

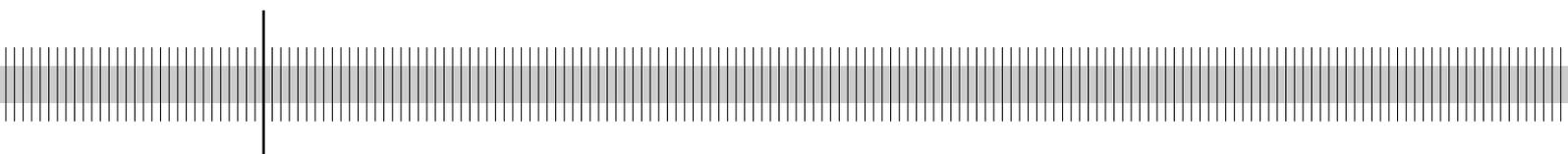
The forecast ability of risk-neutral densities of foreign exchange

Ben Craig

(Federal Reserve Bank of Cleveland)

Joachim Keller

(Deutsche Bundesbank)



Discussion Paper
Series 2: Banking and Financial Studies
No 05/2005

Editorial Board:

Heinz Herrmann
Thilo Liebig
Karl-Heinz Tödter

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-1

Telex within Germany 41227, telex from abroad 414431, fax +49 69 5601071

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Reproduction permitted only if source is stated.

ISBN 3-86558-066-1

Abstract

We estimate the process underlying the pricing of American options by using higher-order lattices combined with a multigrid method. This paper also tests whether the risk-neutral densities given from American options provide a good forecasting tool. We use a nonparametric test of the densities that is based on the inverse probability functions and is modified to account for correlation across time between our random variables, which are uniform under the null hypothesis. We find that the densities based on the American-option markets for foreign exchange do quite well for the forecasting period over which the options are thickly traded. Further, simple models that fit the densities do about as well as more sophisticated models.

KEYWORDS: Risk-neutral densities from option prices, American exchange rate options, Evaluating Density Forecasts, Pentionominal tree, Density evaluation, Overlapping data problem

JEL-CLASSIFICATION: F47, F31, C52, C63

Non technical summary

This paper introduces a new method which allows us to derive state price densities from financial market prices. The method is applied to the prices of US dollar and German D-Mark futures. As a second step, we formally test the forecasting ability of our densities.

One of the first approaches to extracting the information content of option prices was suggested by Longstaff in 1990. However, his method turned out to produce numerically unstable results and even provided negative probabilities in cases where only a few observations were available. Given the limited success of his method, researchers started to develop different types of methods which - broadly speaking - may be classified as parametric or non-parametric methods. Parametric methods are based on known distributions which are mixed or expanded or include more than the two parameters of the (log)normal distribution. Non-parametric methods do not require a parametric form but allow for more general functions. These include kernel estimations, maximum-entropy and curve-fitting methods.

A somewhat different approach is taken by methods that aim at modelling a stochastic process of the underlying asset. Since different processes imply different densities, the underlying process that is consistent with the set of observed market prices of the derivatives is chosen. This is the approach we adopt in this paper. We calculate the instantaneous volatility of the spot as a function of the state of the exchange rate and time to expiration of the contract. A great advantage of obtaining the estimated diffusion process is that this process can be simulated forward to any forecast horizon of interest and that we therefore do not face any time-to-maturity boundary.

Our data comprise over two million transaction prices from the Chicago Mercantile Exchange (CME) for 15 years of US options based on the prices of US dollar and German D-Mark futures. However, the technique we describe could just as well be used with other types of data. We solve the pricing of US options by using higher-order lattices combined with smoothing at the boundaries in order to mitigate the non-differentiability of both the payoff boundary at expiration and the early exercise boundary. The advantage of using higher-order lattices instead of binomial trees is that we need less time steps to be very accurate and that we therefore require less computational time.

Once the process is recovered, we test how well the estimated process performs as a forecasting density over a variety of forecasting horizons. We use a nonparametric

test of the densities that is based on the inverse probability functions and is modified to account for correlation across time between our random variables, which are uniform under the null hypothesis.

We find that the densities based on the US option markets for foreign exchange do quite well for the forecasting period over which the options are thickly traded. Furthermore, simple models that fit the densities do about as well as more sophisticated models.

Nicht technische Zusammenfassung

Im vorliegenden Papier wird ein neues Verfahren vorgestellt, das es erlaubt, Dichtefunktionen von Finanzmarktpreisen abzuleiten. Das Verfahren wird auf amerikanische US-Dollar D-Mark Terminkontrakte angewendet. In einem zweiten Schritt wird die Güte der Dichtefunktionen getestet.

Der erste Ansatz zur Ableitung risikoneutraler Dichtefunktionen wurde zu Beginn der 90iger Jahre von Longstaff vorgeschlagen. Es stellte sich allerdings heraus, dass die dabei implizierten Werte für die Wahrscheinlichkeiten numerisch nicht stabil und sogar teilweise negativ waren. In der Folge wurde eine Vielzahl von präziseren Methoden entwickelt, die grundsätzlich in parametrische und nicht-parametrische Verfahren unterteilt werden. Parametrische Methoden fußen auf bekannten Verteilungen, wie zum Beispiel der logarithmischen Normalverteilung, die über Korrekturterme erweitert, über zusätzliche Parameter flexibler gestaltet oder über Kombinationsvariablen gemischt werden. Nicht-parametrische Verfahren sind frei von Annahmen über die Verteilung und basieren auf allgemeineren Funktionen. Bekannt geworden sind so genannte Kernschätzungen, Maximum-Entropie Methoden und eine Gruppe von Verfahren, die unter dem Begriff "Kurvenanpassungen" zusammengefasst werden, und die auf der Abgreifung impliziter Volatilitätsstrukturen beruhen.

Einen anderen Weg wird mit Verfahren besprochen, die darauf abzielen, den stochastischen Diffusionsprozess des Basiswertes aufzudecken. Da jeder Prozess eine bestimmte Preiszustandsdichte bedingt, wird letztendlich derjenige Prozess ausgewählt, der Optionsprämien impliziert, die möglichst nahe an den marktmäßigen Preisen liegen. Dies ist auch die Vorgehensweise im vorliegenden Forschungspapier. Die jeweilig zu einem Zeitpunkt gegebene Volatilität des Basiswerts ist hier eine Funktion des Niveaus des Wechselkurses und der Restlaufzeit des Kontrakts. Ein großer Vorteil unseres Vorgehens, den Diffusionsprozess zu schätzen, besteht darin, dass dieser Prozess für jeden beliebigen Prognosehorizont simuliert werden kann und dass wir somit keiner Restlaufzeitbeschränkung unterliegen.

Unsere Daten umfassen über zwei Millionen Transaktionspreise der Chicago Mercantile Exchange (CME) für amerikanische Optionen aus fünfzehn Jahren auf Basis der Preise von US-Dollar/D-Mark-Terminkontrakten. Wir lösen das Problem der Bewertung amerikanischer Optionen, indem wir Zustandsbäume höherer Ordnung verwenden, und mildern das Problem der Nicht-Differenzierbarkeit des Wertkurvenverlaufs amerikanischer Optionen dadurch, dass wir an nicht differenzierbaren

Stellen glätten, also zum Fälligkeitszeitpunkt der Optionen bzw. bei Wertentwicklungen des Basiswerts, bei denen sich möglicherweise das Ausüben der Optionen vor Fälligkeit lohnt. Die Verwendung von Zustandsbäumen höherer Ordnung anstelle von binomialen Bäumen hat den Vorteil, dass wir große Preisgenauigkeit mit weniger Zeitschritten erreichen und damit die Berechnungszeit erheblich verkürzen.

Nach der Schätzung der Parameter für den Diffusionsprozess prüfen wir, wie gut sich die aus amerikanischen Optionen abgeleiteten Dichten als Prognosewerkzeug eignen. Dies geschieht mittels eines nichtparametrischen Verteilungstests, der von inversen Wahrscheinlichkeiten abhängt. Wir modifizieren die Tests, um der Korrelation unserer Zufallsvariablen im Zeitverlauf, die sich aus der Überlappung der Kontraktlaufzeit ergibt, Rechnung zu tragen. Unter der Nullhypothese, dass die geschätzten Dichten richtig sind, sind die inversen Wahrscheinlichkeiten gleichverteilt.

Es zeigt sich, dass die auf der Basis amerikanischer Devisenoptionspreise abgeleiteten Dichten für diejenigen Prognosehorizonte recht aussagekräftig sind, in denen der Optionspreismarkt liquide ist. Ferner liefern einfache Modelle für den Diffusionsprozess des Basiswertes nahezu gleich gute Ergebnisse in Bezug auf die risikoneutrale Dichte wie anspruchsvollere Modelle.

Contents

1	Introduction	1
2	The Data	3
3	Estimation of the Processes	4
4	A Higher-Order Lattice	5
5	Evaluating Density Forecasts	13
6	The Results	15
7	Robustness Checks	21
8	Conclusions	25
9	Appendix: Calculation of the Variance for a Pearson Test	27
10	References	28

List of Figures

Higher order lattice	7
Properties of alpha	9
Construction of the lattice	11
State/sigma diagram	17
Five percent quantiles	17
p-values	18
Non-rejection time periods	19
Time periods where complex models outperform lognormal	20
Confidence intervals over $U(0,1)$	25

List of Tables

Stationary bootstrap on v. Mises (American options)	16
Stationary bootstrap on v. Mises (European calculations on American options)	21
American option: no smoothing	21
Bootstrap with exact Black Scholes and with estimation error	22
Stationary bootstrap on von Mises with fixed grid	24

The Forecast Ability of Risk-Neutral Densities of Foreign Exchange*

1 Introduction

It is common to price a new derivative by assuming a continuous process model and computing the value of the derivative given the model. This paper follows the different strategy of Bates (1991) in that it recovers an underlying process that is consistent with the set of observed market prices of the derivatives. Once the process is recovered, we test how well the estimated process performs as a forecasting density over a variety of forecasting horizons. The market derivative studied in this paper is the American-style option on a futures contract written on the U.S. dollar-German mark FX rate. However, the technique we describe could just as well be used with other types of data, including yield curve data for interest rate processes, prices for swaps, barrier options, and so forth, as long as the price of the derivative depends only on a single-state variable. However, the price does not have to be a linear function of the state, and the implied density for the underlying may deviate from a simple lognormal density or it may have no specific parametric form.

It is well known that with complete markets, a sufficiently rich set of European options prices implies a state-price density that one may interpret as a probability density over the price that underlies the derivative contract, if agents are risk neutral. European options have been used to recover the risk-neutral densities for a variety of prices and indices, including oil and the Standard and Poor's 500 Index, using a technique first pointed out by Breeden and Litzenberger (1978). Second differentiation of the call price with respect to the strike price, K , gives the risk-neutral density, $\pi_T(X)$, times a discount factor, $e^{-\rho(T-t)}$. The subsequent literature (e.g. Shimko (1993), Malz (1997), Jackwerth and Rubinstein (1996) and Stutzer (1996)) has concentrated on estimation of the density from noisy or, in the Malz case, extrapolated data on prices by using parametric distributions, mixtures of

*Research Centre, Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, D-60431 Frankfurt/M., e-mail: Ben.R.Craig@clev.frb.org and Joachim.Keller@bundesbank.de. Corresponding author: Ben Craig: Research Department, Federal Reserve Bank of Cleveland, PO Box 6387, Cleveland, Ohio, USA, 44101-1387. Phone: 216-579-2061. Fax: 216-579-3050. We thank Will Melick for helpful comments. The views expressed in this paper represent only the authors' personal opinion and do not necessarily reflect the views of the Deutsche Bundesbank or the Federal Reserve Bank of Cleveland.

parametric distributions, or nonparametric smoothers to fit the second derivative of the option price function with respect to the strike price. Others, like Neuhaus (1995) do not rely on smoothing equations and calculate probabilities at and between strike prices. Once the risk-neutral density is calculated, it can be used to forecast the price of the underlying basis for the option, or it may be used to price other derivatives based on the same sequence.

The method of calculating the risk-neutral density is to first estimate the stochastic process followed by the underlying (here, a futures contract for foreign exchange), based on the set of traded prices of the derivative (here, the American puts and calls reported for the end of the trading day, though our results do not change by using settlement quotes). We solve the pricing of American options by using higher-order lattices combined with smoothing at the boundaries in order to mitigate the non-differentiability of both the payoff boundary at expiration and the early exercise boundary. This paper extends our earlier work (Craig and Keller, 2001) to include a flexible grid scheme that improves the accuracy of the higher-order lattices considerably. By calculating the price of an American option quickly, we can estimate the diffusion process by minimizing the sum of the squares between the calculated prices and the observed prices in the data. A great advantage of obtaining the estimated diffusion process is that this process can be simulated forward to any forecast horizon of interest. We do not face a time-to-maturity boundary, as we would, for example, if we had relied on a Breeden-Litzenberger type procedure and recovered the density only at the terminal date. It is straightforward to generalize our procedure to include other forms of derivatives, risk-aversion parameters, or schemes of uncertainty different than the single-factor models explored in this paper.

This paper also tests whether the densities provided from American options provide a good forecasting tool. We use a nonparametric test of the densities which depends on the inverse probability ideas of Fischer (1930) and others. A problem with the use of these tests in the past has been the time-series nature of the transformed variables when the forecasting windows overlap. The inverse probability of, say, a realized thirty-day-ahead spot at time t is correlated with the corresponding number at time $t - 1$, because the spot shares twenty-nine days of history. We modify the tests based on the inverse probability functions to account for this correlation between our random variables, which are uniform under the null hypothesis. The theory that accounts for time-series dependence in specification tests that are formed from general empirical distribution functions over inverse probabilities is not well developed. We use a bootstrapping procedure, which we test in a separate paper (Craig and Keller, 2003), with more recent tests, which we designed and which are more powerful for short forecasting horizons. We find that the loss of power in

the bootstrapping tests is very small, and that this technique can be employed with confidence.

We find that the densities based on the American-option markets for foreign exchange do quite well for the forecasting period over which the options are thickly traded. Further, simple models that fit the densities do as well as more sophisticated models. This was surprising to us. A simple risk-neutral model does so well in matching the forecast densities that one wonders what the extra sophistication (which does fit the data slightly better) explains.

In the following section we describe our data. We lay out the numerical methods used to calculate the risk-neutral densities implied by American-option prices that are based on a futures contract. Next, we describe the tests we used to evaluate our implied densities, especially those that take into account the time-series nature of the overlapping windows of the forecasts. Our results are detailed in the final section.

2 The Data

American options are exchange-traded, approach a fixed expiration date, and can be exercised before maturity. Our data comprise over two million transaction prices from the Chicago Mercantile Exchange (CME) for fifteen years of options based on the prices of U.S. dollar-German mark futures. The option prices are close-of-day transactions, and they always represent prices that have been used in a trade on that day. While these data are advantageous in that they represent the most liquid exchange-traded market for foreign exchange options and they include more different strike prices each day than all other data sources combined, they have a major disadvantage: the CME options are American-style options based on an underlying futures price, so there is an incentive to exercise the CME option early. One can think of the underlying futures contract as providing a continuous stream of “dividends” as the price of the future to reflect the known expected change of the foreign exchange. As is well known, an American-style option on an underlying stock that provides a continuous stream of dividends does not always provide incentive to hold the option until its expiration date. For some values of the underlying price, a trader can do better by cashing in the option early. This provides a “boundary” of prices, under or over which (depending on whether the option is a call or a put) the trader always exercises the option before the expiration date. This early-exercise boundary is something that we take into account when we calculate our risk-neutral densities.

In addition, some of the data are especially noisy. As a result, we excluded some data points by imposing requirements that all our data had to meet. All options included in the data set had to have both volume and open interest that were positive on the trading day. In addition, because of the historical illiquidity in certain markets, other prices were excluded: options expiring within 10 days of the current trading date, options expiring more than 100 days from the current trading date, and options with strike prices that were greater than 0.05 in relative time-normalized moneyness, following Dumas, Fleming, and Whalley (1998). In other words, options are excluded if $\left| \frac{X_t - K}{K\sqrt{T-t}} \right| > .05$, with K being the strike price, X_t the actual futures rate and $\sqrt{T-t} = \sqrt{\tau}$ the normalizing time factor, which is the difference between the expiration date T and the actual date t . This excludes those options in the extreme tails, where prices are known to be driven more by illiquidity than by market expectations. The time period under investigation runs from January 25, 1984, to December 31, 1998. Days with traded options that did not include at least eight different strike prices were excluded. This left us with 3,728 separate trading days with which to estimate densities. The number of different options sold on the days where densities were estimated ran from a low of 8 to a high of 106. An average day included about 58 usable options prices. Note that all option prices that matched the above filters were used, even those that occasionally did not meet the arbitrage conditions implied by option theory, such as monotonicity with respect to strike price. (In our two million data points, this happened about 20 times.) In the case of our estimation, these anomalies were considered part of the error term in the nonlinear least squares technique.

3 Estimation of the Processes

Following Dumas et al. (1998) and our earlier work, we estimate the parameters of a diffusion process in order to approximate the risk-neutral density for each day. We first calculate the instantaneous volatility of the spot, $\hat{\sigma}_t(X, \tau, \hat{\beta})$, as a function of the state of the exchange rate and time to expiration, τ , of the contract. We estimate the diffusion function, $\hat{\sigma}_t(X, \tau, \hat{\beta})$, parametrically, by minimizing with respect to a parameter vector, $\hat{\beta}$, the sum of the squared deviations of the observed option prices from the prices implied by $\hat{\sigma}_t(X, \tau, \hat{\beta})$. This function is estimated separately for each day for which we have options price data. Each function implies a distinct risk-neutral density for any period in the future for which one wishes to forecast.

As is usual when handling option prices, a trade-off must be made between making

the parameterization of $\hat{\sigma}_t(X, \tau, \hat{\beta})$ rich enough to capture the details of the market's valuation of the risk and overfitting. Following the literature on fitting European options to single-state diffusions, we fit four specifications of $\hat{\sigma}_t(X, \tau, \hat{\beta})$ in this paper:

$$\begin{aligned}\hat{\sigma}_t(X, \tau, \hat{\beta}) &= \beta_1 X, \\ &\beta_0 + \beta_1 X, \\ &\beta_0 + \beta_1 X + \beta_2 X^2 \\ &\beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3.\end{aligned}\tag{1}$$

The first parameterization is the Black-Scholes, lognormal specification. The time series have been centered, therefore the drift term is omitted. The second adds a normal term, which has the effect of allowing for thicker tails on the density. The third specification is a polynomial extension to this, which allows for the standard volatility “smile” and “sneer” often observed in foreign exchange options, and the fourth adds an additional term to account for the possibility of very thick tails. Note that the term $\hat{\sigma}_t$ is not the same as the “implied volatility” reported by the exchanges (and therefore not a function of the strike price), but rather the single-state volatility parameter defining the diffusion process.

By estimating the diffusion process rather than the implied state space density for each expiration date, we allow for tests of forecast densities at a variety of horizons, not just the expiration dates for which we have options data. We obtain forecast densities for one, seven, fourteen, thirty, ninety, and one-hundred-eighty days ahead of the current information by using the separate estimates of $\hat{\sigma}_t(X, \tau, \hat{\beta})$ for each trading day, t . From these densities we acquire the series $\hat{\Pi}_{\theta,t}(X_{t+\theta})$, which is the probability, given the estimated density at t , that the θ -day-ahead forecast is less than or equal to the observed θ -day-ahead outcome, $X_{t+\theta}$. In other words, $\hat{\Pi}_{\theta,t}(X_{t+\theta})$ is the CDF of our estimated forecast densities. For clarity, we drop the θ notation when we refer to an estimated density, so that $\hat{\Pi}_{\theta,t}(X_{t+\theta}) = \hat{\Pi}_t(X_t)$.

4 A Higher-Order Lattice

Our test relies on evaluating the implied option prices quickly and accurately. Our experience suggests that the binomial lattice is not very accurate and requires many time steps to be useful, which requires far too much computational time. To save time, we use higher-order lattices. These work by using higher-order expansions to match the proposed process. They work analogously to the Taylor expansion, as it approximates a function by adding higher-order terms. A set of higher-order

approximations goes more quickly to a smooth function than adding the linear spline terms when both require the same computational effort. This is true as long as the function approximated is smooth.

Indeed, Heston and Zhou (2000) show that the rate of convergence for a lattice that matches q moments in the data converges to the diffusion at a rate $O(\Delta t^{\frac{q-1}{2}})$, if the payoff function is $2q$ times differentiable. Of course, with an American option on a future, the payoff function is not differentiable at two places: at the early-exercise boundary and at the strike price when the option expires. We resolve this problem by using logistical smoothers around the points of nondifferentiability.

The higher-order terms greatly increase the speed with which a function is approximated, but with a cost. The functions must be smooth, and handling the single-factor volatility parameter, which varies over the state space, is problematic. Our value functions are not smooth because they have a point of nondifferentiability at the strike price at the end boundary and (because these are American options) they involve a nonsmooth early-exercise boundary as well. The nonsmoothness is easily seen by looking at the value function of an option at its due date. For values of the underlying less than the strike price on a call option, the value is zero. Above the strike price, the value is the price of the underlying futures contract minus the strike price. Therefore, there is a strict nondifferentiability at the strike price.

We smooth the boundary functions with logistical smoothers that allow infinite differentiability. We describe this method in a previous paper, Craig and Keller (2001). We experimented with a variety of bandwidths for these smoothers and found that for our results they did not matter much. The results reported in this paper use a bandwidth measure of ω of 0.005 for the value function boundary and of 0.003 for the early-exercise boundary.

The second issue is more important and much trickier to handle. The single-factor volatility parameter, $\hat{\sigma}_t(X, \tau, \hat{\beta})$, varies over the parameter state space, X . In other words, the volatility parameter that defines the Brownian process is different for each value of the price. Standard higher-order-lattice procedures rely on the lattices being constant on a variation of the state space (i.e., some transformation of X). To illustrate the problems this causes along with our solution, we must discuss higher-order lattices. We start with a pentinomial lattice that has a constant value of state-space increments Δh , presented in Figure 1a.

In Figure 1a, the time axis goes from t to $t + \Delta t$. The higher-order lattice considers not just the point in the future of X_1 but also the states of one and two increments of Δh , both higher and lower than X_1 , so that each point has five future points factoring into its value at time t . Thus five probabilities are required for this lattice,

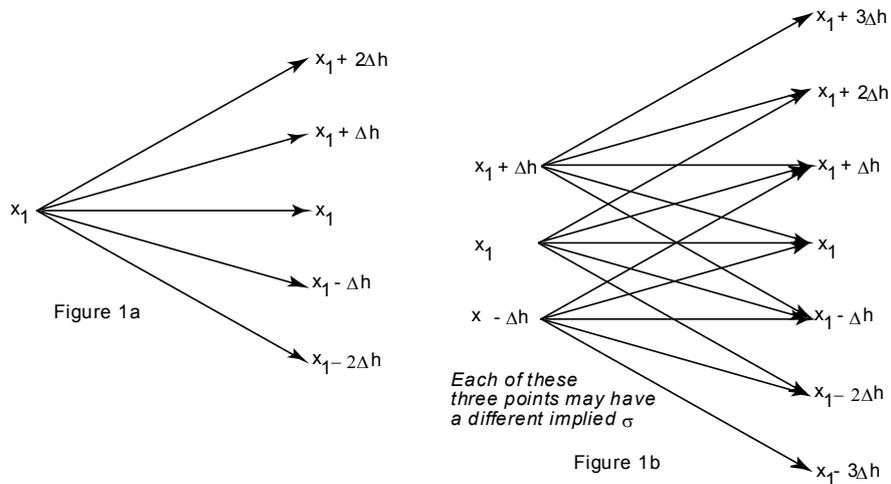


Figure 1: Higher order lattice

which we denote p_0 (for the probability of no change), p_{+1} (for the probability that the state goes up to $X_1 + \Delta h$), p_{+2} (for the probability that the state goes up to $X_1 + 2\Delta h$), and p_{-1} and p_{-2} (for the probabilities that the state goes down by the same amounts, Δh and $2\Delta h$). The lattice points of Figure 1a are combined as in Figure 1b. In this figure, each line represents a separate probability that must be calculated, as the value of $\hat{\sigma}_t(X, \tau, \hat{\beta})$ varies across the state space.

For a Taylor expansion of a function around a point, successively higher derivatives of the function at the point are evaluated. For approximating the stochastic function represented by $\hat{\sigma}_t(X, \tau, \hat{\beta})dz$, with $(dz)^2 = \Delta t$, successively higher moments are set equal to the moments of this local Brownian process. Thus, the odd-numbered moments are set to zero, implying equal probabilities of local up and down increments of the same magnitude (i.e., $\Delta h, -\Delta h$ and $2\Delta h, -2\Delta h$) and allowing the complexity of the system to be reduced. (Note that this is a statement about the instantaneous local behavior of the process. Because the volatility varies over the state space, the density over an outcome at a point in the future can have nonzero odd moments.) Setting the sum of the probabilities equal to unity, along with matching the second and fourth moments, gives a lattice which matches the first five moments of the

local process. Thus, the equations to solve for p_0, p_{+1} and p_{+2} are:

$$\begin{aligned}
2p_{+1} + 2p_{+2} + p_0 &= 1 \\
E[X(t + \Delta t) - X(t)]^2 &= 2p_{+1}\Delta h^2 + 2p_{+2}(2\Delta h)^2 = \hat{\sigma}_t(X, \tau, \hat{\beta})^2 \Delta t \\
E[X(t + \Delta t) - X(t)]^4 &= 2p_{+1}\Delta h^4 + 2p_{+2}(2\Delta h)^4 = 3\hat{\sigma}_t(X, \tau, \hat{\beta})^4 \Delta t^2
\end{aligned} \tag{2}$$

In this simple system, the solution for the implied risk-neutral probabilities is easily calculated as $p_{+2} = \frac{\alpha^2}{8} - \frac{\alpha}{24}$ and $p_{+1} = \frac{2\alpha}{3} - \frac{\alpha^2}{2}$, where $\alpha = \frac{\sigma_t(X, \tau, \hat{\beta})^2 \Delta t}{\Delta h^2}$. For the pentinomial tree, the p_i are positive if and only if $\alpha \in [\frac{1}{3}, \frac{4}{3}]$.

The time step Δt is determined by the size of a chosen state-space increment, Δh , a chosen value of α , and the maximum $\hat{\sigma}_t(X, \tau, \hat{\beta})$ at the end of the lattice, given the initial guess of the diffusion process on day t . In our scheme we used a value for the time step of $\Delta t = \frac{2}{3} \frac{\Delta h^2}{\max(\hat{\sigma}_t(X, \tau, \hat{\beta})^2)}$, which allowed the fourth moments to be matched for the part of the state space where $\hat{\sigma}_t$ is large. The p_i change for each node, according to the diffusion process, $\hat{\sigma}_t(X, \tau, \hat{\beta})$, and the value of the state X , since α is a function of X .

If Δt is constant (as it should be, for doing otherwise would greatly complicate calculating the lattice and would require additional computational time), and Δh is held constant, a problem emerges in the definitions of the probabilities in equation 1. If α , in response to certain state spaces, X , and given Δh increments, becomes less than $\frac{1}{3}$, then the value of p_{+2} will be less than zero, and therefore no longer a probability. Thus, for the fourth moments to be matched, α must be in the interval $[\frac{1}{3}, \frac{4}{3}]$. However, for constant Δt and Δh , the volatility parameter may fall to less than one-quarter of its maximum value over the range of the state space, which places α outside this interval. One possible solution is to simply match only second moments for smaller values of $\hat{\sigma}_t$. However, this destroys the gains from using the higher moments in the lattice. The solution we use in this paper is to vary the grid size, to account for smaller values of $\hat{\sigma}_t$. The problems of the varying $\hat{\sigma}_t$ and the proposed solution are illustrated in Figure 2.

The state space, X , is represented on the vertical axis. On the right side of the graph are plotted hypothetical functions, α_1 and α_2 . While both are functions of $\hat{\sigma}_t$ (and thus of X), α_1 is the function implied with the coarser grid, Δh , whereas α_2 is calculated from a finer grid, $0.5\Delta h$. Given the definition of $\alpha = \frac{\sigma_t(X, \tau, \hat{\beta})^2 \Delta t}{\Delta h^2}$, α can fall to one-quarter of the values of the coarser grid, so that α_2 must be in the interval $[\frac{1}{12}, \frac{1}{3}]$ for the finer grid. At point A, where the coarse grid is barely able to match all four moments with nonnegative probabilities at all nodes, and α_1 equals $\frac{1}{3}$, the fine grid is barely able to match all four moments with nonnegative probabilities,

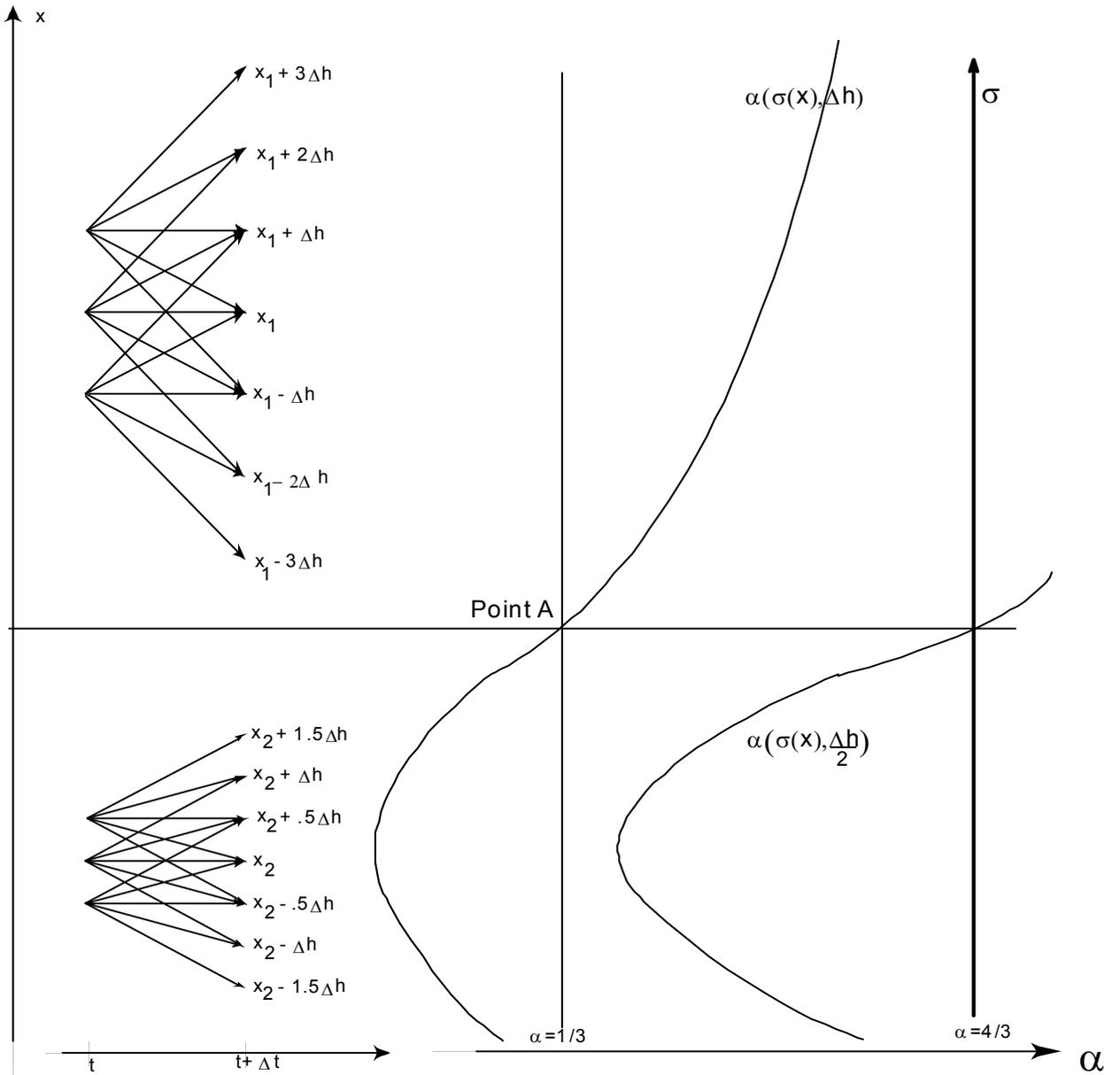


Figure 2: Properties of alpha

while α_2 is near its upper interval end (equal to $\frac{4}{3}$). Thus, point A divides the state space into two regions. Above point A ($\alpha_1 > \frac{1}{3}$), a coarser grid is used. For lattice nodes that completely fall below point A ($\alpha_2 < \frac{1}{3}$), a finer grid is used, as pictured on the left side of Figure 2. Point A therefore defines a grid-shift threshold in terms of the estimated diffusion process $\hat{\sigma}_t(X, \tau, \hat{\beta})$.

Where lattice points use values at $t + \Delta t$ on both sides of point A, the lattice, and its probabilities, must be modified. Figure 3a shows the lattice at point A, which lies between x_3 and $x_3 + \Delta h$. At $x_3 + \Delta h$, $\alpha > \frac{1}{3}$, (α refers to the function of Δh) and the coarser grid is appropriate. At x_3 , $\alpha < \frac{1}{3}$, and so the finer grid is appropriate for this point and points below it. We insert a branch of the lattice between x_3 and $x_3 + \Delta h$, at $x_3 + 0.5\Delta h$. As shown in Figure 3b, for $x_3 + \Delta h$ (and all nodes above it) branches can be found so that a coarse lattice can be constructed. For the node at or below x_3 , branches exist so that a finer lattice may be used. At the middle point, $x_3 + .5\Delta h$, however, a new lattice design must be used.

We use the lattice pictured in Figure 3c. This is a lattice with two inner branches like a finer grid lattice, but with two outer branches that are spaced with wider increments, like the coarser lattice. Matching second and fourth moments yields probabilities $p_{+0.5} = \frac{9\alpha}{4} - 3\alpha^2$, and $p_{+1.5} = \frac{\alpha}{3} - \frac{\alpha^2}{36}$. The parameter α is calculated as a function of a combination of both the fine and the coarse grid and, therefore, Δh , and therefore α must lie in the interval $[\frac{1}{12}, \frac{3}{4}]$. Note that α will be greater than $\frac{1}{12}$, so long as $\hat{\sigma}_t$ does not fall by more than fourfold over the interval Δh .

In contrast to a scheme using approximations that calculate the value of the option only at the trading date, the early-exercise boundary is easily incorporated within the higher-order-lattice framework by adding a maximization operator into the calculation of the discretized value functions at each node and at each time. Thus, for a call option, the value of the node at state X and time $t - \Delta t$, is

$$V(X, t - \Delta t) = \max\{e^{-\rho\Delta t}PV_{X_t}, Fp(X, t - \Delta t) - K\}, \quad (3)$$

where

$$PV_{X_t} \equiv P_u(V(X + \Delta h, t) + V(X - \Delta h, t)) + P_{2u}(V(X + 2\Delta h, t) + V(X - 2\Delta h, t)) + (1 - 2P_u - 2P_{2u})V(X, t), \quad (4)$$

and where K is the strike price, and where $Fp(X, t - \Delta t)$ is the value of the underlying futures price.

For a given diffusion, a higher-order lattice approximates the value function of each option by using the higher-order terms of a moment-generating function for the true value function. Using a higher-order approximation instead of a binomial

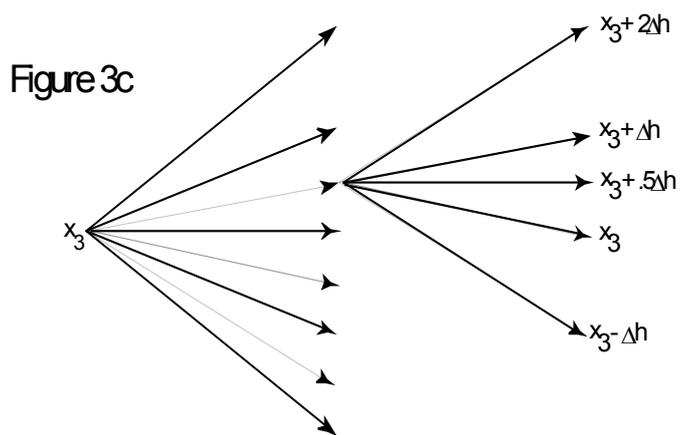
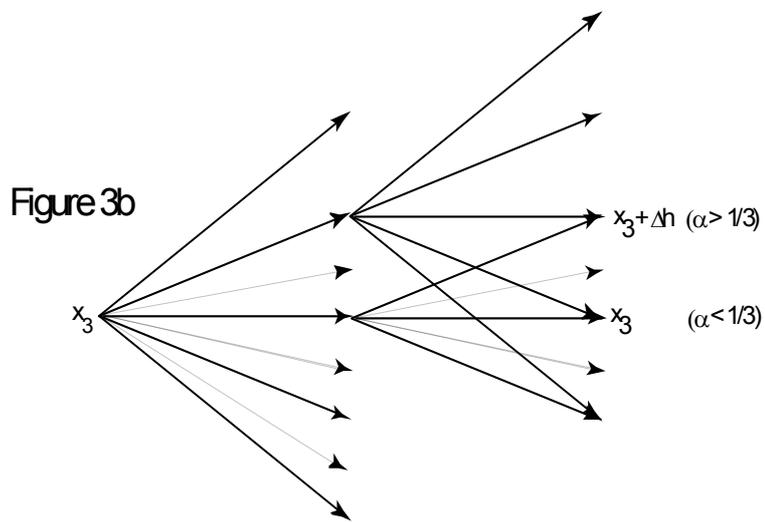
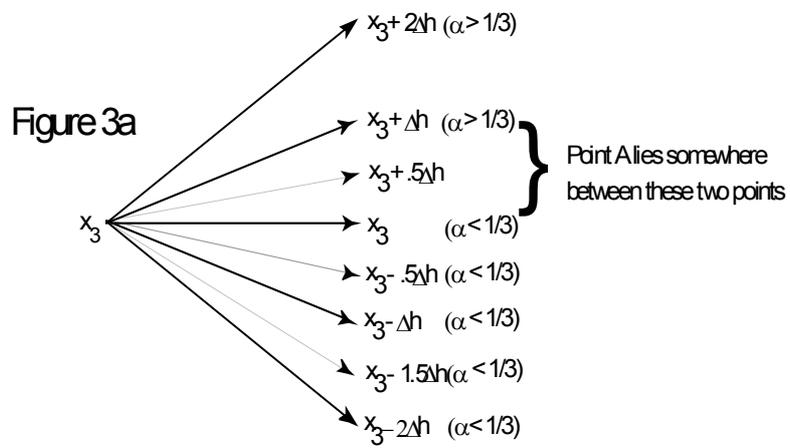


Figure 3: Construction of the lattice

tree to evaluate a diffusion expectation is analogous to using the sum of multinomial variables that have a density that is close to normally distributed to evaluate a normal expectation instead of a sum of binomial variables. Because of central limit theorems, averaging the binomial outcomes does approximate the normal distribution, but it does so more slowly than the sum of variables drawn from a distribution closer to the normal.

Although the kernel smoothing and the variable-grid-size-higher-order lattice both require a lot of computation and complication even for small bandwidths, we find they make a large difference in the calculated theoretical price of an option (and their solution was much closer to the actual value of the option when we had an analytical solution with which to check our numerical performance). As shall be seen in our results, the extra sophistication changes the interpretation of our tests, allowing us to say that the options provide an excellent forecast over the range of horizons for which the options are thickly traded.

Finally, a word is in order about the numerical techniques used to choose the parameters, $\hat{\beta}$, to minimize the distance between the observed and the theoretical prices. (In this case, the distance is measured as the sum of squared deviations.) The estimation had to be conducted over nearly 4,000 separate trading days. An initial problem using standard optimization routines was that for some iterations, the routines either would not converge, or at some point in the iterative process, they would attempt to evaluate the price of options at values of $\hat{\beta}$ for which the price proved difficult to compute. These routines were based on Newton iterations, which depend on the objective function being “well behaved” in the sense of being smooth and having initial values that are close enough to the optimum. The benefit of using these usual routines is that they converge to the exact optimum much more quickly than other routines.

We found this to be a poor trade-off. Part of the problem was that for many of the days we considered, the objective function was flat with respect to $\hat{\beta}$. In other words, the density could be approximated by a wide range of combinations of parameters in the vector, $\hat{\beta}$. Our problem, unlike many estimation problems, was not to identify individual elements of $\hat{\beta}$, but rather to estimate the option-implied density used by the market in its arbitrage-free pricing.

Instead of classical optimization programs, we used simplex-based methods, which although they converged more slowly, were much more likely to search a reasonable area for the optimum. A problem with simplex algorithms is that they often require a bit of art on the part of the operator in terms of choosing useful values for the step-size parameters, as well as initial values, so that the program can find a global optimum. After considerable trial and error, we developed a catalog of ten sets of

initial values and step sizes that were reasonably successful at finding global optima in a set of test prices. Our solution was to run the optimization for all ten sets and then to choose the value of $\hat{\beta}$ which had the lowest distance for the ten runs, which we denote $\hat{\beta}_{10}$.

Several observed phenomena gave us confidence that our methods were finding processes that were close to the true optimal processes. Although often many of the sets did not converge to a value close to the chosen $\hat{\beta}_{10}$, usually, close values were attained in at least four of the sets (though often in different sets for different trading days). On those trading days where only one of the sets attained a value close to $\hat{\beta}_{10}$, at least one other set gave a process that was very close to the process used, although with a different value of $\hat{\beta}$. Each of the ten sets we used had at least one trading day where it gave the chosen value, $\hat{\beta}_{10}$, and each of the ten sets had several days where it gave a result that was much worse than the chosen value. However, as a group, the ten sets attained a process estimate that we were fairly confident was close to the least squares process.

5 Evaluating Density Forecasts

Estimating a process implies a forecasting density, and different methods of estimation lead to different forecasting densities, some of which necessarily must be wrong. The ranking of these incorrect density forecasts is a difficult task, because a ranking depends on the often unknown individual loss function of agents. These loss functions may include more arguments than simply the mean and variance of a distribution. For example, decision makers with nonsymmetric expected loss care about the skewness of the distribution. Moreover, different agents have different loss functions, so that it is often impossible to find a ranking upon which all individuals agree unanimously. However, the correct density is always preferred to false densities in a forecast. Therefore, as a second-best solution one tries to approximate the true density as well as possible. (See Diebold, et al., (1998)).

To assess whether there is significant evidence to indicate that the estimated densities coincide with the true densities, we adopt a two-stage procedure. First, we transform the actual realizations of the foreign exchange rate into probability integrals, given our estimated density. Under the null hypothesis that the true density functions correspond to our estimated densities, the transformed realizations should be uniformly distributed. To assess this property of the transformed realizations we suggest as a second step two different tests, based upon the distance of the observed distribution of the transformed random variables from the uniform distribution. This

distance is in the L_2 topology, and was first suggested by Cramer in the 1920s. These tests are robust to time dependence in the data.

The basic univariate integral transformation theorem originated with Fischer (1930) and has been generalized for the multivariate case by Rosenblatt (1952). A thorough overview of transformation methods in goodness-of-fit techniques is given by Quesenberry (1986). Recently, Diebold et al. (1998) apply this concept to time series, evaluating the densities implied by a MA(1)-t-GARCH(1,1) model. Clements and Smith (2001) use the probability integral transforms for evaluating the density forecast of a self-exciting-threshold autoregressive model.

We report results from a continuous distance statistic, the so-called Cramer-von Mises statistic (von Mises (1931)). This statistic is defined as

$$\widehat{CvM} \equiv \int_0^1 (F(pr_n) - \widehat{F}(pr_n))^2 d(pr_n). \quad (5)$$

Note that this is a distance in the L_2 topology between the empirical distribution function ($Ecdf$) of z_t , $\widehat{F}(pr_n)$, and its theoretical value, $F(pr_n) = pr_n$, representing the uniform null. A similar statistic that was also computed with the same results lies in the L_∞ topology, the so-called Kolmogorov-Smirnov statistic,

$$KS \equiv \sup_{pr_n} |F(pr_n) - \widehat{F}(pr_n)|. \quad (6)$$

In this paper, our statistical inference is based upon bootstrap samples that preserve the time series properties of our original sample, $z_1, \dots, z_t, \dots, z_N$. The notation, z_t is a standard notation for the transformed variables that are distributed uniformly. The stationary bootstrap approach (IFSB) of Politis and Romano (1994) uses a resampling procedure to calculate standard errors of estimators that account for weak data dependence in stationary observations. From this bootstrap we can construct cutoff values for a variety of statistics. In addition, we construct the same values using a limiting-distribution technique, which gets around the ill-posed problem inherent in a variance-covariance matrix that is completely estimated (Craig and Keller (2003b) discuss the limiting distribution technique). This provides a check of the bootstrap technique, particularly for the shorter forecast horizons. We used the data-based choice of $prob$, suggested by Politis and Romano, so that $prob = prob_N \rightarrow N^{-1/3}$, with N equal to the number of observations. With this choice, the mean-squared error of $\widehat{\sigma}_{bt,prob_N}^2$ as an estimator of σ_N^2 is minimal. Fortunately, as long

as $prob \rightarrow 0$ and $Nprob \rightarrow \infty$, fundamental consistency properties of the bootstrap are unaffected by choosing $prob$ suboptimally. As can be directly seen, these requirements are clearly met by the choice of $prob = N^{-1/3}$.

We use the sample sequence $\{z_t\}$ to calculate the Cramer-von Mises statistic, \widehat{CvM} , directly for our sample $Ecdf$, $\widehat{F}(pr_n)$, and then to calculate whether this is a significant distance from the 45° line through the bootstrapped samples. Bootstrapped distribution functions, $F_b(pr_b)$, are also formed, and the CvM_b statistic,

$$CvM_b \equiv \int_0^1 (F(pr_b) - F_b(pr_b))^2 d(pr_b), \quad (7)$$

is evaluated for each bootstrapped sample. Because the sample distribution function \widehat{CvM} and all bootstrapped sample-distribution functions CvM_b are step functions, the integral expression in CvM_b is calculated directly. We computed CvM_b for 10,000 replications and report a number, CvM_b , which is the proportion of bootstrapped distances, CvM_b , that are greater than \widehat{CvM} , the distance between our sample distribution function and the null, the uniform distribution function. A value of CvM_b less than some critical value, α_0 , rejects the hypothesis of $z_1, \dots, z_t, \dots, z_N$ being drawn from a uniform distribution at the α_0 level.

6 The Results

The results for the entire sample of the CvM_b statistics are reported in Table 1. This table presents the probabilities that bootstrapped samples differ from the original sample in the Cramer-von Mises distance by as much as the original sample differs from the null of the 45° line. Lower values than 0.05 imply a rejection of the null at the five percent level. Several things are immediately clear from these tests. First, the data strongly support the option-price densities as useful forecasting densities for forecast horizons that extend to three months. In no case was the model rejected. Second, the simpler models and the more complicated models do about equally well at these horizons, and all do poorly at the very long time horizon. However, one important aspect of our results is that although the statistical procedure did not reject the simple lognormal model in most of the cases when the more sophisticated models were estimated, the simple lognormal model did not do better as a forecasting density. The p-value of the CvM_b statistics falls with the complexity of the estimated process for the shorter-term models and rises with the

complexity of the longer-term models.

Horizon θ in days	1	7	14	30	90	180
Specification*						
$\beta_1 X$	0.1108	0.1418	0.3033	0.3109	0.2050	0.0297
$\beta_0 + \beta_1 X$	0.0870	0.1146	0.2954	0.3404	0.2469	0.0373
$\beta_0 + \beta_1 X + \beta_2 X^2$	0.0773	0.0990	0.2733	0.3273	0.2367	0.0427
$\beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$	0.0811	0.0944	0.2607	0.3146	0.2271	0.0401

*Bold numbers indicate that the hypothesis of an accurate density can't be rejected.

Table 1: Stationary bootstrap on von Mises with a multigrid estimate from American options

Figure 4 plots the process volatility, σ , as a function of δ , which denotes the deviation of the actual state of the underlying from its forward rate, on February 22, 1990, which is just an arbitrary date.

The functions drawn into Figure 4 change their pattern daily depending on the new information coming in each day. If δ equals unity, then the state variable, which is the spot rate of the foreign exchange, is equal to the actual forward rate in the option contracts under consideration. If δ equals 1.1, then the actual state is 10 cents above the relevant forward rate. In Figure 4 four lines are plotted. Each line represents the state-dependent process volatility of a model. BS stands for the Black-Scholes model, where σ depends linearly on the state with a zero constant. Model 1 relaxes this restriction and allows for a positive or negative intercept. In model 2, σ is a quadratic function of the state, and so forth. By visual inspection of Figure 4, it immediately becomes clear that the process volatility is nearly the same for all models, if the actual state is near the ‘‘at the money (ATM)’’ rate, that is, near the forward rate. The more the actual state differs from the ATM point, the bigger are the differences (in terms of the state-dependent volatility) between the more complex models and the Black-Scholes model. Note, however, that the volatility of the more complex models can be above or below the Black-Scholes volatility. In other words, on another day the line of models 1, 2, or 3 can be above the Black-Scholes line. This is true for all quantiles. However, for higher quantiles, the average volatility of the more complex models is slightly above the volatility of the Black-Scholes model, indicating ‘‘fat tails.’’ Figure 5 makes this point clear.

In Figure 5, the 5 percent quantiles of the density implied by the Black-Scholes model are indicated on the abscissa. The 5 percent quantiles of the models are indicated on the ordinates. Several observations can be made immediately: First,

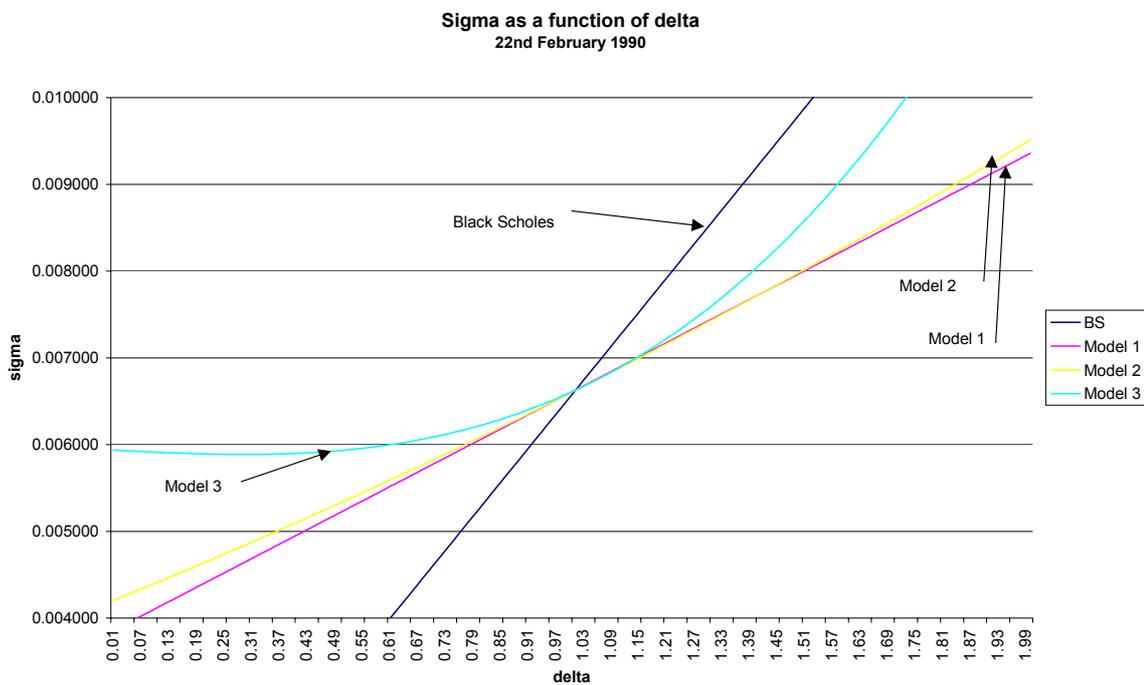


Figure 4: State/sigma diagram

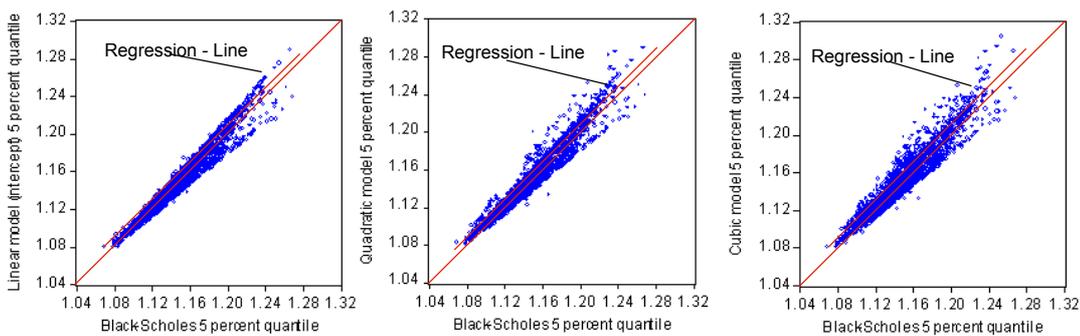


Figure 5: 5 percent quantiles

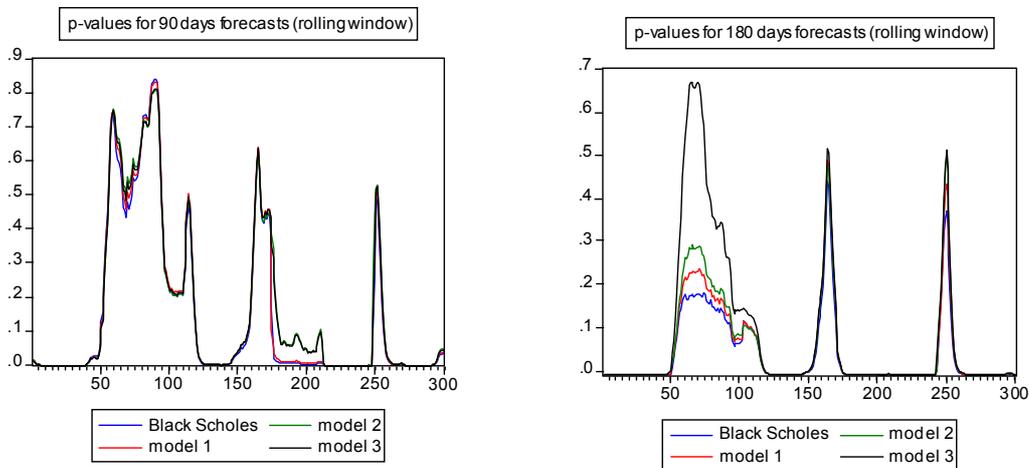


Figure 6: p-values

the estimated σ volatility fans out with higher values for the 5 percent quantile. Second, the average of the estimated σ of the more complex models lies above the 45 degree line, indicating a slightly higher 5 percent quantile for these models. In other words, the more complex models imply a higher probability mass in the outer regions of the density. In light of this observation, it is clear why the more complex models do better with longer-term forecasts (90 days and 180 days). This is because states which are further away from the ATM point become more relevant only in the long run. Simply speaking, the state needs some time to be driven into regions where model differences matter.

The outcome is more complex when we examine our risk-neutral densities on the basis of a rolling window. We chose to look at the sample as if the researcher were continually monitoring the forecasting performance of the simulated risk-neutral densities from 1000 trading days after the start of our sample period. We then stepped continuously forward, testing the forecasting properties of our densities with a step size of 20 trading days, until the end of the sample period. This led to about 300 p-values for each model and forecasting horizon. The results are depicted in Figure 6.

In Figure 6 each number represents a specific date and encompasses the p-value statistics, summarizing therefore their actual performance, of 1000 tested risk-neutral densities at different forecasting horizons before that date. Again, it turns out that

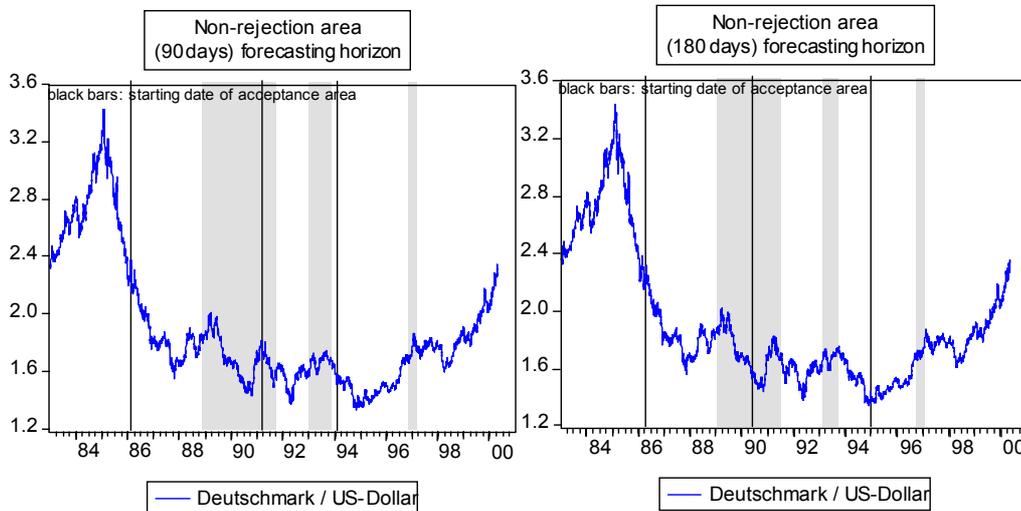


Figure 7: Non-rejection time periods

the differences among the four models seem to be negligible, aside from some periods in which the more complex models clearly outperform the Black-Scholes model. We first turn to the time periods in which the Black-Scholes model cannot be rejected and then turn to the time periods in which we observe substantial differences between the different models. Figure 7 is a plot of the U.S. dollar-German mark exchange rate. The shaded areas indicate the time periods in which the Black-Scholes model for the 90- and 180-day forecasting horizons are not rejected on the basis of 1000 tested densities (with p-values above 5 percent). Since one value in Figure 7 corresponds to 1000 days in Figure 7, the affected areas are quite large. The shaded areas for the more complicated models look similar, aside from some specific time periods we will discuss later. As it turns out, much of the time we do not accept the Black-Scholes model (and the other models) as good forecasting densities for the longer horizons. What is the reason for this? Since all densities are centered around the forward rate, which differs only by the interest rate differential from the spot rate, the forecasting densities do not capture the steep run-up in the value of the U.S. dollar in 1984 and the subsequent tumbling after early 1985. Therefore, we are able to reject the null of a good forecasting horizon until the end of 1988. Then, in late November 1988, our test starts to indicate that our densities are better forecasting tools. The reason for this is that the area from the first black bar to the

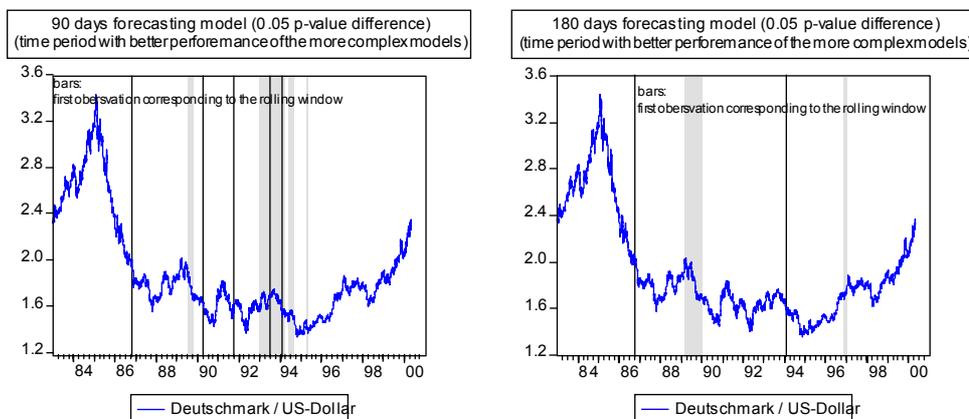


Figure 8: Time periods where complex models outperform lognormal

beginning of the shaded area represents a comparatively balanced period, in which we find no long sequences of one-sided errors. Since our test is very sensitive to this kind of failure and it does not materialize during this period, we start to accept the null. We find similar explanations for the other nonrejection periods.

Figure 8 (90- and 180-day horizon) also plots the U.S. dollar-German mark exchange rate. The shaded areas indicate where the most complex cubic model (Model 4) performs at least 5 percent (in terms of the p-value) better than the Black-Scholes model. Here again, the values in Figure 6 correspond to 1000 days in Figure 8, and therefore the time periods affected by several large forecasting mistakes are quite large. The black lines in Figure 6 indicate the first observation of the shaded area that leads to a p-value difference bigger than 0.05. Remember that the more complex models, that is, those whose densities allow for fat tails, become relevant only in the long run, as we have argued on the basis of Figures 4 and 5. One of the reasons we rejected the simpler models at these forecasting horizons was because of the frequency of one-sided errors. The fat-tail property of the more complex models 3 and 4 gives more leeway for large, one-sided errors under the null.

7 Robustness Checks

In this section we check the extent to which our results are affected by the option design (European vs. American), by the numerical method of smoothing nondifferentiable areas and by the bootstrap-test procedure.

Horizon θ in days	1	7	14	30	90	180
Specification*						
$\beta_1 X$	0.1096	0.1385	0.2928	0.3080	0.2033	0.0238
$\beta_0 + \beta_1 X$	0.0776	0.1207	0.3088	0.3549	0.2596	0.0308
$\beta_0 + \beta_1 X + \beta_2 X^2$	0.0635	0.1550	0.2306	0.2275	0.1714	0.0301
$\beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$	0.0960	0.1168	0.2858	0.3160	0.2228	0.0326

*Bold numbers indicate that the hypothesis of an accurate density can't be rejected.

Table 3: von Mises with a multigrid estimate
from European calculations on American options

Table 3 reports results from the estimates that are obtained when we assume that the options are European even when they are not. The false European options do a very good job of approximating the foreign exchange distributions. They were consistently poorer than the true American options only for the very long time horizon of 180 days, or for the more complex models in the mid-range horizons. The European options even did marginally better, in some ranges, than the correct American options. This surprised us. We had thought that the American early exercise would matter more in setting the density. However, this market seems to be one where simplicity rules the probability distributions. Standard assumptions do a good job of expaining foreign exchange beliefs.

Horizon θ in days	1	7	14	30	90	180
Specification*						
$\beta_1 X$	0.1155	0.1471	0.3072	0.3097	0.2029	0.0295
$\beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$	0.0839	0.1004	0.2650	0.3157	0.2244	0.0398

*Bold numbers indicate that the hypothesis of an accurate density can't be rejected.

Table 4: American option: no smoothing

We also conducted studies of the importance of the various components of our numerical methods in computing accurate forecasting densities. For example, Table 4 reports results from estimates that did not smooth the node for the American options. Note that this had little effect on the estimates, in spite of the reduced accuracy. This is less surprising than it might have been, had taking into account the American option mattered more. The estimates, however, do fare less well at the very place that American options matter, which is at the long horizons. The correct method, of course, only presents a marginal improvement because using the correct American method only represents a marginal improvement over the European method.

Horizon θ in days	1	7	14	30	90	180
Black Scholes*						
Exact Spec.	0.1032	0.1063	0.2751	0.4215	0.3225	0.0507
Local Bootstrap Once	0.1009	0.1088	0.2625	0.4610	0.3269	0.0703
Local Bootstrap Many	0.0998	0.1059	0.2586	0.4113	0.3152	0.0492

*Bold numbers indicate that the hypothesis of an accurate density can't be rejected.

Table 5: Bootstrap with exact Black Scholes and with estimation error

Table 5 reports robustness checks for slightly different assumptions about the densities and our estimates of the transformed variables. We first checked the accuracy using the simulated paths to estimate the transformed variables, z_t . This could be an issue because simple Monte-Carlo simulations indicated that the \widehat{CvM} statistic contained noticeable noise even when the number of simulations exceeded 10,000 and the number of time steps exceeded 10,000. The immediate question was what effect the noise had on the size of our test: Did the inaccuracy of the transformation to z_t cause enough of a change in the power of the test to affect our interpretation? The first line of Table 5 reports the bootstrap results for z_t 's that are generated by the closed-form solution of a simple lognormal density rather than by Monte-Carlo methods. Clearly, the results are not affected in this case. The only difference is that at long ranges, the approximation error in generating the z_t 's seems to have made our original tests (in Table 1) more powerful than they should have been, so that the density at the 180-day horizon is now accepted at the five percent level.

The second two lines of Table 5 concern our statistical assumptions about the densities generated by the options. Our maintained hypothesis is that the coefficients, $\beta_0, \beta_1, \beta_2$ and β_3 , are known to both the market and the researcher. In

other words, we, the researchers, have the densities, or the specification, that we are testing. The usefulness of stating the hypothesis this way is that it means we are testing whether the transformed variables, z_t , are distributed as $U[0, 1]$. If, however, we view the densities as estimates, then poorly measured values of $\beta_0, \beta_1, \beta_2$ and β_3 will mean that the z_t 's are no longer uniformly distributed, since the "estimated" coefficients imply a distribution that differs from the forecasting distribution implied by the options, even under the risk-neutral null. One possible solution is a localized bootstrap, suggested by Corradi and Swanson. Here the bootstrap samples under the alternative are generated from a different distribution: Instead of generating the z_t 's from the exact distribution implied by β_1 , say, in the case of Black-Scholes, we generate them from a distribution that samples from a sample of β_1 's. In other words, we acknowledge that the various estimates are generated by an error-ridden estimation process, so we bootstrap the estimates in generating our series of z_t 's. The distribution of β_1 's is "locally" normal, because the nonlinear least squares estimates of them are assumed to be well enough behaved that they are asymptotically normal.

Thus the procedure which yields the second and third lines of Table 5 is to first draw a sample of β_1 with a known distribution, but with a sample error generated by the estimation error.¹ We compute the \widehat{CvM} distance of our sample from this sample. Then we bootstrap the \widehat{CvM} distance of our original set of z_t 's from a series of z_t 's that are bootstrapped. We applied two forms of bootstrapping. In one, reported on the second line, a series of β_1 's was bootstrapped only once to generate a new set, z_{1t} . This set was then bootstrapped in the method of Politis and Romano to give a critical value. The third line reports the results of the second bootstrap procedure, where each replication of the Politis and Romano bootstrap was also computed from a separate bootstrapped value of β_1 . Note that for this table, the individual elements, z_t , are all calculated from the closed-form solution.

It is clear from the table that estimation error and how it is handled has little to do with our results. This is not so surprising in that options allow one to estimate the densities fairly exactly. In many cases we had over 100 options to work with. Our estimates for the Black-Scholes, for example, typically had standard errors of about one percent of the estimate. For the more complicated models, the standard errors were higher at times, but the 95 percent confidence ellipsoid encompassed densities that were all quite close to one another. As a result, the bootstrap results,

¹Statistical interference (i.e., $\widehat{\sigma}_i(\widehat{\beta}_i)$), of $\widehat{\beta}_i$ is based on NLLS. The nonlinear regression model is $p_j = h(X^i, \beta_i) + \varepsilon_j$, with p_j being the j observed American option prices at t . h is the nonlinear function that links the diffusion process to the prices.

which take into account this small estimation error, do not differ from the results of Table 1, which ignore it.

Table 6 reports results from an earlier study (Craig and Keller, 2001), which used bootstrap estimates for estimates that relied on a less accurate lattice method, which although of high order, had a fixed grid. Notice that the reliability of the option densities over the forecast horizons for which they are thickly traded is much less striking. Indeed, even the simplest model, the lognormal, performs poorly at the two-week horizon, whereas the more accurate flexible grid system reported in Table 1 delivers estimates that are fine forecasting densities. This is also true, though less striking, at the three-month horizon. More accurate evaluation of the models delivered more reliable forecasts.

Horizon θ in days	1	7	14	30	90	180
Specification*						
$\beta_1 X$	0.0000	0.0000	0.0006	0.0301	0.2103	0.0937
$\beta_1 X$ with more smoothing	0.0011	0.0042	0.0493	0.2254	0.2045	0.0151
$\beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$	0.0000	0.0000	0.0004	0.0099	0.1159	0.0446

*Bold numbers indicate that the hypothesis of an accurate density can't be rejected.

Table 6: Stationary bootstrap on von Mises with fixed grid

Figure 9 shows the results of the much weaker test, which observes the individual probabilities one by one. We report this to show where the long-term forecasts are failing. The figure plots the sample cumulative distribution functions. The 45 degree line represents the null, and the two bold lines represent confidence intervals of two standard errors above and below the estimated CDF. This is a weaker test than the bootstrapped \widehat{CvM} distance, as is clear from the fact that the null is always within the confidence interval. Even so, Figure 9 allows one to observe where the forecast is failing. What seems to be missing in the option-based forecast densities is a skewness in the distribution to the left. The actual distribution of z_t 's has too many large deviations to the left that are not accounted for in the forecast densities, and too few large deviations to the right. During this sixteen-year period of our data, then, taking option-based densities that are calculated from the one-week to three-month options and extending these densities to the half-year horizon fails because this approach overpredicts larger declines in the dollar, and underpredicts larger depreciations.

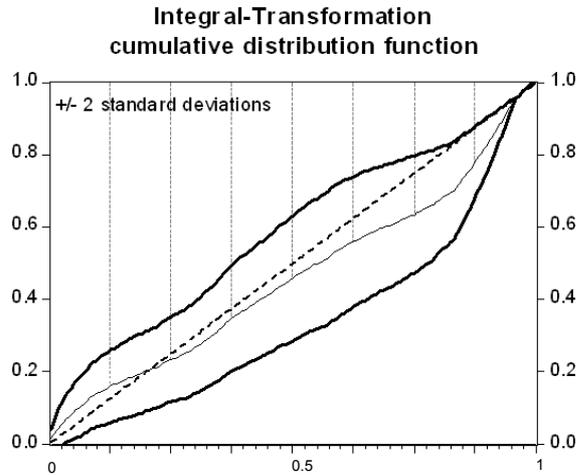


Figure 9: Confidence intervals over $U(0,1)$

8 Conclusions

This paper investigates the risk-neutral densities of foreign exchange options with respect to their application to estimation and forecast testing. Because of our modeling strategy, which expands the densities in a single state space, we can evaluate the implied densities for a single day at any forecast horizon that we wish. In order to expand our densities from the lognormal to a richer space of densities (along the polynomial expansions proposed by Dumas, et al.), we invented a rich apparatus which allows us to estimate the processes for a wide variety of options, such as Bermuda options or (as in this case) American options on a future.

While it is reassuring that our planned research was correct for these particular options, in the end, the expansions had very little effect on forecast performance. In only a few periods of our sample did the expanded density perform better than a simple lognormal density. This is certainly more true of our U.S. dollar-German mark foreign exchange option than of other options for which tests indicate that the lognormal performs less well than more skewed and thick-tailed densities. The same could be said of our correctness in handling the early-exercise boundary implicit in the American option. While over some of the periods and some forecasting horizons (particularly the longer ones), our apparatus created marginally better densities than one which treated the American option as if it were a simpler European option, it

is surprising that for these options, taking the early exercise into account in the estimation had little effect.

Far more important was the approximation error due to lattices that were of too low an order, or instances where the lattices did not match enough moments of the process over the entire grid. Indeed, it was remarkably easy to see when the grid size was too coarse, or when it did not encompass a large enough state-space for the estimates that were computed under these conditions gave uniformly horrible forecast properties. We found that computing successively finer grids until the results changed by only a small amount was a good strategy for ensuring an accurate estimate. This was also true when we checked our solution against estimates for which we had analytical solutions for the true value.

Compared to making sure that the numerical approximation of the lattice was close, all other modifications to our forecast approximations or tests of the forecasts were of second order. Accounting for estimation error in the process, for example, had only a tiny effect. Again, much of the reason for this centered on the particular properties of the market we study. The lognormal is not a terrible approximation for forecasts of the U.S. dollar-German mark exchange rate for this time period, as long as the volatility is updated on the day the forecast is being made. The options market has many observations for each day, so that estimation error is not large compared to forecast error. Most of the options are around the money, and volatility is fairly low, so that most of the time, the early-exercise ability does not change the estimate from one calculated by treating the options as European options.

9 Appendix: Calculation of the Variance for a Pearson Test

In the case of independence across the N cells and across time, entries $V(i, k) = 0$ if $i \neq k$ and $V(i, k) = \widehat{cov}(\widehat{p}_i, \widehat{p}_i) = \widehat{var}(\widehat{p}_n) = \widehat{\gamma}^n(0)$ if $i = k$ and the time overlap $h = 0$.

$$\widehat{cov}(\widehat{p}_i, \widehat{p}_k) = \frac{1}{T} \left[\widehat{\gamma}^{i,i}(0) + \sum_{j=1}^h \left(1 - \frac{j}{T}\right) \left(\widehat{\gamma}^{i,k}(j) + \widehat{\gamma}^{k,i}(j)\right) \right]$$

$$\widehat{\gamma}^{i,k}(j) = \frac{1}{T} \sum_{t=j+1}^T \left(I_t^i - \bar{I}^i\right) \left(I_{t-j}^k - \bar{I}^k\right) \quad \forall \quad i = 1, \dots, N \wedge k = i + 1, \dots, N$$

$\widehat{\gamma}^{k,i}(j)$ is calculated analogously. Note, that $\widehat{\gamma}^{k,i}(j) \neq \widehat{\gamma}^{i,k}(j)$, since $\widehat{p}_i \neq \widehat{p}_k$. For $i = n$, $\widehat{\gamma} = \widehat{\gamma}^{i,i}(j) = \widehat{\gamma}^n(j)$. In our case there is data dependence across the N cells and across time. Therefore, entries of $V(i, k) \neq 0$ for $i \neq k$. For each $\widehat{cov}(\widehat{p}_i, \widehat{p}_k)$ time dependence would be calculated up to order $\widehat{\gamma}^{i,k}(h)$.

10 References

BATES, D. S. [1991]: "The Crash of 87: Was it Expected? The Evidence from Options Markets," *Journal of Finance*, 46-3, pp. 1009-1044.

BREEDEN, D. AND R.H. LITZENBERGER [1978]: "Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business*, 51, pp. 621-651.

CLEMENTS, M.P. AND J. SMITH [2001]: "Evaluating Forecast from SETAR Models of Exchange Rates," *Journal of International Money and Finance*, 20, pp. 133-148.

CORRADI, V. AND N. SWANSON [2001]: "Bootstrap Specification Tests with Dependent Observations and Parameter Estimation Error," unpublished M.S, Department of Economics, University of Exeter, UK.

CRAIG, B. AND J. KELLER [2001]: "A Comparison of Densities Implied by Over-the-Counter and Exchange Traded Options in Foreign Exchange," unpublished M.S.

CRAIG, B. AND J. KELLER [2003]: "A Specification Test of Densities where the Data Overlap," unpublished M.S.

CRAMER, H. [1928]: "On the Composition of Elementary Errors," *Skand. Ak-tuartidskr.*, pp. 141-180.

CSÓRGÓ, M. AND L. HORVÁTH [1993]: "Weighted Approximations in Probability and Statistics," Wiley, New York.

DIEBOLD, F.X. AND T.A. GUNTHER AND A.S. TAY [1998]: "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39, 4, pp. 863-883.

DUMAS, B., J. FLEMING AND R.E. WHALEY [1998]: "Implied Volatility Functions: Empirical Tests," *Journal of Finance*, 56, pp. 2059-2106.

FIGLEWSKI, S. AND B. GAO [1999]: "The Adaptive Mesh Model: A New Approach to Efficient Option Pricing," *Journal of Financial Economics*, 53, pp. 313-351.

FISCHER, R.A. [1930]: "Inverse Probability," Proceedings of the Cambridge Philosophical Society, 36, 4, pp. 528 - 535.

HESTON, S., AND ZHOU, G. [2000]: "On the Rate of Convergence of Discrete-Time Contingent Claims", Mathematical Finance.

JACKWERTH, C.J. AND M. RUBINSTEIN [1996]: "Recovering Probability Distributions," Journal of Finance, 51, pp. 1611-1631.

MALZ, A.M. [1997]: "Estimating the Probability Distribution of the Future Exchange Rate from Option Prices," The Journal of Derivatives, 4, pp. 18-36.

NEUHAUS, H. [1995]: "The Information Content of Derivatives for Monetary Policy", Discussion Paper 3/95, Economic Research Group of the Deutsche Bundesbank.

PEARSON, K. [1905]: "On the General Theory of Skew Correlation and Non-linear Regression," Draper's Non-linear Regression, Draper's Company Memoirs, Biometric Series II.

POLITIS, D.N. AND J.R. ROMANO [1994]: "The Stationary Bootstrap," American Statistical Association, 89, 428, pp. 1303-1313.

QUESENBERY, C.P. [1986]: "Some Transformation Methods in Goodness-of-Fit", in: Goodness-of-Fit Techniques, eds. Ralph B. D'Agostino and Micheal A. Stephens, pp. 235-275.

ROSENBLATT [1952]: "Remarks on a Multivariate Transformation," An. Math. Stat. 23, pp. 470-472.

SHIMKO, D.[1993]: "Bounds of Probability," Risk, 6, 4, pp. 33-37.

STUTZER, M. [1996]: "Maximum Entropy and Options Prices," Journal of Finance, 51, pp. 1588-1610.

VON MISES, R. [1931]: "Wahrscheinlichkeitsrechnung", Wien, Leipzig.

The following Discussion Papers have been published since 2004:

Series 1: Economic Studies

1	2004	Foreign Bank Entry into Emerging Economies: An Empirical Assessment of the Determinants and Risks Predicated on German FDI Data	Torsten Wezel
2	2004	Does Co-Financing by Multilateral Development Banks Increase “Risky” Direct Investment in Emerging Markets? – Evidence for German Banking FDI	Torsten Wezel
3	2004	Policy Instrument Choice and Non-Coordinated Monetary Policy in Interdependent Economies	Giovanni Lombardo Alan Sutherland
4	2004	Inflation Targeting Rules and Welfare in an Asymmetric Currency Area	Giovanni Lombardo
5	2004	FDI versus cross-border financial services: The globalisation of German banks	Claudia M. Buch Alexander Lipponer
6	2004	Clustering or competition? The foreign investment behaviour of German banks	Claudia M. Buch Alexander Lipponer
7	2004	PPP: a Disaggregated View	Christoph Fischer
8	2004	A rental-equivalence index for owner-occupied housing in West Germany 1985 to 1998	Claudia Kurz Johannes Hoffmann
9	2004	The Inventory Cycle of the German Economy	Thomas A. Knetsch
10	2004	Evaluating the German Inventory Cycle Using Data from the Ifo Business Survey	Thomas A. Knetsch
11	2004	Real-time data and business cycle analysis in Germany	Jörg Döpke

12	2004	Business Cycle Transmission from the US to Germany – a Structural Factor Approach	Sandra Eickmeier
13	2004	Consumption Smoothing Across States and Time: International Insurance vs. Foreign Loans	George M. von Furstenberg
14	2004	Real-Time Estimation of the Output Gap in Japan and its Usefulness for Inflation Forecasting and Policymaking	Koichiro Kamada
15	2004	Welfare Implications of the Design of a Currency Union in Case of Member Countries of Different Sizes and Output Persistence	Rainer Frey
16	2004	On the decision to go public: Evidence from privately-held firms	Ekkehart Boehmer Alexander Ljungqvist
17	2004	Who do you trust while bubbles grow and blow? A comparative analysis of the explanatory power of accounting and patent information for the market values of German firms	Fred Ramb Markus Reitzig
18	2004	The Economic Impact of Venture Capital	Astrid Romain, Bruno van Pottelsberghe
19	2004	The Determinants of Venture Capital: Additional Evidence	Astrid Romain, Bruno van Pottelsberghe
20	2004	Financial constraints for investors and the speed of adaption: Are innovators special?	Ulf von Kalckreuth
21	2004	How effective are automatic stabilisers? Theory and results for Germany and other OECD countries	Michael Scharnagl Karl-Heinz Tödter

22	2004	Asset Prices in Taylor Rules: Specification, Estimation, and Policy Implications for the ECB	Pierre L. Siklos Thomas Werner Martin T. Bohl
23	2004	Financial Liberalization and Business Cycles: The Experience of Countries in the Baltics and Central Eastern Europe	Lúcio Vinhas de Souza
24	2004	Towards a Joint Characterization of Monetary Policy and the Dynamics of the Term Structure of Interest Rates	Ralf Fendel
25	2004	How the Bundesbank really conducted monetary policy: An analysis based on real-time data	Christina Gerberding Andreas Worms Franz Seitz
26	2004	Real-time Data for Norway: Challenges for Monetary Policy	T. Bernhardsen, Ø. Eitrheim, A.S. Jore, Ø. Røisland
27	2004	Do Consumer Confidence Indexes Help Forecast Consumer Spending in Real Time?	Dean Croushore
28	2004	The use of real time information in Phillips curve relationships for the euro area	Maritta Paloviita David Mayes
29	2004	The reliability of Canadian output gap estimates	Jean-Philippe Cayen Simon van Norden
30	2004	Forecast quality and simple instrument rules - a real-time data approach	Heinz Glück Stefan P. Schleicher
31	2004	Measurement errors in GDP and forward-looking monetary policy: The Swiss case	Peter Kugler Thomas J. Jordan Carlos Lenz Marcel R. Savioz

32	2004	Estimating Equilibrium Real Interest Rates in Real Time	Todd E. Clark Sharon Kozicki
33	2004	Interest rate reaction functions for the euro area Evidence from panel data analysis	Karsten Ruth
34	2004	The Contribution of Rapid Financial Development to Asymmetric Growth of Manufacturing Industries: Common Claims vs. Evidence for Poland	George M. von Furstenberg
35	2004	Fiscal rules and monetary policy in a dynamic stochastic general equilibrium model	Jana Kremer
36	2004	Inflation and core money growth in the euro area	Manfred J.M. Neumann Claus Greiber
37	2004	Taylor rules for the euro area: the issue of real-time data	Dieter Gerdesmeier Barbara Roffia
38	2004	What do deficits tell us about debt? Empirical evidence on creative accounting with fiscal rules in the EU	Jürgen von Hagen Guntram B. Wolff
39	2004	Optimal lender of last resort policy in different financial systems	Falko Fecht Marcel Tyrell
40	2004	Expected budget deficits and interest rate swap spreads - Evidence for France, Germany and Italy	Kirsten Heppke-Falk Felix Hüfner
41	2004	Testing for business cycle asymmetries based on autoregressions with a Markov-switching intercept	Malte Knüppel
1	2005	Financial constraints and capacity adjustment in the United Kingdom – Evidence from a large panel of survey data	Ulf von Kalckreuth Emma Murphy

2	2005	Common stationary and non-stationary factors in the euro area analyzed in a large-scale factor model	Sandra Eickmeier
3	2005	Financial intermediaries, markets, and growth	F. Fecht, K. Huang, A. Martin
4	2005	The New Keynesian Phillips Curve in Europe: does it fit or does it fail?	Peter Tillmann
5	2005	Taxes and the financial structure of German inward FDI	Fred Ramb A. J. Weichenrieder
6	2005	International diversification at home and abroad	Fang Cai Francis E. Warnock
7	2005	Multinational enterprises, international trade, and productivity growth: Firm-level evidence from the United States	Wolfgang Keller Steven R. Yeaple
8	2005	Location choice and employment decisions: a comparison of German and Swedish multinationals	S. O. Becker, K. Ekholm, R. Jäckle, M.-A. Muendler
9	2005	Business cycles and FDI: evidence from German sectoral data	Claudia M. Buch Alexander Lipponer
10	2005	Multinational firms, exclusivity, and the degree of backward linkages	Ping Lin Kamal Saggi
11	2005	Firm-level evidence on international stock market comovement	Robin Brooks Marco Del Negro
12	2005	The determinants of intra-firm trade: in search for export-import magnification effects	Peter Egger Michael Pfaffermayr

13	2005	Foreign direct investment, spillovers and absorptive capacity: evidence from quantile regressions	Sourafel Girma Holger Görg
14	2005	Learning on the quick and cheap: gains from trade through imported expertise	James R. Markusen Thomas F. Rutherford
15	2005	Discriminatory auctions with seller discretion: evidence from German treasury auctions	Jörg Rocholl
16	2005	Consumption, wealth and business cycles: why is Germany different?	B. Hamburg, M. Hoffmann, J. Keller
17	2005	Tax incentives and the location of FDI: evidence from a panel of German multinationals	Thiess Buettner Martin Ruf
18	2005	Monetary Disequilibria and the Euro/Dollar Exchange Rate	Dieter Nautz Karsten Ruth
19	2005	Berechnung trendbereinigter Indikatoren für Deutschland mit Hilfe von Filterverfahren	Stefan Stamford
20	2005	How synchronized are central and east European economies with the euro area? Evidence from a structural factor model	Sandra Eickmeier Jörg Breitung
21	2005	Asymptotic distribution of linear unbiased estimators in the presence of heavy-tailed stochastic regressors and residuals	J.-R. Kurz-Kim S.T. Rachev G. Samorodnitsky
22	2005	The Role of Contracting Schemes for the Welfare Costs of Nominal Rigidities over the Business Cycle	Matthias Pastian

Series 2: Banking and Financial Studies

1	2004	Forecasting Credit Portfolio Risk	A. Hamerle, T. Liebig, H. Scheule
2	2004	Systematic Risk in Recovery Rates – An Empirical Analysis of US Corporate Credit Exposures	Klaus Düllmann Monika Trapp
3	2004	Does capital regulation matter for bank behaviour? Evidence for German savings banks	Frank Heid Daniel Porath Stéphanie Stolz
4	2004	German bank lending during emerging market crises: A bank level analysis	F. Heid, T. Nestmann, B. Weder di Mauro, N. von Westernhagen
5	2004	How will Basel II affect bank lending to emerging markets? An analysis based on German bank level data	T. Liebig, D. Porath, B. Weder di Mauro, M. Wedow
6	2004	Estimating probabilities of default for German savings banks and credit cooperatives	Daniel Porath
1	2005	Measurement matters – Input price proxies and bank efficiency in Germany	Michael Koetter
2	2005	The supervisor's portfolio: the market price risk of German banks from 2001 to 2003 – Analysis and models for risk aggregation	Christoph Memmel Carsten Wehn
3	2005	Do banks diversify loan portfolios? A tentative answer based on individual bank loan portfolios	Andreas Kamp Andreas Pfingsten Daniel Porath
4	2005	Banks, markets, and efficiency	F. Fecht, A. Martin

5 2005 The forecast ability of risk-neutral densities of foreign exchange Ben Craig
Joachim Keller

