Internet Appendix to

"Are Options on Index Futures Profitable for Risk-Averse Investors? Empirical Evidence"*

*Citation format: Constantinides, George M., Michal Czerwonko, Jens Carsten Jackwerth, and Stylianos Perrakis, 2011, Internet Appendix to "Are Options on Index Futures Profitable for Risk-Averse Investors? Empirical Evidence," *Journal of Finance* 66, 1407-1437, http://www.afajof.org/supplements.asp.

- IA.A demonstrates the ability of the Constantinides and Perrakis (2007) lower bounds to identify good buys for calls and puts. The results in Section III of the published paper were derived for the upper bounds since violations of the lower bounds were too infrequent for statistical inference.
- IA.B provides details on the construction and calibration of the index return tree described in Section II.B of the published paper.
- IA.C validates the applicability of the Davidson and Duclos (2000) and Davidson and Duclos (2006) tests to portfolios containing options by verifying the tests' performance with simulated data of known characteristics. These tests were applied to derive the results in Section III of the published paper and are described in Appendix C of that paper.
- IA.D extends the results in Section III of the published paper to selling put options via straddles positions triggered by violations of the call upper bound.
- IA.E to IA.H provide robustness checks in addition to the results in Section IV of the published paper with respect to the following characteristics of the tests: initial portfolio composition; risk aversion coefficient; futures basis risk; assumed equity risk premium.
- IA.I verifies the performance of the Davidson and Duclos (2006) tests for a restricted moneyness range of the call options included in the sample for reasons described in Appendix C of the published paper.

Appendix A: Demonstration of the Ability of the Constantinides and Perrakis (2007) Lower Bounds to Identify Good Buy Options

We construct portfolios with long positions in calls and puts bought at artificial prices equal to their lower bound, as determined in Constantinides and Perrakis (2007), and test the hypotheses $H_0 : IT \succ_2 OT$ and $H_0 : OT \nsucc_2 IT$. The results are reported in Table IA.I. Both hypotheses are rejected. One exception is the case in which the volatility input is the unconditional volatility, which, as Table IA.I shows, has large prediction errors of future volatility. The results demonstrate the ability of the lower bounds to identify good buy call and put options.

Table IA.I
Demonstration of the Ability of the Lower Bounds to Identify Good Buy Options

The equally weighted average of all artificial options equal to their corresponding put lower bound given by equations (14) and (15) of Constantinides and Perrakis (2007) and the corresponding call lower bound given by equation (31) of the same paper, and equivalent to one option per share, was traded at each date. The symbols * and ** denote a difference in sample means of the *OT* and *IT* traders significant at the 5% and 1% levels in a one-sided bootstrap test with 9,999 trials. Maximal *t*-statistics for the Davidson and Duclos (2000) test are compared to critical values of the Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for nominal levels of 1%, 5%, and 10% with $k = 20$ and $\nu = \infty$. The *p*-values for $H_0 : OT \succ_2 IT$, which are greater than 10%, the highest nominal level available in the Stoline and Ury (1979) tables, are not reported here. The *p*-values for the Davidson and Duclos (2006) test are based on 999 bootstrap trials. The *p*-values for $H_0 : IT \nsucc_2 OT$ are equal to one and are not reported here.

Volatility prediction mode	#months with viol. (# months)	$\hat{\mu}_{OT} - \hat{\mu}_{IT}$ (annualized)	DD (2000) *p*-value $H_0 : IT \succ_2 OT$	DD (2006) *p*-value $H_0 : OT \nsucc_2 IT$ 10% trimming in left tail, trimming in right tail as below: 0%	5%	10%
Panel A: Options Purchased at the Call Lower Bound						
Unconditional	247 (247)	0.0063	<0.05	0.271	0.199	0.083
90-day	247 (247)	0.0160*	<0.01	0.050	0.022	0.005
Adjusted IV	226 (226)	0.0134*	<0.01	0.095	0.042	0.009
EGARCH	247 (227)	0.0151*	<0.01	0.064	0.025	0.012
Panel B: Options Purchased at the Put Lower Bound						
Unconditional	247 (247)	0.0014	>0.1	0.462	0.427	0.315
90-day	247 (247)	0.0093*	<0.01	0.083	0.028	0.007
Adjusted IV	226 (226)	0.0125**	<0.01	0.026	0.007	0.000
EGARCH	247 (227)	0.0062	<0.05	0.173	0.067	0.015

Appendix B: Calibration of the Index Return Tree

For every month, we model the path of the daily index return until the option expiration on a T- step recombining tree, where T is the number of trading days in that particular month. For example, if the third Friday of July is on July 27, we record the price of the July option on June 27, which is 30 calendar days earlier. (If June 27 is a holiday, we record the price on June 26.) If there are 21 trading days between June 27 and July 27, we model the path of the daily index return until the option expiration on a 21-step tree.

The paths of the daily index return emanate with m branches from each node. The objective is to match as closely as possible the first four moments of the daily return distribution. As explained in Section II.A of the published text, we fix the mean and use the estimated volatility from one of our four methods. We use as the third and fourth moments the observed sample moments over the 90 preceding calendar days.

In the first step of our algorithm, we pick an odd value for the number of branches m and group the sample of daily returns in a histogram with m bins of equal length (on the log scale) such that the extreme bins are centered on the extreme observed returns. The center of each bin then becomes a state in the lattice, with the ordered states and the corresponding probabilities denoted respectively as x_i and p_i, $i = 1...m$. Note that this equidistant log scale and an odd value for the number of branches m are necessary for the lattice to recombine.

We do not build our lattice by discretizing a kernel-smoothed distribution because this method requires a substantially larger lattice. We do not adopt the Edgeworth/Gram-Charlier binomial lattice methodology, as in Rubinstein (1998), because it sometimes results in negative probabilities.

In a second step, we match our moments by fixing the number of branches *m* and matching the first three moments by changing the spacing (via parameters *a* and *b*) and the probabilities (via parameter *c*). The fourth moment is then matched by changing the number of branches, *m*.

We derive the required parameters *a*, *b*, and *c* by solving the following set of nonlinear equations, which are simply three moment conditions for the constants *a*, *b*, and *c*:

$$\begin{aligned} &\sum_{i=1}^{m} p_i^* \exp\left(ax_i + b\right) - \exp\left(\hat{\mu}\right) = 0 \\ &\sum_{i=1}^{m} p_i^* \left[\exp\left(ax_i + b\right)\right]^2 - \exp\left(\hat{\mu}\right)^2 - \hat{\sigma}^2 = 0 \ , \\ &\sum_{i=1}^{m} p_i^* \left[\exp\left(ax_i + b\right) - \exp\left(\hat{\mu}\right)\right]^3 - \hat{\mu}_3 \hat{\sigma}^3 = 0 \end{aligned} \tag{IA.1}$$

where $\exp\left(\hat{\mu}\right)$ and $\hat{\sigma}^2$ are the first and second target moments, respectively; $\hat{\mu}_3$ is the sample skewness; and $p_i^* \equiv \dfrac{p_i + c\underline{1}_{\left(i \geq n^*\right)}\underline{1}_{\left(p_i \neq 0\right)}}{\sum_{i=1}^{m}\left(p_i + c\underline{1}_{\left(i \geq n^*\right)}\underline{1}_{\left(p_i \neq 0\right)}\right)}$, where $\underline{1}_{(.)}$ is the indicator function and n^* is the index to x_i, which brackets from above the target expected log return $\hat{\mu}$. The first indicator function ensures that the constant *c* is added only to the probabilities in the right tail of the distribution; the second one ensures that the constant *c* is added only to the positive probabilities. Note that the affine transformation of the log states x_i preserves the equal distance between the adjacent states. The constant *a* ensures the desired scale of the log states x_i, the constant *b* ensures the desired location of these states, while the constant *c* increases or decreases the probabilities in the right tail relative to the left one to match the desired skewness. Note that the presented adjustment of the probabilities in the right tail may not yield an admissible solution, that is, we may end up with some negative probabilities. If this is the case, we introduce an analogous adjustment in the left tail of the distribution.

To match the fourth sample moment $\hat{\mu}_4$, we search over *m*, the number of nodes in the lattice. With each new *m* the initial distribution derived from a histogram changes, providing some variability in the fourth moment after the adjustments resulting from solving (IA.1). After a search over a range of *m*'s, we pick the distribution that has the lowest absolute difference between its kurtosis and the sample kurtosis $\hat{\mu}_4$. This search procedure results in very small errors in matching $\hat{\mu}_4$ for the data that we use while we obtain the exact match in the first three moments. For the four volatility prediction modes that we apply in our work, the relative error on the fourth moment has the following characteristics: median 0.003%, 99th percentile 0.105%, maximum 1.659% across 973 observations while we constrain the lattice size *m* to be no larger than 201. This lattice size appears unattractive to derive recursive conditional expectations. However, the use of fast Fourier transforms results in a fairly short processing time. See Cerny (2004).

Appendix C: The Performance of the Tests in Simulated Data

We investigate via simulation the Type I and II errors of the DD (2000) and DD (2006) tests. We find that the latter test has rejection probabilities much lower than one when the null of nondominance is false. Therefore, our results are conservative.

We independently draw monthly index log returns from a normal distribution with mean such that the *arithmetic* annual return has mean 0.0870 and standard deviation 0.1522. These moments are the same as the mean capital gain and standard deviation in our sample. We set the dividend yield of the index equal to zero. We generate 1,000 histories of length 250 months each, roughly equal to the length of our sample; we also generate 1,000 histories of length 1,000 months each. We set the annualized, continuously compounded interest rate to 0.0470. When we investigate Type I errors, we generate prices of one-month calls with a range of moneyness by the Black-Scholes-Merton formula (BSM prices). By construction, BSM prices are within the bounds and do not present an opportunity to build an OT portfolio that stochastically dominates the IT portfolio. When we investigate Type II errors, we generate prices (violating prices) of one-month calls with a range of moneyness by the Black-Scholes formula but setting the interest rate equal to the arithmetic return on the index. By construction, violating prices do present an opportunity to build an OT portfolio that stochastically dominates the IT portfolio.

In Table IA.II, we present simulated rejection probabilities of $H_0 : IT \succ_2 OT$ by the DD (2000) test for moneyness *K/S* = 0.96-1.08 and level of significance $\alpha = 0.01, 0.05, 0.10$. The null is false both for BSM and violating call prices and the test does a good job in rejecting the null. In Table IA.III, we present simulated rejection probabilities of $H_0 : OT \succ_d IT$ by the DD (2000) test. The null is false for BSM call prices but the test only rarely rejects the null. The null is true for violating call prices and the test only rarely rejects the null. We conclude that Type I errors are infrequent but Type II errors are frequent. Increasing the sample size from 250 to 1,000 does not help. Therefore, we rely more heavily on the DD (2006) test, described next.

The DD (2006) test requires that one specify the range of the outcomes. DD (2006) and Davidson (2007) demonstrate that rejecting the null of nondominance is not feasible for the entire support of the joint distribution since the leftmost *t*-statistic is approximately equal to one by construction and the rightmost *t*-statistic corresponds to the difference between sample means, whose significance is a stronger condition than necessary for second-order stochastic dominance. In the case in which the tested samples are uncorrelated, the trimming in the tails simply discards extreme observations until the desired degree of trimming is reached. In our case of correlated (coupled) samples, the trimming is symmetrical with respect to either distribution. To trim in the left tail, we first discard a couple characterized by the lowest value for the first sample, and then a couple characterized by the lowest value for the second sample until the desired proportion of all couples is discarded. We proceed analogously in the right tail with the couples characterized by the highest values for either sample. In all cases presented below, we trim 10% of coupled observations in the left tail while we vary the amount of trimming in the right tail. Note that we may expect DD (2006) to be more conservative than DD (2000) since a pre-condition to the former test is finding nonnegative *t*-statistics in the entire joint support of the two compared distributions (i.e., without trimming).

In Tables IA.IV to IA.VI, we present simulated rejection probabilities of $H_0 : OT \not\succeq_2 IT$ by the DD (2006) test. The null is true for BSM call prices and false for violating prices. In all tables, we trim 10% of the paired outcomes in the left tail. In Table IA.IV, we do not trim the paired outcomes in the right tail. The probability of rejection of the null hypothesis when it is true is very low at all moneyness levels and improves dramatically when the sample size goes from 250 to

1,000. On the other hand, the probability of rejecting the null when it is false is unacceptably low and does not improve when the sample size increases. Tables IA.V and IA.VI present the same information but with trimming of the right-hand tail of the data by 5% and 10%, respectively. With 5% trimming the probability of rejecting the null hypothesis when it is false improves significantly, but still the test is very conservative in its rejection probabilities. It shows some improvement when the sample size increases to 1,000, but the rejection probabilities remain low. On the other hand, the probability of rejecting the null hypothesis when it is true depends strongly on the degree of moneyness of the mispriced option. It is at acceptable levels when moneyness K/S is less than 1.02 but rises for higher numbers at a sample size of 250, but improves in a major way in all cases when the sample size increases to 1,000 and becomes acceptable for all but the highest degree of moneyness. The results are similar for 10% trimming: the probabilities of rejection of H_0 when it is false improve for all but the highest degree of moneyness but the test remains very conservative in its rejections in all cases, with the sample size playing a relatively modest role; the probabilities of rejecting H_0 when it is true at a sample size of 250 are acceptable only for in-the-money calls and improve dramatically at a sample size of 1,000, becoming acceptable for all but the highest degree of moneyness.

We repeat the simulations by drawing returns from the empirical distribution instead of a lognormal distribution and obtain almost identical results for both tests.

Table IA.II
Simulated Rejection Probabilities of $H_0: IT \succ_2 OT$ by the DD (2000) Test

K/S	Writing at Black-Scholes Price			Writing at Call Upper Bound		
	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
			Panel A: Sample Size 250			
0.96	1	1	1	1	1	1
0.98	1	1	1	1	1	1
1.00	1	1	1	1	1	1
1.02	1	1	1	1	1	1
1.04	1	1	1	1	1	1
1.06	1	1	1	1	1	1
1.08	1	1	1	1	1	1
			Panel B: Sample Size 1,000			
0.96	0.863	0.863	0.863	1	1	1
0.98	0.892	0.892	0.892	1	1	1
1.00	0.935	0.935	0.935	1	1	1
1.02	0.971	0.971	0.971	1	1	1
1.04	0.983	0.983	0.983	1	1	1
1.06	0.995	0.995	0.995	1	1	1
1.08	1	1	1	1	1	1

Table IA.III
Simulated Rejection Probabilities of $H_0: OT \succ_2 IT$ by the DD (2000) Test

K/S	Writing at Black-Scholes Price			Writing at Call Upper Bound		
	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
			Panel A: Sample Size 250			
0.96	0.007	0.003	0.002	0	0	0
0.98	0.004	0.003	0	0	0	0
1	0.001	0	0	0	0	0
1.02	0	0	0	0	0	0
1.04	0	0	0	0	0	0
1.06	0	0	0	0	0	0
1.08	0	0	0	0	0	0
			Panel B: Sample Size 1,000			
0.96	0.100	0.067	0.028	0.001	0.001	0
0.98	0.079	0.051	0.018	0.001	0	0
1	0.050	0.036	0.010	0	0	0
1.02	0.021	0.015	0.003	0	0	0
1.04	0.012	0.005	0.001	0	0	0
1.06	0.003	0.003	0	0	0	0
1.08	0	0	0	0	0	0

Table IA.IV
Simulated Rejection Probabilities of $H_0 : OT$ f/$_2$ IT by the DD (2006) Test without Trimming

K/S	Writing at Black-Scholes Price			Writing at Call Upper Bound		
	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
			Panel A: Sample Size 250			
0.96	0.008	0.003	0	0.099	0.052	0.014
0.98	0.009	0.005	0	0.107	0.059	0.016
1	0.013	0.007	0.002	0.110	0.062	0.015
1.02	0.020	0.013	0.002	0.123	0.065	0.021
1.04	0.029	0.018	0.005	0.136	0.078	0.028
1.06	0.048	0.032	0.010	0.156	0.093	0.040
1.08	0.099	0.066	0.031	0.187	0.136	0.072
			Panel B: Sample Size 1,000			
0.96	0	0	0	0.090	0.051	0.007
0.98	0	0	0	0.098	0.053	0.010
1	0	0	0	0.111	0.056	0.013
1.02	0.001	0	0	0.110	0.060	0.020
1.04	0.005	0.003	0	0.124	0.073	0.028
1.06	0.009	0.004	0	0.124	0.069	0.025
1.08	0.024	0.010	0.001	0.145	0.088	0.028

Table IA.V
Simulated Rejection Probabilities of $H_0 : OT$ f/$_2$ IT by the DD (2006) Test with 5% Trimming in the Right Tail

K/S	Writing at Black-Scholes Price			Writing at Call Upper Bound		
	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
			Panel A: Sample Size 250			
0.96	0.020	0.009	0	0.188	0.120	0.044
0.98	0.028	0.013	0.003	0.221	0.142	0.053
1	0.043	0.023	0.007	0.274	0.179	0.081
1.02	0.119	0.075	0.027	0.341	0.268	0.130
1.04	0.149	0.106	0.046	0.468	0.406	0.286
1.06	0.267	0.240	0.191	0.508	0.462	0.371
1.08	0.366	0.365	0.365	0.544	0.544	0.544
			Panel B: Sample Size 1,000			
0.96	0.002	0	0	0.319	0.207	0.083
0.98	0.005	0	0	0.375	0.263	0.115
1	0.010	0.005	0.000	0.454	0.357	0.183
1.02	0.039	0.039	0.022	0.485	0.460	0.316
1.04	0.069	0.060	0.033	0.489	0.489	0.483
1.06	0.113	0.113	0.113	0.505	0.505	0.502
1.08	0.209	0.209	0.209	0.509	0.509	0.509

Table IA.VI

Simulated Rejection Probabilities of $H_0 : OT \succcurlyeq_2 IT$ by the DD (2006) Test with 10% Trimming in the Right Tail

K/S	Writing at Black-Scholes Price			Writing at Call Upper Bound		
	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
Panel A: Sample Size 250						
0.96	0.046	0.022	0.005	0.360	0.255	0.122
0.98	0.062	0.037	0.010	0.416	0.313	0.156
1	0.094	0.060	0.022	0.470	0.414	0.262
1.02	0.196	0.195	0.167	0.489	0.481	0.420
1.04	0.225	0.221	0.203	0.512	0.512	0.512
1.06	0.288	0.288	0.288	0.536	0.536	0.536
1.08	0.366	0.366	0.366	0.544	0.544	0.544
Panel B: Sample Size 1,000						
0.96	0.010	0.005	0	0.453	0.453	0.371
0.98	0.015	0.008	0.004	0.479	0.479	0.465
1	0.025	0.023	0.010	0.478	0.478	0.478
1.02	0.039	0.039	0.039	0.485	0.485	0.485
1.04	0.072	0.072	0.072	0.489	0.489	0.489
1.06	0.113	0.113	0.113	0.505	0.505	0.505
1.08	0.209	0.209	0.209	0.509	0.509	0.509

Appendix D: Returns to Straddles/Strangles Triggered by Call Upper Bound Violations

We examine the existence of good sell put options by testing the policy of shorting straddles and strangles triggered by observing call options violating their upper bounds at the same or similar strike price. The results are reported in Table IA.VII and show that the portfolio of the OT trader stochastically dominates the portfolio of the IT trader.

Table IA.VII
Returns of Straddles/Strangles Trader and Index Trader

The equally weighted average of all violating options equivalent to one call and one put per share was traded at each date. Trades were executed whenever there was a call violating the upper bound and a put traded at the same strike (for straddles) or within 0.98 to 1.02 moneyness bound (for strangles) for the same date. The symbols * and ** denote a difference in sample means of the *OT* and *IT* traders significant at the 5% and 1% levels in a one-sided bootstrap test with 9,999 trials. Maximal *t*-statistics for the Davidson and Duclos (2000) test are compared to critical values of the Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for nominal levels of 1%, 5%, and 10% with k = 20 and $\nu = \infty$. The *p*-values for $H_0 : OT \succ_2 IT$, which are greater than 10%, the highest nominal level available in the Stoline and Ury (1979) tables, are not reported here. *p*-values for the Davidson and Duclos (2006) test are based on 999 bootstrap trials. The *p*-values for $H_0 : IT \nsucc_2 OT$ are equal to one and are not reported here.

Volatility prediction mode	#months with viol. (# months)	$\hat{\mu}_{OT} - \hat{\mu}_{IT}$ (annualized)	DD (2000) *p*-value $H_0 : IT \succ_2 OT$	DD (2006) *p*-value $H_0 : OT \nsucc_2 IT$ 10% trimming in left tail, trimming in right tail as below:		
				no trimming	5% trimming	10% trimming
			Panel A: Straddles			
Unconditional	34 (247)	0.0058	<0.1	0.290	0.171	0.066
90-day	66 (247)	0.0068	<0.05	0.262	0.157	0.040
Adjusted IV	71 (226)	0.0165**	<0.05	0.048	0.016	0.018
EGARCH	40 (247)	0.0158**	<0.1	0.034	0.039	0.042
			Panel B: Straddles and Strangles			
Unconditional	40 (247)	0.0081	<0.1	0.231	0.138	0.054
90-day	80 (247)	0.0143*	<0.01	0.126	0.042	0.011
Adjusted IV	94 (226)	0.0235**	<0.01	0.020	0.023	0.025
EGARCH	54 (247)	0.0172**	<0.05	0.053	0.012	0.014

Appendix E: Robustness to Initial Portfolio Composition

We investigate the robustness of the results to the initial portfolio composition by including various open option positions in the stock account portion of the IT portfolio. The results are reported in Table IA.VIII and show that our main conclusion that the portfolio of the OT trader stochastically dominates the portfolio of the IT trader is robust to the initial portfolio composition.

Table IA.VIII
Returns of Options Trader and Index Trader with Open Options Positions in the IT Portfolio

The equally weighted average of all violating options within the indicated moneyness range equivalent to 0.8 options per share in Panel A and one option per share in Panels B and C was traded at each date. The symbols * and ** denote a difference in sample means of the *OT* and *IT* traders significant at the 5% and 1% levels in a one sided bootstrap test with 9,999 trials. Maximal *t*-statistics for the Davidson and Duclos (2000) test are compared to critical values of the Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for nominal levels of 1%, 5%, and 10% with k = 20 and $\nu = \infty$. The *p*-values for $H_0 : OT \succ_2 IT$, which are greater than 10%, the highest nominal level available in the Stoline and Ury (1979) tables, are not reported here. The *p*-values for the Davidson and Duclos (2006) test are based on 999 bootstrap trials. The *p*-values for $H_0 : IT \not\succ_2 OT$ are equal to one and are not reported here.

Volatility prediction mode	#months with viol. (# months)	$\hat{\mu}_{OT} - \hat{\mu}_{IT}$ (annualized)	DD (2000) *p*-value $H_0 : IT \succ_2 OT$	DD (2006) *p*-value $H_0 : OT \not\succ_2 IT$ 10% trimming in left tail, trimming in right tail as below: no trimming	5%	10%
Panel A: 0.2 ATM Calls Short, Equivalent of 0.8 Calls in Violation Written						
Unconditional	43 (247)	0.0028	<0.01	0.244	0.026	0.000
90-day	100 (247)	0.0042	<0.01	0.149	0.005	0.002
Adjusted IV	120 (226)	0.0055*	<0.01	0.133	0.032	0.000
EGARCH	65 (247)	0.0055**	<0.01	0.071	0.000	0.000
Panel B: One ATM Call Long, Equivalent of One Call in Violation Written						
Unconditional	43 (247)	0.0035	<0.01	0.244	0.062	0.000
90-day	100 (247)	0.0052	<0.01	0.149	0.020	0.002
Adjusted IV	120 (226)	0.0069*	<0.01	0.133	0.047	0.000
EGARCH	65 (247)	0.0068**	<0.01	0.071	0.000	0.000
Panel C: 0.5 ATM Puts Long, Equivalent of One Call in Violation Written						
Unconditional	43 (247)	0.0031	<0.01	0.244	0.023	0.001
90-day	100 (247)	0.0048	<0.01	0.149	0.007	0.000
Adjusted IV	120 (226)	0.0065*	<0.01	0.133	0.036	0.000
EGARCH	65 (247)	0.0065**	<0.01	0.067	0.002	0.003

Appendix F: Robustness to the Risk Aversion Coefficient

We estimate the returns of the IT portfolio by optimally rebalancing it according to the procedure described in Section II.C of the paper, assuming that the risk aversion coefficient is equal to 10 rather than two. The results are shown in Table IA.IX. As noted in Section IV.C of the paper, the results are virtually indistinguishable from those of Table V in the main paper.

Table IA.IX
Returns of Options Trader and Index Trader with Risk Aversion Coefficient 10

The equally weighted average of all violating options equivalent to one option per share was traded at each date. The symbols * and ** denote a difference in sample means of the *OT* and *IT* traders significant at the 5% and 1% levels in a one sided bootstrap test with 9,999 trials. Maximal *t*-statistics for the Davidson and Duclos (2000) test are compared to critical values of the Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for nominal levels of 1%, 5%, and 10% with k = 20 and $\nu = \infty$. The *p*-values for $H_0 : OT \succ_2 IT$, which are greater than 10%, the highest nominal level available in the Stoline and Ury (1979) tables, are not reported here. The *p*-values for the Davidson and Duclos (2006) test are based on 999 bootstrap trials. The *p*-values for $H_0 : IT \not\succ_2 OT$ are equal to one and are not reported here.

Volatility prediction mode	#months with viol. (# months)	$\hat{\mu}_{OT} - \hat{\mu}_{IT}$ (annualized)	DD (2000) *p*-value $H_0 : IT \succ_2 OT$	DD (2006) *p*-value $H_0 : OT \not\succ_2 IT$ 10% trimming in left tail, trimming in right tail as below: no trimming	5%	10%
			Panel A: Call Upper Bound			
Unconditional	43 (247)	0.0006	<0.01	0.275	0.041	0.001
90-day	100 (247)	0.0008	<0.01	0.201	0	0.004
Adjusted IV	120 (226)	0.0013*	<0.01	0.124	0.031	0
EGARCH	65 (247)	0.0012**	<0.01	0.083	0	0
			Panel B: Put Upper Bound			
Unconditional	23 (247)	0.0002	>0.1	0.427	0.210	0.157
90-day	16 (247)	-0.0002	>0.1	1	1	1
Adjusted IV	4 (226)	n/a	n/a	n/a	n/a	n/a
EGARCH	9 (247)	n/a	n/a	n/a	n/a	n/a

Appendix G: Robustness to the Futures Basis Risk

We investigate the robustness of the results to the futures basis risk by estimating the bounds under the assumption that futures basis risk is zero, $\bar{\varepsilon} = 0$, in equation (2) of the paper. The results are reported in Table IA.X and are very similar to those in Table V in the paper. We conclude that the results are robust to the futures basis risk.

Table IA.X
Returns of Options Trader and Index Trader without Futures Basis Risk

The table differs from Table V only in that the basis risk is set at zero, $\bar{\varepsilon} = 0$, instead of bounding the risk by $\bar{\varepsilon} = 0.5\%$. The equally weighted average of all violating options equivalent to one option per share was traded at each date. The symbol ** denotes a difference in sample means of the *OT* and *IT* traders significant at the 5% level in a one-sided bootstrap test with 9,999 trials. Maximal *t*-statistics for the Davidson-Duclos (2000) test are compared to critical values of the Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for nominal levels of 1%, 5%, and 10% with $k = 20$ and $\nu = \infty$. The *p*-values for $H_0 : OT \succ_2 IT$, which are greater than 10%, the highest nominal level available in the Stoline and Ury (1979) tables, are not reported here. The *p*-values for the Davidson-Duclos (2006) test are based on 999 bootstrap trials. The *p*-values for $H_0 : IT \nsucc_2 OT$ are equal to one and are not reported here.

Volatility prediction mode	#months with viol. (# months)	$\hat{\mu}_{OT} - \hat{\mu}_{IT}$ (annualized)	DD (2000) *p*-value $H_0 : IT \succ_2 OT$	DD (2006) *p*-value $H_0 : OT \nsucc_2 IT$ 10% trimming in left tail, trimming in right tail as below:		
				0%	5%	10%
Panel A: Call Upper Bound						
Unconditional	67 (247)	0.0012	<0.01	0.412	0.128	0.005
90-day	156 (247)	0.0083**	<0.01	0.083	0.011	0.000
Adjusted IV	195 (226)	0.0032	<0.01	0.337	0.255	0.076
EGARCH	112 (247)	0.0037	<0.01	0.261	0.074	0.000
Panel B: Put Upper Bound						
Unconditional	36 (247)	-0.0015	<0.1	1	1	1
90-day	52 (247)	-0.0003	<0.01	1	1	1
Adjusted IV	64 (226)	-0.0012	<0.01	1	1	1
EGARCH	38 (247)	0.0014	<0.01	0.374	0.199	0.004

Appendix H: Robustness to the Assumed Equity Risk Premium

In the paper, we set the expected premium on the index to 4%. We investigate the robustness of the results to the assumed equity risk premium. Here we set the expected premium on the index at 2% instead of 4%. The results are reported in Table IA.XI, Panel A. Since the call and put upper bounds are higher, the options trader is more selective than before in writing options that violate these bounds. In Panel B, we report the results when we set the premium to 6%. In both cases, the stochastic dominance results in writing calls are as strong as in Table V. We conclude that the results in Table V are robust to the assumption that the expected premium on the index is 4%.

Table IA.XI
Returns of Options Trader and Index Trader with Different Risk Premium

The equally weighted average of all violating options equivalent to one option per share was traded at each date. The symbol ** denotes a difference in sample means of the *OT* and *IT* traders significant at the 5% level in a one-sided bootstrap test with 9,999 trials. Maximal *t*-statistics for the Davidson-Duclos (2000) test are compared to critical values of the Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for nominal levels of 1%, 5%, and 10% with $k = 20$ and $\nu = \infty$. The *p*-values for $H_0 : OT \succ_2 IT$, which are greater than 10%, the highest nominal level available in the Stoline and Ury (1979) tables, are not reported here. The *p*-values for the Davidson-Duclos (2006) test are based on 999 bootstrap trials. The *p*-values for $H_0 : IT \not\succ_2 OT$ are equal to one and are not reported here.

Volatility prediction mode	#months with viol. (# months)	$\hat{\mu}_{OT} - \hat{\mu}_{IT}$ (annualized)	DD (2000) *p*-value $H_0 : IT \succ_2 OT$	DD (2006) *p*-value $H_0 : OT \not\succ_2 IT$ 10% trimming in left tail, trimming in right tail as below:		
				no trimming	5%	10%
Panel A: Equity Risk Premium 2%						
Unconditional	48 (247)	0.0048	<0.01	0.176	0.016	0.000
90-day	114 (247)	0.0057	<0.01	0.168	0.005	0.000
Adjusted IV	140 (226)	0.0058*	<0.01	0.199	0.120	0.000
EGARCH	77 (247)	0.0045	<0.01	0.214	0.045	0.000
Panel B: Equity Risk Premium 6%						
Unconditional	38 (247)	0.0009	<0.01	0.434	0.052	0.004
90-day	85 (247)	0.0035	<0.01	0.228	0.012	0.002
Adjusted IV	96 (226)	0.0052*	<0.01	0.156	0.042	0
EGARCH	58 (247)	0.0051*	<0.01	0.118	0	0

Appendix I: Restricting the Moneyness Range of Violating Calls

In Internet Appendix C, using simulated data with characteristics that mirror our sample, we compute the rejection probabilities of the null hypothesis when it is true as well as when it is false. DD (2006) is a weak test without trimming, since it has very low probabilities of rejection of the nondominance null even when it is false. With 5% trimming, the test is still conservative as far as rejecting the false nondominance null. Problems with rejection of the null when it is true occur only for deep OTM options. For this reason, we repeat the stochastic dominance tests for the call upper bound in Panel A of Table V for a restricted moneyness range, that is, by removing violating OTM calls outside the range from the sample. The results are reported in Table IA.XII and remain essentially unchanged.

Table IA.XII
Returns of Call Trader and Index Trader When Restricting the Moneyness Range of Violating Calls

The equally weighted average of all violating calls within the indicated moneyness range equivalent to one call per share was traded at each date. The symbols * and ** denote a difference in sample means of the *OT* and *IT* traders significant at the 5% and 1% levels in a one sided bootstrap test with 9,999 trials. Maximal *t*-statistics for the Davidson and Duclos (2000) test are compared to critical values of the Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for nominal levels of 1%, 5%, and 10% with $k = 20$ and $\nu = \infty$. The *p*-values for $H_0 : OT \succ_2 IT$, which are greater than 10%, the highest nominal level available in the Stoline and Ury (1979) tables, are not reported here. The *p*-values for the Davidson and Duclos (2006) test are based on 999 bootstrap trials. The *p*-values for $H_0 : IT \not\succ_2 OT$ are equal to one and are not reported here.

Volatility prediction mode	#months with viol. (# months)	$\hat{\mu}_{OT} - \hat{\mu}_{IT}$ (annualized)	DD (2000) *p*-value $H_0 : IT \succ_2 OT$	DD (2006) *p*-value $H_0 : OT \not\succ_2 IT$ 10% trimming in left tail, trimming in right tail as below:		
				no trimming	5%	10%
			Panel A: K/F < 1.04			
Unconditional	42 (247)	0.0034	<0.01	0.184	0.004	0.004
90-day	87 (247)	0.0051	<0.01	0.164	0.000	0.000
Adjusted IV	108 (226)	0.0089**	<0.01	0.106	0.002	0.004
EGARCH	58 (247)	0.0079**	<0.01	0.124	0.000	0.000
			Panel B: K/F < 1.03			
Unconditional	42 (247)	0.0026	<0.01	0.303	0.068	0.002
90-day	80 (247)	0.0045	<0.01	0.199	0.018	0.000
Adjusted IV	98 (226)	0.0088*	<0.01	0.088	0.016	0.004
EGARCH	50 (247)	0.0082**	<0.01	0.088	0.000	0.000
			Panel C: K/F < 1.02			
Unconditional	39 (247)	0.0027	<0.01	0.362	0.094	0.002
90-day	75 (247)	0.0047	<0.01	0.256	0.050	0.000
Adjusted IV	90 (226)	0.0071*	<0.01	0.174	0.072	0.000
EGARCH	46 (247)	0.0079**	<0.01	0.088	0.008	0.000

References

Cerny, Ales, 2004, Introduction to Fast Fourier Transform in Finance, *Journal of Derivatives* 12, 73-88.

Constantinides, George M., and Stylianos Perrakis, 2007, Stochastic Dominance Bounds on American Option Prices in Markets with Frictions, *Review of Finance* 11, 71-115.

Davidson, Russell, and Jean-Yves Duclos, 2000, Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality, *Econometrica* 68, 1435-1464.

Davidson, Russell, and Jean-Yves Duclos, 2006, Testing for Restricted Stochastic Dominance, Working Paper, McGill University.

Stoline, Michael R., and Hans K. Ury, 1979, Tables of the Studentized Maximum Modulus Distribution and an Application to Multiple Comparisons among Means, *Technometrics* 21, 87-93.