

Recovering Risk Aversion from Option Prices and Realized Returns

Jens Carsten Jackwerth

University of Wisconsin, Madison

A relationship exists between aggregate risk-neutral and subjective probability distributions and risk aversion functions. We empirically derive risk aversion functions implied by option prices and realized returns on the S&P500 index simultaneously. These risk aversion functions dramatically change shapes around the 1987 crash: Precrash, they are positive and decreasing in wealth and largely consistent with standard assumptions made in economic theory. Postcrash, they are partially negative and partially increasing and irreconcilable with those assumptions. Mispricing in the option market is the most likely cause. Simulated trading strategies exploiting this mispricing show excess returns, even after accounting for the possibility of further crashes, transaction costs, and hedges against the downside risk.

A relationship exists between aggregate (i.e., marketwide) risk-neutral and subjective probability distributions and risk aversion functions across wealth. In each state of the world the following relationship holds:

$$\begin{aligned} \text{risk-neutral probability} &= \text{subjective probability} \\ &\times \text{risk aversion adjustment.} \end{aligned}$$

The risk-neutral probability is the price, multiplied by the riskfree return, that an investor would pay for receiving \$1 in that state. We need to multiply by the riskfree return in order to have a proper probability distribution that sums to one. The subjective probability is simply the assessment of an investor of how likely a state is to occur. These two probabilities would be identical if the investor were indifferent to risk. However, the investor might value a dollar more highly in certain states, namely ones where wealth is low. The risk aversion adjustment indicates these preferences of the investor. Once we know both risk-neutral and subjective probability distributions, we can derive those preferences empirically.

While the theoretical relationship is well known, this article is the first to recover risk aversion empirically from the two distributions. This methodology is also used in Ait-Sahalia and Lo (2000). Other, less closely related

I would like to thank Kerry Back (the editor), Ian Cooper, Julian Franks, Nils Hakansson, Hayne Leland, Anthony Neuberger, Mark Rubinstein, two anonymous referees, and seminar participants at Berkeley, Michigan, McGill, Yale, Harvard, London Business School, Long Term Capital Management (London), University of Southern California, Inquire (UK) at Newcastle, Davis, City University London, Warwick, Bank of England (London), the Computational Finance Conference at London, and the European Financial Management Conference at Lisbon for their helpful discussions. Address correspondence to Jens Carsten Jackwerth, 975 University Ave., Room 5279, Madison, WI 53706. Email: jjackwerth@bus.wisc.edu.

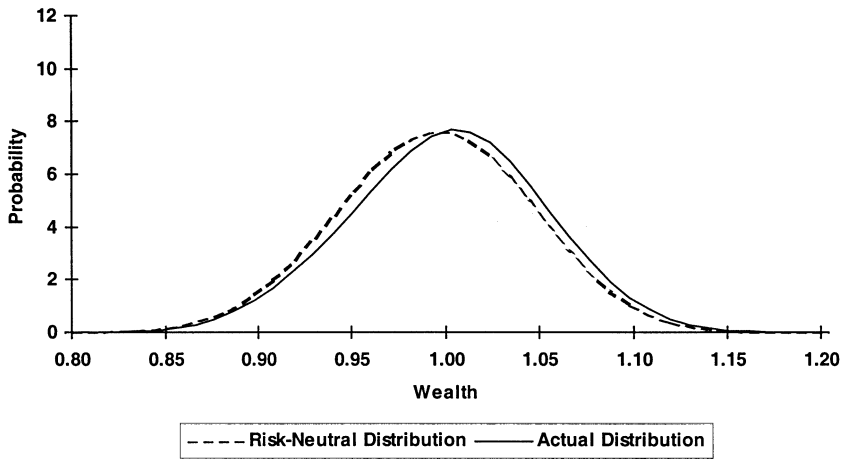


Figure 1
Pre-crash risk-neutral and actual distributions.

The risk-neutral distribution is the 31-day option implied distribution on October 22, 1986. The actual distribution is a kernel density of the 31-day nonoverlapping returns over the 4 years prior to October 22, 1986.

studies are Derman et al. (1997) and Rosenberg and Engle (1997). In our study, the risk-neutral distributions are recovered from option prices by a variation of the method of Jackwerth and Rubinstein (1996).¹ For the subjective distributions, we assume that they are approximated by the actual (historical) return distribution of a broad index such as the S&P500.

At this point, we are confronted with a potential puzzle which is depicted in Figures 1 and 2. During the pre-crash period both the risk-neutral and the actual distributions look about lognormal. Figure 1 depicts the distributions on October 22, 1986, and is typical of the pre-crash period. The actual distribution is located slightly to the right due to the risk premium. We know that in this case a power utility function of moderate risk aversion is likely to provide a reasonable risk aversion adjustment for translating one distribution into the other.

Next, we turn to the post-crash period where the actual distribution looks about lognormal again. However, Jackwerth and Rubinstein (1996) document that the risk-neutral distribution is now left-skewed and leptokurtic (more peaked). Figure 2 depicts the distributions on April 15, 1992, and is typical of the post-crash period. If we conclude that the risk-neutral distributions changed in shape around the crash and that the actual distributions (which proxy for the subjective distributions) did not, then we could conclude that

¹ For alternative methods see Shimko (1993), Rubinstein (1994), and the extensions in Brown and Toft (1999) for polynomial interpolation of the smile. Melick and Thomas (1997) use lognormal mixtures and Ait-Sahalia and Lo (1998) use kernels. Earlier, theoretical articles are Ross (1976), Banz and Miller (1978), and Breeden and Litzenberger (1978).

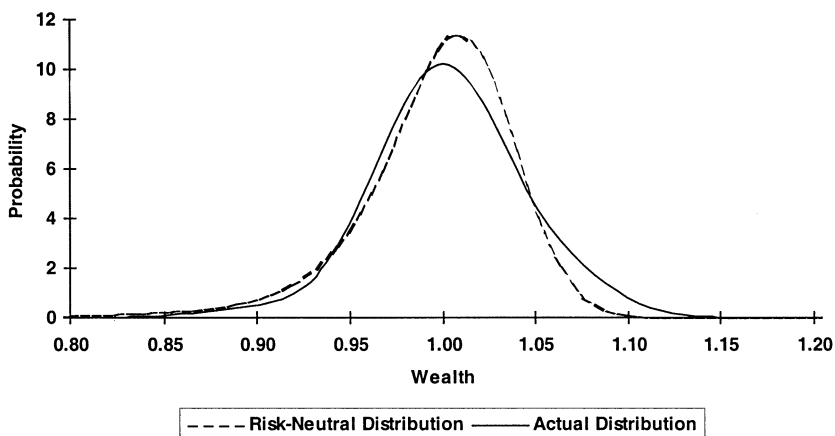


Figure 2
Postcrash risk-neutral and actual distributions.
 The risk-neutral distribution is the 31-day option implied distribution on April 15, 1992. The actual distribution is a kernel density of the 31-day nonoverlapping returns over the 4 years prior to April 15, 1992.

the third component (risk aversion functions) changed, too, around the crash. This article sets out to empirically investigate this possibility.

We find indeed that implied risk aversion functions dramatically change shapes around the 1987 crash and thereafter exhibit negative and increasing risk aversion across parts of the wealth dimension. These findings are remarkably stable with respect to changes of almost all model parameters. Moreover, a number of potential criticisms of our methodology are discussed and rejected. One potential explanation is that some options are mispriced during the postcrash period. A simulated trading strategy exploiting this mispricing mean variance dominates holding the market. This result would not surprise us since the trading strategies involve selling put options and there were no more significant crashes between 1988 and 1995. However, we simulate crashes as well and still find strategies that strictly mean variance dominate holding the market even in the presence of transaction costs. A final concern is that the trading strategies can bankrupt the investor. Nonetheless, we are able to hedge the downside risk and still achieve excess returns.

Section 1 presents the model for obtaining the implied risk aversion functions. Section 2 describes the data and procedures used for the empirical work. Section 3 presents and analyzes the empirical results. Section 4 concludes.

1. Implied Risk Aversion Functions

In this section we introduce a model of the economy within which we derive implied risk aversion functions across wealth. We consider a complete

market economy, which implies the existence of a representative investor [Constantinides (1982)]. The representative investor has an endowed wealth of 1 unit and a fixed time horizon t . The problem is to maximize utility across future wealth subject to the budget constraint:

$$\max_W \int Q(W)U(W)dW - \lambda \left(\frac{1}{r^t} \int P(W)WdW - 1 \right), \quad (1)$$

where

W = future wealth

Q = subjective probability across states

U = state-independent utility function across states

λ = shadow price of the budget constraint (incremental utility associated with increasing the available investment marginally)

r = 1 + riskfree interest rate

t = time horizon

P = risk-neutral probability across states

By differentiating with respect to wealth, we find the first-order conditions which have to hold in equilibrium. In equilibrium we know that the aggregate investor has to hold the market portfolio. Thus, if S is the return (cum dividends) on the market portfolio across states, then the following relation has to hold in equilibrium:

$$U'(S) = \frac{\lambda P(S)}{r^t Q(S)}. \quad (2)$$

The approach to find the implied utility functions by using Equation (2) directly suffers from the fact that we cannot measure the value of the shadow price λ accurately. Its formula involves the probability in the tails of the subjective distribution where we have little confidence in our estimates. However, there exists an interesting way around this problem. We can find a simple expression for the absolute risk aversion, which can be expressed in terms of the subjective and risk-neutral distributions after differentiating Equation (2) with respect to S once more:

$$U''(S) = \frac{\lambda P'(S)Q(S) - P(S)Q'(S)}{r^t Q^2(S)}. \quad (3)$$

In our economy, we would like to ensure that utility functions are increasing ($U' > 0$) and concave ($U'' < 0$). U' is guaranteed to be positive since all terms on the right-hand side of Equation (2) are positive. However, the concavity of the utility function is not guaranteed by Equation (3), and we will check for violations in the empirical section. Finally, we can write an expression for the absolute risk aversion in terms of the subjective and risk-neutral distributions [Leland (1980)]:

$$RA = -\frac{U''(S)}{U'(S)} = -\frac{\frac{\lambda P'(S)Q(S) - P(S)Q'(S)}{r^t Q^2(S)}}{\frac{\lambda P(S)}{r^t Q(S)}} = \frac{Q'(S)}{Q(S)} - \frac{P'(S)}{P(S)}. \quad (4)$$

The expression for the risk aversion functions no longer depends on the shadow price λ and is thus much more tractable. In particular, with Equation (4) we are able to obtain the risk aversion for any state by using only local information. Thus we can concentrate our attention on the center states with the highest probability of occurring and neglect the tails of the distributions. In particular, relative errors around the probability distributions are lower in the center than in the tails.

2. Data and Estimation

For our economy with a representative investor, we would like to use a broad index resembling the market in our empirical work. Thus the empirical research in this article is based on a database which contains minute-by-minute trades and quotes covering S&P500 European index options, S&P500 index futures, and S&P500 index levels from April 2, 1986, through December 29, 1995 [Jackwerth and Rubinstein (1996)]. Moreover, we use daily closing prices and dividends on the S&P500 from January 2, 1928, through December 29, 1995. Quotes with excessive spreads (200 cents for options and 20 cents for futures) are eliminated. Option quotes violating general no-arbitrage conditions are eliminated. Option quotes violating no-arbitrage conditions in vertical or butterfly spreads were tentatively dropped and then reintroduced in such a way as to keep as many quotes as possible.

We assume that the future dividends can be approximated by a dividend yield, which is based on the actual payments throughout the life of the option. Thereby we ignore the lumpiness of the dividend payments; but since we are dealing with European index options, this matters little as we are only interested in the total dividend paid over the life of the option.

For the index level, we do not use the observed index itself but a futures-based index. The reasons are that the index can be stale due to infrequent trading and that traders use the futures contract for hedging rather than the underlying. In order to obtain the futures-based index, we calculate the implied interest rates between all futures quotes and trades and the corresponding index level. Next, we compute the median daily interest rates for all available maturities and deflate all futures quotes and trades by these implied repo-rates. For each minute, the median of all deflated quotes and trades is calculated and used as the futures-based index for that minute. As we are interested in working with representative daily sets of options, we use for the daily index level the median of the intraday levels of the futures-based index.

For the interest rate, we calculate the put-call parity implied borrowing and lending rates for all feasible pairs of options of the same maturity on a given day where we use the futures-based index. We then compute the median daily borrowing and lending rates. Finally, we use the average rate as our daily interest rate for that maturity.

We are now ready to look at the representative daily set of options. As further safeguards, we omit the first and last 30 minutes of quotes each day and all quotes with bids of 1/16 or less. Also, we only consider options with strike price:index level (moneyness) ratios between 0.84 and 1.12. This procedure eliminates far-away-from-the-money observations, which are potentially unreliable due to their low volume and low sensitivity to volatility. The remaining bid and ask quotes are translated into their Black–Scholes implied volatilities,² and we use the midpoint (average) implied volatilities thereafter. For every strike price we obtain the mean daily implied midpoint volatility. If there are fewer than 20 quotes for a strike price during the day then we omit these quotes from the data. Finally, using the daily futures-based index level, we translate the implied midpoint volatilities back into option prices. This is our representative daily set of option prices.

Every 31 days from April 1986 through December 1995, we estimate the subjective and risk-neutral distributions for the 31-day index return and infer from these the absolute risk aversion functions for a representative investor with a 1-month investment horizon. Out of our total sample we are left with 114 days for which we have nonoverlapping 31-day risk aversion functions.

2.1 Risk-neutral probability distribution

In order to find the 31-day risk-neutral probability distribution, Jackwerth and Rubinstein (1996) propose a method where they search for the smoothest risk-neutral distribution, which at the same time explains the option prices. The trade-off between the two contradicting goals is exogenously specified. Three problems arise with this approach which leads us to deviate slightly from the Jackwerth–Rubinstein method. First, matching the option prices by minimizing the squared error puts more weight on in-the-money compared to out-of-the-money options. Instead, we match implied volatilities, which are of the same order of magnitude across striking prices. Second, the Jackwerth–Rubinstein method does not account for the fact that at-the-money option prices vary less throughout the day than away-from-the-money options. Therefore we measure the standard deviation of the implied volatilities during the day for each strike price. This information is then used to force a tighter fit around at-the-money options (where standard deviations are lower) and to allow for a looser fit for away-from-the-money options. Third, the Jackwerth–Rubinstein method uses as a measure of smoothness the integral of squared curvature of the probability distribution. Due to the general bell-shape of probability distributions there will always be some curvature present. Here we minimize the curvature of the implied volatilities (volatility smile), which in a Black–Scholes economy would be flat. In the

² The implied volatility (σ^*) causes the Black–Scholes formula to accurately price the option in the market: $C = sd^{-t}N(x) - Kr^{-t}N(x - \sigma^*\sqrt{t})$, where s is the index level, d is the dividend yield, t is the time, $N(\bullet)$ is the cumulative normal distribution, $x = (\ln(sd^{-t}/Kr^{-t}))/\sigma^*\sqrt{t} + \frac{1}{2}\sigma^*\sqrt{t}$, and K is the strike price.

limiting case, where we observe only one option, the method coincides with the Black–Scholes model and backs out a lognormal risk-neutral distribution.

Our new method thus achieves a similar objective as the Jackwerth–Rubinstein method, namely finding a smooth risk-neutral distribution, which at the same time explains the option prices, but avoids the above problems. In order to implement the method, we discretize the future index value such that the spacing coincides with the strike prices and the midpoints between the strike prices:

$$S_j = S_0 + j\Delta, j = 0, \dots, 99 \quad (5)$$

where

S_0 = lowest future index value; evenly divisible by Δ and at least 50 Δ
smaller than today's index value
 $\Delta = 2.50$

In order to find the risk-neutral probability distribution, we minimize the following objective function for which the first-order conditions can be solved in closed form:

$$\min_{\sigma_j} (1-p) \sum_{j=0}^{99} (\sigma_j'')^2 + p \sum_i \left(\frac{\sigma_i - \bar{\sigma}_i}{STD_i} \right)^2 \quad (6)$$

where

σ_j = implied volatility associated with the strike price $K_j = S_j$
 p = trade-off parameter for balancing smoothness versus fit
 σ_j'' = second derivative of implied volatility with respect to strike prices, numerically approximated by
 $\sigma_j'' = (\sigma_{j-1} - 2\sigma_j + \sigma_{j+1})/\Delta^2$
 σ_i = implied volatility associated with the strike price K_i
 $\bar{\sigma}_i$ = implied volatility of observed option with strike price K_i
 STD_i = standard deviation of $\bar{\sigma}_i$

Alternatively, we could employ the related method of kernel regression [Härdle (1990)] which yields very similar volatility functions. Once we have solved Equation (6) for the optimal implied volatility function, we calculate the Black–Scholes option prices across strike prices and differentiate twice with respect to the strike price. After multiplying by the riskfree return (r^t), we write down the risk-neutral probability distribution directly:

$$P(S_j) = r^t \left[\frac{r^{-t} n(d2_j)}{S_j \sigma_j \sqrt{t}} [1 + 2S_j \sqrt{t} d1_j \sigma_j'] \right. \\ \left. + sd^{-t} \sqrt{t} n(d1_j) \left[\sigma_j'' + \frac{d1_j d2_j}{\sigma_j} (\sigma_j')^2 \right] \right] \quad (7)$$

where

$$\begin{aligned}
 n(x) &= \text{standard normal density function} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \\
 d1_j &= (\ln(sd^{-t}/S_j r^{-t})/\sigma_j \sqrt{t}) + \frac{1}{2}\sigma_j \sqrt{t} \\
 d2_j &= d1_j - \sigma_j \sqrt{t} \\
 d &= 1 + \text{dividend yield} \\
 s &= \text{index value today} \\
 \sigma'_j &= \text{first derivative of implied volatility with respect to strike prices,} \\
 &\quad \text{numerically approximated by } \sigma'_j = (\sigma_{j+1} - \sigma_{j-1})/2\Delta
 \end{aligned}$$

Our approach nests the Black–Scholes economy with its lognormal risk-neutral distribution as a limiting case when the trade-off parameter (p) goes to zero. Empirically we set the trade-off parameter to 7×10^{-12} . This choice ensures that all risk-neutral probability distributions are nonnegative and thus arbitrage-free. Only a few options are priced outside their empirical bid/ask spreads, and those tend to be options with extreme moneyness.

2.2 Subjective probability distributions

In order to estimate the subjective distribution, we use the historical time series of the index. The length of the sample is an issue, and we select 4-year time frames. This particular choice and a number of parameter choices below can be criticized as somewhat arbitrary. However, we show that varying these initial choices does not change the basic findings.

We calculate 31-day, nonoverlapping returns from our 4-year sample and derive the kernel density with a Gaussian kernel and bandwidth:³

$$h = \frac{1.8\sigma}{\sqrt[5]{n}} \quad (8)$$

where

$$\begin{aligned}
 h &= \text{kernel bandwidth} \\
 \sigma &= \text{standard deviation of the sample returns} \\
 n &= \text{number of observations}
 \end{aligned}$$

The kernel density, which approximates the subjective probability distribution, is a continuous function and for our computations we have to discretize it onto the same spacing (S_j), which we employed for the risk-neutral distribution. We achieve this by numerically integrating the kernel density across the intervals around each state S_j .

We measure the actual distribution with a historical mean, which can be rather different from the mean of the future period associated with the risk-neutral probability distribution. We know that the historical mean is very

³ See Silverman (1986, p. 15). This choice of a bandwidth leads to slight oversmoothing but removes spurious multimodalities. These multimodalities are due to the small sample size of only 48 returns over the past 4 years. Solving an optimization problem, which searches for the smoothest distribution consistent with the actual returns, gives very similar results but is significantly slower.

difficult to estimate. Merton (1980) argues that one should explicitly model positive risk premia, and thus we de-mean the actual distribution and then reset the mean to the riskfree interest rate +8% per year for the risk premium. So far, we made a few modeling choices to obtain our risk aversion functions. However, we will show shortly that our empirical results are not sensitive to reasonable changes in any of these choices.

3. Empirical Results

We are now in the position to calculate the risk aversion functions across wealth. The empirically observed risk aversion functions are unfortunately not nearly as well behaved as standard assumptions made in economic theory suggest, namely, that risk aversion functions should be positive and monotonically downwards sloping. Figure 3 shows our empirical results, which are aggregated but still indicative of the respective periods covered. We also report the empirical standard deviations by showing the mean risk aversion levels across wealth ± 0.5 times their standard deviations.⁴

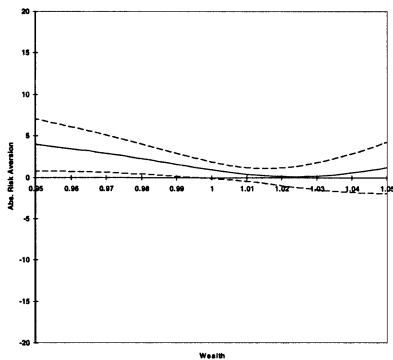
Pre-crash, the risk aversion function is largely consistent with standard assumptions made in economic theory: it is positive and decreasing across wealth except that it increases again for wealth levels greater than 1.03, which corresponds to an increase in wealth of 43% over the year. However, in the post-crash period two disconcerting features arise and become progressively worse as we move through time. First, the post-crash risk aversion functions are negative around the center. Second, the risk aversion functions rise for wealth levels greater than about 0.99. Before we analyze these features in detail, we want to address the potential concern that our particular modeling choices caused them.

3.1 Stability of the empirical risk aversion functions

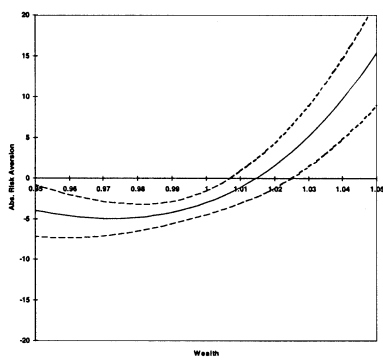
We find that the general shape of the functions in Figure 3 is remarkably stable with respect to different modeling choices. Using a different length for the historical sample of 2 or 10 years instead of 4 years does not change Figure 3 significantly. Also, some option traders are concerned that the hedging of out-of-the-money put options has become more difficult after the crash as they became aware of the possibility of large jumps. We therefore investigate the impact of option prices with low strike prices (moneyness between

⁴ We also investigate bootstrapped error bounds. They are very similar and only somewhat tighter than the empirical bounds and therefore we do not graph them separately. For those error bounds, we assume that the nonoverlapping historical returns are independent and identically distributed, and we then sample 100 times with replacement. Each new set of nonoverlapping returns can be viewed as another sample, and we recalculate the risk aversion functions. For each day, we calculated the mean risk aversion levels across wealth and the mean levels ± 0.5 times their standard deviations. For each of our periods, we finally calculate the means of the risk aversion functions and their bounds. See Press et al. (1992, p. 689) for a very clear introduction and Härdle (1990, p. 105) for further details on the bootstrap method.

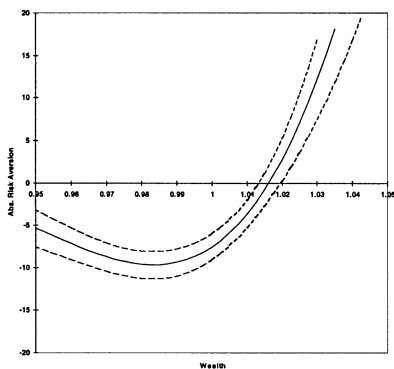
Panel A: 860402 – 871018



Panel B: 881019 - 910318



Panel C: 910319 – 930818



Panel D: 930819 - 951229

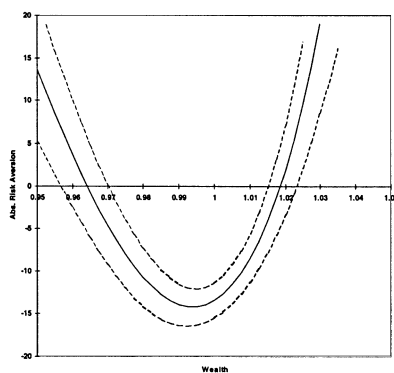


Figure 3

Risk aversion functions across wealth.

For 114 days from April 2, 1986, through December 29, 1995, we calculate the absolute risk aversion functions across wealth. A wealth level of one corresponds to the forward price of the index where the forward contract has the same maturity as the options used for obtaining the risk-neutral probability distribution. For each wealth level we calculate the mean risk aversion across the periods 860402–871018 (panel A), 881019–910318 (panel B), 910319–930818 (panel C), and 930819–951229 (panel D). In addition, we calculate their empirical standard deviations and graph the means ± 0.5 times the standard deviations.

0.84 and 0.89) on our empirical findings by omitting them from our calculations. Only the lower left tails of the risk-neutral probability distributions are somewhat affected and Figure 3 does not change significantly altogether.

Small changes in the risk premium in the interval where most financial economists would expect it to lie over the long run (5–10% per year) do not change Figure 3 other than through slight shifts (downward and upward,

respectively). However, very large risk premia cause all risk aversion functions to shift upward while the shape of the functions is preserved. A risk premium of more than 40% per year would finally cause all functions to be nonnegative throughout.

Also, we changed the bandwidth for the kernel density from 0.5 to 2 times its value in Equation (8). Small bandwidths lead to spurious multimodalities, which cause all risk aversion functions to fluctuate more. However, the basic features of Figure 3 are still preserved in that the precrash risk aversion function is mostly positive and the postcrash ones are mainly negative.

Next, we use the actual distributions straight with their means as calculated from the historical data. Figure 3 does not change significantly except that the precrash risk aversion function is now even higher and the postcrash ones are even lower.

Finally, we change the time-to-maturity to 15 and 64 days and, again, Figure 3 is unaffected. We cannot investigate longer horizons since the number of nonoverlapping returns in our historical sample dwindles dramatically, and we would like to observe at least 25 return observations (corresponding to at most 64-day returns in a 4-year sample) for the kernel density estimation.

3.2 Discussion of the literature

Aït-Sahalia and Lo (2000) and Rosenberg and Engle (1997) also have studied empirical risk aversion functions. Less closely related, Derman et al. (1997) ask if option prices are “fair” given the actual distribution. Most of the above studies find positive and rather “well-behaved” risk aversion functions in the postcrash period. There are a number of possible reasons why their results might differ from ours.

Aït-Sahalia and Lo (2000) present a positive risk aversion function in their Figure 4, but the risk aversion function associated with the marginal utility in their Figure 3 is partially negative. A number of possible reasons could explain their result. First, the sign of the risk aversion function changes according to the bandwidth they use for their kernel estimation; and the sign is evidently rather sensitive to their particular methodology. Second, the study provides only one snap-shot in time (based on option prices during the year 1993) where the distributions and the risk aversion function are estimated and not, as we do, a time series of the risk aversion functions over 10 years. We know that we can observe “well-behaved” risk aversion functions on individual days postcrash and that the pattern of partially negative risk aversion functions only emerges as we check many dates over a longer time series. Third, Aït-Sahalia and Lo (2000) use overlapping returns to estimate the actual distribution. We recalculated their actual distributions with and without overlapping returns, and the two distributions look rather different. Since they only show one risk aversion function, the shape of that function can be affected by the choice of overlapping versus nonoverlapping returns.

Rosenberg and Engle (1997) only present results on marginal utilities, but the associated risk aversion functions are all positive. The reason appears to be the particular way in which the risk-neutral distributions are estimated. Their “sigma-shape” polynomial, which is related to implied volatilities, seems to be a rather rigid and inflexible function. As a result, the empirical option prices may not be very well matched, as the implied risk-neutral distribution is surprisingly close to a lognormal distribution. In particular, the risk-neutral distribution is only slightly left-skewed and less peaked than the lognormal distribution. One interpretation is that the positive risk aversion functions could be a result of their estimation technique.

Derman et al. (1997) implicitly assess the risk aversion function even if they do not estimate it formally. They observe that on one particular date the actual distribution of overlapping returns is left-skewed and thus similar to the risk-neutral distribution. They interpret this observation as being consistent with a “well-behaved” risk aversion function. Two possible reasons could explain their observation, which we have discussed already in relation to Aït-Sahalia and Lo (2000). First, their study provides only one snapshot in time and not a time series. Second, Derman et al. (1997) use overlapping returns, too. We recalculated their actual distributions with and without overlapping returns and, again, the two distributions look rather different. Their observation is thus sensitive to the way the actual distribution is estimated.

3.3 Discussion of the empirical risk aversion functions

In our discussion of Figure 3 we want to point out that the results for the precrash period are rather encouraging. The implied risk aversion function is positive throughout; and both absolute and relative risk aversion functions are decreasing with increasing wealth until wealth levels of about 1.03, as one might deem reasonable behavior for the economy.

The problems occur postcrash where risk aversion functions are now negative throughout wealth levels from about 0.96 through 1.01. Moreover, risk aversion is increasing for wealth levels greater than about 0.99, that is, for ending up wealthier at the end of 31 days than at the beginning of those 31 days.

Inspection of Equation (4) yields insight into this behavior. The risk aversion function is the vertical distance of two downward sloping functions (Q'/Q and P'/P). The respective intersections with the wealth axis occur at the modes (that is, the maxima) of the probability distributions. In the precrash period, depicted in Figure 1, with about lognormal distributions for the subjective and risk-neutral probabilities, the mean, median, and mode of each distribution tend to lie rather close together compared to the risk premium. This changes in the postcrash period, depicted in Figure 2, where the risk-neutral distribution suddenly exhibits a mode to the right of the mean. The mode of the risk-neutral distribution often appears even to the right of the mean (and mode) of the subjective distribution, causing negative risk

aversion functions. The steep drop in probability to the right of the mode of the risk-neutral distribution causes the P'/P function to fall rapidly after intersecting the wealth axis. This leads to increasing absolute risk aversion for wealth levels greater than about 0.99.

One way to interpret Figure 3 would simply be to believe what it suggests, namely that the risk aversion functions changed dramatically around the crash. If all the assumptions for the representative investor model hold true, then concavity in the utility functions implies all positive risk aversion functions. In turn, since we observe negative risk aversion functions, we would now have to believe that the utility functions are convex. This would lead to corner solutions where the representative investor with a convex utility function would avoid returns in all states of the world where the utility function is convex and would seek returns in the adjacent states where the utility function is concave. This cannot be the case since the representative investor has to hold the market. If we are not willing to accept the implications of Figure 3, then we might suspect that one or more of the assumptions are violated. Here we are faced with the two usual concerns that the S&P500 index might not be a good proxy for the market as a whole (Roll's critique) and that the assumptions needed for the existence of a representative investor, namely a complete market and a state-independent utility function, might be violated. Both these questions arise in numerous studies and have been investigated in the literature. However, advancing our understanding of these questions takes us beyond the scope of the present article and we will not spend time on addressing them here.

More closely related to the present study are concerns that we might omit volatility clustering when we compute the actual distribution. Or the subjective distribution might not be well approximated by the actual distribution. Specifically, investors could make mistakes in deriving their subjective distributions from the actual distribution, which could then lead to options being mispriced. We will now look into each of these explanations in turn.

So far we estimated the actual distribution nonparametrically as a kernel density. Alternatively, in order to capture volatility clustering, we also estimated the index return process as a GARCH(1,1) model, and we simulated 10,000 31-day returns (22 trading days). The resulting simulated distributions are very close to lognormal distributions, and the implied risk aversion functions exhibit the same patterns as in Figure 3, only more pronounced.

A further explanation for the empirical findings might be that the subjective distribution is not well approximated by the actual distribution due to the so-called peso problem. In that case, the investor could observe a historical sample of returns, without a crash, which looks about lognormal. However, in determining the subjective distribution the investor might well incorporate the possibility of a crash. The investor could then arrive at a subjective distribution, which is much more similar in shape to the risk-neutral distribution than the actual distribution would suggest. However, the 4-year period after

the crash of 1987 allows us to rule out a peso problem: during those 4 years, all historical samples include the crash. Therefore the subjective distributions reflect a crash of about 20% with a chance of one in every 4 years. Nonetheless, the risk aversion function in panel B of Figure 3 during the 4 years right after the crash (October 1988–March 1991) is U-shaped and negative around the center and thus incompatible with standard assumptions made in economic theory. In addition, we also used historical samples spanning 2 and 10 instead of 4 years. These samples weigh the crash differently in the subjective distributions, but the risk aversion functions are still U-shaped and negative around the center during the 2 and 10 years after the crash. We conclude that a peso problem cannot explain the empirical risk aversion functions.

However, the discussion of the peso problem leaves one more unpalatable possibility open: investors might be overly concerned with the possibility of further crashes and therefore could have updated their subjective beliefs about the distribution of market returns too much. The resulting subjective distribution would in this case be left-skewed which, given any well-behaved risk aversion function [Rubinstein (1994)], results in a similarly left-skewed risk-neutral distribution, just as we observe empirically. The options based on this distribution would then exhibit too steep of a volatility smile and out-of-the-money puts would be priced too high. If we now try to extract the empirical risk aversion function from the option prices (which are affected by the mistakes of investors) and the realized returns (which are unaffected by the mistakes) then it is to little surprise that the risk aversion functions exhibit perverse shapes. In order to investigate this possibility further, we note that we should be able to make money from this mispricing, and we employ the following methodology.

3.4 Trading strategies

During the 7 years starting in October 1988 and ending in December 1995, we select 86 trading dates 1 month apart. The sample is extended from an earlier version of this article. The additional data for 1995 provides for a true out-of-sample test, which fully supports all our findings. Each trading period, we start out with \$100 to invest, and we can invest into the riskfree asset or into the S&P500 or we can select one of two option strategies. These involve lending \$100 and simultaneously selling puts with moneyness of either 0.95 or 1.00. The number of puts sold is chosen such that the betas of the options strategies are about 1. We can now directly compare the option strategies against holding the market which otherwise would be difficult as excess returns depend on the riskiness (beta) of the strategies. Also, we ensure that the margin requirements for index options of the Chicago Board Options Exchange are always maintained.

All positions are closed out 1 month later, and we earn the riskfree return on our lending. For each of the 86 one-month periods we bootstrap 1000 sample returns from the historical monthly index returns throughout the whole

7-year sample. Our sample does not include the crash, but we suspect that the options are priced to reflect that possibility. We therefore compute five different runs where we introduce artificial crashes with a probability of one crash every 4, 8, 16, and 32 years and one run without any crashes. With a crash every 4 years, we lose about 5% risk premium per year due to crashes. Therefore we did not model more frequent crashes, since the security market line would then slope downward. A crash is created by replacing one monthly return during the specified period with that return times 0.8. We finally regress the returns for each strategy against the market returns. We show in Figure 4 the expected, annualized excess returns (alphas) across betas.

We find excess returns for both option trading strategies as long as the probability of a crash does not exceed one in 8 years, and even a 20% crash every 8 years seems to be a rather pessimistic outlook. For the strategy where we sell puts with moneyness of 1.00, we would even come out ahead with crashes as frequent as one every 4 years. Above, we noticed that the perverse shape of the risk aversion functions is not affected much by the deep-out-of-the-money puts. Here we find the result supported, in that the

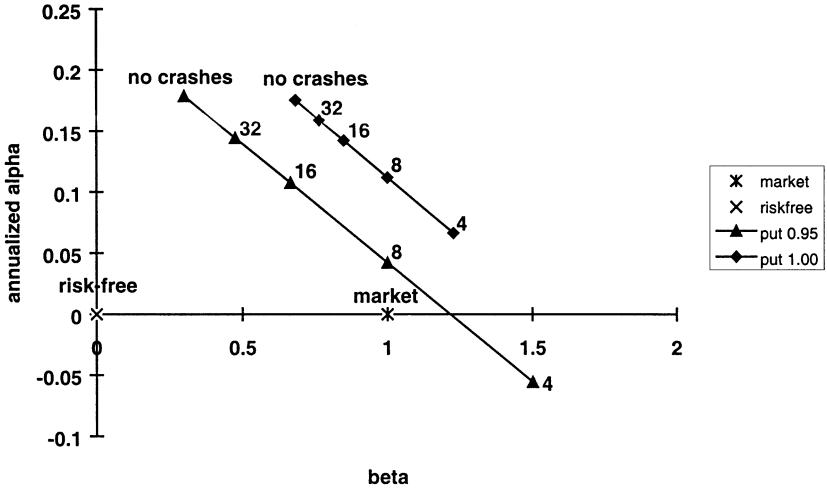


Figure 4
Risk-adjusted expected excess returns of selling puts.

We depict the strategy of investing \$100 in the riskfree asset and in the S&P500. Returns are calculated for 86 31-day periods starting in October 1988 until December 1995. The market return for each 31-day period is bootstrapped from the complete 7-year sample. In four runs, artificial crashes of 20% during 1 month are added every 4, 8, 16, and 32 years on average and in a fifth run, no crashes are added. We report the average of the 86 expected excess returns (alphas) across betas. The expected excess returns for the riskless asset and the market are zero with betas of 0 and 1 respectively. We also report two strategies of lending \$100 and selling puts with moneyness levels of 0.95 and 1.00. The number of puts sold is chosen such that the strategies have a beta of about 1 if there is a crash every 8 years.

most profitable strategy involves selling at-the-money puts. Also, if the same analysis is conducted for longer-term options (say 180 days to expiration), then the option trading strategies are generally more profitable. The reason for this finding is that, over the next half year, it is more likely for the market to recover from a crash than over the next month only.

A number of objections can be made. Namely, we ignore transaction costs, employ a bootstrap method for the index returns, and use a mean variance framework to assess the option trading strategies. We will now address each objection in turn.

We recalculated Figure 4 with transaction costs, and the expected excess returns per year are only 4% (2%) lower for the trading strategy involving puts of moneyness 0.95 (1.00). With transaction costs, all positions are entered at the current bid prices observed on the trading date and charged a 2% commission. The commission is mainly relevant for the initial sale since the puts are often not exercised, and we therefore normally do not pay commission at expiry.

Another objection might be that the bootstrap approach changes the original sequence of returns. We therefore use the above analysis, including the artificial crashes, in combination with the observed returns in their historical sequence for each trading period. Figure 4 changes only insignificantly.

Finally, in the above analysis we do not properly adjust for the risk of higher-order moments beyond mean and variance such as the negative skewness which is created by selling puts. Recent work by Leland (1999), based on Rubinstein (1976), allows us to remedy this problem. If the market exhibits lognormal returns (which approximately holds true for the actual returns), then the usual beta measure should be replaced with an adjusted beta measure. Next, we can replace the usual alpha measure with an adjusted alpha to measure the adjusted expected excess return. However, Figure 4 does not change much when we use the adjusted measure.

The above strategies of selling puts can lose all the money invested if there are a sufficient number of crashes and the margin is exhausted. But there are two interesting hedging strategies available. For one, we can limit the downside risk by setting up a vertical bull spread: in addition to every one put sold in our above positions, we buy one out-of-the-money put with moneyness of 0.90. We thereby cap our maximum losses at 5% (10%) of the index value for each put with moneyness of 0.95 (1.00) sold. Given the number of puts typically sold, this amounts to a maximum loss of about \$25 on our initial investment of \$100. We show the results of the hedged trading strategies in Figure 5. The excess returns turn out to be similar to the excess returns of the unhedged strategies, but the hedged strategies are significantly less sensitive to crashes.

Finally, Toft and Prucyk (1997) report that individual stock option smiles are significantly flatter than the index option smiles in the postcrash period. Another potential hedging strategy is therefore to sell puts on the index and

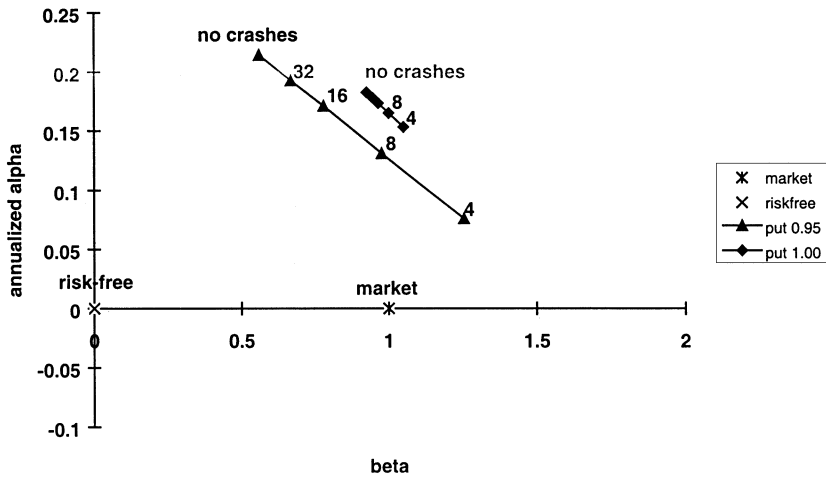


Figure 5
Risk-adjusted expected excess returns of hedged strategies.

We depict the strategy of investing \$100 in the riskfree asset and in the S&P500. Returns are calculated for 86 31-day periods starting in October 1988 until December 1995. The market return for each 31-day period is bootstrapped from the complete 7-year sample. In four runs, artificial crashes of 20% during one month are added every 4, 8, 16, and 32 years on average and in a fifth run, no crashes are added. We report the average of the 86 expected excess returns (alphas) across betas. The expected excess returns for the riskless asset and the market are zero with betas of 0 and 1, respectively. We also report two strategies of lending \$100, buying a put with moneyness of 0.90, and selling puts with moneyness levels of 0.95 and 1.00. The number of puts bought and sold is chosen such that the strategies have a beta of about 1 if there is a crash every 8 years.

to buy puts on (only a few) individual stocks. The hedge will perform better the more we need it since in the case of a crash, correlations across different stocks (and the index) tend to go up [Kelly (1994)]. Further research will investigate the economic causes of negative skewness as the smile for index options is very steep compared to the smiles for individual stock options.

4. Conclusion

This article models a representative investor economy in which we can empirically determine implied risk aversion functions through simple expressions of risk aversion across wealth in terms of the subjective and risk-neutral probability distributions only. The resulting risk aversion functions are largely consistent with standard assumptions made in economic theory during the precrash period, but we find partially negative and partially increasing risk aversion functions during the postcrash period. These findings are remarkably stable with respect to varying input parameters. Changes in the empirical methodology do not seem to affect the results.

Thus we are left with either believing that risk aversion functions did indeed change dramatically around the crash or questioning the assumptions

of the model in order to explain their radical shapes or concluding that the market misprices some options. The most likely explanation is mispricing of options in the market, since negative risk aversion functions are irreconcilable with a representative investor and since we can defend most other assumptions of the model rather well. A simulated trading strategy exploiting such mispricing yields risk-adjusted expected excess returns during the postcrash period. These excess returns persist even when we account for transaction costs and hedge the downside risk.

References

- Ait-Sahalia, Y., and A. W. Lo, 1998, "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," *Journal of Finance*, 53, 499–547.
- Ait-Sahalia, Y., and A. W. Lo, 2000, "Nonparametric Risk Management and Implied Risk Aversion," *Journal of Econometrics*, 94, 9–51.
- Banz, R., and M. Miller, 1978, "Prices for State-Contingent Claims: Some Estimates and Applications," *Journal of Business*, 51, 653–672.
- Breedon, D., and R. Litzenberger, 1978, "Prices of State-Contingent Claims Implicit in Options Prices," *Journal of Business*, 51, 621–651.
- Brown, G., and K. B. Toft, 1999, "Constructing Binomial Trees from Multiple Implied Probability Distributions," *Journal of Derivatives*, 7, 83–100.
- Constantinides, G. M., 1982, "Intertemporal Asset Pricing with Heterogeneous Consumers and without Demand Aggregation," *Journal of Business*, 55, 253–267.
- Derman, E., M. Kamal, I. Kani, and J. Zou, 1997, "Is the Volatility Skew Fair?," Quantitative Strategies Research Notes, Goldman Sachs, New York.
- Hardle, W., 1990, *Applied Nonparametric Regression*, Cambridge University Press, Cambridge, MA.
- Jackwerth, J. C., and M. Rubinstein, 1996, "Recovering Probability Distributions from Option Prices," *Journal of Finance*, 51, 1611–1631.
- Kelly, M. A., 1994, "The U.S. Equity Market Has Become More Diversified," J P Morgan Equity Derivative Strategies, New York.
- Leland, H., 1980, "Who Should Buy Portfolio Insurance?," *Journal of Finance*, 35, 581–594.
- Leland, H., 1999, "Beyond Mean-Variance: Risk and Performance Measurement in a Nonsymmetrical World," *Financial Analyst Journal*, 55, 27–36.
- Melick, W. R., and C. P. Thomas, 1997, "Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil During the Gulf Crisis," *Journal of Financial and Quantitative Analysis*, 32, 91–115.
- Merton, R. C., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1992, *Numerical Recipes in C* (2nd ed.), Cambridge University Press, Cambridge, MA.
- Rosenberg, J. V., and R. F. Engle, 1997, "Option Hedging Using Empirical Pricing Kernels," working paper, New York University.
- Ross, S., 1976, "Options and Efficiency," *Quarterly Journal of Economics*, 90, 75–89.
- Rubinstein, M., 1976, "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell Journal of Economics and Management Science*, 7, 407–425.

Rubinstein, M., 1994, "Implied Binomial Trees," *Journal of Finance*, 49, 771–818.

Shimko, D., 1993, "Bounds on Probability," *RISK*, 6, 33–37.

Silverman, B. W., 1986, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.

Toft, K. B., and B. Prucyk, 1997, "Options on Leveraged Equity: Theory and Empirical Tests," *Journal of Finance*, 52, 1151–1180.