</td>
<td>0.684</td>
<td>0.681</td>
<td>0.686</td>
<td>0.686</td>
</tr>
<tr>
<td></td>
<td>1/2</td>
<td>0.617</td>
<td>0.647</td>
<td>0.654</td>
<td>0.655</td>
<td>0.649</td>
<td>0.655</td>
<td>0.656</td>
</tr>
</tbody>
</table>

Table 5: Linear factor regressions.

This table reports the coefficients (with Newey-West t-statistics in parentheses) for a variety of linear factor regressions, estimated using the realized excess return for our option portfolio (for various choices of the tuning parameter λ) as the dependent variable. We report the estimated constant (alpha) and R² statistic for all three values of λ, and (to conserve space) slope coefficients for only λ = 1/3. Factor returns are aggregated from daily data to match the holding period for each portfolio. The excess return on the S&P 500 futures contract is our market factor. SMB, HML, and MOM are the size, book-to-market, and momentum factors from Kenneth French’s website. ∆VIX is the change in the CBOE VIX index over the holding period.

are very close to rejection in the latter case). In column 2 we remove the constraint that the coefficient of excess market returns is equal to one. The estimated coefficient falls to around 0.8 for all values of λ, while the estimated alphas jump to 36, 43 and 39 basis points when λ = 1/5, 1/3, 1/2 respectively. These alphas are all significantly greater than zero at the 0.01 level. The increase in alpha values indicates that our option portfolios appear to have value as a hedge against modest market fluctuations.

In columns 3-7 we introduce various combinations of other factors from the literature as
additional regressors. We include the Fama and French (1993) size and book-to-market factors, the Jegadeesh and Titman (1993) momentum portfolio and/or the change in the CBOE’s VIX index. With the possible exception of the book-to-market factor, the coefficients of all additional regressors are statistically indistinguishable from zero. Their inclusion has little impact on the $R^2$ of each regression, and leads to estimates of alpha that lie between those reported in columns 1 and 2. The estimated alphas in columns 3-7 range from 31 to 41 basis points, or from 3.8% to 5.0% in terms of annualized return, and are in all cases significantly positive at the 0.05 level. In sum, it is clear that the apparently superior returns generated by our option portfolios cannot be attributed to covariation with the factors considered here.

5 Final remarks

In this paper we have provided further empirical evidence of pricing kernel nonmonotonicity in the US market by showing that it is often possible, in principle, to identify option portfolios whose return distributions stochastically dominate that of the market portfolio. The existence of such portfolios is incompatible with pricing kernel monotonicity, as demonstrated by Dybvig (1988) and Beare (2011). Our empirical results may seem startling, but are entirely consistent with the recent empirical literature on the pricing kernel puzzle, as reviewed by Hens and Reichlin (2012). In particular, our results complement those of Constantinides et al. (2009, 2011), who documented extensive violations of stochastic dominance bounds in the prices of S&P 500 index options over the period 1983-2006.

What are the practical implications of our findings for investors? We certainly do not expect a naïve application of our selection algorithm to provide a reliable guide to option

---

14The regression in column 4, with the Fama-French factors plus the momentum portfolio, was proposed by Carhart (1997) and is a standard benchmark for mutual fund performance evaluation. Inclusion of the VIX factor was proposed by Ang et al. (2000).

15In addition to the factor regressions reported in Table 5, we ran an alternative set of regressions in which the selected option portfolio was augmented with offsetting positions in the market index and risk-free bonds so as to induce zero return covariation with the market portfolio under the estimated physical measure. The estimated alphas for this portfolio were similar to those reported in columns 2-7 of Table 5, ranging from 32 to 43 basis points. Further details are available on request.
portfolio choice in practice. Our empirical analysis has excluded consideration of transaction costs, which are likely to be very high for option portfolios composed of several offsetting long and short positions across a range of strikes. Restrictions associated with short selling, such as margin requirements, may also be prohibitive for small investors.

For large financial institutions such as banks, it may be possible to mitigate the transaction costs associated with forming a portfolio of options with offsetting long and short positions. Such a portfolio could then be issued to investors as a structured product. Structured products are financial derivatives targeted to retail investors that deliver a payoff at maturity contingent on the value of some stock or stock index; see e.g. Hens and Rieger (2014) for a recent discussion. Many structured products are in fact nothing more than a bundle of long and short positions in options on a market index, coupled with a long position in the index. A number of studies have compared the prices of structured products set by issuers to the prices (net transaction costs) implied by market rates for individual options. In a study of the US market for structured products, Benet et al. (2006) found that the prices set by issuers exceeded the implied market prices by 3–5%. Burth et al. (2001) found an average price differential of 1.4% for coupon-free structured products in the highly developed Swiss market, while Stoimenov and Wilkens (2005) found an average differential of 3.67% for structured products based on vanilla options in the German market. By comparison, the estimated alphas for our option portfolios reported above (see Table 5) are in the range 3.4–5.3% at the annual level. Based on these numbers, it is plausible that there may be a profitable market for structured products designed specifically to exploit pricing kernel nonmonotonicity. We leave the investigation of this possibility to future research.

References


36


