Inferring Risk-Averse Probability Distributions
From Option Prices using Implied Binomial Trees

Tom Arnold
The Robins School of Business
Department of Finance
1 Gateway Road
University of Richmond
Richmond, VA 23173
Tel: 804-287-6399
Fax: 804-289-8878
tarnold@richmond.edu

Timothy Falcon Crack
Dept. Finance and Quant. Analysis,
Otago University,
PO Box 56, Dunedin, New Zealand.
Tel: +64(0)3-479-8310
Fax: +64(0)3-479-8193
tcrack@business.otago.ac.nz

Adam Schwartz
Associate Professor of Business Administration
Washington and Lee University
Williams School of Commerce, Economics, and Politics
Lexington, VA 24450
Tel: 540-458-8254
Fax: 540-458-8639
schwartza@wlu.edu

This version: July 02, 2008

Keywords: Binomial Option Pricing, Implied Binomial Trees, VAR, Value at Risk, Hedge Funds, Physical Probabilities, Risk-Neutral Probabilities, Risk-Averse Probabilities, Subjective Probabilities, Calibration.

*We thank David Alexander, Ravi Bansal, Alex Butler, Scott Chaput, Robin Grieves, Robert Hauswald, Jimmy Hilliard, Stewart Mayhew, Susan Monaco, Sanjay Nawalkha, Mark Rubinstein, Louis Scott, Richard Shockley, seminar participants at the University of Georgia and University of Otago, and an anonymous referee. This paper subsumes an earlier working paper by Arnold and Crack, “Option Pricing in the Real World: A Generalized Binomial Model with Applications to Real Options” http://ssrn.com/abstract=240554. The earlier paper also contains alternative proofs of the RA-BT model. Any errors are ours.
Abstract

We generalize the Rubinstein (1994) risk-neutral implied binomial tree (R-IBT) model by introducing a risk premium. Our new risk-averse implied binomial tree model (RA-IBT) has both probabilistic and pricing applications. We use the RA-IBT model to estimate the pricing kernel (i.e., marginal rate of substitution) and implied relative risk aversion for a representative agent; we are the first to use wholly implied methods for this. We also use the RA-IBT to explore the differences between risk-neutral and risk-averse moments of returns. We discuss practical applications of the model to Value at Risk and stochastic volatility option pricing models.
We generalize the Rubinstein (1994) risk-neutral implied binomial tree model (R-IBT) to a physical-world risk-averse implied binomial tree model (RA-IBT). The R-IBT and RA-IBT trees are bound together via a no-arbitrage-driven relationship requiring a risk premium (or a risk-adjusted discount rate) on the underlying asset at any node. The RA-IBT provides a powerful numerical platform for many empirical financial option and real option applications. These include probabilistic inference, pricing, and utility theory applications.

For ease of exposition, we have estimated a constant risk premium RA-IBT here using S&P500 index options. In our implied tree this is consistent with assuming a representative agent with a power-like utility function where the CRRA parameter is allowed to vary across states and through time.\(^2\) With these assumptions we also estimate the pricing kernel (marginal rate of substitution) and implied relative risk aversion (RRA) of our agent and compare and contrast our results with other authors’ findings. We also demonstrate that using our RA-IBT to infer risk-averse probabilities of a decline in the S&P500 relative to the current spot can give different results to the risk-neutral probabilities of a decline relative to the spot inferred from a standard R-IBT. There are direct implications for Value at Risk estimates in leveraged portfolios (e.g., hedge funds). We also use the RA-IBT to examine differences in volatility, skewness, and kurtosis between the risk-neutral and risk-averse implied return distributions. We discuss practical applications of the RA-IBT model to stochastic volatility option pricing models.

Section 1 motivates the paper and reviews the literature. In Section 2, a simple (non-implied) generalized risk-averse binomial tree (RA-BT) is introduced. In Section 3, the implied version of the generalized tree is derived (i.e., the RA-IBT). Section 4 contains empirical

\(^2\) In a technical appendix (available upon request) we show how to relax the constant risk premium assumption and generate a risk premium at any node using an assumed utility function of a representative agent. Examples are given for power utility and negative exponential utility.
estimation and analysis of the RA-IBT model using S&P500 options. Section 5 concludes the paper. Most derivations/proofs are relegated to the appendices.

1. Motivation and Literature Review

In this section we first motivate and introduce our risk-averse models, and then we review the literature and compare and contrast our risk-averse implied binomial tree (RA-IBT) with other research.

1.1 Motivation/Introduction of Risk-Averse Trees

The traditional binomial tree model of Cox, Ross, and Rubinstein (CRR, 1979) is very powerful, but it is constrained in many respects. For example, as step sizes tend to zero, the distributions of stock prices and of continuously-compounded returns in a CRR tree are constrained to be lognormal and normal, respectively. CRR also assume that both the up-jump probability at each step and the local volatility structure are fixed throughout the tree. In addition, all risk-averse probabilities have fallen out of the CRR model and have been replaced by risk-neutral probabilities. The CRR model cannot, therefore, reproduce some well known empirical results (e.g., fat tails, skewness, volatility smiles, etc), and any probabilistic inferences from a CRR tree must be risk-neutral inferences, not risk-averse probability inferences from the physical economy.

The Rubinstein (1994) implied binomial tree (R-IBT) generalizes the CRR model to fit the prices of a series of traded options of the same maturity. The R-IBT allows significant deviations from lognormality of prices (and from normality of continuously-compounded

---

3 There are of course other restrictions in the CRR tree: at any node in the tree there are only two possible values for the one step ahead stock price; many steps are required to get accurate pricing; and the traditional CRR tree does not price path-dependent options. Fixes or extensions exist for all of these, but they are outside of our area of interest for this paper.
returns), allows the up-jump probability to vary throughout the tree, and allows the local 
volatility structure to vary throughout the tree. The R-IBT model is still constrained, however, in 
that the probability structure is risk neutral, and therefore, like the CRR model, it does not allow 
probabilistic inferences about the physical economy.

In this paper, we begin with a generalization of the risk-neutral CRR model to a risk-
averse binomial tree (RA-BT). We then present a generalization of the risk-averse RA-BT to a 
risk-averse implied binomial tree (RA-IBT). Like the R-IBT model, the RA-IBT model 
captures volatility smiles and varying local volatility. It should be noted, however, that both the 
RA-BT and the RA-IBT have one extra input compared to the CRR and R-IBT models, 
respectively: the risk-averse trees need to be supplied with a risk premium (or a risk-adjusted 
discount rate) at every node to make them estimable. This risk premium feeds into a no-arbitrage 
relationship that drives a transformation from the risk-neutral trees to the risk-averse trees.

For empirical ease, a constant risk premium is imposed in the RA-IBT estimations in this 
paper. Whether the risk premium is imposed directly or derived from utility assumptions for a 
representative agent, to each such risk premium function over the nodes of an RA-IBT there 
corresponds a different implied risk-averse probability distribution for the future prices of the 
underlying asset, which in turn implies a unique fully specified stochastic process for the 
underlying asset prices. Sensitivity analysis is thus essential for any inferences.

The CRR, RA-BT, R-IBT and RA-IBT models can each be used to both value and hedge 
both European-style and American-style options. Neither the R-IBT model nor the RA-IBT 
model (under the conditions developed in this paper) can be calibrated using American-style

---

4 Alternatively, the RA-IBT may be viewed as a generalization to risk-averse probabilities of the risk-neutral R-IBT, 
which in turn is a generalization to implied trees of the non-implied CRR model.
options.5

1.2 Literature Review

Rubinstein (1994, Footnote 25, p793) describes implied trees developed by Hayne Leland (1980) that use “subjective probabilities” (i.e., an individual investor’s non-risk-neutral probabilities). Leland’s trees are used to assess what sort of investors would buy exotic options or buy portfolio insurance. Like us, Leland attempts to infer subjective probabilities. Unlike us, his trees are for individuals whereas our focus is the aggregate market, and his trees use constant asset level move size at each step whereas ours are implied trees with variable move sizes.

Stutzer (1996) also infers “subjective” (i.e., risk-averse) probability densities from options data. Stutzer differs from us, however, in that he uses diffusions rather than binomial trees, he requires historical data that are not needed here, and he uses the risk-averse density to estimate the risk-neutral density for risk-neutral pricing (the focus of his paper), whereas our focus is the risk-averse density itself.

Jackwerth and Rubinstein (1996) infer probabilities from option prices using binomial trees. However, Jackwerth and Rubinstein differ from us because they use risk-neutral probabilities, whereas, we use risk-averse probabilities. We do, however, use their “smooth” objective function in estimating our R-IBT.

Jackwerth (1999) is an excellent systematic classification and review of implied risk-neutral distributions and of implied binomial trees. It contains little mention of risk-averse implied distributions, however, except for discussion of the subsequently published Jackwerth (2000). In that paper Jackwerth recovers risk-neutral densities from S&P500 options, and uses

---

5 We can, however, generalize the implied risk-averse model so that it can be calibrated using American-style options. We can do this by applying our transformation directly to a Jackwerth (1997) generalized R-IBT. We omit this here for ease of exposition.
historical realized index returns to approximate subjective (i.e., risk averse) densities. Jackwerth is thus able to infer aggregate absolute risk-aversion functions for different states. When comparing pre- and post-Crash of 1987 data, changes in the implied risk-neutral densities that are not accompanied by changes in the risk-averse densities imply changes in absolute risk-aversion functions that appear inconsistent with any sensible economic theory (e.g., Jackwerth finds significantly negative absolute risk aversion). Jackwerth suggests that overpriced put options may explain the inconsistency and he is able to construct profitable trading strategies that appear to exploit this potential mis-pricing. Jackwerth acknowledges that another explanation for inconsistent risk-aversion functions is that his risk-averse distributions built using historical observations may differ from true ex-ante risk-averse distributions (Jackwerth [1999, pp445–446]); he rejects this as an explanation because his trading strategies appear profitable (supporting the hypothesis that his inconsistent risk-aversion functions are driven by exploitable mis-priced options, not by poor estimates of subjective probability). We show, however, that we get wholly positive risk aversion functions for the same time period as Jackwerth and using the same underlying index, but using implied techniques for both the risk-averse and risk-neutral densities (see our Section 4.4 for details).

Jackwerth (2004) is an expanded and updated version of Jackwerth (1999). Jackwerth (2004) also has an expanded section on implied risk aversion and useful summary tables of categorized literature. Jackwerth (2004) labels his earlier findings (i.e., those in Jackwerth [2000]) as a “pricing kernel puzzle.” The puzzle is that although the implied marginal utility of wealth function should be monotonically decreasing in wealth, Jackwerth’s empirical estimates of it after the Crash of 1987 are locally increasing in wealth for wealth levels near the initial wealth level. This suggests that in these ranges of wealth the representative agent is risk seeking
not risk averse. Ait-Sahalia and Lo (2000) find a similar locally-humped plot of the scaled marginal rate of substitution (their Figure 3, p.36) using S&P500 futures prices. Our paper is similar to Jackwerth (2000, 2004) and Ait-Sahalia and Lo (2000) in that we recover both risk-neutral and risk-averse distributions and compare them, but we differ in that we infer our ex-ante risk-averse distributions from option prices using the RA-IBT rather than from a backward-looking historical time series of the underlying as these authors do.

Another strand of the literature includes Bliss and Panigirtzoglou (2004) and Alonso et al. (2006) and several other papers they cite. Like Ziegler (2003), these authors exploit the Breeden and Litzenberger (1978) result, rather than implied binomial trees, to derive risk-neutral densities. They then calibrate the parameters of a chosen utility function that is used to risk-adjust the risk-neutral density. The objective of the calibration is that the risk-averse density should best explain subsequently realized returns (see further discussion in Section 4.5). This allows them to discuss implied risk aversion. Their work is closely related to our own, except that rather than use implied binomial trees, they use numerical smoothing techniques to account for volatility smiles over a range of option strikes and they use the Breeden and Litzenberger (1978) result. Their are two advantages to our approach over theses approaches. The first advantage is that our implied tree is guaranteed to be arbitrage free (assuming there are no arbitrage opportunities amongst the quoted option prices), whereas the Bliss and Panigirtzoglou (2004) numerical smoothing techniques are not guaranteed to be arbitrage free. The second advantage of our approach is that we use a simple numerical estimation without splines or smoothing (we estimate an R-IBT and then apply a simple direct transformation to it).

Blackburn (2006) focuses on the time series properties of risk aversion and whether the representative agent’s utility is time separable or not. Like Bliss and Panigirtzoglou (2004) and
Alonso et al. (2006) he exploits the Breeden and Litzenberger (1978) result and uses splines to derive the risk-neutral density. Unlike these authors, Blackburn argues that a calibration that maximizes the forecast ability of the risk-averse density is inappropriate (because it looks ahead in a way not possible when the agent is making decisions). Instead, Blackburn obtains a risk aversion estimate using five years of historical data. Blackburn (2006) thus differs from us in that he does not use trees at all, and he uses historical data to estimate the risk aversion parameter.

Finally, a recent strand of literature that explores other techniques for resolving pricing kernel anomalies includes, for example, Garcia et al. (2005) and Benzoni et al. (2007). Garcia et al. (2005) introduce a latent state variable representing states of the economy. Covariance risk with the state variable means that options prices depend upon preferences. Their generalized option pricing models capture some known option pricing biases. Benzoni et al. (2007) introduce a persistent stochastic growth variable with a jump and use it to drive aggregate dividend and consumption processes. Their general equilibrium framework captures some of the stylized facts of equity and equity options markets.

2. A Risk-Averse Binomial Tree (RA-BT) Model

As mentioned previously, the RA-BT model is a risk-averse generalization of the CRR model and it is not an implied tree. We now derive the RA-BT model so that the implied version (i.e., the RA-IBT) can be developed in Section 3. We relegate some details to the Appendices.

We begin by noting that generation of an RA-BT or RA-IBT tree relies upon three interrelated technical steps: first, we derive the functional form of a no-arbitrage-driven transformation between the risk-neutral and risk-averse trees; second, we need to generate a risk premium or risk-adjusted discount rate at every node of the RA-IBT to feed into the transformation in the first step; third, we need to combine the first two steps and propagate risk-
averse probabilities through the risk-averse tree.

To work our way toward establishing the first two steps in the case of the RA-BT, recall that a continuous-time option pricing model using the risk-adjusted probability measure requires a stochastic path-dependent risk-adjusted discount rate; no single risk-adjusted discount rate can capture the changes in the option’s risk associated with the moneyness of the option. Black and Scholes recognize this with their “instantaneous CAPM” approach to deriving the Black-Scholes PDE (Black and Scholes [1973, pp. 645–646], Ingersoll [1989, pp. 323–324]). However, the (Black-Scholes) model that emerges is difficult to interpret with respect to the physical world because the risk-averse probability parameters fall out of the calculation.

The risk-averse RA-BT model is similar to a discretized version of the original Black-Scholes instantaneous CAPM derivation that allows for changing risk-adjusted discount rates. The numerical discretization allows us to infer physical-world parameters from the tree—an inference not explicitly available in the *closed form* continuous-time (i.e., Black-Scholes World) limit of the RA-BT pricing model.7

To generate the RA-BT model, begin with the assumptions of the CRR model as follows. Consider an asset with spot price $S_t$ at time $t$. From time $t$ to time $t + \Delta t$, the asset price either moves up by a multiplicative growth factor $u = e^{\sigma \sqrt{\Delta t}}$, or moves down by a multiplicative growth factor $d = e^{-\sigma \sqrt{\Delta t}}$. Assume a constant continuously-compounded riskless interest rate $r$, so the riskless growth factor is $R = e^{r \Delta t}$ over the time step. Let $V_t$ be the time-$t$ price of a European-style derivative that at time $t + \Delta t$ has the value $V_u$ in the up state and $V_d$ in the down state. Let

---

6 Cox and Rubinstein (1985, p324) discuss a related problem with a discount rate that is correct on average.
7 Baz and Strong (1997) do look at the closed form limits of some elements of the Black-Scholes model transformed to risk-averse probabilities, but they are very limited in what they can say.
\( q = (R - d)/(u - d) \) denote the fixed CRR risk-neutral probability of an up move. Then, to avoid arbitrage, the CRR model says that Equation (1) holds for the value of the European-style derivative over the time step \( \Delta t \).

\[
V_t = \frac{1}{R} E_{RN} \left( V_{t+\Delta t} \right) \\
= \frac{1}{R} \left[ q V_u + (1 - q) V_d \right] \\
= \frac{1}{R} \left[ \left( \frac{R - d}{u - d} \right) V_u + \left( \frac{u - R}{u - d} \right) V_d \right],
\]

where \( E_{RN}(\cdot) \) is the risk-neutral probability expectations operator.\(^8\)

We show in the proof in Appendix A that if \( K = e^{\Delta t} \) is a risk-adjusted compounding factor over time step \( \Delta t \), then \( p = (K - d)/(u - d) \) is the risk-averse probability of the up state and Equation (2) holds for the value of the European-style derivative over the time step \( \Delta t \).

\[
V_t = \frac{1}{R} \left[ E(V_{t+\Delta t}) - \left( \frac{V_u - V_d}{u - d} \right) (K - R) \right] \\
= \frac{1}{R} \left[ p V_u + (1 - p) V_d - \left( \frac{V_u - V_d}{u - d} \right) (K - R) \right],
\]

where \( E(\cdot) \) is the risk-averse probability expectations operator. That is, given an estimated CRR tree, and given a risk-adjusted discount rate \( k \) (or a risk premium \( k-r \)), we can deduce the risk-averse probability \( p \) for any up move in the CRR tree. The derivative pricing is identical between Equations (1) and (2), but the probabilities are risk-neutral in (1) and risk-averse in (2).

Equation (2) is, essentially, a certainty equivalent formula.\(^9\) It is derived from the no-

---

\(^8\) Note that much of CRR (1979) works with the physical tree rather than the risk-neutral tree. For example, their moment-matching argument used to deduce \( u \) and \( d \) is executed using the physical tree. Note also that CRR use “\( p \)” and “\( q \)” to denote what we label “\( q \)” and “\( p \),” respectively.

\(^9\) Similar/related formulae appear, for example, in Black and Scholes (1973, p645), Bogue and Roll (1974, p605), Constantinides (1978, p603), and Cox and Rubinstein (1985, Sections 5-5 to 5-7).
arbitrage-driven relationship in Equation (1) and it forms the basis for our first technical step: it provides a means of deducing risk-averse probabilities to overlay on the price structure of an underlying risk-neutral tree. Our RA-BT tree is, thus, a CRR tree transformed by replacing risk-neutral probabilities with risk-averse probabilities (see Appendix A for more details).

The second step is to generate a risk premium or risk-adjusted discount rate to feed into each node of the no-arbitrage transformation between risk-neutral and risk-averse trees. The risk premium can be derived from an asset pricing model such as the CAPM. In the case of a macro-asset (e.g., the S&P500 index portfolio), a representative agent argument frees us of the CAPM assumptions, and the risk-adjusted discount rates that feed into each node of this transformation can be derived using general assumed utility functions for the representative agent (power, negative exponential, etc).\(^\text{10}\)

Note that for any given node, each admissible \(k\) produces the same option valuation at that node.\(^\text{11}\) That is, the risk-adjusted discount rate \(k\) determines the risk-averse probability \(p\) and, by construction, \(k\) and \(p\) offset each other within the pricing equation to leave the option value unchanged. In fact, one must be cautious in interpreting \(p\) as the risk-averse probability of an up move at a node, because \textit{any} admissible \(k\) produces the same option valuation at that node (see Appendix A for more details). This leads us to a clause: The fidelity with which \(p\) reproduces the true risk-averse probability of an up move at a given node of a one-step RA-BT tree depends upon the accuracy in the choice of \(k\) as the true risk-adjusted discount rate for the underlying security.

\(^{10}\)Bogue and Roll (1974) suggest that repeated use of a one-period CAPM in a multi-step tree is admissible (e.g., Constantinides (1978, p613)), but being able to justify the RA-BT or RA-IBT without CAPM assumptions is a powerful alternative. In a technical appendix (available upon request) we show how to generate the risk premium at any node assuming power utility or negative exponential utility for a representative agent.
This *fidelity clause* is both a major strength and a slight weakness of our paper. It is a major strength because once supplied with the risk-adjusted discount rate $k$ at any node we can take that node on an existing risk-neutral tree and transform it into a node on a risk-adjusted tree while retaining the derivatives pricing at that node. That is, we get existence of a solution to the risk-averse tree driven by existence of a solution to the no-arbitrage-driven risk-neutral tree, and we get it via a no-arbitrage transformation between the two trees at that node. The fidelity clause is a slight weakness because the risk-averse probabilities inferred from the RA-BT (and subsequently from the RA-IBT as discussed below) are only as good as the discount rate fed into the transformation between the trees.

The third step is to propagate probabilities through the tree. It is almost trivial in the case of the RA-BT. The multi-period RA-BT model follows immediately from the single-period model in Equation (2) simply by applying the single-period model iteratively backwards through the tree. The values for the underlying asset price in the risk-averse tree are identical to the values for the underlying asset price in the CRR tree. In the special case where $k = r$ at every node, or, equivalently, $K = R$ at every node, the risk-averse model reduces to the CRR model. In the limit where step size tends to zero, Black-Scholes-World pricing is obtained.

Assuming a constant risk-premium in the multi-period RA-BT places a restriction on the most general form of the RA-BT where the risk premium varies freely with state and time. A constant risk premium in the RA-BT is consistent with assuming power utility for a representative agent.

### 3. The Risk-Averse Implied Binomial Tree (RA-IBT)

In this section, existence of a solution to the RA-IBT is established. That is, we show that

---

\[^{11}\text{Admissible } k \text{ requires } -\sigma \sqrt{\Delta t} < k \Delta t < \sigma \sqrt{\Delta t} , \text{ or equivalently } d < K < u .\]
there exists a set of ending nodal risk-adjusted probabilities and an associated unique fully specified stochastic process for price under the risk-averse probabilities. We also discuss the relationships between the RA-IBT, RA-BT, CRR, and the R-IBT models.

We need to establish the same three technical steps in building an RA-IBT that we established for the RA-BT. The first two steps are established by starting with an R-IBT. Rubinstein’s R-IBT is a multi-step binomial tree. Each step in the R-IBT is a one-step CRR model. The R-IBT (and thus each of the one-step CRR trees of which it is composed) is estimated using a calibration to market prices of traded European-style options. This calibration allows the parameters of each one-step CRR model to be different at every node within the R-IBT. In particular, the R-IBT allows the risk-neutral probability of an up move \( q \) to differ at each node in the R-IBT. By doing so, the final distribution of asset prices in an R-IBT need not be lognormal even when step sizes tend to zero. The R-IBT is well behaved and robust (Chriss, 1997, p431). In practice, as long as no arbitrage violations exist amongst the option prices, then a solution exists for the R-IBT (Rubinstein [1994, p783]).

We can both establish that a solution to the risk-averse implied binomial tree (RA-IBT) model exists, and demonstrate how the solution is related to the other binomial trees, by asking what happens if at each node within an R-IBT we transform the one-step CRR model there to a one-step RA-BT model. That is, given risk-adjusted discount rate \( k \) at that node, and focusing on just one internal node of the tree, we replace risk-neutral probability of an up move \( q = (R - d)/(u - d) \) with risk-averse probability of an up move \( p = (K - d)/(u - d) \), where \( R \) and \( K \) are as defined earlier. We do not change \( u \) or \( d \), or the values of the underlying or the value of the derivative at this node, just the probabilities. The valuation formula at this node, looking
ahead to the next two nodes, then changes from Equation (1) to Equation (2). This is the first step we needed to establish.

If we do exactly the same transformation for every internal node in the tree we create a new tree full of one-step risk-averse tree (RA-BT) models. The new tree provides the same pricing as the R-IBT at each node. That is, the underlying asset and any derivative have the same values at each node on the new tree as they had in the R-IBT we started with. That is, we have a new tree that has risk-averse probabilities of an up jump at any step, risk-adjusted discount rates for the underlying, and a new pricing formula (Equation (2)) at each node. This new tree is our RA-IBT. If a solution exists for the R-IBT (and we note above that it does in the absence of arbitrage opportunities between the options), then we can build our new RA-IBT tree using the transformation described above and in detail in Appendix A.

The second step is generation of the risk premium at each node to feed into the transformation at the first step. To simplify the exposition, we assume a constant risk premium throughout the tree. This is not a requirement of the model, but rather an empirical restriction imposed here for ease of exposition. This restriction is consistent with a representative agent who possesses a power-like utility function (i.e., power utility where the CRRA parameter varies with the state).

This new RA-IBT tree captures volatility smiles and excess skewness and kurtosis. The probability structure in this tree is, however, no longer risk-neutral, but risk-averse. It has all the benefits of the R-IBT, but without the restriction to risk-neutral probabilities. Of course, where the risk-premium is zero, the RA-IBT reduces to an R-IBT.

Although implicit, we have not yet mentioned explicitly how the ending nodal risk-averse

---

12 To be precise, note that the one-step CRR trees embedded within the R-IBT do not necessarily possess the
probabilities are generated in our RA-IBT. This third step is much more complicated in the case of the implied RA-IBT tree than it was for the RA-BT. In the original Rubinstein R-IBT an optimization is performed, and the ending nodal risk-neutral probabilities are the choice variables that are estimated. Rubinstein (1994, p. 790) supplies a backward recursion that starts at these ending nodal probabilities, and works backward through the tree to deduce all the $u$, $d$, and $q$ parameters which typically vary at each node (with constant discount rate $r$ for both the underlying and the option). Indeed, most binomial trees are characterized by the need to propagate variables \textit{backward} through the tree. In our RA-IBT something quite different is needed. Having already solved for an R-IBT and propagated its ending nodal probabilities backward through the R-IBT, we then apply the transformation described above to arrive at the $u$, $d$, and $p$ parameters which typically vary at each node through the RA-IBT. Recall that the transformation needs a risk-adjusted discount rate $k$ for the underlying at each node either from an assumed utility function or imposed (with utility consequences). The $u$ and $d$ parameters in the RA-IBT tree are unchanged from those in the R-IBT tree. We may then propagate the up-step probabilities $p$ \textit{forwards} through the RA-IBT to obtain nodal probabilities at each node out to the ending nodes of the RA-IBT.\footnote{We do not require this property here, so it does not matter.} This completes the derivation of the RA-IBT.

We now discuss another distinction between the R-IBT and the RA-IBT. Rubinstein’s R-IBT possesses binomial path independence (BPI). That is, each path leading to any given node arrives there with equal probability (Chriss, 1997, p417). The nodal probability at any node in an R-IBT is thus simply the path probability times the number of paths arriving at that node. Our transformation from the risk-neutral R-IBT to the risk-adjusted RA-IBT does not, however, preserve BPI. It is not true that paths through our risk-averse RA-IBT tree have equal path
probability. Rubinstein forces his R-IBT tree to have BPI so as to reduce the degrees of freedom enough to be able to solve the problem and arrive at a solution that propagates naturally backwards through his R-IBT tree. The structure in our RA-IBT tree is so similar to the structure in his R-IBT that the lack of BPI may, at first glance, appear to be a problem. In fact, this is not a problem for us because the fact that Rubinstein enforces BPI in his tree yields his solution which guarantees the existence of ours. Our direct transformation of Rubinstein’s tree to ours is followed by a forward propagation of the probabilities without ever needing to explicitly work out path probabilities. The bottom line is that we do not need BPI to arrive at our solution to the RA-IBT.

Let us emphasize again that we rely upon three distinct but interrelated technical steps. The first technical step is the derivation of the functional form of the no-arbitrage transformation that binds the risk-neutral and risk-averse trees, as demonstrated in Appendix A. At any given node in the tree, this transformation is a function of the risk-adjusted discount rate or risk premium at that node. The second technical step is the ability to derive this risk-adjusted discount rate at any given node. The third technical step is our demonstration of how to combine and implement these first two steps and propagate risk-adjusted probabilities through the risk-averse binomial tree.

4. Empirical Analysis

4.1 The Options Data

We use data on S&P500 Index Options (SPX) over the period from January 1993 to September 1995. 20 out of 33 possible sets are usable after we eliminate data that do not provide

---

13 In a technical appendix (available upon request) we demonstrate the algorithm for propagation of up-step probabilities forwards through the tree to obtain nodal probabilities at each node of the RA-IBT.
an adequate cross-section of prices. Although dated, the SPX data is intra-day and is preferred to closing data due to the significant possibility of stale quotes at the close of trading; intra-day option data is currently not easily available for purchase. The SPX options are European-style options following a March cycle. Expiration is on the Saturday following the third Friday of the month. These data were purchased from the Chicago Board Options Exchange (CBOE). Our data sample terminates at 1995 because the CBOE stopped selling the data, but also because we want to compare our results to other researchers’ results using the same data.

Each month ten call option quotes are used with bid and ask prices in excess of $0.50. The options are of different strikes, but the same maturity (two months, but varying between 59 and 61 days throughout the time period). The options are selected closest to-the-money and as close as possible to 11:00AM CST. For a given calibration, all of the quotes are usually collected within a quarter of an hour. The objective is to collect a large and reasonably synchronous sample of option prices that mitigates the potential for large price changes in the underlying security. Correspondingly, the index level is sampled as close as possible to 11:00AM CST. The index level is adjusted for dividends (i.e., the discounted value of future dividends during the life of the option is subtracted from the index price based on historic dividend payouts collected from the S&P 500 Information Bulletin). Given the short maturity of the options and the stability of the index over this time period, using the actual dividends as a substitute for anticipated dividends appears reasonable. Further, for the same reason, the riskless rate is used to discount the dividends. The effect of using an assumed higher discount rate corresponding to the index for discounting the dividends is negligible. Finally, option quotes are screened for arbitrage violations and the associated risk-free rate is inferred from closing quote midpoints of U. S. Treasury Securities that straddle the option maturity date (collected from the Wall Street
Given the criteria for the data, 20 sets of options are available for empirical analysis. Using a 200-step binomial tree, the RA-IBT model is estimated using imposed risk premiums of 0.0% (this is the R-IBT), 3.7%, 7.5%, and 11.3% (i.e., RA-IBTs consistent with power utility with varying CRRA). In total, 20 binomial trees are estimated via optimization using ten option quotes each (these are the R-IBT trees); an additional 60 RA-IBT trees are derived as transformations of the R-IBTs (three different risk premiums for each R-IBT). A CRR tree for each of the 20 sets of options data is also computed for comparison purposes.

Table 1 displays, as an example, the R-IBT prices vs. the bid and ask prices for the first set of options (1/19/1993). The dividend-adjusted level of the S&P 500 around 11AM CST was 433.78. The model is constrained to price within the bid and ask prices of each option. Our RA-IBT trees produce option pricing that is identical to the R-IBT pricing to at least ten decimal places—see Appendix B for further details.

4.2 Tail and Crash Probabilities

Jackwerth and Rubinstein (1996) report crash/decline probabilities inferred from an R-IBT and they show that after the Crash of 87 these probabilities are quite different from those inferred from the more restrictive lognormal CRR model. In this section we show that although these R-IBT probabilities are superior to those inferred from a CRR model, the risk-neutral nature of the R-IBT means that the risk-neutral decline probabilities inferred from an R-IBT can

---

14 Dimson et al. (2003) report an arithmetic mean US equity risk premium of 7.5% per annum with a standard deviation of 19.7% (Table 1). Their sample size is 102 years of annual data. The standard error of the mean is thus
be misleading when compared to the risk-averse probabilities inferred from our RA-IBT.

In Table 2, averages are reported across the data to produce percentage crash/decline CDFs for each model relative to the current dividend-adjusted price—arguably the most important measure of likelihood of decline. Table 3 reports the same information, but each of the returns is first standardized (continuously-compounded, de-meaned and scaled by standard deviation) before averaging across the data.\footnote{The log returns analyzed in Table 3 are normally distributed in the CRR model in the limit as step size goes to zero. So, we expect a value of 50\% for the CDF at a decline of 0 standard deviations or worse, and 2.5\% for a decline of 1.96 standard deviations or worse. A 200-step tree is not enough to get these results precisely, but the reader can see in Table 3 that they are very close. The simple net returns discussed in Table 2 have no such simple}

This is the same standardizing transformation that Jackwerth and Rubinstein (1996) perform when reporting crash probabilities. Table 3 thus reports crash/decline probabilities relative to the mean of the distribution of log returns rather than relative to the current dividend-adjusted spot. The average of the implied volatility based on the at-the-money Black-Scholes value across the data sets is 12.0\%, thus, a one-standard-deviation or worse decline for the standardized log return in Table 3 corresponds roughly to a 12\% or worse decline for the simple net return relative to the mean of the two-month forecast distribution of prices.

Looking at Tables 2 and 3 we see that the CRR crash/decline probabilities differ significantly from any of the implied trees; this is what Jackwerth and Rubinstein (1996) report. Looking across implied trees only, we see many differences in Table 2, but few differences in Table 3. This means that the implied densities are of relatively similar shape, but have quite

\begin{align*}
given by \frac{0.197}{\sqrt{102}} &= 1.95\%.
\end{align*}
We use a risk premium of 7.5\% plus or minus 1.96 standard errors to get 7.5\% plus or minus 3.8\%.

\footnote{The log returns analyzed in Table 3 are normally distributed in the CRR model in the limit as step size goes to zero. So, we expect a value of 50\% for the CDF at a decline of 0 standard deviations or worse, and 2.5\% for a decline of 1.96 standard deviations or worse. A 200-step tree is not enough to get these results precisely, but the reader can see in Table 3 that they are very close. The simple net returns discussed in Table 2 have no such simple}
different means. The former is hardly surprising given that the RA-IBTs are inferred from R-IBTs via a transformation, and the latter is expected given that the RA-IBTs are constructed to capture different means. Looking at the probability of a 5% or worse decline in Table 2, we see that the CRR model says it is 12.377%, the R-IBT says it is 8.295%, and the RA-IBT with median risk premium says it is only 5.086%. For declines beyond 10%, the implied trees all roughly agree, but for any level of modest decline, say 1–10%, the risk-averse RA-IBT implied trees give quite different probabilities to the risk-neutral R-IBT implied tree or the CRR tree. A modest decline in the index can easily produce an extreme decline in a leveraged portfolio, so there are direct implications for Value at Risk (VAR) analysis of leveraged portfolios (e.g., hedge funds). Thus, although the risk-neutral R-IBT results are a step in the right direction compared to the risk-neutral CRR results, the R-IBT probabilities may be misleading for inferring the risk-averse probabilities of large moves in leveraged portfolios.

For crash/decline probabilities beyond a 15% decline in Table 2 or beyond a one standard deviation decline in Table 3, we are into the extreme left tail of the distribution. We do not place much credence in these extreme estimates. Although smoothly estimated, the tails of an IBT are not as reliable as the central portion (Jackwerth and Rubinstein [1996], Jackwerth [2004], Jackwerth and Rubinstein [2004]).

Note that although our CRR crash/decline probabilities in Table 3 are virtually identical to those reported in Jackwerth and Rubinstein (1996, Column 1 of their Table 4), and although our R-IBT probabilities match roughly Jackwerth and Rubinstein’s post-1987 R-IBT probabilities for any decline (zero or more standard deviations) or for a big crash (three standard deviations or more), our R-IBT probabilities for a one- and two-standard deviation decline or

property; they are lognormal in the CRR case, so the 0% decline or worse is not with probability 50%, even in the
worse are only a quarter of Jackwerth and Rubinstein’s numbers. Probability mass has been carved out of the left flank of the distribution and pushed toward the median, making the PDF steeper and more peaked than it was in Jackwerth and Rubinstein’s sample. Our data start in 1993, and their data (used to construct their Table 4) end in 1992 but we suspect the differences are because we are using two-month options and Jackwerth and Rubinstein used four- to eight-month options. We think market participants fear a big crash (three standard deviations or worse) or any decline at all (zero standard deviations or worse) with roughly the same probability regardless of the horizon being two-months or four to eight months, but one- or two-standard deviation declines are less likely and less feared in two months than in four to eight months.

Ait-Sahalia and Lo (2000) discuss the use of implied densities for VAR. They use kernel regressions rather than IBTs to estimate risk-neutral state price densities (SPDs). Their computational routines use pooled data to estimate a SPD that is consistent with the time series and cross section of S&P500 options prices—like Ait-Sahalia and Lo (1998). Their research goal in respect to VAR is quite different from ours. They argue that existing VAR techniques that use risk-averse probability densities are inferior to VAR techniques that use risk-neutral probability densities. Their basic tenet is that SPDs are continuous time representations of Arrow-Debreu prices and that as such they incorporate a measure of the economic, rather than just the statistical, value of dollar losses; they refer to this as E-VAR (Economic VAR). They argue, essentially, that a dollar loss in one state of the world is not necessarily of the same significance as a dollar loss in another state of the world, and so a state-price density needs to be used to assess them relative to each other.

We leave practitioners to decide the merits of Ait-Sahalia and Lo’s E-VAR. We argue

\[ \text{limit as step size goes to zero.} \]
here only that the risk-neutral and risk-adjusted tail probabilities are different and that if you want a statistical measure of VAR then it is important that you know how different the risk-neutral and risk-adjusted tail probabilities are (as in Table 2).

Finally, practitioners currently using historical data to implement ex-post VAR estimates may benefit from using the ex-ante risk-averse tail probabilities derived using our RA-IBT. If you are going to use risk-averse probabilities to forecast future VAR, then we think it makes sense to use forward looking probabilities rather than backward looking ones.

4.3 Risk-Neutral versus Risk-Averse Moments

Table 4 displays the annual volatility of continuously-compounded returns, the skewness coefficient of these returns, and the kurtosis coefficient of these returns with a 7.5% risk premium. Tests of differences between implied moments (i.e., volatility, skewness, or kurtosis) for each pair of risk premiums (taken from 0.0%, 3.7%, 7.5%, and 11.3%) were performed using a non-parametric sign test. Although all differences were statistically significant, none of the differences appears economically significant to us. As such, we report the implied moments in only the medium 7.5% risk-premium case in Table 4, and we do not report the statistics.

The skewness and kurtosis numbers reported in Table 4 are quite different from the numbers reported in Jackwerth and Rubinstein (1996, Figure 6, p. 1269). Jackwerth and Rubinstein report skewness of roughly -1, and kurtosis of roughly 1.25 at the end of their sample (which corresponds to the beginning of our sample). Our numbers are almost double theirs. The explanation is that Jackwerth and Rubinstein use options of all maturities and report only the
median skewness and kurtosis numbers whereas we use only two-month options.

Based on our data and on our (unreported) moment estimations for each risk premium level, an under-estimation of the risk premium leads to an over estimation of volatility and an under-estimation of the magnitude of both the negative skewness and the positive kurtosis—but the differences do not seem economically significant. Thus, the selection of the risk premium (or the risk aversion coefficient in the utility function) affects the mean of the risk-averse returns distribution, but not the shape of the distribution of returns.

[INSERT TABLE 4 HERE]

Another application of the RA-IBT model is for the estimation of stochastic volatility option pricing models. Practitioners may be tempted to calibrate their models by forcing the ex-ante second through fourth risk-neutral moments to match the risk-averse moments based on historical data. Our implied risk-averse model can be used to test the differences between ex-ante risk-neutral moments, ex-ante risk-averse moments, and historical risk-averse moments to indicate whether any such calibration is justified or not.

4.4 Marginal Rate of Substitution and Implied Relative Risk Aversion

We now consider the relationships between utility functions and probability densities. There are many reasons for doing this. For example, stochastic volatility type and jump diffusion type option pricing models typically have skewed and kurtotic return distributions for the underlying security. In these models, a utility function is necessary to map from the theoretical risk-averse probability return distribution of the underlying security to the associated risk-neutral probabilities.

---

16 A normal distribution has a skewness coefficient of zero and a kurtosis coefficient of zero. In the CRR case, not shown, the skewness and kurtosis are very close to zero, because the continuously-compounded returns distribution in the CRR 200-step tree is very nearly normal.
return distribution. Generally, the utility function is chosen with convenience in mind and not necessarily due to a particular economic rationale. Empirical work that harvests information about aggregate utility from option prices can help to shape the assumptions made in such models.17

To use our estimated density functions for the S&P500 to make inferences about representative agent utility functions, we assume that some unspecified equilibrium asset pricing model holds, that it applies in a representative agent setting, and that the S&P500 as a broad market index serves as a proxy for aggregate consumption. For related discussion see Ait-Sahalia and Lo (2000) who in turn cite Brown and Gibbons (1985) and also discuss the limits of these assumptions; see also Bliss and Panigirtzoglou (2004).

Let \( f_{IBT}(S_T) \) and \( f_{RA-IBT}(S_T) \) denote the implied risk-neutral and risk-averse density functions respectively that are inferred from our IBTs for the future level \( S_T \) of the S&P500.18 We exploit two relations between utility functions and probability densities. The first relation is that the ratio of the risk-neutral density to the risk-averse density gives, up to a constant that is independent of the index level, the marginal rate of substitution (MRS) of the representative agent between consumption at time \( T \) and time \( t \), as shown in Equation (3).19

\[
\frac{f_{IBT}(S_T)}{f_{RA-IBT}(S_T)} \propto \frac{U''(S_T)}{U'(S_T)} = MRS
\]  

The MRS is the “pricing kernel.” We will use the terms “MRS” and “pricing kernel”

---

17 Ingersoll (1989, p38) shows that the utility function can be recovered, up to a linear transformation, by twice integrating the Arrow-Pratt absolute risk aversion function (i.e., the relative risk aversion divided by wealth).
18 Strictly speaking, our discrete trees yield discrete probability masses associated with ending discrete nodal values of the index. For our 200-step trees we associate the probability mass with the width of a range of index values about the node, and we deduce the density \( f \) as the constant value of mass/width over that range about that node.
interchangeably.

If the representative agent is risk-neutral then we expect the MRS to be unity. If the representative agent is risk averse, we expect the MRS to be downward sloping as a function of wealth. Jackwerth (2000, 2004) reports several authors finding that the MRS is locally upward sloping for some wealth levels near the initial wealth: a “pricing kernel puzzle” because it means the representative investor is locally risk seeing.

We find that the MRS is downward sloping and well behaved for each level of risk premium fed into the RA-IBT model (see Figure 1 for the 3.7% and 7.5% risk premium cases). The time period for our data sample encompasses that of Ait-Sahalia and Lo (2000). Our results for the MRS are quite similar to those in Ait-Sahalia and Lo (2000, Figure 3, p36), though their confidence interval is bordering on negative territory at high values of the index whereas ours is clearly positive everywhere. Our results are, however, quite unlike the oddly shaped, locally increasing, pricing kernels in Jackwerth (2004, Figure 11, p57). The Jackwerth (2004) data are from a much later time period than ours, and this could partially explain differences in results.

The Ait-Sahalia and Lo (2000) MRS is calculated using non-parametric techniques for the numerator, and historical data for the denominator in Equation (3). The Jackwerth (2004) MRS is calculated using IBTs for the numerator and historical data for the denominator in Equation (3). We differ from each of these authors in that we are the first to use wholly implied techniques for both numerator and denominator. Our MRS results are much smoother than Ait-Sahalia and Lo’s and wholly downward sloping unlike Jackwerth (2004).

The second relation between utility and probability densities we exploit is that we can estimate the implied Arrow-Pratt measure of relative risk aversion (RRA) using Equation (4)
Typical empirical estimates of the RRA range from about 0 to 55 (see good summaries of prior findings in Ait-Sahalia and Lo [2000, Table 5, p 39] and Jackwerth [2004, pp53–54]). Jackwerth finds however, clearly negative values for absolute (and thus also for relative) risk aversion near initial levels of wealth (Jackwerth [2000, Figure 3, p442]). Negative RRA ties in with locally upward sloping MRS and forms his pricing kernel puzzle—inconsistent with economic theory.

Our implied RRA numbers are wholly positive across all levels of wealth, and are of the order of 3–9, 7–18, and 9–27 (not shown) for the 3.7%, 7.5%, and 11.3% risk premium cases, respectively. Our implied RRA numbers (Figure 2) are similar to Ait-Sahalia and Lo (2000, Figure 4, p38) both in sign, size, and behaviour across states, but our plots are again much smoother than Ait-Sahalia and Lo’s. Like Ait-Sahalia and Lo we find economically and statistically significant evidence against CRRA and we find that RRA increases with increasing wealth beyond current levels. Of course, varying CRRA is consistent with our assumption of an imposed constant risk premium. Our results are quite different from the clearly negative results in Jackwerth (2000, Figure 3, p442). Unlike our MRS results, the time period that Jackwerth uses to calculate his risk aversion in Panel D of his Figure 3 (Jackwerth [2000, p442]) overlaps

\[ RRA = S_T \left[ f'_{RA-IBT} (S_T) / f_{RA-IBT} (S_T) - f''_{R-IBT} (S_T) / f_{R-IBT} (S_T) \right] \]  

(4)

20 See Leland (1980, Equation 3, p587) but set his \( Y(W_t) = W_t \) for aggregate consumption and then note that in this case \( \lambda p(W_t) = e^{r(T-t)} u'(W_t) f^*(W_t) \), say, where the “*” denotes the risk neutral density. Then his \( Y' \equiv 1 \) and the result follows. See also Jackwerth (1999, p72), Jackwerth (2000, p436), and Ait-Sahalia and Lo (2000, p27).

21 In fact, we use \( RRA = (1 + \rho)[f'_{RA-IBT} (\rho) / f_{RA-IBT} (\rho) - f'_{R-IBT} (\rho) / f_{R-IBT} (\rho)] \), where \( \rho = S_t/S^*-1 \) is the simple net return so that we can put each of the 20 months’ estimations on an equal footing and average across.
substantially with the time period we use. So, different time periods are not the explanation.

Perhaps it is hardly surprising that we find MRS>0 and RRA>0 because we assume that the risk premium in our RA-IBT trees is positive. If we put a negative risk premium of -15% into our RA-IBT, then for values of relative wealth close to 1, we can reproduce Jackwerth’s RRA number of approximately -15. We obtain, however, a MRS function that is wholly upward sloping and a RRA function that is wholly negative and concave down (i.e., qualitatively the mirror images in the horizontal axis of Figure 2 but with a different scale). It follows that our assumption of a positive risk premium does not explain the differences between what we find and what Jackwerth finds.

The major difference between our analysis and Jackwerth’s is that we use implied trees for the risk-averse density estimation and he uses historical data. Although Jackwerth (2000, pp445–446) dismisses his use of historical data to estimate the risk-averse density as a cause of the puzzle, we suspect that this may be at least in part responsible for his economically unintuitive results.

Ait-Sahalia and Lo (1998) criticise IBTs as possessing inherently non-stationary estimates relative to non-parametric techniques, but we certainly do not see that in the results we present. In particular, in comparing our MRS and RRA estimations to those in Ait-Sahalia and Lo (2000), we note that our plots are much smoother and less “bouncy.” This difference in smoothness may be because their non-parametric kernel estimations use a band width that is too small, though Jackwerth suggests that they may in fact have over-smoothed their results (Jackwerth [2004, p54]). Alternatively, the difference in smoothness may be because Ait-Sahalia

them to get the mean RRA and its empirical standard error. This form of the RRA is mathematically identical to
and Lo estimate their plots as a snapshot over one year, whereas we re-estimate our trees 20
times over three years and then average the results. Our averaging may produce smoother results
than their non-parametric technique, but we would not expect this if our IBTs were inherently
non-stationary, as suggested by Ait-Sahalia and Lo.

Bliss and Panigirtzoglou (2004) approach the stationarity issue from yet another angle.
Rather than using implied binomial trees, they use Breeden and Litzenberger’s (1978) well
known result to deduce risk-neutral densities from option prices. They then assume a utility
function (either power or exponential) and use it to obtain a risk-adjusted density function. By
fixing the utility function and allowing the density functions to be time-varying, they avoid
having to assume that the density functions are stationary. The aim of their paper is to calibrate
the parameters of the utility function so as not to be able to reject the risk-adjusted implied
densities as forecasts of subsequently realized returns distributions. They find RRA estimates
that decrease with increasing horizon, and little evidence of pricing kernel anomalies.

The bottom line is that we find no pricing kernel puzzle using wholly implied techniques.
So, the representative agent is risk-averse with no local risk seeking behavior. We do find
significant variation in RRA across states that is inconsistent with an assumption of CRRA. We
also find that risk aversion increases with increasing wealth beyond current wealth. Our
empirical estimates of MRS and RRA are different enough from many articles in the literature
that we can comment on possible problems with the estimation techniques used by others.

**4.5 Directions for Future Research**

In the spirit of Bliss and Panigirtzoglou (2004) and Alonso *et al.* (2006), the risk-averse

---

Equation (4)—although we have never seen it published.
RA-IBT and the risk-neutral R-IBT can be used to calibrate the parameters of an assumed utility function so that the RA-IBT best predicts distributions of future realized returns to the S&P500, say. For example, assume we have estimated an R-IBT based on traded SPX options. We could then choose an exponential utility function (with CARA but IRRA), say, and estimate the parameter of the utility function by using the utility function to determine the risk premium node by node in the RA-IBT so that the RA-IBT best represents subsequently realized return densities. This sort of analysis is, however, subject to the look-ahead problems identified by Blackburn (2006)—as mentioned previously.

5. Conclusion

Our risk-averse implied binomial tree (RA-IBT) model generalizes Rubinstein’s risk-neutral implied binomial tree (R-IBT) model by allowing for a non-zero risk premium on the underlying asset. The RA-IBT accommodates a risk premium that is time varying and/or state dependent depending upon either the assumed utility function of the representative agent (in the case of a macro asset like the S&P500 index portfolio) or a CAPM beta (in the case of an individual stock). We have imposed a constant risk premium in our empirical work with S&P500 index options (consistent with assumed power utility with varying CRRA for a representative agent). Our S&P500 Index options data run between 1993 and 1995. We demonstrate that risk-averse decline probabilities can differ significantly from the risk-neutral decline probabilities. We also estimate the pricing kernel (marginal rate of substitution) and implied relative risk aversion and compare and contrast our results with other researchers’ results. In particular, we are the first to use implied techniques for both the risk-neutral and risk-averse densities and we find no “pricing kernel puzzle” using these techniques (compared to other authors who use
historical data to generate the risk-averse density). We also find that the moments of the risk-
neutral return distribution are statistically different from the moments of the corresponding risk-
averse return distribution, but these differences do not seem economically significant to us. Our
results have immediate and direct applications to Value at Risk and stochastic volatility model
estimation.

Suppose that you have a one-step CRR tree, with parameters as denoted in the paper and pricing over time step $\Delta t$ given by Equation (1). In this case we know that $q = (R - d)/(u - d)$ is the risk-neutral probability of an up move over the time step because this is the only value of $q$ that solves $E(S_{t+\Delta t}) = qS_u + (1 - q)S_d = S_t e^{r(\Delta t)}$ when $S_u = u \times S$ and $S_d = d \times S$. We may expand the no-arbitrage-driven Equation (1) as shown in Equation (1A).

\[
V_t = \frac{1}{R} [qV_u + (1 - q)V_d] \\
= \frac{1}{R(u - d)} [(R - d)V_u + (u - R)V_d] \\
= \frac{1}{R(u - d)} [(K - d)V_u + (u - K)V_d - (V_u - V_d)(K - R)] \quad (1A) \\
= \frac{1}{R} \left\{ \left( \frac{K - d}{u - d} \right) V_u + \left( \frac{u - K}{u - d} \right) V_d - \left( \frac{V_u - V_d}{u - d} \right) (K - R) \right\} \\
= \frac{1}{R} \left\{ pV_u + (1 - p)V_d - \left( \frac{V_u - V_d}{u - d} \right) (K - R) \right\}
\]

where $p = (K - d)/(u - d)$. Note that we have tautologically added and subtracted a term $K$ in Equation (1A); it follows that Equation (1A) hold for any $K$. Even if $K$ is average annual rainfall in London during the 17th century, Equation (1A) are still tautologically valid. Of course, in that case, $p$ has no sensible interpretation with respect to the asset price $S_t$.

Now, suppose that we are exogenously given a particular $K = e^{k\Delta t}$ where $k$ is the risk-adjusted annualized discount rate on the security at this price level and over this time step (i.e., $k$ satisfies $E(S_{t+\Delta t}) = S_t e^{k(\Delta t)}$). What does this value of $K$ imply about the value of $p$ with respect to the asset price $S_t$? Simple algebra yields Equation (2A) directly.
\[
[pS_u + (1 - p)S_d] = \left(\frac{K - d}{u - d}\right)S_u + \left[1 - \left(\frac{K - d}{u - d}\right)\right]S_d
\]

\[
= \left(\frac{K - d}{u - d}\right)(u \times S_t) + \left[1 - \left(\frac{K - d}{u - d}\right)\right](d \times S_t)
\]

\[
= \left(\frac{K - d}{u - d}\right)(u \times S_t) + \left(\frac{u - K}{u - d}\right)(d \times S_t)
\]

\[
= KS_t = S_te^{\Delta t} = E(S_{t+\Delta t}),
\]

where the last two equalities follow from the definition of the given \( K \). That is,

\[
[pS_u + (1 - p)S_d] = S_te^{\Delta t} = E(S_{t+\Delta t}) \text{ in this case. However, we know that in a binomial framework there are only two possible outcomes for } S_{t+\Delta t} : S_u \text{ or } S_d. \text{ If these outcomes occur with risk-averse probabilities } p' \text{ and } (1 - p') \text{, say, then it follows that } E(S_{t+\Delta t}) = [p'S_u + (1 - p')S_d],
\]

where \( p' \) is the risk-adjusted probability of an up move. There is no choice but to conclude that \( p = (K - d)/(u - d) \) is the risk-averse probability \( p' \) of the up move at this price level and over this time step.

In summary, then, given \( K = e^{k\Delta t} \) where \( k \) is the risk-adjusted annualized discount rate on the security at this price level and over this time step (i.e., \( k \) satisfies \( E(S_{t+\Delta t}) = S_te^{k\Delta t} \)), then \( p = (K - d)/(u - d) \) is the risk-averse probability of an up move in the one-step tree and the last line in Equation (1A) is the valuation algorithm for discounting backward the future value of a European-style derivative over this time step. The only assumptions needed to obtain these results are the no-arbitrage assumptions driving the CRR one-step tree and the assumption that \( k \) is the correct risk-adjusted annualized discount rate on the security at this price level and over this time step. In particular, the functional form of the transformation in the one-step tree makes no explicit assumptions about utility functions.

Note that in our particular numerical implementation of the RA-IBT, we impose that \( k \) is
constant throughout a multi-step implied tree. Under our representative agent assumptions, this constant risk premium implementation is consistent with power utility or power-like utility. So, although the one-step transformation makes no explicit assumptions about utility, restrictions placed upon its implementation in a multi-step tree do have utility implications. An alternate proof based on the CAPM is available upon request.

Appendix B: Estimation using a transformed R-IBT

All the risk-averse RA-IBT models in the paper are estimated by transforming the solution to a risk-neutral Rubinstein R-IBT (see discussion of the transformation in Section 3 and Appendix A). The Rubinstein R-IBT is estimated using an optimization routine that minimizes the Jackwerth and Rubinstein (1996) smooth objective function subject to pricing the traded options within the spread. A lower bound of 0.0000005 is enforced on the ending nodal probabilities to assist estimation. In practice, the pricing is very good. Of the 200 options we price in calibrating the R-IBTs (i.e., ten options priced in each of 20 periods), only two are not priced within the spread. One is priced at 1/1000th of a penny below the bid, and the other is 12/1000ths of a penny below the bid. Otherwise, all other options are priced strictly within the spread. The RA-IBTs, by construction as direct transformations of the R-IBTs, produce identical pricing to the R-IBTs from which they are derived.
References


Table 1: Model Price Comparison with Bid and Ask Quotes of 60-Day Options (1/19/1993)

<table>
<thead>
<tr>
<th>Strike Price</th>
<th>Bid Price:</th>
<th>Model Price:</th>
<th>Ask Price:</th>
</tr>
</thead>
<tbody>
<tr>
<td>405</td>
<td>32.0000</td>
<td>32.7518</td>
<td>34.0000</td>
</tr>
<tr>
<td>410</td>
<td>27.3750</td>
<td>27.9949</td>
<td>29.3750</td>
</tr>
<tr>
<td>415</td>
<td>23.0000</td>
<td>23.4068</td>
<td>25.0000</td>
</tr>
<tr>
<td>420</td>
<td>19.0625</td>
<td>19.0629</td>
<td>19.8125</td>
</tr>
<tr>
<td>425</td>
<td>15.0000</td>
<td>15.0581</td>
<td>16.0000</td>
</tr>
<tr>
<td>430</td>
<td>11.2500</td>
<td>11.4714</td>
<td>12.2500</td>
</tr>
<tr>
<td>435*</td>
<td>8.1250</td>
<td>8.3688</td>
<td>8.6250</td>
</tr>
<tr>
<td>440</td>
<td>5.1250</td>
<td>5.7877</td>
<td>6.1250</td>
</tr>
<tr>
<td>445</td>
<td>3.3750</td>
<td>3.7434</td>
<td>3.8750</td>
</tr>
<tr>
<td>450</td>
<td>1.7500</td>
<td>2.2497</td>
<td>2.2500</td>
</tr>
</tbody>
</table>

* Indicates the at-the-money option. The risk-free rate is 2.88% per annum. The annual implied volatility is 11.2%. The objective function (see Appendix B.1) is minimized to a value of 0.0000425. These model prices are from a 200-step R-IBT using a lower bound on nodal probabilities of 0.0000005. Identical model prices (to more than 10 decimal places) are obtained from the three RA-IBT models with different risk premiums.
### Table 2: Implied Crash/Decline Probabilities by Model and Risk Premium
Using the Log Return Based on the Dividend-Adjusted Spot Price

<table>
<thead>
<tr>
<th>Crash/Decline Magnitude:</th>
<th>CRR Model R-IBT RP = 0</th>
<th>RA-IBT RP = 3.7%</th>
<th>RA-IBT RP = 7.5%</th>
<th>RA-IBT RP = 11.3%</th>
</tr>
</thead>
<tbody>
<tr>
<td>-40%</td>
<td>0.000%</td>
<td>1.044%</td>
<td>1.014%</td>
<td>0.984%</td>
</tr>
<tr>
<td>-30%</td>
<td>0.000%</td>
<td>1.108%</td>
<td>1.090%</td>
<td>1.071%</td>
</tr>
<tr>
<td>-20%</td>
<td>0.001%</td>
<td>1.111%</td>
<td>1.091%</td>
<td>1.072%</td>
</tr>
<tr>
<td>-10%</td>
<td>1.386%</td>
<td>1.248%</td>
<td>1.184%</td>
<td>1.134%</td>
</tr>
<tr>
<td>-9%</td>
<td>2.279%</td>
<td>1.516%</td>
<td>1.372%</td>
<td>1.263%</td>
</tr>
<tr>
<td>-8%</td>
<td>3.541%</td>
<td>2.057%</td>
<td>1.749%</td>
<td>1.521%</td>
</tr>
<tr>
<td>-7%</td>
<td>5.643%</td>
<td>3.329%</td>
<td>2.685%</td>
<td>2.194%</td>
</tr>
<tr>
<td>-6%</td>
<td>8.507%</td>
<td>5.316%</td>
<td>4.191%</td>
<td>3.310%</td>
</tr>
<tr>
<td>-5%</td>
<td>12.377%</td>
<td>8.295%</td>
<td>6.522%</td>
<td>5.086%</td>
</tr>
<tr>
<td>-4%</td>
<td>16.925%</td>
<td>12.023%</td>
<td>9.555%</td>
<td>7.491%</td>
</tr>
<tr>
<td>-3%</td>
<td>22.625%</td>
<td>16.876%</td>
<td>13.615%</td>
<td>10.797%</td>
</tr>
<tr>
<td>-2%</td>
<td>29.646%</td>
<td>23.044%</td>
<td>18.951%</td>
<td>15.291%</td>
</tr>
<tr>
<td>-1%</td>
<td>37.263%</td>
<td>30.099%</td>
<td>25.227%</td>
<td>20.722%</td>
</tr>
<tr>
<td>0%</td>
<td>45.783%</td>
<td>38.613%</td>
<td>33.076%</td>
<td>27.761%</td>
</tr>
</tbody>
</table>

The means are calculated over a time series using our 20 sets of options data calibrations from 1993 to 1995. More specifically, the mean is calculated over values of the CDF of the simple net return based on the dividend-adjusted spot price and the ending nodal values in the tree. So, this is decline relative to current dividend-adjusted spot price. Medians are qualitatively very similar and are omitted. Standard errors are small; in all but the first three CRR entries, the reported number is at least five times its empirical standard error.
**Table 3: Implied Crash/Decline Probabilities by Model and Risk Premium Using Number of Standard Deviations of Standardized Log Returns**

<table>
<thead>
<tr>
<th>Crash/Decline SD:</th>
<th>CRR Model</th>
<th>R-IBT RP = 0</th>
<th>RA-IBT RP = 3.7%</th>
<th>RA-IBT RP = 7.5%</th>
<th>RA-IBT RP = 11.3%</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.00</td>
<td>0.136%</td>
<td>1.110%</td>
<td>1.091%</td>
<td>1.072%</td>
<td>1.054%</td>
</tr>
<tr>
<td>-2.00</td>
<td>2.303%</td>
<td>1.113%</td>
<td>1.095%</td>
<td>1.083%</td>
<td>1.069%</td>
</tr>
<tr>
<td>-1.00</td>
<td>15.951%</td>
<td>3.447%</td>
<td>3.785%</td>
<td>3.912%</td>
<td>3.932%</td>
</tr>
<tr>
<td>-0.75</td>
<td>23.336%</td>
<td>7.469%</td>
<td>7.647%</td>
<td>7.662%</td>
<td>7.731%</td>
</tr>
<tr>
<td>-0.50</td>
<td>31.055%</td>
<td>15.413%</td>
<td>14.988%</td>
<td>15.020%</td>
<td>14.883%</td>
</tr>
<tr>
<td>-0.25</td>
<td>39.140%</td>
<td>27.214%</td>
<td>26.478%</td>
<td>26.113%</td>
<td>25.479%</td>
</tr>
<tr>
<td>0.00</td>
<td>49.995%</td>
<td>42.090%</td>
<td>41.648%</td>
<td>41.346%</td>
<td>40.746%</td>
</tr>
</tbody>
</table>

Crash/Decline is the number of standard deviations of the *standardized log return* which is {ln(ending node price ÷ dividend-adjusted spot price) – maturity-adjusted mean} ÷ maturity-adjusted standard deviation. This is the same definition used by Jackwerth and Rubinstein (p.1627, 1996). So, this is decline relative to the future expected level of the spot, not relative to the current dividend-adjusted level shown in Table 2. For each options set, the return and volatility differ (see Table 4 for a volatility comparison). The averages are calculated in time series using our 20 sets of options data calibrations from 1993 to 1995. Medians are qualitatively very similar and are omitted. In all cases the reported number is at least seven times its empirical standard error.
### Table 4: Model Implied Annual Moments (Using Dividend-Adjusted Log-Returns and a risk-premium of 7.5%)

<table>
<thead>
<tr>
<th>Date:</th>
<th>Days to Maturity:</th>
<th>CRR: Implied σ Annual:</th>
<th>Implied σ Annual:</th>
<th>Implied Skewness:</th>
<th>Implied Kurtosis:</th>
</tr>
</thead>
<tbody>
<tr>
<td>930119</td>
<td>60</td>
<td>11.24%</td>
<td>18.42%</td>
<td>-1.81158</td>
<td>2.59636</td>
</tr>
<tr>
<td>930322</td>
<td>61</td>
<td>12.64%</td>
<td>15.85%</td>
<td>-1.75812</td>
<td>2.71996</td>
</tr>
<tr>
<td>930419</td>
<td>61</td>
<td>12.78%</td>
<td>21.53%</td>
<td>-1.81760</td>
<td>2.57990</td>
</tr>
<tr>
<td>930621</td>
<td>61</td>
<td>11.64%</td>
<td>19.46%</td>
<td>-1.78284</td>
<td>2.54899</td>
</tr>
<tr>
<td>930720</td>
<td>60</td>
<td>11.58%</td>
<td>16.57%</td>
<td>-1.80459</td>
<td>2.68638</td>
</tr>
<tr>
<td>930921</td>
<td>60</td>
<td>13.78%</td>
<td>21.09%</td>
<td>-1.84357</td>
<td>2.67693</td>
</tr>
<tr>
<td>931018</td>
<td>61</td>
<td>11.35%</td>
<td>20.45%</td>
<td>-1.81766</td>
<td>2.53593</td>
</tr>
<tr>
<td>931220</td>
<td>61</td>
<td>10.41%</td>
<td>18.14%</td>
<td>-1.82601</td>
<td>2.56703</td>
</tr>
<tr>
<td>940322</td>
<td>60</td>
<td>12.03%</td>
<td>20.15%</td>
<td>-1.82811</td>
<td>2.59643</td>
</tr>
<tr>
<td>940419</td>
<td>60</td>
<td>14.36%</td>
<td>25.14%</td>
<td>-1.80667</td>
<td>2.53870</td>
</tr>
<tr>
<td>940621</td>
<td>60</td>
<td>12.22%</td>
<td>20.36%</td>
<td>-1.84598</td>
<td>2.62520</td>
</tr>
<tr>
<td>940823</td>
<td>60</td>
<td>9.66%</td>
<td>15.76%</td>
<td>-1.84285</td>
<td>2.63700</td>
</tr>
<tr>
<td>940921</td>
<td>59</td>
<td>13.32%</td>
<td>23.19%</td>
<td>-1.83481</td>
<td>2.57655</td>
</tr>
<tr>
<td>941122</td>
<td>60</td>
<td>15.66%</td>
<td>23.82%</td>
<td>-1.85866</td>
<td>2.69788</td>
</tr>
<tr>
<td>941219</td>
<td>61</td>
<td>13.09%</td>
<td>21.04%</td>
<td>-1.84492</td>
<td>2.64867</td>
</tr>
<tr>
<td>950222</td>
<td>59</td>
<td>10.04%</td>
<td>15.42%</td>
<td>-1.82316</td>
<td>2.66229</td>
</tr>
<tr>
<td>950321</td>
<td>60</td>
<td>10.98%</td>
<td>17.05%</td>
<td>-1.85282</td>
<td>2.68936</td>
</tr>
<tr>
<td>950523</td>
<td>60</td>
<td>9.70%</td>
<td>12.72%</td>
<td>-1.72001</td>
<td>2.65395</td>
</tr>
<tr>
<td>950822</td>
<td>60</td>
<td>12.44%</td>
<td>16.28%</td>
<td>-1.82257</td>
<td>2.77527</td>
</tr>
<tr>
<td>950919</td>
<td>60</td>
<td>11.69%</td>
<td>16.17%</td>
<td>-1.79143</td>
<td>2.70329</td>
</tr>
</tbody>
</table>

Normalized skewness is the signed 1/3 power of the absolute value of the third moment of the standardized log return (see pages 1627–1628 of Jackwerth and Rubinstein [1996]). For a normal distribution, it is equal to zero. If we had presented a more traditional skewness measure as the third central moment of standardized log return, then the skewness number reported in the first row on 930119 would instead read -5.94530. Normalized kurtosis is the signed 1/4 power of the absolute value of the fourth moment of the standardized log return less 3. For a normal distribution, it is equal to zero. If we had presented a more traditional kurtosis measure as the fourth central moment of standardized log return less 3, then the kurtosis number reported in the first row on 930119 would instead read 45.44221.
Figure 1: For the 3.7% and 7.5% risk premium cases, and for each of the 20 months’ optimizations from 1993 to 1995, we estimate the scaled marginal rate of substitution (MRS) (i.e., the pricing kernel) as the ratio $f_{R-IBT}(S_T)/f_{RA-IBT}(S_T)$. We then average these MRS numbers across the 20 months’ estimations by first associating each with the percentage change in the S&P500 relative to that month’s dividend adjusted index level $S^*$. We calculate the average only where each individual density possesses a contiguous range of values that does not use the optimization’s lower bound on the density of 0.0000005. Each tree’s ending nodes are different, so for each return level on the plot we linearly interpolate between the individual estimates before taking the average. We show the average plus and minus two empirical standard errors of the mean.
Figure 2: For the 3.7% and 7.5% risk premium cases, and for each of the 20 months’ optimizations from 1993 to 1995, we estimate the implied Arrow-Pratt RRA for each two-month-ahead level $S_T$ of the S&P500 as $S_T[f'_{RA-IBT}(S_T)/f_{RA-IBT}(S_T) - f'_{R-IBT}(S_T)/f_{R-IBT}(S_T)]$. We average these estimates over the 20 months’ estimations by first putting them on an equal footing by associating each with the percent change in the S&P500 relative to that month’s dividend adjusted index level $S^*$. We calculate the mean only where each month’s density (and its slope estimate) possesses a contiguous range of values that does not use the optimization’s lower bound on the density (0.0000005). Each tree’s ending nodes are different, so for each return level on the plot we linearly interpolate between the individual estimates before taking the average. We show the average plus and minus two empirical standard errors of the mean.