

104

Reihe Ökonomie
Economics Series

Extracting Risk-neutral Probability Distributions from Option Prices Using Trading Volume as a Filter

Dominique Y. Dupont

104

Reihe Ökonomie
Economics Series

Extracting Risk-neutral Probability Distributions from Option Prices Using Trading Volume as a Filter

Dominique Y. Dupont

September 2001

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Contact:

Dominique Y. Dupont
EURANDOM – TUE
P.O. Box 513
NL-5600 MB Eindhoven
The Netherlands
☎: +31 / 40 / 247 8123
fax: +31 / 40 / 247 8190
email: dupont@eurandom.tue.nl

Founded in 1963 by two prominent Austrians living in exile – the sociologist Paul F. Lazarsfeld and the economist Oskar Morgenstern – with the financial support from the Ford Foundation, the Austrian Federal Ministry of Education and the City of Vienna, the Institute for Advanced Studies (IHS) is the first institution for postgraduate education and research in economics and the social sciences in Austria. The **Economics Series** presents research done at the Department of Economics and Finance and aims to share “work in progress” in a timely way before formal publication. As usual, authors bear full responsibility for the content of their contributions.

Das Institut für Höhere Studien (IHS) wurde im Jahr 1963 von zwei prominenten Exilösterreichern – dem Soziologen Paul F. Lazarsfeld und dem Ökonomen Oskar Morgenstern – mit Hilfe der Ford-Stiftung, des Österreichischen Bundesministeriums für Unterricht und der Stadt Wien gegründet und ist somit die erste nachuniversitäre Lehr- und Forschungsstätte für die Sozial- und Wirtschaftswissenschaften in Österreich. Die **Reihe Ökonomie** bietet Einblick in die Forschungsarbeit der Abteilung für Ökonomie und Finanzwirtschaft und verfolgt das Ziel, abteilungsinterne Diskussionsbeiträge einer breiteren fachinternen Öffentlichkeit zugänglich zu machen. Die inhaltliche Verantwortung für die veröffentlichten Beiträge liegt bei den Autoren und Autorinnen.

Abstract

This paper introduces a new technique to infer the risk-neutral probability distribution of an asset from the prices of options on this asset. The technique is based on using the trading volume of each option as a proxy of the informativeness of the option. Not requiring the implied probability distribution to recover exactly the market prices of the options allows us to weight each option by a function of its trading volume. As a result, we obtain implied probability distributions that are both smoother and should be more reflective of fundamentals.

Keywords

Implied risk-neutral probability distribution, implied-tree method

JEL Classifications

G13, G14

Comments

The author thanks James Clouse, Mark Hooker, Vincent Reinhart, Brian Sack, Casper de Vries, and seminar participants at Eurandom. Part of this work was done when the author was an economist in the Division of Monetary Affairs at the Board of Governors of the Federal Reserve System.

Contents

1 Current Methods to Compute Risk-Neutral Probability Distributions	2
1.1 Definitions	2
1.2 Retrieving risk-neutral distributions when there are as many states of the world as traded assets	3
1.3 Retrieving risk-neutral distributions when there are more states of the world than traded assets	5
2 A New Non-Parametric Method to Compute Risk-Neutral Probability Distributions from the Prices of Traded Assets	7
3 Early Exercise	10
4 Application	12
4.1 Data	12
4.2 Results	14
5 Conclusion	18
Appendix	20
A Retrieving risk-neutral distributions when there are as many marketable assets as states of the world	20
B More states than marketable assets	21
References	22
Tables and Figures	24

Trading activity in derivatives products, whether on organized exchanges or on over-the-counter (OTC) markets has increased dramatically in the past fifteen years and derivatives have become a mainstay of financial markets.¹ Standard, plain-vanilla contracts are very liquid and trading data on exchange-traded products are publicly available. Derivatives prices can be used to extract information on the probability distributions of the underlying assets because the payoffs of derivatives contracts depend on the future values of the underlying assets. This information is of use not only to derivatives traders but also to a wider public, including holders of the underlying securities and policy makers.

Different instruments are used to infer different characteristics of the distribution of the underlying asset. Quotes on futures contracts give some information on the market's expectations of the underlying asset price; option prices provide a measure of the dispersion of the asset price around its mean, the implied volatility, which can be used to gauge the degree of uncertainty prevalent in the markets.² If the Black-Scholes theory were true, the implied volatility would be constant across strikes and would suffice to describe the whole distribution of the underlying asset. In practice, implied volatilities are not constant across strikes. Fortunately, combinations of options can yield information about the distribution of the underlying asset without the need for a priori assumptions. Bates (1991) devises a technique using option prices to estimate the skewness of the distribution of the underlying asset. If quotes on European options were available for a continuum of strike prices, one could use the method of Breeden and Litzenberger (1978) to retrieve the entire probability density function (thereafter, "pdf") of a given asset. In this case, the implied probability distribution would subsume the Bates skewness measure.

In practice, strike prices, or strikes for short, are available only at discrete intervals. If one assumes that all possible future values of the underlying asset coincide with the available strikes, the probability distribution of the asset could be uniquely recovered from the available option prices. Alternatively, the researcher could assume that the asset can take values other than the available strikes and strive to obtain a distribution consistent with the observed option prices. Because multiple distributions are consistent with the option prices in that case, some additional selection criteria is needed. The researcher may conjecture a functional form for the implied pdf and choose the parameters to replicate as well as possible the market prices of the options, or, alternatively, use non-parametric methods. Whatever the techniques used to construct them, implied probability distributions have become popular analysis tools. Some economists, like Campa et al. (1997), Melick and Thomas (1997, 1998), and Söderling and Svensson (1997) use

¹The notional amount of interest rate, currency, and equity futures and options has increased almost twentyfold between 1986 and 1997 (IMF, 1998). More recent data point to a 30-percent increase in the notional amount of OTC derivatives between June 1998 and June 2000 (BIS, 2000)

²The implied volatility of an option at a given strike is the volatility at which the price of the option assigned by the Black-Scholes (1973) formula coincides with the market price of the option.

option-based probability distributions to gauge the immediate market impact of some particular event (Table 1). Others, like Rubinstein (1994), Jackwerth and Rubinstein (1996), and Stutzer (1996) use option prices to compare implied probability distributions across longer time horizons, for example several years before and after the crash of 1987, or to build time-series of some characteristics of the distribution, like skewness and kurtosis.

Current non-parametric approaches typically try to force the probability distribution to perfectly replicate the market prices of the options. Instead, we choose to only impose a non-replication cost in the objective function and trade off increased smoothness of the implied pdf against lower goodness of fit with the market prices of the options. The advantage of this method is that one can weight each option price by a proxy of its informativeness. We select trading volume as such a proxy because there are strong economic arguments to think that asset prices contain more information when these assets are more actively traded.

1 Current Methods to Compute Risk-Neutral Probability Distributions

The following section introduces a simple two-period framework to review the concept of risk-neutral distribution and presents several methods designed to retrieve such probability distributions from option prices.

1.1 Definitions

Risk in the economy is generated by the random payoff of a single asset, S , which can take n possible values in the second period and is referred to as the "underlying asset." Agents can trade several assets in the first period that all pay out in the second period: m European call options on the underlying asset, a forward contract on the same asset, and a riskless bond. Each of the realizations of S defines a "state of the world." There is no transaction costs or restrictions to trade, and the market is free of arbitrage. We assume without loss of generality that the options are calls; prices of European puts can always be translated into equivalent call prices using the put-call parity. The following theorem presents a well-known result in financial economics.

Theorem 1

If market prices are arbitrage-free, there exists a probability distribution, P , such that, if $\pi(x)$ is the price of a claim paying out x in the last period,

$$\pi(x) = E^P[e^{-r\tau} x] \tag{1}$$

where τ is the time between the two periods, r is the (fixed) interest rate, and $E^P[\cdot]$ denotes the expectation against P . This probability distribution is called "risk-neutral."

Proof: Call x_i the realization of asset x when $S = S_i$, z_i the random variable that pays \$1 when $S = S_i$ and \$0 otherwise (z_i is the "state-contingent

claim" for the i^{th} state), and let $p_i = e^{r\tau} \pi(z_i)$. Then, because the function π is linear and $x = \sum_{i=1}^n x_i z_i$,

$$\pi(x) = \sum_{i=1}^n e^{-r\tau} x_i p_i. \quad (2)$$

The p_i 's are positive because π is assumed to admit no arbitrage opportunities and they sum up to 1 because the riskless bond, which pays \$1 in every state and is worth by definition $e^{-r\tau}$ can be replicated by holding all contingent claims. Hence, the p_i 's define a probability distribution, P , and $\pi(x) = E^P[e^{-r\tau}x]$. *Q.E.D.*

In the following, we review several techniques to construct (risk-neutral) probability distributions consistent with the prices of selected assets.

Notations: $C(K_j)$ is the price of a call option struck at K_j , $j = 1, \dots, m$, and F is the forward price of the underlying asset.

Corollary: A series of non-negative scalars p_i , $i = 1, \dots, n$, defines a risk-neutral probability distribution consistent with the prices of the traded assets if and only if

$$\begin{cases} \sum_{i=1}^n p_i e^{-r\tau} \max(0, S_i - K_j) & = C(K_j), \\ \sum_{i=1}^n p_i S_i & = F, \\ \sum_{i=1}^n p_i & = 1. \end{cases} \quad (3)$$

Eq. (3) defines a system of $m + 2$ equations in n unknowns (the p_i 's) and is equivalent to the matrix equality

$$M \cdot P = Q \quad (4)$$

where $P = (p_i)_{i=1}^n$, Q gathers the $m + 2$ terms in the right-hand side of the equation, and M is a $m + 2 \times n$ matrix obtained from the left-hand side of the equation.

The "risk-neutral" distribution combines characteristics of the "real-world" probability distribution and the market participants' appetite for risk. Nonetheless, quotes on derivative instruments taken at two moments in time can give information on the evolution of the real-world probability distribution of the underlying asset if attitudes toward risk are assumed to remain constant in the interval.

If there are as many derivatives assets as states of the world, the risk-neutral probability distribution is unique and can be inferred from the option prices in an intuitive manner. If there are more states of the world than derivatives, the researcher can use parametric or non-parametric methods to pin down one risk-neutral probability distribution.

1.2 Retrieving risk-neutral distributions when there are as many states of the world as traded assets

Setting the number of states of the world equal to the number of assets ($n = m + 2$) and ignoring the positivity constraint on the risk-neutral prob-

abilities, one can easily recover the unique risk-neutral probability distribution. Assume that the spacing between two strike prices is constant and equal to Δ . Define the possible realizations of S as follows

$$\begin{cases} S_1 &= K_1 - \Delta, \\ S_i &= K_{i-1} \text{ for } i = 2, \dots, m+1, \\ S_{m+2} &= K_m + \Delta. \end{cases} \quad (5)$$

With $n = m + 2$, the system defined in Eq. (3) admits a unique and simple solution. The implied probabilities of states in the heart of the distribution match those obtained using the "butterfly spread" strategy of Breeden and Litzenberger (1978). More precisely, if K_i is not the lowest or the highest strike, a trader can replicate the claim paying out \$1 when $S = K_i$ and nothing otherwise by taking a long position in $1/\Delta$ calls struck at K_{i-1} and K_{i+1} , and a short position in $2/\Delta$ calls struck at K_i (see Appendix). The procedure is illustrated graphically in Figure 1. The risk-neutral probability of the state " $S = K_i$ " is

$$P(S = K_i) = p_{i+1} = e^{r\tau}(C(K_{i-1}) - 2C(K_i) + C(K_{i+1}))/\Delta. \quad (6)$$

The probability in the tails of the distribution defined in Eq. (4) can be recovered using a simple trading strategy and the fact that probabilities sum up to 1.

$$\begin{cases} P(S \geq K_m) &= p_{m+1} + p_{m+2} = e^{r\tau}(C(K_{m-1}) - C(K_m))/\Delta, \\ P(S \leq K_1) &= p_1 + p_2 = 1 - \sum_{i=3}^{m+2} p_i. \end{cases} \quad (7)$$

Taking into account the underlying asset and the riskless bond enables us to extend the range of the distribution beyond the lowest and the highest strikes.

The density function at $S = K$ is $f(K) = P(S = K)/\Delta$ and hence

$$f(K) = e^{r\tau} (C(K - \Delta) - 2C(K) + C(K + \Delta))/\Delta^2. \quad (8)$$

$f(K)$ is the finite-difference approximation of $e^{r\tau} \frac{\partial^2 C(K)}{\partial K^2}$. More generally, as shown in Breeden and Litzenberger (1978), if S takes a continuum of possible values and each of those is the strike price of a call option, the risk-neutral probability density function away from the boundaries of the distribution is

$$f(K) = e^{r\tau} \frac{\partial^2 C(K)}{\partial K^2}. \quad (9)$$

No positivity constraint need to be imposed on the probability density functions defined in Eqs. (8) and (9) because the price of an option is a convex function of the strike.³ However, option prices observed in the market are not always convex in their strikes because of transaction costs or non-synchronous trading.

³This is the case when the underlying price process is a one-dimensional diffusion and in some stochastic-volatility models (Bergman et al., 1996).

In the present framework, one can replicate the payoffs of the contingent claims, and hence of any asset, by forming portfolios of the options: the market is "complete" and a unique risk-neutral distribution is inferred from option prices. In practice, m is given and the researcher decides n , that is, makes markets complete by constraining the possible realizations of the underlying asset. The fewer the strike prices, the more constraining this approach is.

Alternatively, we can consider that there are more states than traded assets and construct a risk-neutral probability distribution consistent with the prices of the traded options.

1.3 Retrieving risk-neutral distributions when there are more states of the world than traded assets

Creating more states than traded assets enables one to refine the grid of the support of the probability distribution and to extend the tails of the distribution farther away from the lowest and highest strikes. With more states than assets, markets are incomplete and multiple risk-neutral distributions can fit the option prices. In other words, when $n > m + 2$, there are more unknowns than equations in the linear system defined by Eq. (3) and the system admits multiple solutions. One needs some additional criteria to pin down a unique probability distribution. One solution is to impose a functional form for the probability distribution and to estimate its parameters using the option data. Alternatively, one can choose non-parametric methods, which can deliver implied distributions that perfectly match the option prices observed in the market. Even though not all non-parametric methods aim at perfectly replicating the market prices of derivatives, their extra flexibility allows a closer fit with market prices.

Non-parametric methods

One such method consists in minimizing the Euclidean distance between the distribution that fits the option prices, P , and an initial distribution, P_0 . This is based on the "implied-tree method" of Rubinstein (1994).⁴ The implied-trees of Rubinstein (1994) and Jackwerth and Rubinstein (1996) only require the option prices obtained from the risk-neutral pdf to be between the bid and ask prices observed in the market. When only closing or settlement prices are available, the implied-tree procedure typically constrains the pdf to exactly recover the market prices of the options (Campa et al., 1998).

Mathematically, the objective is to

$$\min_{p_i} \sum_{i=1}^n (p_i - p_i^0)^2 \quad (10)$$

subject to the constraints of Eq. (3) and

$$p_i \geq 0 \text{ for } i = 1, \dots, n. \quad (11)$$

⁴The original article by Rubinstein used binomial trees. However, the method is called "implied trees" even if no tree is used.

Eq. (3) imposes $m + 2$ conditions on the n risk-neutral probabilities p_i 's. Consequently, the n -dimensional vector P can be expressed as a linear function of $n - m - 2$ of its components and of the prices of the marketed assets. Using this relation and the first-order conditions implied in Eq. (10) while ignoring the non-negativity constraints on the p_i 's, one can easily solve for P as a linear function of the price of the marketed assets (see Appendix).

The quadratic distance function in Eq. (10) can be replaced by the "entropy" criterion function of Stutzer (1996), and Buchen and Kelly (1997):

$$\min_{p_i} \sum_{i=1}^n p_i \log(p_i/p_i^0). \quad (12)$$

Eq. (12) guarantees the positivity of the risk-neutral probabilities. Using Eq. (12) or using Eq. (10) with $p_i \geq 0$ result in a non-linear optimization problem requiring the use of numerical solution techniques. For the sake of computational simplicity, we use the quadratic distance function without imposing positivity of the risk-neutral probabilities.⁵

Parametric methods

An alternative to these non-parametric methods is to impose a flexible but parametric form on the probability distribution and determine its parameters by maximizing the fit of the option prices implied by the probability distribution to the prices observed in the market. (The objective function to minimize typically is the sum of the square differences between the option prices implied by the distribution and those observed in the market.) The additional structure provided by parametric methods may reduce the risk of overfitting the data but option prices based on the risk-neutral probability distribution fail to coincide exactly with those observed in the market. However, the researcher can improve the fit by allowing more flexibility in the parametric function (at the cost of an increase in the numerical complexity of the optimization problem).

Melick and Thomas (1997) assume that the price of the underlying asset follows a mixture of N lognormals, that is, that the risk-neutral density function g is defined by

$$g(t) = \sum_{i=1}^N \theta_i g_i(t), \quad (13)$$

where

$$g_i(t) = \frac{1}{\sqrt{2\pi}\sigma_i t} \exp \left[\frac{1}{2} \left(\frac{\log(t) - \mu_i}{\sigma_i} \right)^2 \right], \quad (14)$$

and

$$\sum_{i=1}^N \theta_i = 1. \quad (15)$$

⁵When such constraints are imposed, perfect replication of the option prices is not guaranteed.

The more lognormal distributions are included, the better the fit with the observed option prices but the more difficult the estimation procedure may be. Numerical difficulties seem to limit the choice of N to 2 to 3 in practice. Anecdotally, a change-of-regime model where the asset price at maturity follows distribution g_i with probability θ_i yields the density function of Eq. (13) but reference to this model is not necessary and the mixture-of-distribution method can be seen in its own right as a way to obtain a parametric probability distribution that closely matches options prices.

Söderling and Svensson (1997) point out that the mix-of-distribution approach sometimes generates implied pdfs with sharp spikes, and that the loss function (the minimization of which yields estimates of the parameters) is often very flat. This could cause the implausibly large changes in the shape of the pdfs between consecutive days observed by Clews et al. (2000). Jackwerth (1999) confirms that the mixture-of-distribution method tends to overfit the data when many distributions are used in the mix. Even if the overall shape of the distribution is roughly stable across short periods of time, individual distributions in the mix can change drastically, which invalidates the use of the method in estimating a change-of-regime model.

2 A New Non-Parametric Method to Compute Risk-Neutral Probability Distributions from the Prices of Traded Assets

The method introduced below is both an extension and a simplification of implied trees. The implied-tree method above forces the risk-neutral probability distribution to exactly replicate the option prices. However, one may not wish to match all quoted options prices exactly if some poorly reflect economic fundamentals. This paper innovates by substituting a penalty for not matching the option prices instead of the obligation of recovering them exactly. Not fitting the option prices perfectly allows us to weight each option price by its trading volume, which we use as a proxy for the informativeness of the option's price. In other words, highly traded options are assumed to contain more information than low-volume options.

Trading volume is typically very unequally distributed across strike prices. Clews et al. (2000) note that at- and near-the-money options account for most of trading volume in options on Financial Times Share 100 Index (FTSE 100) and on short-term sterling interest-rate futures. Recent data on options on Eurodollar, Standard and Poors 500 Index (S&P 500) futures reveals the same pattern. Exchanges typically publish daily settlement prices for options that have not been traded on the day. Such quotes are model based and should not be used to construct implied probability distributions. It is also doubtful whether prices of options with very low trading volume are meaningful.

Economic motivations for using trading volume as a filter

There are several reasons to weight each option by its trading volume. The equilibrium price of an asset averages the different values assigned to the

asset by the participating traders. Intuitively, if buying and selling orders contain some noise, the latter should have a lesser impact on highly traded options because the idiosyncratic components of the order flow are averaged across a larger pool of traders. This should lead us to give more weight to high-volume options, following the example of generalized least-squares regressions, which deal with variables that are means of other random variables by weighting each observation by the size of the sample it averages.

Another reason to use trading volume to discriminate among the different options is that quotes on high-volume options are more likely to correspond to simultaneous transactions. Transaction volume typically displays a U-shape pattern across the trading day (Jain and Joh, 1988). Since settlement prices of exchange-traded derivatives most often coincide with their prices at the close, quotes on high-volume options are likely to correspond to transactions occurring during the peak-volume time at the end of the trading session. Put another way, options with low cumulative trading volume on the day are more likely to have stopped trading during the relative low-activity period in the middle of the day. Scaling down the importance of low-volume options adapts end-of-day data to a technique, risk-neutral probability inference, that assumes that all option prices are determined simultaneously.

One justification for using trading volume to discriminate among options on the same underlying asset is grounded in economic theory: Models developed by Huffman (1992) and Blume, Easley, and O'Hara (1994) suggest that trading volume is higher when the agents' information about the true value of the traded asset is more precise. Although the lack of reliable indicator of the precision of traders' information hinders the testing of this proposition, other predictions of these models are clearly borne out by the data. For example, there is strong empirical support for the positive correlation between trading volume and the absolute difference of the price of an asset predicted by both these models, as documented by Karpoff (1987), and Gallant, Rossi, and Tauchen (1992). Using a general-equilibrium economy where differently-informed agents exchange productive capital, Huffman (1992) obtains a strong relationship between the transaction volume of capital and a measure of its informativeness about the future state of the economy. Blume, Easley, and O'Hara (1994) investigate the informational role of trading volume in a noisy rational-expectations exchange economy where the precision of the private signals observed by the traders is random. The equilibrium trading volume is an increasing function of the precision of the private signals and conveys information not revealed by the equilibrium price. Dupont (1997) uses a noisy rational-expectations framework where the precision of the informed traders' signals is deterministic but the signals' correlation structure is richer than that in Blume, Easley, and O'Hara (1994). The mean trading volume and the informativeness of the equilibrium market price, measured by its correlation with the true value of the traded asset, display very similar patterns to changes in the precision of the informed traders' private signals or in the correlation between their signals. Both the mean trading volume and the informativeness of the market price are increasing in the precision of the private signals and decreasing in the

correlation between those signals. The latter pattern confirms the conclusion of Shalen (1993) whose model generates a positive relation between the dispersion of the traders' beliefs.

A simple method to use trading volume as a filter

The method presented in this paper weights every option by a function of its trading volume. To simplify, we neglect the early exercise clause of the (American) options used and do not impose non-negativity constraints on the implied probabilities. As shown in the next section, the early-exercise feature does not seem to have much impact on the prices of options that are not very far in the money. Ignoring positivity constraints allows us to solve the optimization problem in closed form, using only linear algebra, because the loss function is quadratic and the constraints linear in the implied probabilities, so that the implied probabilities are linear in the option prices. In contrast, Jackwerth and Rubinstein (1997) impose positivity constraints on the implied probabilities and use an advanced algorithm (Broyden-Fletcher-Goldfarb-Shanno) to compute them. Even those who prefer to use non-linear optimization techniques to guarantee non-negative probabilities may be interested in having a simpler, faster procedure ready, especially if they have to recover implied probabilities daily or on demand.

The option prices can be weighted by their trading volume or a function of their trading volume. We impose that options with zero volume be given zero weight. We experiment with two specifications. Let θ_i^s be the weight allocated to the option struck at K_i by specification s and $V(i)$ be the trading volume of this option.

$$\begin{aligned}\theta_i^1 &= V(i) / \sum_{k=1}^m V(k), \\ \theta_i^2 &= \log(1 + V(i)) / \sum_{k=1}^m \log(1 + V(k)).\end{aligned}\tag{16}$$

The second specification gives high-volume options less predominance than the first. One could also exclude options whose trading volume is below a threshold.

Moreover, we fully use the flexibility offered by the non-parametric character of the implied-tree method. For example, we choose different distributions to build the initial p_i^0 's probabilities. In the heart of the distribution, that is, for prices above the minimum strike and below the maximum strike, the p_i^0 's are generated by a lognormal distribution with a volatility that minimizes the average square deviation between the asset prices implied by the distribution and the asset prices observed in the market. In contrast, the p_i^0 's in the tails of the distribution, are generated by a lognormal distribution with a higher volatility. The increase in the volatility of the lognormal distribution which serves as a reference in the estimation of the implied risk-neutral distribution matches the observation that implied volatilities tend to be higher for strikes prices further away from the center of the distribution. We also impose a quadratic smoothness criteria on the implied probabilities. Using this smoothness criteria, we can dispense with imposing that the implied probabilities be close to initial probabilities in the heart of the distribution for the options on Eurodollar futures (but not for those on S&P 500 futures). However, it is still advisable to impose closeness to a well-behaved

distribution in the tails. A major advantage of implied-tree methods over parametric approaches is that the user can easily tailor the method to his own needs.

The new objective function is

$$\begin{aligned}\Omega(P) &= \lambda_1 \sum_{i \in I} (p_i - p_i^0)^2 \\ &+ \lambda_2 \left[\sum_{j=1}^m \theta_j \left(\sum_{i=1}^n p_i e^{-r\tau} \max(0, S_i - K_j) - C(K_j) \right) \right. \\ &+ \left. \sum_{i=1}^n p_i S_i - F + \sum_{i=1}^n p_i - 1 \right] \\ &+ (1 - \lambda_1 - \lambda_2) \sum_{i=2}^{n-1} (p_{i-1} - 2p_i + p_{i+1})^2,\end{aligned}\tag{17}$$

where the λ_i 's are the weights on the constraints, I indexes the p_i 's for which a cost of deviating from the initial probabilities p_i^0 's is imposed, and the θ_i 's are the weights on the prices of the options observed in the market.

To express Eq. (17) in matrix form, write

$$\begin{aligned}S_1 &= \text{Diag}(I[i \in I]), \\ \theta &= \text{Diag}(\theta_1, \dots, \theta_m, 1, 1),\end{aligned}\tag{18}$$

and let R the matrix so that $P'RP = \sum_{i=2}^{n-1} (p_{i-1} - 2p_i + p_{i+1})^2$. The objective function becomes

$$\Omega(P) = \lambda_1 (P - P_0)' S_1 (P - P_0) + \lambda_2 (MP - Q)' \theta (MP - Q) + (1 - \lambda_1 - \lambda_2) P' R P,\tag{19}$$

where the matrices M and Q follow from Eq. (4). The solution is, assuming invertibility of the relevant matrix,

$$P = (\lambda_1 S_1 + \lambda_2 M' \theta M + (1 - \lambda_1 - \lambda_2) R)^{-1} (\lambda_1 S_1 P_0 + \lambda_2 M' \theta Q).\tag{20}$$

3 Early exercise

In this section, we review the impact of the early-exercise feature on the prices of American-style options and the special character of options that are very far in the money.⁶

Many options traded on financial exchanges can be exercised at any time before expiry. Melick and Thomas (1997) take this early-exercise feature into account by letting the model price of the option be a weighted average of the upper and lower bounds on American options and by making the weight a parameter of the objective function. As an alternative, Pirkner, Weigend, and Zimmermann (1999) combine the mixture-of-distribution method and implied binomial trees to capture the path-dependency generated by the early-exercise clause. Clews et al. (2000) use the mixture-of-distribution technique but adjust for the early exercise premium using the method developed by Barone-Adesi and Whaley (1987).

⁶The underlying asset is now assumed to be a futures contract, so that the prices of American and European calls differ. In contrast, these prices coincide when the underlying asset is a stock that pays no dividend.

However, recent research suggests that the early-exercise feature of traded options could be ignored. Söderling and Svensson (1997) compare the implied pdfs obtained using the method of Melick and Thomas (1997) to those obtained using the mixture-of-distribution technique while treating American options as though they were European options. They report that the two pdfs are indistinguishable to the eye and that "the correction for American-style is not important, at least not compared to pricing errors (which possibly reflect some other type of misspecification)" [p. 405].

To gauge the impact of the early exercise feature on option prices, we computed the implied volatilities of options on S&P 500 index futures traded on the Chicago Mercantile Exchange (CME) neglecting and taking account of the Americanness of these options (Figure 2). Except for options very far in-the-money, the implied volatilities are identical across the two methods. We show below why far-in-the-money options should be treated differently than other options (or excluded entirely).

Since American-style options can be exercised at any time, no-arbitrage lower bounds on their values are immediate. If F_t is the price of the underlying asset (a futures contract), $C_t(F_t, K)$ the price of a call option struck at K , and $P_t(F_t, K)$ the price of a put option struck at K , then

$$\begin{aligned} C_t(F_t, K) &\geq F_t - K, \\ P_t(F_t, K) &\geq K - F_t. \end{aligned} \tag{21}$$

The right-hand side of Eq. (21) represent the payment obtained from exercising the option when it is cash-settled or the net profit of exercising the option and taking a counterbalancing trade in the underlying asset otherwise. Little information about the future realizations of the underlying asset value can be inferred from options whose prices coincide with their lower bounds. For example, assume the underlying asset follows a geometric Brownian motion, as in Black-Scholes. If the theoretical price of an American call coincides with the market price of the option for a volatility σ when the market price hits its lower bound, any volatility lower than σ will also be consistent with the observed option price.

This relates to the trade-off faced by the holder of an in-the-money American call option. He can either exercise it immediately and pocket the difference between the underlying asset price and the strike price or wait for the underlying asset price to reach even higher levels. When the price of an American option reaches its lower bound, investors are indifferent between exercising and selling the option. Let's study further the case when exercising the option becomes optimal and the price of the option consequently reaches its lower bound.⁷ First, holding constant the volatility of the asset, exercising the option is more attractive the higher the realization of the asset price compared to the strike. As a result, the price of an American option is more likely to hit its lower bound when the option is far in the money.

⁷If the price of such an option, a call for example, were above its lower bound when exercising the option is the optimal strategy, traders would realize arbitrage profits by writing calls (which would be immediately exercised) and booking a profit of $C_t - (F_t - K)$ per option.

Second, holding constant the price of the underlying asset and lowering its volatility decreases the likelihood that the asset price attains even higher levels in the future. This tilts the balance toward exercising the option. Moreover, when the volatility reaches the level that makes exercising the option the best alternative, any volatility lower than this threshold will result in the same outcome. Consequently, only upper bounds on the implied volatility can be inferred from American options whose prices equal their lower bounds.

The overvaluation of the implied volatility linked with neglecting the early exercise clause of American options tends to be the highest when the option's price reaches its lower bound. Solving for the implied volatility of an American-style option while ignoring its early-exercise clause should bias the volatility estimate upwards because, all things equal, an American option is more valuable than a European option. When the early-exercise clause is ignored, the difference between these two values shows up in the implied volatility. However, as shown in Figure 2, the effect is small when the option price is strictly above its lower bound and, because the implied volatilities are only numerically-obtained approximations, they are sometimes slightly lower when the early-exercise clause is ignored.

Even if some information can be extracted from options whose price equal their lower bounds, they should not be handled in the same way as options with prices strictly above their lower bounds. We decide to exclude lower-bound option prices entirely because they typically do not correspond to any transactions (trading volume for such options is zero). Using the butterfly spread on such data would be particularly misleading. When it coincides with its lower bound, the option price is linear in the strike; the butterfly spread (which is determined by the curvature of the option price as a function of the strike) would affect a zero probability of the underlying asset value reaching the strike.

4 Application

In this section, we present the data used in the paper and the impact of weighting the market price of each option by its trading volume when extracting implied probability distributions.

4.1 Data

We use daily, publicly available, data on Eurodollar and S&P 500 futures and options maturing in March 2001 and traded on the Chicago Mercantile Exchange (CME) on February 28, 2001. End-of-day quotes have some clear disadvantages compared to intraday data. Prices of options with different strikes may correspond to transactions occurring at different times over the day. The implied-pdf procedure assumes that all prices are simultaneously determined and the lack of information about actual transaction times makes the validity of this assumption hard to assess. CME also makes available intraday trading information updated every ten to twenty minutes. However, these data yield a snapshot of trading activity at a point in time and much

work would be necessary to reconstitute an approximate history of market activity over the day. In contrast, Jackwerth and Rubinstein (1997) use transaction-by-transaction data on the S&P 500 index options traded on the Chicago Board Options Exchange (CBOE). Their rich data set also allows them to distinguish whether the trade occurred at the bid or at the ask, a distinction we cannot make with end-of-day data. (CME intraday data sometimes indicate whether the transaction price is a bid or an ask.) However, end-of-day data are much more available than high-frequency data and answer the practical need for daily implied pdfs in banking or regulatory institutions.

We exclude options with prices equal to their lower bounds. When calls and puts are both available for the same strike, we select the option with the highest trading volume. One should not use multiple options for the same strike when attempting to perfectly replicate the option prices. If the market prices of the put and the call verify the put-call parity, then the price of one can be deduced from the other and the prices of the underlying asset and the riskless bond, so that only one option contains information. If the prices of the put and the call options at the same strike fail to verify the put-call parity, no risk-neutral probability distribution can price simultaneously the underlying asset, the riskless bond and the two options.⁸ Including prices of calls and puts with the same strike is possible when attempting to only approximate the option prices, as is the case in the volume-weighted implied trees introduced above. However, including multiple same-strike options on Eurodollar futures does not have any distinguishable impact on implied pdfs, possibly because departure from the put-call parity is minimal and trading volumes of out-the-money options are typically much lower than that of in-the-money option with the same strikes.

Eurodollar futures and options are quoted according to the International Money Market (IMM) conventions and a simple transformation of the data is necessary before computing implied volatilities or implied probability distributions. A Eurodollar futures contract pays out 100 minus the spot three-month Eurodollar rate at expiry expressed in percentage points and the Eurodollar futures rate is defined as 100 minus the corresponding futures price. (Futures and options contracts used in the paper mature simultaneously, at which time the Eurodollar spot and futures rates coincide.) We need to transform options based on futures prices to options based on futures rates because the Eurodollar implied volatilities pertain to the volatility of the Eurodollar rate. A call on a Eurodollar futures contract struck at K translates into a put on the Eurodollar futures rate struck at $100 - K$.

Table 2 presents some summary statistics on the trading volume for options on the S&P 500 and the Eurodollar March futures traded on February 28, 2001.⁹ The number of options and the average trading volume per op-

⁸The put-call parity holds exactly for European options and only approximately for American options. However, our pdf-extraction technique treats all options as if they were European-style, so that, including multiple same-strike options in the data set could still result in numerical failure when the American option prices do not verify the put-call parity.

⁹This day is chosen as an example. The trading volume patterns that day are fairly

tion greatly differ across underlying assets. The S&P 500 futures support the most options (73), about ten times as many as the Eurodollar futures (7). The average trading volume, in number of contracts, is higher for options on the Eurodollar than on the S&P 500 futures (but the average price of options on Eurodollar futures is lower than that of options on S&P 500 futures). Trading activity is low or nil for many options on S&P 500 futures. About 23 per cent of the options on S&P 500 futures have zero volume, and, in total, about half of the options have trading volumes of 10 contracts or less, a trading volume more than 10 times lower than the mean. In contrast, none of the options on Eurodollar futures used in the paper has zero volume.

Figure 3 shows the cumulative distribution of trading volume and the trading volume per strike for options on S&P 500 futures and Eurodollar futures. The distance between the distribution function and the diagonal line (which joins the lowest to the highest trading volumes) measures the concentration of the trading volume because the cumulative distribution function of a uniformly distributed random variable would coincide with the diagonal line. Extreme volumes dominate the distribution of trading volume on S&P 500 futures options whereas that on Eurodollar futures options is more balanced.

Although, low volumes tend to be more frequent at strikes farther away from the heart of the distribution, low volumes can occur at any strike (as shown by the options on S&P 500 futures) and strikes in the tail of the distribution can support high trading volumes (as shown by the options on Eurodollar futures). This confirms the idea of this paper that one should take into account trading volume option by option and not merely select options according to their moneyness.

4.2 Results

The following section presents the results of extracting risk-neutral probability distributions alternatively striving to perfectly replicate the market prices of the derivatives on the underlying asset and being satisfied with only an approximate fit, a choice that allows the use of the options' trading volumes to weight their prices.

Implied pdf perfectly replicating the option prices

Figure 4 shows the implied volatilities and the implied probability distributions using the butterfly-spread method and the implied-tree method obtained from options on S&P 500 and Eurodollar futures. The high proportion of zero-volume quotes in S&P 500 futures options does not seem to affect the overall smoothness of the smile, suggesting that zero-volume quotes may be the results of interpolating the smile or the option prices.

The implied probability distributions in Figure 4 are designed to perfectly match the market prices of the options and no positivity constraint is

representative of those on other days. Statistics below pertain to data excluding options whose prices coincide with their lower bounds and options with a lower trading volume than options with identical strike and maturity.

imposed. The implied pdf obtained using the butterfly method on the S&P 500 futures options shows multiple local peaks and troughs and is sometimes negative. The implied probabilities are negative when a triplet of consecutive option prices is not convex in the strike price. Hence, the implied butterfly distribution would not perfectly replicate the market prices of the options if positivity constraints were imposed. Introducing more states of the world than derivatives assets and letting the probability distribution be close to a reference probability distribution while striving to match exactly the option prices—the principle behind the implied-tree method—does not improve the shape of the implied pdf obtained from options on S&P 500 futures. Every spike in the butterfly-spread pdf is matched by a spike in the implied-tree pdf of similar or greater magnitude (the vertical axis in the graph in the lower panel is truncated). Imposing positivity constraints, if possible, would only hide the basic fact that some triplets of option prices are not convex in the strike price.

Using the butterfly method to construct the implied probability distribution yields better results when dealing with options on Eurodollar futures. However, the implied-tree method assigns negative implied probabilities to regions where the butterfly-spread probabilities were positive. Imposing non-negativity constraints would address this problem. The advantage of the implied-tree method is that it allows extending the support of the distribution further in the tails. In the butterfly-spread framework, the probabilities of the lowest and the highest possible values of the underlying asset are qualitatively different from the other implied probabilities because these two probabilities measure the probability mass in the tails. In this regard, the pickup in the implied pdf at both ends of the Eurodollar rate range suggests that the Eurodollar rate at expiry is likely to equal values outside this range. The implied-tree method spreads the probability mass over a greater range of values of the Eurodollar rate and yields longer and thinner tails. In this region, the constraint that the implied pdf be close to a reference pdf significantly affects the implied probability distribution because of the relative paucity of option data in the tails.

Alternative to the implied-tree method

Instead of implied trees, one could use the method of Breeden and Litzenberger (1978), summarized in Eq. (9), on previously smoothed option prices. Figure 5 shows the result of fitting a smooth function through the prices of all put options on the S&P 500 futures and taking the first and the second derivatives of this function. As shown in the upper panel, the interpolation scheme delivers an apparently increasing and convex function. However, as shown in the lower panels, deviations from monotonicity and convexity that are barely visible in the top panel translate into slight bumps in the first derivative and much sharper spikes in the second derivative. Choosing a cubic spline for the option-price function like in Figure 5—an interpolating function that is cubic on the segments between two strikes and passes smoothly through the market prices of the options— results in linear second derivatives and coincides with the butterfly spread in the heart of the distri-

bution. Choosing a higher order for the polynomial interpolating function often results in even more pronounced swings in the second derivatives. Increasing the order of the interpolating polynomial may allow the imposition of smoothness conditions on the second derivative of the price function at every strike but typically causes the second derivative function to behave erratically between some strikes.

Shimko (1993) and Campa et al. (1997) fit the implied volatilities using a polynomial function of the strike price. We do not apply their methods to our data because the volatility smile obtained from options on the S&P 500 futures is even less smooth than the put price function in Figure 5. More generally, interpolation schemes that work well on data set of small size and good quality may deliver less acceptable results when applied to more numerous and lower-quality data. One may suspect that some options on S&P 500 futures with consecutive strikes correspond to transactions that occurred at different times over the trading day. A way of improving the quality of our end-of-day data would be to filter out options that most likely traded at a different time than the others. Trading volume can be used as such a filter when no information on transaction time is available.

The undesirable behavior of the pdf also seems to be due more to the constraint of replicating the price of every option than to the selected interpolation scheme. To remedy this, one could smooth the prices of the options using a spline with fewer knot points than option prices and obtain the implied pdf by taking the second derivative of this function (and scaling results by $\exp(-r\tau)$). The procedure is simple but the resulting pdf depends on the order of the polynomial used and on the choice of knot points. Those willing to forgo the perfect replication of the market prices of the options may prefer the method outlined in the paper (where a non-matching cost replaces the requirement to perfectly recover the option prices) because of the greater transparency between the parameters (the weight affected to the non-matching cost, etc.) and the resulting pdf.

Deleting from the data the observations that cause the price of the options not to be convex in the strike price, or slightly modifying the data to convexify the option prices, would guarantee the non-negativity of the implied probabilities when the butterfly spread is used and diminish the likelihood of negative probabilities when the implied tree method is chosen. However, as shown in Figure 6, there is no unique way to select the data to make the call price a convex function of the strike price because the butterfly-spread based implied probability at any point in the asset-price range depends on the price of three options. In Figure 6, the prices of the put options, represented by the dots joined by the solid line, do not define a convex function of the strikes, because a line joining points B and E would pass below D. This can be remedied by pulling down D, shifting up B or E, for example towards C and F, respectively, or deleting B or E altogether and using A or G instead in the butterfly spread. Each of these choices results in a different implied probability distribution.

Implied pdf imperfectly replicating the option prices

The effect of not requiring the implied distribution to perfectly match the market prices of the options and of weighting each option price by a function of its trading volume is shown in Figure 7. The upper panels show the implied probability distribution when only an approximate fit with the option prices is acceptable and every option price has the same weight. The constraint that the implied pdf replicate the market prices of the option is more significantly lessened for options on S&P 500 futures than for those on Eurodollar futures to insure a relatively smooth pdf for the S&P 500. The weight on the cost of non matching (λ_2 in equation (19)) is 0.5 for the S&P 500 futures options but 0.98 for Eurodollar futures options. However, weights on the non-matching cost are not directly comparable across different types of assets. For example, there are more options on the S&P 500 futures than on the Eurodollar futures, and hence relatively more probable discrepancies between the market prices of the options and those implied by the pdf. Even with the lessened matching constraint, the pdf obtained from the S&P 500 futures options is fairly erratic. Using trading volume to weight each option price smoothes some of the wiggles, especially when the level of the trading volume is used instead of its logarithm. The pdf obtained from options on Eurodollar futures is hardly affected by weighting each option by the logarithm of its trading volume but is significantly smoother when the level of the trading volume is chosen.

The two weighting schemes used affect options with zero trading volume a zero weight. Figure 8 shows the implied pdfs based on options on S&P 500 futures excluding zero-volume options while affecting the same weight to all remaining options, whatever their trading volumes. The implied pdfs are computed using the butterfly spread method, the implied-tree method imposing a perfect match with the option prices, and the implied-tree method requiring only a partial match with those prices (the weight on the non-matching cost is unchanged). Methods with perfect match results in wiggly and sometimes negative pdfs while relaxing this constraint improves the smoothness of the implied pdf. Weighting each observation by its trading volume improves further the smoothness of the pdf.

The improvement in the overall smoothness of the implied probability distribution does not come at significant costs for the fit between the implied distribution and the market prices of the derivatives. For options on the Eurodollar futures, the mean absolute deviation between the prices implied by the distribution and those observed in the market is about \$11 when a lognormal distribution is fitted to the option prices, \$4 when the implied-tree with imperfect matching but no volume-weighting is used, and \$5 when the volume-weighting scheme is used. The respective figures for options on S&P 500 futures are \$633, \$7 and \$9.¹⁰

¹⁰Options and futures used in the paper are quoted in "points" and 1 point is worth \$25 for derivatives on the Eurodollar and \$2.5 for those on the S&P 500 index. The mean prices are \$54 for options on the Eurodollar futures and \$2,257 for options on the S&P 500 futures. The sample excludes options whose prices equal their lower bounds or with zero trading volume. When both a call and a put option are traded at a given strike, the option with the highest trading volume is selected.

The weighting scheme as bandwidth

As shown in Figure 7, choosing the logarithm or, alternatively, the level of the trading volume of each option significantly impacts the implied pdf. Weighting the option prices by the logarithm of the trading volume has roughly the same result as not weighting the option prices at all for options on Eurodollar futures but has a significant effect for options on the S&P 500 futures. Figure 9 shows the weights on the options for both underlying assets. Choosing to weight the option prices by the logarithm of their trading volume limits the influence of high trading volumes. It nearly affects each option on Eurodollar futures the same weight but preserves some variation across weighted traded volumes for options on S&P 500 futures, while still significantly reducing the weight of highly traded options. Figure 9 suggests that it is advisable to graph the weight given to each option by potential weighting schemes before computing implied pdfs. As shown in Figure 10, which displays the relative weights as functions of the option's trading volume when the weights are based on the level and on the logarithm of trading volume, the weights based on the logarithm of the trading volume flatten out as soon as trading volume increases away from zero so that this weighting scheme differs little from not weighting the options. (Trading volume is assumed to be uniformly distributed across strikes, so that the cumulative probability distribution of the trading volume is the 45-degree line.) The volume-based weighting scheme plays a role somewhat similar to that of the bandwidth in a kernel regression. Favoring options with high trading volume reduces the effective number of options used to compute the implied pdf; weighting options more equally includes less relevant data.

5 Conclusion

The paper presents a new way of inferring the probability distribution of an asset from the prices of derivatives based on this asset. The approach recognizes that market prices of options and futures are inherently noisy and introduces two changes. First, instead of requiring that the implied probability distribution perfectly replicate the market prices of the derivatives, the new procedure only imposes a cost of non-matching those prices. Second, it uses trading volume as an indicator of how relevant the price data are and weight each option by a function of its trading volume.

The user selects such a function according to how much he wants to favor high-volume options against low-volume options in the weighting scheme. Comparing the implied probability distributions obtained with the new method to those obtained using the butterfly-spread or the implied-tree methods reveals significant improvements in the degree of smoothness of the distributions with a minimal deterioration of the fit between the derivatives prices obtained from the implied probability distributions and the corresponding market prices.

Although we only use the trading volume-weighting scheme to implied trees, the method can be applied to any other technique mapping market quotes on derivatives contracts to the risk-neutral probability distribution of

the underlying asset. Computational changes are minimal and trading volume data are as easily available as price data. Users of parametric methods do not even need to consider the appropriate weight on the non-matching cost to be able to use trading volume to weight option prices because parametric methods typically cannot replicate observed option prices. (In contrast, when perfect replication of the option prices is possible, one needs to loosen the matching constraint for a weighting scheme on the option prices to make any difference.) Weighting each option by its trading volume could reduce the problems linked with overfitting the data and produce implied probability distributions that are both smoother and more reflective of fundamentals.

Appendix

A Retrieving risk-neutral distributions when there are as many marketable assets as states of the world

Setting the number of states of the world equal to the number of assets and ignoring the positivity constraint on the risk-neutral probabilities, one can easily recover the (unique) risk-neutral probability distribution. From Eq. (5), it follows that

$$\begin{cases} S_i &= S_1 + \Delta(i-1) & \text{for } i = 1 \dots, n, \\ K_j &= K_1 + \Delta(j-1) & \text{for } j = 1 \dots, m, \\ S_i - K_j &= \Delta(i-1-j). \end{cases} \quad (22)$$

Consequently, Eq. (3) is equivalent to

$$\begin{cases} \sum_{i=1}^{m+2} (i-1) p_i &= (F - S_1)/\Delta, \\ \sum_{i=1}^{m+2} \max(0, i-1-j) p_i &= e^{r\tau} C(K_j)/\Delta, \\ \sum_{i=1}^{m+2} p_i &= 1. \end{cases} \quad (23)$$

Eq. (23) is equivalent to

$$A \cdot P = Z \quad (24)$$

with the components of A and Z given by

$$\begin{cases} a_{j,i} &= \max(0, i-j) & \text{for } j = 1, \dots, m+1, i = 1, \dots, m+2, \\ a_{m+2,i} &= 1 & \text{for } i = 1, \dots, m+2, \\ z_1 &= (F - S_1)/\Delta, \\ z_j &= e^{r\tau} C(K_{j-1})/\Delta & \text{for } j = 2, \dots, m+1, \\ z_{m+2} &= 1. \end{cases} \quad (25)$$

$A^{-1} = B$ with the components of B equal to 0 except

$$\begin{cases} b_{1,1} = -1, & b_{1,2} = 1, & b_{1,m+2} = 1, \\ b_{i,i-1} = 1, & b_{i,i} = -2, & b_{i,i+1} = 1 & \text{for } i = 2, \dots, m, \\ b_{m+1,m} = 1, & b_{m+1,m+1} = -2, \\ b_{m+2,m+1} = 1. \end{cases} \quad (26)$$

To prove that $B \cdot A$ is the identity matrix, let $c_{i,j} = \sum_{k=1}^{m+2} b_{i,k} a_{k,j}$ and show that $c_{i,j} = 1$ if $i = j$ and 0 otherwise.

1. Case $2 \leq i \leq m$:

For $2 \leq i \leq m$, $b_{i,m+2} = 1$ if $i = 1$ and $b_{i,m+2} = 0$ if $i \geq 2$ and for $k \leq m+1$, $a_{k,j} = \max(0, j-k)$. Hence,

$$\begin{aligned} c_{i,j} &= \sum_{k=1}^{m+1} b_{i,k} a_{k,j}, \\ &= \max(0, j-i+1) - 2 \max(0, j-i) + \max(0, j-i-1). \end{aligned} \quad (27)$$

Eq. (27) implies that, when $2 \leq i \leq m$, $c_{i,j} = 1$ if $i = j$ and 0 otherwise.

2. Case $i = 1, m + 1, m + 2$:

$$\begin{cases} c_{1,j} &= -\max(0, j - 1) + \max(0, j - 2) + 1, \\ c_{m+1,j} &= \max(0, j - m) - 2\max(0, j - m - 1), \\ c_{m+2,j} &= \max(0, j - m - 1). \end{cases} \quad (28)$$

Eq. (28) shows that, when $i = 1, m + 1, m + 2$, $c_{i,j} = 1$ if $i = j$ and 0 otherwise.

Likewise, $A \cdot B$ is the identity matrix.

B More states than marketable assets

Assume that the matrix M in Eq. (4) is full-rank and let S be the $n \times n$ selection matrix such that the first $m + 2$ columns of $\tilde{M} = MS'$ are linearly independent. Eq. (3) is equivalent to $\tilde{M}\tilde{P} = Q$ with $\tilde{P} = SP$. Let $\tilde{M} = [\tilde{M}_1 \tilde{M}_2]$ and $\tilde{P} = [\tilde{P}_1 \tilde{P}_2]$. Then,

$$\tilde{P}_1 = (\tilde{M}_1)^{-1} [Q - \tilde{M}_2\tilde{P}_2].$$

Let S_1 and S_2 be the selection matrices so that $\tilde{M}_1 = \tilde{M}S_1$ and $\tilde{M}_2 = \tilde{M}S_2$. Using the inverse of the selection matrix S , we write $P = S'\tilde{P}$ and developing terms,

$$\tilde{P} = AU + B,$$

where

$$U = \tilde{P}_2, \quad A = \begin{bmatrix} -[MS'S_1]^{-1}[MS'S_2] \\ I_{n-m-2} \end{bmatrix}, \quad B = \begin{bmatrix} [MS'S_1]^{-1}Q \\ O_{n-m-2} \end{bmatrix},$$

where I_{n-m-2} is the $(n - m - 2) \times (n - m - 2)$ identity matrix, and O_{n-m-2} is a $n - m - 2$ dimensional vector of zeroes.

Let $P_0 = (p_i^0)_{i=1}^n$, then, Eq. (10) is equivalent to

$$\min(P - P_0)'(P - P_0). \quad (29)$$

Hence, combining this equation with the constraints, we get

$$\Omega(U) = (AU + B - P_0)'(AU + B - P_0).$$

The solution to this problem, assuming the relevant matrix is non-singular is

$$U^* = (A'A)^{-1}A'(P_0 - B). \quad (30)$$

Finally, the resulting probability vector is

$$P^* = AU^* + B.$$

References

- Bank for International Settlements, 2000. Press release on global derivatives market activity. Basle, Switzerland.
- Barone-Adesi, G. and Whaley, R., 1987. Efficient analytic approximation of American option values. *Journal of Finance*, 42-2, 301-320.
- Bates, D., 1991. The crash of 87: Was it expected? The evidence from options markets. *Journal of Finance*, 46, 1009-1044.
- Bergman Y., Grundy, B., and Wiener, Z., 1996. General properties of option prices. *Journal of Finance*, 51-5, 1573-1610.
- Black, F., and Scholes, M., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 637-659.
- Blume, L., Easley, D., and O'Hara, M., 1994. Market statistics and technical analysis: the role of volume. *Journal of Finance*, 49-1.
- Breedon, D., and Litzenberger, R., 1978. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51-4, 621-651.
- Buchen, P., and Kelly, M., 1996. The maximum entropy distribution of an asset inferred from option prices. *Journal of Financial and Quantitative Analysis*, 31-1, 143-159.
- Campa, J., Chang, P., and Reider, R., 1997. ERM bandwidths for EMU and after: evidence from foreign exchange options. *Economic Policy*, 24, 55-89.
- Campa, J., Chang, P., and Reider, R., 1998. Implied exchange rate distributions: evidence from OTC option markets. *Journal of International Money and Finance*, 17, 117-160.
- Clews, R., Panigirtzoglou, N., and Proudman, J., 2000. Recent developments in extracting information from option markets. *Bank of England Quarterly Bulletin*, 50-60.
- Dupont, D., 1997. Trading volume and information distribution in a market-clearing framework. FEDS 1997-41, Board of Governors of the Federal Reserve System.
- Gallant, R., Rossi, P., and Tauchen, G., 1992. Stock prices and volume. *Review of Financial Studies*, 5-2, 199-242.
- Jackwerth, J. and Rubinstein, M., 1996. Recovering probability distributions from option prices. *Journal of Finance*, 51-5, 1611-1632.

Jain, P. and Joh, G., 1988. The dependence between hourly prices and trading volume. *Journal of Financial and Quantitative Analysis*, 23-3, 269-283.

Huffman, G., 1992. Information, asset prices, and the volume of trade. *Journal of Finance*, 47-4, 1575-1590.

International Monetary Fund, 1998. International capital markets. IMF, Washington, DC.

Karpoff, J., 1987. The relation between price changes and trading volume: a survey. *Journal of financial and quantitative analysis*, 22-1, 109-126.

Melick, W. and Thomas, C., 1997. Recovering an asset's implied pdf from option prices: an application to crude oil during the Gulf war. *Journal of Financial and Quantitative Analysis*, 32-1, 91-115.

Melick, W. and Thomas, C., 1998. Confidence intervals and constant-maturity series for probability measures extracted from option prices. *Proceedings of the conference on information in financial asset prices*. Bank of Canada, 293-330.

Pirkner, C., Weigend, A., and Zimmermann, H., 1999. Extracting risk-neutral densities from option prices using mixture binomial trees. *IEEE/IAFE Conference on Computational Intelligence for Financial Engineering*.

Rubinstein, M., 1994. Implied binomial trees. *Journal of Finance*, 69-3, 771-818.

Shalen, C., 1993. Volume, volatility, and the dispersion of beliefs. *Review of Financial Studies*, 6-2, 405-434.

Shimko, D., 1993. Bounds on probability. *Risk*, 6, 33-37.

Söderlind, P. and Svensson, L., 1997. New techniques to extract market expectations from financial instruments. *Journal of Monetary Economics*, 40, 383-429.

Stutzer, M., 1996. A simple nonparametric approach to derivative security valuation. *Journal of Finance*, 51-5, 1633-1652.

	Underlying Asset	Event	Period
1	Crude oil	Persian Gulf crisis	1990-91
2	10-year Bunds	M3 growth release	1994
3	US Dollar/Canadian Dollar	Quebec sovereignty referendum	1995
4	Deutsche Mark/French Franc	French Budget	1996
5	3-month eurodollar	Employment data release	1997
6	3-month euromark	Bundesbank comments	1997

Table 1: **Event studies using implied distributions.** The implied pdfs are based on options on the assets mentioned above or on futures on these assets. Sources: 1. Melick and Thomas (1997) ; 2. Söderling and Svensson (1997); 3. *ibid.* p. 405; 4. Campa et al. (1997); 5. and 6. Melick and Thomas (1998)

variable	Eurodollar	S&P 500	S&P 500 <i>excludes obs. with zero trading volume</i>
number of observations	7	73	56
minimum	100	0	1
maximum	17,000	1,112	1,112
mean	6,962	114	149
median	6,350	10	30

Table 2: **Trading volume of options on Eurodollar and S&P 500 futures traded on February 28, 2001 and maturing in March 2001.** The data exclude the options whose prices coincide with their lower bounds. When both a put and a call option are traded at a given strike, the option with the highest trading volume is selected.

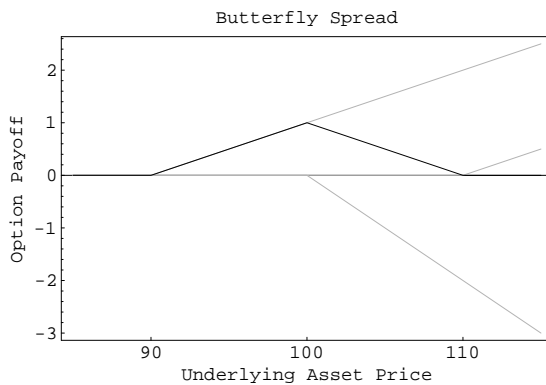


Figure 1: **Butterfly spread.** If the asset price, S , takes values 90, 100, 110 on the interval $[90,110]$, the portfolio constituted by a long position in $1/10$ calls struck at 90 and 110, and two short positions in $1/10$ calls struck at 100 pays out \$1 if $S = 100$ and nothing otherwise.

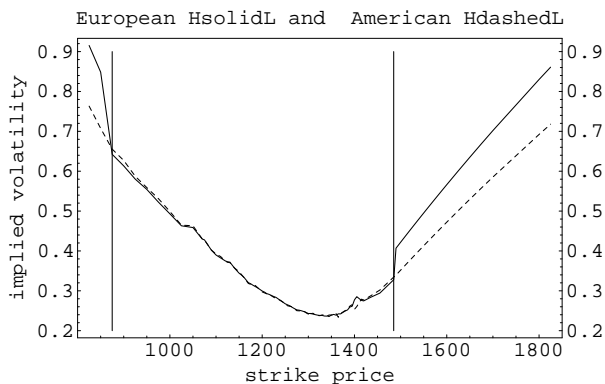


Figure 2: **Implied volatility smiles obtained from options on S&P 500 futures traded on February 28, 2001 and maturing in March 2001.** Implied volatilities are based on Black-Scholes (solid line) and on numerical schemes taking into account the option's early-exercise clause (dashed line). Outside the two vertical lines the market prices of options coincide with their lower boundaries; inside the two vertical lines options whose market prices coincide with the lower boundaries are excluded and, when both a call and a put are available at the same strike, the option with the highest trading volume on the day is selected.

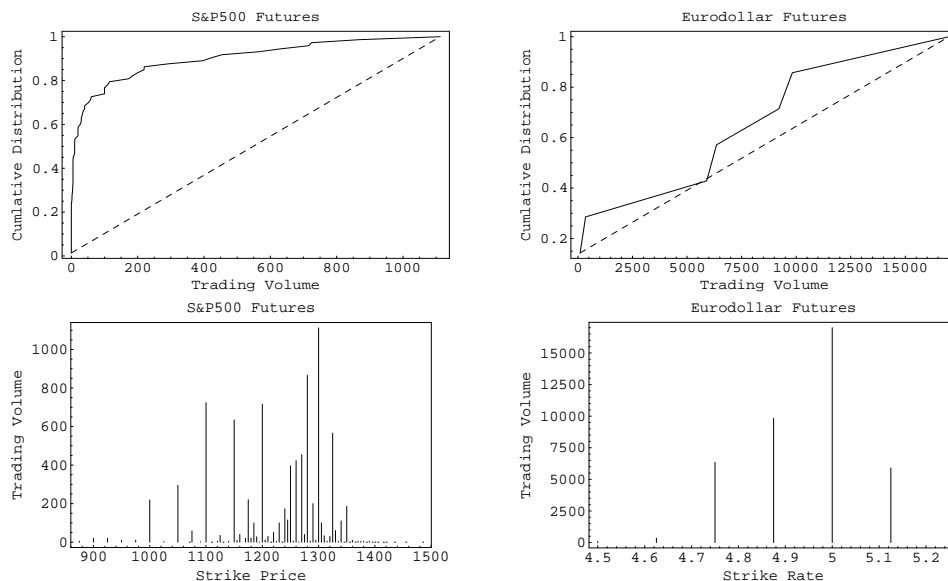


Figure 3: Cumulative distribution of trading volume (upper panels) and trading volume per strike (lower panels) of options on S&P 500 futures (left panels) and Eurodollar futures (right panels) traded on February 28, 2001 and maturing in March 2001. Zero-volume options are included in the data set. When both a call and a put are available at the same strike, the option with the highest trading volume on the day is selected.

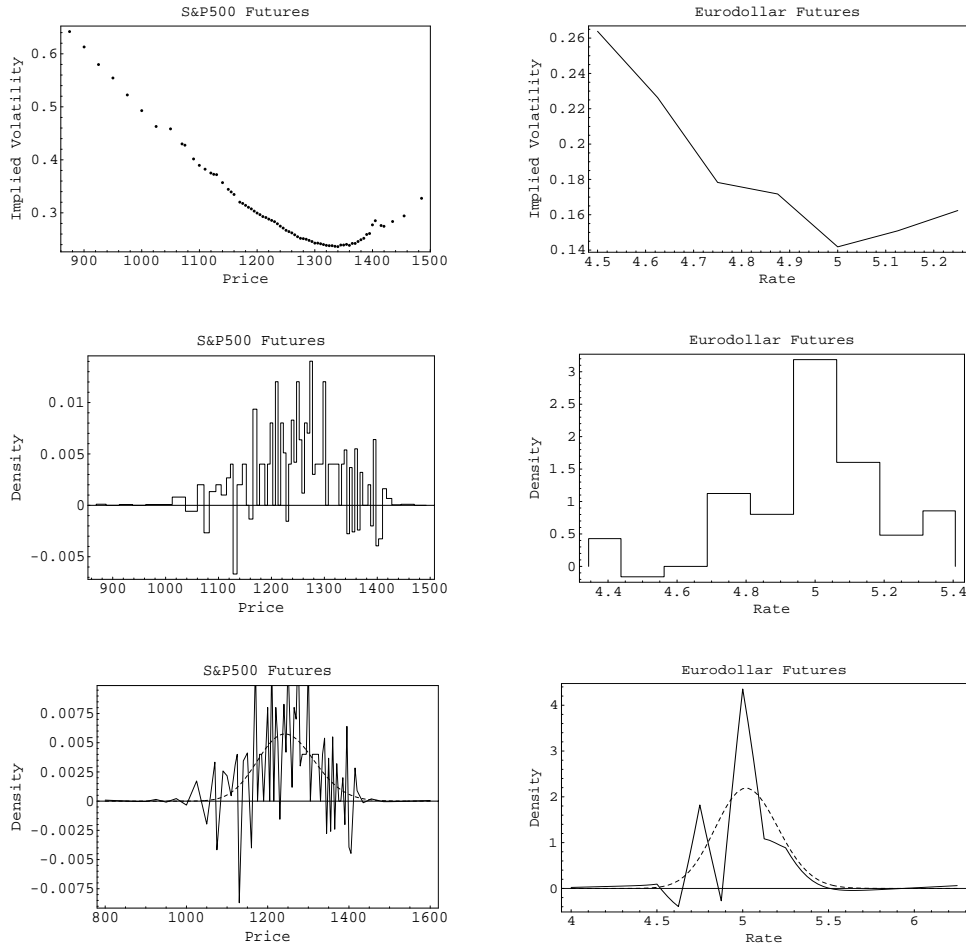


Figure 4: Implied volatilities (upper panels), implied probability distributions using the butterfly-spread method (middle panels) and the implied-tree method (lower panels) obtained from options on S&P 500 futures (left panels) and Eurodollar futures (right panels) traded on February 28, 2001 and maturing in March 2001. The dashed lines in the lower panels show the lognormal pdfs that best match the market prices of the options. The option prices computed from the implied distributions perfectly fit market prices. No positivity constraint is imposed on the implied probabilities. Options whose market prices coincide with the lower boundaries are excluded. The trading volume of some options may be zero. When both a call and a put are available at the same strike, the option with the highest trading volume on the day is selected. The futures rate on the Eurodollar was 5.0075 percent and the futures price on the S&P 500 was 1242.

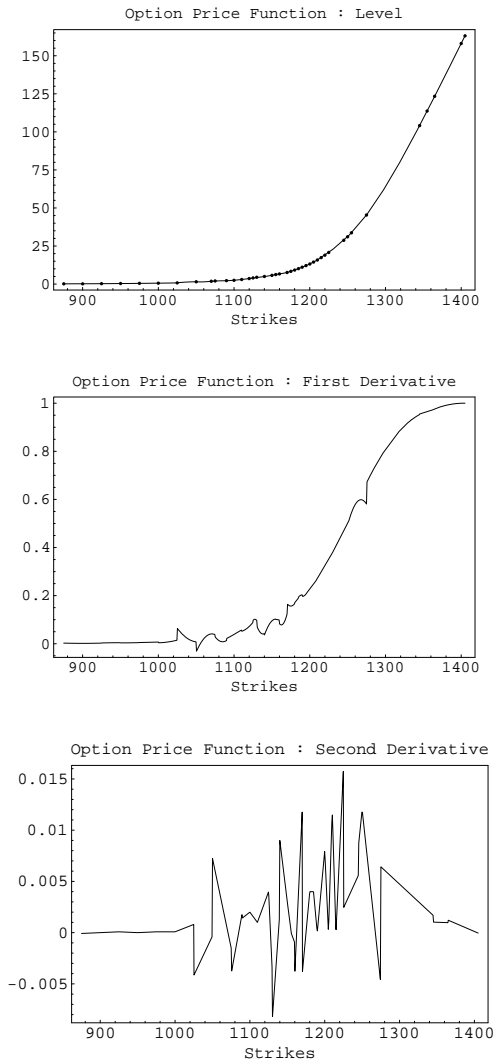


Figure 5: **Smooth interpolation of the prices of the puts on S&P 500 futures (the upper panel) and its first and second derivatives (the lower panels).** The interpolating function is a cubic spline that passes smoothly through all the option prices. Nearly imperceptible deviations from convexity by the interpolating function translates into large spikes in its second derivative.

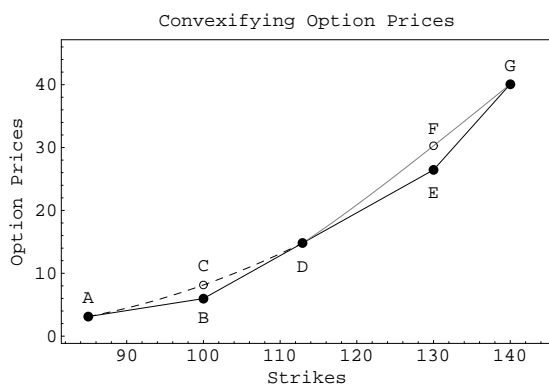


Figure 6: **Convexifying non-convex option prices.** Put prices, represented by the solid circles, do not define a convex function of the strikes because D lies above the line joining B and E. There exist multiple ways of convexifying the option price function, each yielding a different implied probability distribution.

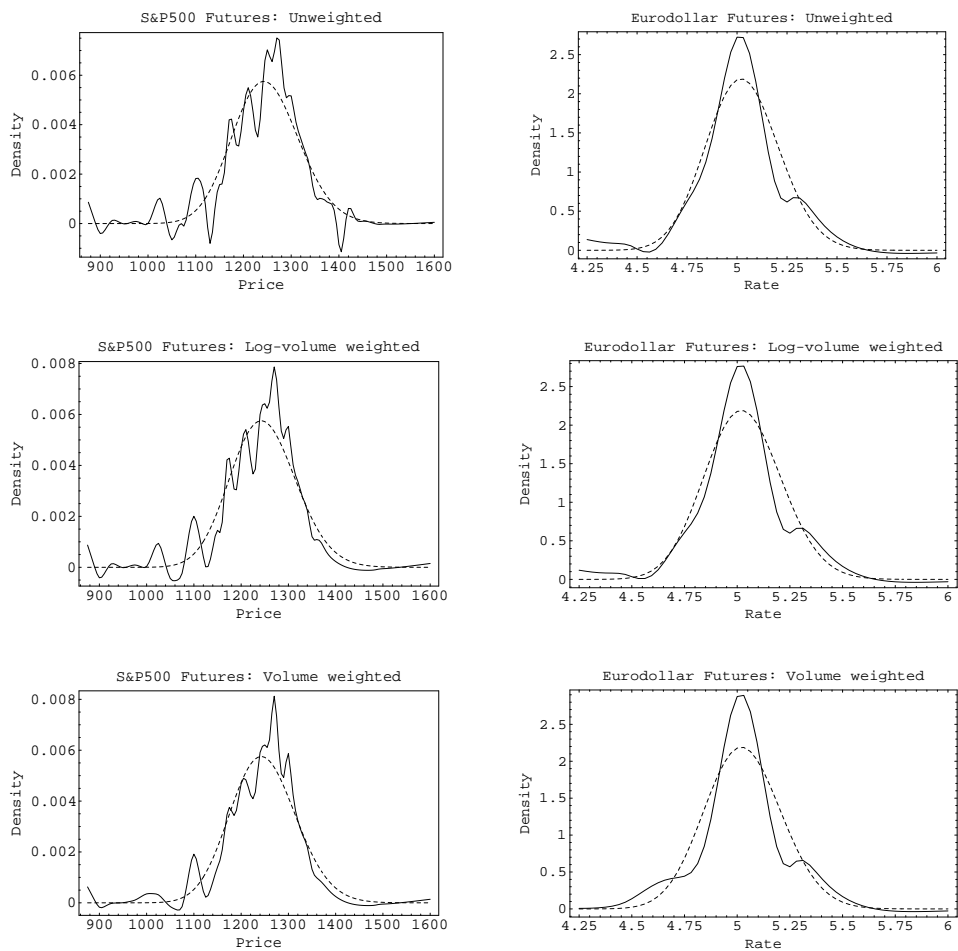


Figure 7: Implied probability distributions using variants of the implied-tree method obtained from options on S&P 500 futures (left panels) and Eurodollar futures (right panels) traded on February 28, 2001 and maturing in March 2001. Option prices computed based on the implied distributions match market prices only partially. The implied distributions minimize a loss function penalizing the following three features: deviations between the market prices of the options and the prices based on the implied distributions; deviations between the implied probabilities and a reference probability function; and roughness of the implied probability density. The weighting of the options in the loss function is as follows: options are unweighted in the upper panels, they are weighted by the logarithm of their trading volumes in the middle panels, and by their trading volumes in the lower panels. Except in the upper panels, options with zero volume have zero weight in the loss function.

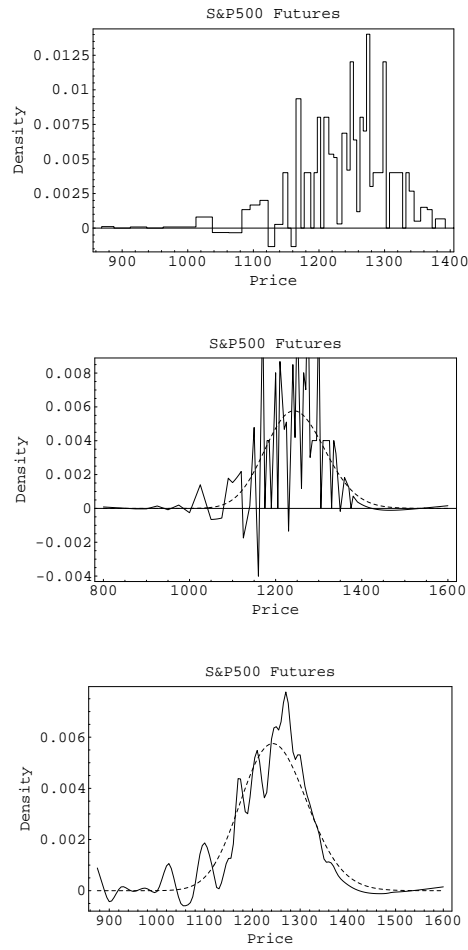


Figure 8: Implied probability distributions obtained from options on S&P 500 futures traded on February 28, 2001 and maturing in March 2001 excluding zero-volume options. The upper panel uses the butterfly-spread method; the other panels use variants of the implied-tree method. The method used in the middle panel imposes a perfect fit between the market prices of the options and those based on the implied probability distribution; the method used in the lower panel only imposes a cost of not matching the market prices of the options. Every option is given the same weight.

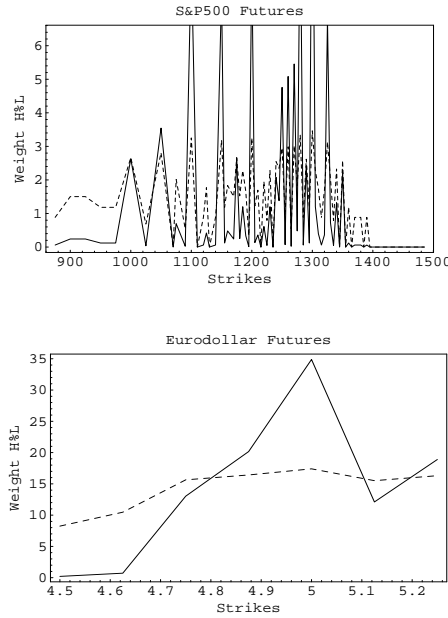


Figure 9: Weights on each option on S&P 500 futures (upper panel) and on Eurodollar futures (lower panel) traded on February 28, 2001 and maturing in March 2001 based on the option's trading volume (solid line) and on the logarithm of its trading volume (dashed line).

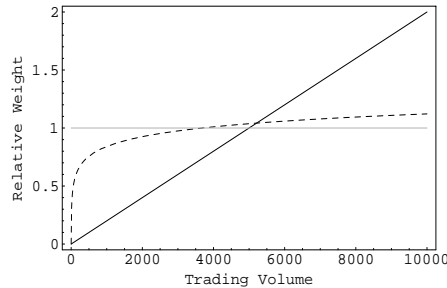


Figure 10: Relative weights as functions of the option's trading volume when the weights are based on the option's trading volume (solid line), on the logarithm of its trading volume (dashed line), and when every option is given the same weight (light-gray line). Trading volume is assumed to be uniformly distributed across strikes (so that the cumulative probability distribution of the trading volume is the 45-degree line). When the upper bound on the trading volume is high, the weights based on the logarithm of the trading volume flatten out as soon as trading volume increases away from zero so that this weighting scheme yields differs little from not weighting the options. The relative weight of an option is its weight multiplied by the number of options.

Author: Dominique Y. Dupont

Title: Extracting Risk-Neutral Probability Distributions from Option Prices Using Trading Volume as a Filter

Reihe Ökonomie / Economics Series 104

Editor: Robert M. Kunst (Econometrics)

Associate Editors: Walter Fisher (Macroeconomics), Klaus Ritzberger (Microeconomics)

ISSN: 1605-7996

© 2001 by the Department of Economics and Finance, Institute for Advanced Studies (IHS),
Stumpergasse 56, A-1060 Vienna • ☎ +43 1 59991-0 • Fax +43 1 5970635 • <http://www.ihs.ac.at>
