SVERIGES RIKSBANK
WORKING PAPER SERIES

146

SVERIGES
RIKSBANK

Evaluating Implied RNDs by some New Confidence Interval Estimation Techniques

Magnus Andersson and Magnus Lomakka

JANUARY 2003

WORKING PAPERS ARE OBTAINABLE FROM

Sveriges Riksbank • Information Riksbank • SE-103 37 Stockholm
Fax international: +46 8 787 05 26
Telephone international: +46 8 787 01 00
E-mail: info@riksbank.se

The Working Paper series presents reports on matters in the sphere of activities of the Riksbank that are considered to be of interest to a wider public.
The papers are to be regarded as reports on ongoing studies and the authors will be pleased to receive comments.

Evaluating Implied RNDs by Some New Confidence Interval Estimation Techniques

Magnus Andersson and Magnus Lomakka*

Stockholm School of Economics

Stockholm University

Sveriges Riksbank Working Paper Series
No. 146
January 2003

Abstract

This paper evaluates the precision of the parametric double lognormal (DLN) and the nonparametric smoothing spline method (SPLINE) for estimating risk-neutral distributions (RNDs) from observed option prices. By using a bootstrap technique confidence bands are estimated for the riskneutral distributions (RNDs) and the width is used as the criterion when evaluating the precision of the two. Previous literature on estimating confidence bands has to a large extent been estimated by Monte Carlo methods. We argue that the bootstrap technique is to be preferred due to the non-normality of the error structure. Our findings favour the SPLINE method, yielding tighter confidence bands. An example showing how the confidence intervals could be used for practical purposes is also provided.

Keywords: Implied risk-neutral distribution, Confidence intervals, Bootstrap

JEL Classification: C14, G14, E59

---

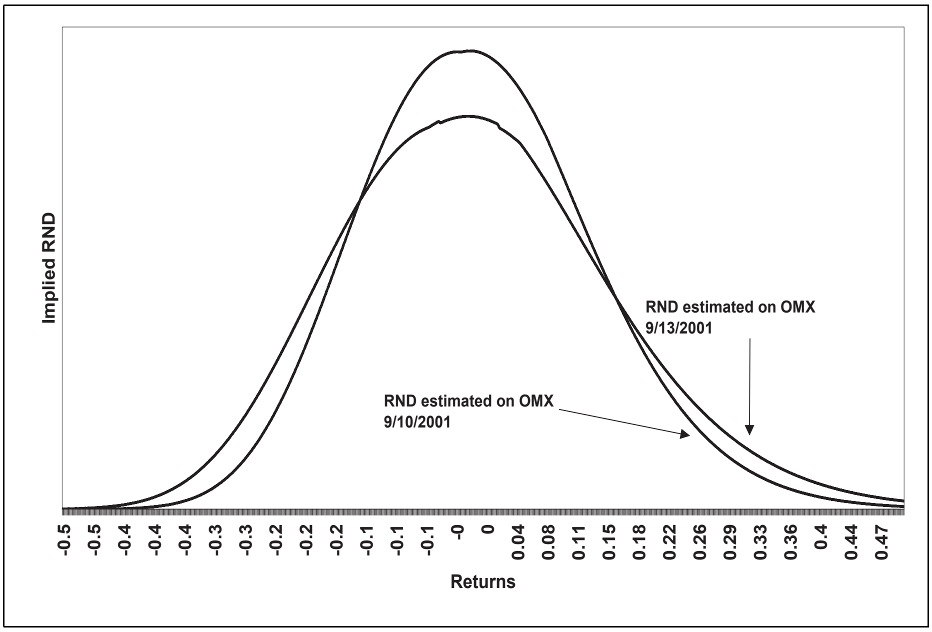
*We are grateful for useful discussions with and suggestions from Paul Söderlind, Hans Wijkander, Henrik Amilon, Anders Vredin, Jorge Barros Luis, John Fell, Niklas Nordman and Peter Hördahl. Any remaining errors are the authors responsibility. Correspondence: Magnus Andersson, European Central Bank, Capital Market and Financial Structure Division, Postfach 16 03 19, D-60066 Frankfurt am Main, Germany, e-mail magnus.andersson@ecb.int; or Magnus Lomakka, AP-fund 1, Skeppsbron 2, SE-103 25 Sweden, e-mail magnus.lomakka@ap1.se. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Riksbank, the ECB or the AP-fund.

1 Introduction

Information in financial prices is a valuable tool for policymakers in order to grasp market participants' perception of the future development of asset prices. For example, forward contracts may in some cases provide a point estimate of expected future asset prices. Using option prices, this methodology has been refined making it possible to extract a whole probability density function surrounding the expected mean, hence quantifying the uncertainty measured by higher moments such as skewness and kurtosis. Due to the risk-neutral feature of option pricing models these distributions are classified as risk-neutral distributions (RND). RNDs have many fields of application within the empirical finance area. For instance, it is important for central banks when performing inflation forecasts to monitor not only the development of commodity prices (e.g. oil) but also possible asymmetries in expectations displayed in the estimated RNDs in order to judge uncertainty about future inflation impulses. Moreover, market participants use RNDs as tool in their investment decisions. For instance, some investors measure the fatness of the tails (the kurtosis) and use this as an indicator of risk-appetite among market participants, reasoning that the fatness is negatively correlated with the degree of risk-appetite. The fact that option prices are forward-looking makes them tractable as such indicators. Two RND estimation techniques have been discerned as standard methods by users such as central banks and market participants—the nonparametric smoothed implied volatility smile (SPLINE) method and the parametric double log-normal (DLN) technique. Thus it is of great importance to test the reliability of these two methods to estimate RNDs.

One way to accomplish such an endeavour is simply to compare confidence intervals around the estimated RNDs. However, earlier work within this field has generated either extremely wide and spurious-looking confidence bands or quite conspicuously narrow confidence intervals. We show that the confidence bands in the previous literature to a large extent rely on assumptions that are not valid, for instance the normality assumption underlying the Monte Carlo method. We therefore try to extend the previous literature in a direction that brings about confidence bands that more closely reflect the true uncertainty of the estimated RNDs. *The main aim of this paper is consequently to evaluate the precision of the two standard RND estimation techniques. This is conducted by estimating confidence bands. The width of the bands is used as the main criterion when determining which RND estimation technique that is to be preferred. Two bootstrap methods are used when estimating the confidence bands which enable us to take the non-normality of the*

Figure 1: RNDs estimated before and after the events of September 11, 2001



pricing error into account.

From a practitioner's viewpoint it is also interesting to evaluate whether a change in the shape of the RND in connection with a macro event could be deemed as significant. Up until now RND densities have been estimated before and slightly after new information has hit the market, and conclusions have been drawn by comparing the two RNDs visually. For instance, the effect on the Swedish stock market (OMX), of the extreme events of September 11, 2001, are reflected in the RNDs shown in Figure 1.

Such a procedure does not provide any information as to whether there has been any statistically significant change in the shape of the RND. Estimating confidence bands around the RNDs would therefore provide policymakers such as central banks with a better indicator as to whether investors' perceptions regarding risk have changed.

This paper is structured in the following manner: Section 2 provides an exposition of the theoretical background of RNDs and confidence band estimation. Section 3 includes a discussion on how to improve confidence band estimation. The results of the confidence interval estimations are presented in Section 4 together with a case study. In Section 5 some conclusions are provided.

2 Theoretical Background

In this section an overview of the methodology is provided with reference to RND and confidence interval estimation in general. First we derive the estimation procedure for the two RND estimation methods, i.e. DLN and the SPLINE technique. Second a discussion on why and how confidence bands for RNDs should and can be estimated is supplied. Third, earlier work on RND confidence interval estimation are presented.

2.1 RND Estimation

The starting point of RND estimation methods was a paper by Breeden and Litzenberger (1978), who showed that relation between option prices and risk-neutral densities given by

$$q(S_T) = e^{-r\tau} \frac{\partial^2 C(S_t, X, \tau)}{\partial X^2}, \tag{1}$$

where $q(S_T)$ is the RND density of the underlying asset at time T, C is the call price function, S_t is the value of the underlying asset at time t, X is the strike price of the option and τ is the time until expiration of the option.[1] Since Breeden and Litzenberger's pioneering work a number of methods have been discussed in the literature—see for example Bliss and Panigirtzoglou (2001a). Two methods have now been discerned as standard techniques when estimating RNDs: the double lognormal (DLN) and the smoothed implied volatility smile (SPLINE) methods. The two methods differ in the sense that the DLN is parametric while the SPLINE method is non-parametric. These two techniques are standard among many practitioners and the fact that the estimation techniques are considerably different is the motivation for choosing these two methods for our evaluation. The details of the estimation procedure for the two methods are outlined in the following two subsections.

2.1.1 Double Lognormal Method

The Black-Scholes formula involves the assumption that the price of the underlying asset at maturity is lognormally distributed. However, investors generally attach higher probabilities for extreme outcomes than is suggested by the Black-Scholes formula. This implies that the terminal RND derived from observed option prices will have "fatter" tails compared with the lognormal distribution. Richey (1990), who examined the impact of non-normal underlying returns densities, found a way

[1] The formula is valid for put option functions as well.

to capture this specific characteristic by assuming that the functional form of the terminal density of the underlying asset is a mixture of two or more lognormal densities. In order to minimize the number of parameters to be estimated, a mixture of two lognormal densities (DLN) is employed in this paper—which has become a standard procedure in the literature. The estimation procedure chooses parameters to minimize the square of the deviation between the observed call/put prices and the theoretical prices obtained from lognormal distributions.

The prices of European call and put options can be written as

$$C(S_t, X, \tau) = e^{-r\tau} \int_X^{\infty} q(S_T)(S_T - X) dS_T, \tag{2}$$

and

$$P(S_t, X, \tau) = e^{-r\tau} \int_0^{X} q(S_T)(X - S_T) dS_T. \tag{3}$$

The DLN approach assumes that $q(S_T)$ is a weighted sum of two lognormal density functions, i.e.,

$$q(S_T) = \pi \times \mathrm{LnD}\left(\alpha_1, \beta_1; S_T\right) + (1-\pi) \times \mathrm{LnD}\left(\alpha_2, \beta_2; S_T\right), \tag{4}$$

where 'LnD' is the two lognormal densities, α_i and β_i correspond to the mean and the standard deviation of density i, and π and $(1-\pi)$ are the respective weights put on the two densities. Replacing the expression for the density from (4) into the call and put price formulas in (2) and (3) makes it possible to estimate the theoretical call and put prices, given the parameter vector $\Theta = [\pi, \alpha_1, \alpha_2, \beta_1, \beta_2]$.

Finally, the implied RND can be extracted given the observed option prices and theoretical prices provided by the parameter vector Θ. Moreover, as Bahra (1997) proposed, additional information could be exploited by including the future price as an extra observation in the minimization problem. In absence of arbitrage possibilities, the future price should equal the mean of the RND, which is represented by the last term in equation (5). The estimation is performed using non-linear least squares:

$$\begin{aligned} \min_{\Theta} \quad & \sum_{i=1}^{n} \left[c(X_i, \tau)\hat{c}_i\right]^2 + \sum_{i=1}^{n} \left[p(X_i, \tau) - \hat{p}_i\right]^2 \\ & + \left[e^{r\tau} S_t - \pi e^{\alpha_1 + \frac{1}{2}\beta_1^2} + (1-\pi) e^{\alpha_2 + \frac{1}{2}\beta_2^2}\right]^2 \end{aligned} \tag{5}$$

where $\hat{c}_i$ and $\hat{p}_i$ are the observed call and put prices and we have restricted β_1 and

β_2 to fulfil $0.25 < \frac{\beta_1}{\beta_2} < 4$.[2]

2.1.2 Smoothed Implied Volatility Method

The smoothed implied volatility smile method (SPLINE) for estimating RNDs originates from Schimko (1993), and it explicitly utilizes the results of Breeden and Litzenberger (1978). This nonparametric method involves approximating of a function, often a cubic spline, to some discrete observations. Schimko backed out the implied volatility from the Black-Scholes formula, taking the call and put prices of the options as given. A cubic spline was then approximated to the implied volatility observations. Transforming the discrete call/put price observations to a continuum of prices makes it possible to obtain the RND by differentiating the price function twice with respect to the strike price. The cubic spline can be fitted in spaces other than implied volatility/strike price; Bates (1991), for example, discusses fitting in the price/strike space, while Malz (1997) suggests a transformation to the implied volatility/delta space. The "delta" of call and put prices measures its sensitivity to changes in the underlying asset and hence, the implied volatility/delta space will have the effect that options close to at-the-money will be less closely grouped together compared to options far away from at-the-money. Thus more shape will be permitted at the center of the PDF than in the tails. The latter approach is implemented in this paper due to the attractive curvature feature at the center of the distribution. However, fitting the SPLINE function in the implied volatility/delta space might complicate the comparison with the DLN method since the minimization problems are solved in different spaces. Firstly, the data is transformed from the price/strike space into the implied volatility/delta space. The delta equals

$$\frac{\partial C(S_t, X, \tau)}{\partial S_t} \text{ and } \frac{\partial P(S_t, X, \tau)}{\partial S_t}, \tag{6}$$

i.e., the first derivative of the observed call and put prices with respect to the value of the underlying asset. The implied volatility is obtained by backing out the volatility from the Black-Scholes formula taking the price of the option as given.

A continuous function is then approximated around the discrete observations

[2]These parameter restrictions are used in order to produce densities which are not spurious looking. Ignoring this restriction in the optimization procedure can in some cases lead to a situation in which the shape of one of the two lognormal densities becomes a "spike" or in which the final RND is bimodal in its shape. Another drawback of ignoring this restriction is that the optimization procedure sometimes fails to find any solution and hence no RND can be extracted from that particular dataset. It worth noticing, however, that our choice of restriction is somewhat arbitrary and might marginally effect the results.

using so-called smoothing splines. The SPLINE function is extracted by solving the following minimization program

$$\min_{\omega} \lambda \sum_{i} \left\{ w(i) \left[y(i) - s(x(i;\omega)) \right]^2 \right\} + (1-\lambda) \int f''(x;\omega)^2 dx, \tag{7}$$

where $y(i)$ and $s(x(i;\omega))$ are the actual observations and the observations generated from the SPLINE function, respectively, $f''(x;\omega)$ is the second derivative of the SPLINE function (in this case with respect to delta), $w(i)$ is the weighting parameter[3] for the observations, $\lambda \in [0,1]$ is the weighting parameter deciding the degree of smoothness in the SPLINE function,[4] and ω is the matrix of parameters of the cubic spline. The value of the smoothing parameter λ is of great importance. A small value of λ will have the effect that the minimization program primarily minimizes the curvature of the SPLINE function. In the extreme case of $\lambda = 0$, the SPLINE function is equivalent to linear least squares estimation. A high value on λ instead has the opposite effect; the minimization program puts greater weight on minimizing the sum of the squared errors between the actual observations and the observations obtained from the SPLINE function.

A problem with the SPLINE method is to choose the value of the smoothing parameter λ, which is not related to economic theory. To avoid choosing the value of λ arbitrarily, various methods can be employed. Bliss and Panigirtzoglou (2001a) for example, compared the robustness between the DLN and the SPLINE techniques, choosing the value of λ by computing the sum of squared errors obtained from the DLN method. This involves carrying out a line search over the range of possible λ and choosing the value of the smoothing parameter which generates a sum of squared errors equal to that from the DLN method. This approach has the disadvantage that one has to rely on the DLN estimation when deciding λ. Therefore we use another procedure choosing the value of the smoothing parameter by implementing a cross-validation score (CVS) method in line with Craven and Wahba (1979). Using this approach, the problem of choosing the value of the smoothing parameter λ can be solved within the SPLINE method estimation. The outline of the procedure is as follows:

[3]The observations are equally weighted in our estimations.

[4]The minimizing problems of the DLN and the SPLINE methods have one common feature, namely, that the squared errors (defined as the observed value less the theoretical value) are taken into account. However, the SPLINE minimizing problem needs an additional feature to make sense, i.e., a restriction to what extent the curvature may fluctuate. Otherwise the SPLINE method would always render a perfect fit to the observed values, which brings about that the estimation becomes extremely sensitive to data-outliers.

- Loop over all possible values of λ.[5] In each loop observations are deleted one by one and a SPLINE function is estimated from the remaining observations.[6] Then the squared difference between the deleted observations and the values generated by the SPLINE function is computed, i.e, the cross-validation score. Hence, the CVS can be as

$$\text{CVS}(\lambda) = \sum_{i=1}^{n} (y(i) - g_\lambda(t_i))^2$$

  where $y(i)$ is the actual observation as a function of t_i and $g_\lambda(t_i)$ is the smoothing spline estimated from the data pairs excluding $(t_i, y(i))$.

- The sum of squared errors for each λ is compared, and the smoothing parameter with the smallest sum of squared errors is selected.

The CVS procedure generally yields a high value of λ in the range 0.98–0.9999, which is in line with Bliss and Panigirtzoglou (2001a), who test different smoothing parameters when employing RNDs as a forecasting tool. The fact that the CVS procedure yields a high λ will have a significant impact on our results because the width of the confidence bands is used as the main criterion when deciding on which method to prefer. A high λ gives small error terms and because we draw with replacement from these error terms (bootstrap) when estimating the confidence bands the resulting bands will obviously be quite narrow. Having said that, we believe that the CVS method is the most appropriate way to proceed since no arbitrary a priori assumption about the λ is necessary.

After choosing the smoothing parameter and estimating the SPLINE function in the implied volatility/delta space, the data are transformed back to the price/strike price space using the Black-Scholes formula[7]. The approximated function outside the observed range of strike prices is estimated assuming constant implied volatility.[8]

[5] We choose to loop over values of λ starting at 0.1 with a distance of 0.1. For higher values of λ (close to 1) the distance decreases. The reason for this is that the global minimum for the CVS appears close to 1.

[6] In order to avoid extrapolating the SPLINE function, the end observations are excluded in the CVS procedure.

[7] Note that the Black-Scholes formula is only used as a tool when transforming data between different spaces. This may seem a bit odd since a volatility smile implies that the Black-Scholes formula is not valid. However, we use the Black-Scholes merely as a computing tool between spaces, keeping the characteristics of the observed price quotations.

[8] Thus, we extrapolate the implied volatility from the end nods. This assumption is crucial in general RND estimation because it determines the shape of the tails of the RND. However, it is of minor importance when measuring the width of the confidence interval bands.

The RND is then found by a direct application of the Breeden and Litzenberger (1978) results, see equation (1).

2.2 Confidence Intervals

The option prices that are used as inputs in the RND estimations may not always be correct due to a number of potential sources of error. The need to quantify the uncertainty surrounding the estimated RND can be solved by estimating confidence bands, see Bliss and Panigirtzoglou (2001a) for a further discussion. Some important sources of error are:

- Lack of liquidity for options deep-in-the-money and deep-out-of-the-money may create "mis-pricing", since option pricing models do not take any liquidity premium into account.
- The utilization of settlement prices. If there is no trade in an option during a trading day the settlement price is a theoretical price calculated by the exchange. Therefore, if settlement prices are employed, the sample will most likely include strikes that were not actually traded.
- A narrow spectrum of strike prices. In most cases there are no observations far from at-the-money. Hence the tails of the RND are totally dependent on the estimation method.
- Pure data errors stemming from erroneous recording of prices.

The problem of lack of liquidity could be handled by using only the most liquid strikes. However, such a procedure would emphasize the problem concerning a narrow spectrum of strikes. If the estimation is based on settlement prices, there will be quite a large number of non-traded strikes in the data. Hence, in order to better reflect reality, non-traded strikes should be excluded. If the estimation is based on real-time snapshot quotations, the problem of large bid-ask spreads has to be carefully monitored. For instance, using the average of bid and ask quotations might generate fictitious arbitrage opportunities. An inherent problem in all options markets is that a rather narrow spectrum of strikes is actually traded. Therefore excluding strikes on the basis of liquidity considerations may introduce even more uncertainty about the estimated RND. Since there is a vast amount of option price data to be recorded, pure data errors are not unlikely to appear. Screening for theoretical arbitrage opportunities that exceed what could be deemed as reasonable from a transaction cost point of view is one way to reduce the influence of pure data errors.

Nevertheless, far from all sources of error can be eliminated, and thus there will always be some uncertainty about the estimated RND. All in all, the uncertainty should be quantified in an appropriate manner. The most straightforward way is probably to define the theoretically correct price and compare it to the observed price. The theoretically price should be decided by the model underlying the RND estimation. The pricing error could be extracted by taking the observed price less the theoretical price. Confidence bands can be estimated in several ways, for instance by Monte Carlo simulation or by bootstrap. As discussed later, two new bootstrap procedures will be used in this paper. The first method involves resampling from historical pricing errors, while the second method involves resampling from current pricing errors. The details of the bootstrap procedures are discussed Section 3.1.

2.3 Previous Literature on Confidence Interval Estimation and the Way to Proceed

Earlier work on the uncertainty of estimated RNDs has mainly focused on inflicting some kind of perturbations on the data or the estimated parameters and then repeating this procedure until it is possible to extract confidence bands. Söderlind and Svensson (1997) assumed that the correct model was a mixture of lognormal distributions and that the discrepancies between observed prices and theoretical prices were due to random error terms. Thus the sum of squared pricing errors was minimized by non-linear least squares estimation in order to retrieve the parameters of the distribution functions. A heteroscedastic-consistent estimator was implemented to calculate the covariance matrix taking the heteroscedastic price errors into account. Finally, a confidence interval was estimated by applying the delta method. Even though the problem with heteroscedastic price errors was addressed by implementing a heteroscedastic-consistent estimator of the covariance matrix, the asymptotic normality assumption underlying the delta method may lead to underestimation of the width of the confidence bands.

Melick and Thomas (1998) proceeded in the same realm as Söderlind and Svensson, constructing a 95 percent confidence interval for implied RNDs of the Euromark future. The parameters were estimated by constrained maximum likelihood. A Monte Carlo experiment was performed where 500 pseudo-distributions were created from the estimated DLN. The resulting confidence interval conveyed the message that there was little uncertainty about the estimated RNDs. Melick and Thomas also found, however, that the error terms were not likely to be independent. Hence the tight confidence bands might originate from the fact that the normality assumption underlying the Monte Carlo technique was not fulfilled. To avoid imposing any structure on the error terms Melick and Thomas created pseudo samples by drawing with replacement from the available data (bootstrap). The resulting confidence bands were quite wide, thus indicating large uncertainty about the estimated RNDs. As the pseudo samples were drawn with replacement, many pseudo samples had relatively large gaps between strikes, which caused the DLN method to run into some rather serious estimation problems. Furthermore, Melick and Thomas (1998) put a lower bound of 0.02 for the dispersion parameters, thereby avoiding running into problems associated with the optimization procedure. This lower bound keeps the estimation procedure from problems afflicted with discerning between observationally equivalent distributions. However, it might be the case that it is necessary to impose more far-reaching restrictions in order to arrive at reasonably shaped RNDs. Altogether the strange-looking confidence bands of the Melick and Thomas (1998) bootstrap simulation are mainly due to estimation problems. Hence, the resulting confidence interval is a poor approximation of the true uncertainty of the implied RND. Melick and Thomas (1998) concluded that estimation problems stemming from relatively large gaps between the strike prices in the data was the main reason for the wide confidence interval.

Söderlind (2000) took a somewhat different approach when estimating the confidence bands of the implied RNDs resulting from the DLN. By adding error terms to the theoretical option and forward prices implied by the original parameter estimates, 100 simulated data sets were created and utilized to estimate a confidence interval. Söderlind chose two methods to generate the error terms: the first method involved generating pseudo-random numbers from an i.i.d. normal distribution, and in the second method bootstrapped the original data set. The simulated 90 percent confidence intervals of both methods for the underlying contract (short sterling) were found to be very narrow. In the Monte Carlo experiment by Söderlind (2000), the error terms were drawn from an i.i.d. normal distribution with second moment equal to the estimated variance for the corresponding dataset. Thus any heteroscedasticity and/or non-normality of the pricing error distribution was unaccounted for, which might be one reason for the narrow confidence bands.

The bootstrap experiment by Söderlind (2000) was conducted in such a manner that the observed error terms were resampled and added to the theoretical prices. This solves one of the problems with the Melick and Thomas (1998) approach, namely, the generation of large gaps between strike prices. However, the problem of heteroscedastic error terms still remains unsolved.

Bliss and Panigirtzoglou (2001a) went down yet another road and perturbed the observed option prices rather than the theoretical prices. The simulated prices were attained by adding a uniformly distributed random perturbation of between plus one half and minus one half of the contract's tick size. In the light of the price perturbations the robustness of the DLN and the smoothing spline method were investigated. Bliss and Panigirtzoglou (2001a) concluded that there was strong evidence of superior stability of the SPLINE method over the DLN and that the confidence intervals for both methods were sometimes so large that the RND estimates became useless. All in all the DLN was deemed to be inferior to the SPLINE method.

To sum up, earlier work has recognized that the error terms are unlikely to be normally distributed. Still, the Monte Carlo simulations have been conducted under the assumption that the normality assumption for the price errors is valid. Thus, one of the tasks of this article is to check whether the error terms are normally distributed by conducting Bera-Jarque tests. If the null hypothesis of normality can be rejected, then a non-parametric method will be used to construct the confidence interval, taking the non-normality into account (a thorough description follows in Section 3).

3 Confidence Band Estimation

The previous literature describing different procedures of how to compute confidence bands, outlined in Section 2, has a number of drawbacks that could be improved upon. Hence this section is devoted to describing, in detail, how this could be accomplished by employing two different bootstrap techniques. The first bootstrap method employs historical pricing errors, while the second method uses current pricing errors. The former approach has the advantage of capturing the non-normal and heteroscedasticity characteristics of the error terms. On the other hand, the estimated confidence bands will not differ much between different data sets. This disadvantage can be resolved by bootstrapping the actual error terms, which motivates including the latter bootstrap method. This approach has its drawbacks as well, since it not fully taking the heteroscedastic property into consideration. Finally, the DLN and SPLINE methods are examined and the width of the generated confidence bands is utilized as a tool when deciding which one of the two methods to prefer from a precision point of view. The width of the confidence bands is chosen since we believe this measurement is best suited for our purposes when we want to quantify and visualize the uncertainty connected with the RND estimation methods.

3.1 Bootstrap from Historical Errors

The first bootstrap method proposed is much in line with Melick and Thomas (1998), who estimated confidence bands using Monte Carlo simulations of the DLN method. Melick and Thomas (1998) recognized that the error terms were not independent of option type (call/put) or strike price. This is not unique for the data set of Melick and Thomas (German short-term interest rates), but is also present in the Swedish stock market options (OMX) employed in our study. Another non-appealing property of the error structure is the non-normality. Melick and Thomas chose not to take any of these features into consideration in their Monte Carlo simulations. This paper takes the observed dependence into account by utilizing the historical pattern of the error terms, generated by running DLN estimations from January 1993 until June 2001 with 30 days until expiration.[9]

To deal with the issue that the strike price range differs over time (when the value of the underlying index varies) the data were grouped[10] by relative strike price where

$$\text{relative strike price } (X) = \left(\frac{\text{strike price}}{\text{future price}} - 1\right). \tag{8}$$

The historical pattern of the DLN error structure is illustrated in Figure 2.

Clearly, volatility varies not only across strike prices but also depending on option type (call/put). In the SPLINE method the data were grouped in the space where the smoothing SPLINE function was estimated, i.e., the implied volatility/delta space. The data were grouped over the range of deltas and Figure 3 shows that for the SPLINE method the error structure varies depending on option type and whether the option is in-the-money or not.[11]

One way to estimate confidence bands is to draw errors for each strike price, assuming normality with zero mean and a variance equal to its corresponding group variance in line with Söderlind (2000). However, the normality assumption under-

[9] There are generally between 10 and 15 different strikes available for each day.

[10] The grouping procedure was conducted in a somewhat arbitrary manner. The error terms were plotted and put together so that a sufficiently large number of observations were included in each group. This yielded twelve groups for the DLN method with an average of 118 error observations, and eleven groups for the SPLINE method with an average of 147 observation.

[11] The same type of non-normal pattern, with regards to both the DLN and the SPLINE method, was observed for the put errors.

Figure 2: Historical errors, DLN method, call

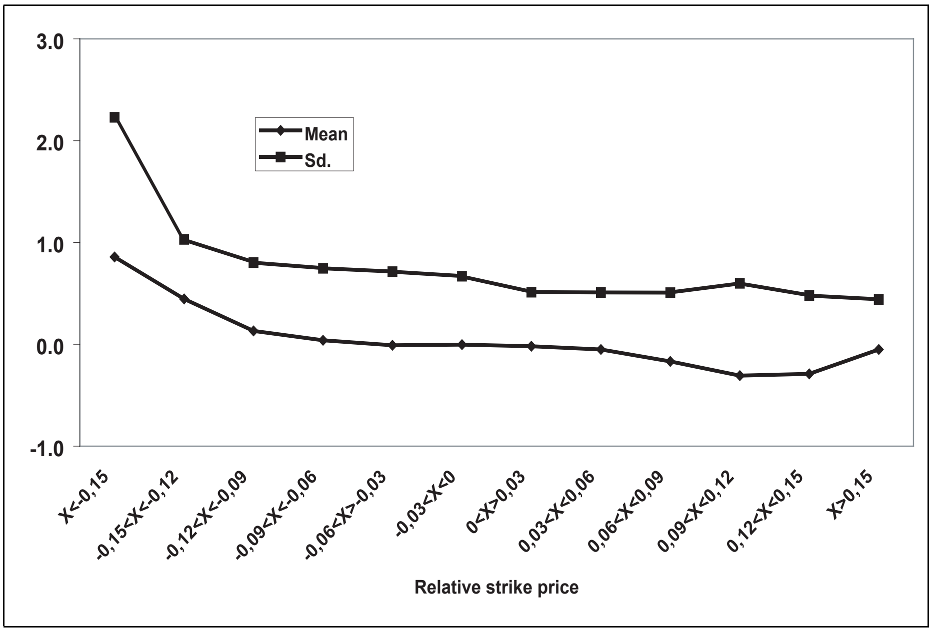
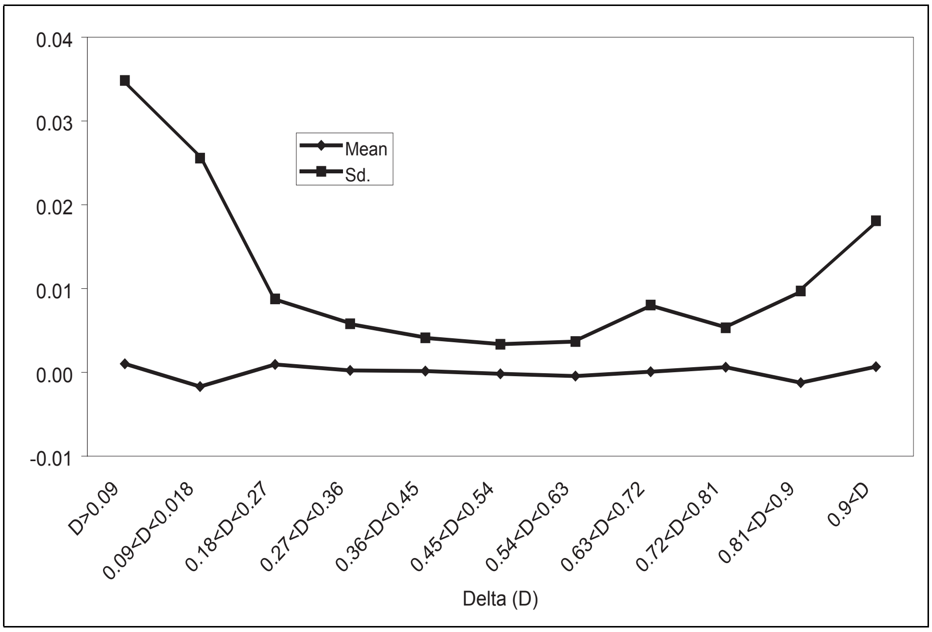


Figure 3: Historical errors, SPLINE method, call



lying the Monte Carlo approach is tested by a Bera-Jarque test. This test rejects normality, at the five percent significance level, for all groups except two for the DLN case[12] and all groups for the SPLINE case. We therefore abstain from assuming normality and instead propose an alternative procedure where the following steps are performed for each pseudo distribution in order to estimate the confidence bands and take the non-normality of the error structure into account:

- Compute the relative strike price/delta for each call and put price depending on whether the DLN or SPLINE is used in the estimation.
- For each observation, draw an error term with replacement from its corresponding group (bootstrap within each group).
- Add the errors to the theoretical prices in order to get pseudo prices and run the RND estimation.

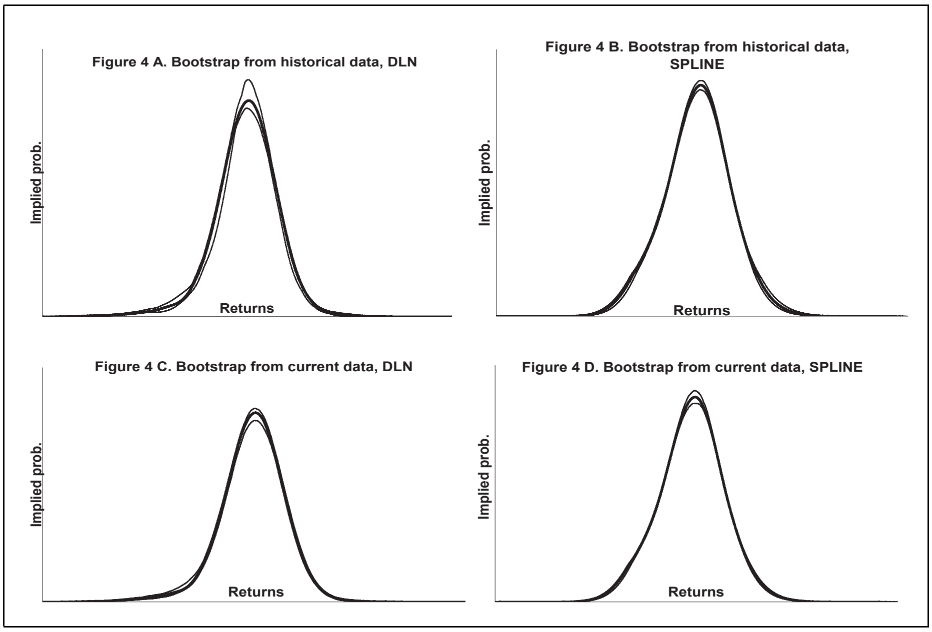
These steps are repeated 500 times yielding 500 pseudo RND distributions which are used to extract a 95 percent confidence interval.

3.2 Bootstrap from Actual Error Terms

The second bootstrap experiment is conducted in line with Söderlind (2000), drawing with replacement from the current error terms. The pseudo-samples are created by adding the bootstrapped error terms to the theoretical prices. However, the fact that the structure of the error terms differs across strike prices/delta depending of estimation method implies that the problem with heteroscedasticity still remains. The small number of observations in this approach puts a restriction on the number of groups since sub-samples with too few observations invalidate the assumptions underlying the bootstrap methodology. Hence the error terms are grouped depending on option type and whether they are in-the-money or not. This will have the effect that when an error term is drawn, for instance, for a call price that is deep in-the-money the error term will have the characteristics reflecting call prices deep in-the-money. The steps of estimating the confidence interval, apart from the error terms on which to base the confidence interval, follow the same steps as described in Section 3.1.

[12] These were the groups with relative strike price between $(-0.15, -0.12)$ and $(0.12, 0.15)$ respectively.

Figure 4: Confidence interval based on OMX options 1/24/2001



4 Results

In this section the results of the confidence band estimation will be described and displayed visually. Furthermore, we measure the width of the confidence bands and conclude that the SPLINE method is to be favored from a precision point of view, as it yields tighter bands than the DLN method. The section ends with a case study showing how estimated confidence intervals can be used for practical purposes.

As mentioned earlier, two different methods of estimating confidence bands are employed: the bootstrap using historical data, and the bootstrap on current data with the DLN and SPLINE as the RND estimation methods. In order to illustrate how the confidence bands might look like for the different approaches, a randomly chosen data set (trade date 1/24/2001) has been utilized as the basis for the illustration and the resulting Figures can be examined below. The procedure was repeated for a wide number of data sets yielding the same conclusions as those drawn below.

4.1 Bootstrap from Historical Data

Figures 4A and 4B display the estimated RNDs for the DLN and SPLINE methods, and the confidence intervals based on historical data. As shown in Figure 4A the confidence bands for the DLN method appear to be fairly narrow. Nonetheless, the confidence bands are wider than in the Monte Carlo experiments performed by Melick and Thomas (1998), Söderlind and Svensson (1997) and Söderlind (2001).

This is likely to be explained by the fact that these experiments did not address the non-normal and heteroscedastic features of the pricing errors.

The confidence bands stemming from the SPLINE method seem to be narrower than the DLN counterparts, see Figure 4B. This is much in line with the results of Bliss and Panigirtzoglou (2001a), which state that the SPLINE method is more precise than the DLN method. Nevertheless, the confidence interval is still wider than the confidence interval estimated using the Monte Carlo technique. Hence the non-normality and heteroscedastic features appear to matter for the SPLINE method as well.

4.2 Bootstrap from Current Data

Figures 4C and 4D show the estimated RNDs for the DLN and SPLINE methods, and the confidence intervals based on current data. The confidence bands for the DLN method are quite narrow compared with the results of Melick and Thomas (1998), see Figure 4C. This is preponderantly due to the fact that the DLN estimation in this paper are based on more restrictive constraints on the dispersion parameters for the log-normal distributions, thus not allowing for spurious-looking RNDs in the sample. On the other hand it appears as the confidence interval is wider than the counterpart of Söderlind (2000). This might be due to the fact that Söderlind (2000) did not consider the problem with heteroscedastic pricing errors. The confidence interval for the SPLINE method seems to be more narrow than for the DLN equivalence, thus again validating the results of Bliss and Panigirtzoglou (2001a), see Figure 4D.

To sum up, earlier work on confidence bands estimation that yielded quite narrow confidence bands is problematic since the non-normal and the heteroscedastic features were not addressed. Moreover, the bootstrap experiment conducted by Melick and Thomas (1998) suffered from the inclusion of spurious looking RNDs in the pseudo-sample. Söderlind (2000) addressed all problems but one: the heteroscedastic nature of the pricing errors. The results of this paper indicate that both heteroscedasticity and non-normality of pricing errors should be accounted for in order to, as accurately as possible, quantify the uncertainty of the estimated RNDs. Finally, the SPLINE method seems to be more precise than the DLN method.

4.3 Quantitative Results

Thus far, this paper has tentatively suggested that the SPLINE method is more robust than the DLN method since it seems to produce tighter confidence bands.

Table 1: Evaluation of estimation methods

	Method	Average width of the confidence bands	Average number of spurious densities
Bootstrap 1 (historical data)	Spline	1	0
	DLN	2	247.6
Bootstrap 2 (current data)	Spline	1	0
	DLN	4.1	269.7

Note: The average width is in relative terms, i.e., the SPLINE is normalized to one.

In order to verify this we choose the following procedure: Seven randomly chosen datasets[13] were estimated for the two estimation methods and the two different bootstrap techniques. The mean of the width of the confidence bands is used this as the evaluation tool and the results are presented in Table 1. This study confirms the tentative suggestion that the SPLINE method is to be preferred from a precision point of view. In this context it is worth mentioning that the bootstrap techniques adopted in this study may, in some cases, have the inconvenient feature that the RND will become negative very far out in the tails. However, since these probabilities are extremely small it does not affect the interpretation and evaluation of our results.

Table 1 shows that the SPLINE method consistently produces tighter confidence bands than the DLN method, both for the bootstrap from historical data and the bootstrap from current data. The bootstrap from historical data yields confidence bands for the DLN method that are twice as wide as the SPLINE counterpart. The difference is even more noticeable in the bootstrap from current data: the DLN method delivers bands that are approximately four times wider. The width of the confidence bands, we believe, ought to be the main criterion when choosing between DLN and SPLINE as the estimation method.

However, for the practitioner other criteria such as computational efficiency and the rate of convergence should be taken into consideration as well before deciding upon estimation method. The SPLINE method optimization routine always converges and hence the whole pseudo sample is intact when estimating the confidence interval. This differs considerably from the DLN method, where a great many RNDs in the pseudo sample have to be omitted in order to avoid spurious-looking confidence intervals. In addition, the SPLINE method also consistently requires less computing time than the DLN method. Thus from a practitioner's point of view

[13]One data set was drawn from a uniform distribution for each year during the period 1995–2001.

there should be little doubt that the SPLINE method is to be preferred.

A further issue is what bootstrap technique to apply. We generally believe that the bootstrap from actual error terms is superior. Estimating confidence bands from historical error terms has the drawback of not being able to capture the specific characteristics of the data set at hand. On the other hand, if only a few strike prices are available, there will be hardly no variation from pseudosample to pseudosample when drawing with replacement, thus invalidating the assumptions underlying the bootstrap methodology. In that case the bootstrap from historical data is to be preferred to the bootstrap from current data.

5 Case Study: Did the ECB and the BoE Interest Rate Cuts on November 8, 2001, Alter Stock Market Expectations?

In this section an example is provided of how the estimated confidence bands may be applied as a practical tool in order to quantify whether the perception of risk has changed significantly or not when new information hits the market.

The intention of this exercise is to examine if investors changed their perception, or "true" probabilities, for the future outcome concerning the Swedish stock market after the 50 basis point interest rate cut by the ECB and Band of England (BoE) on November 8, 2001. Moreover, if the market's perception towards risk remains unchanged one can interpret the cuts to have been well foreseen by market participants. Hence it is possible by this approach to measure the degree of central bank transparency with respect to the market's risk assessment.

It is important to note that the estimated implied RND derived from observed option prices reflects the market's "true" perception concerning the future density if and only if investors are risk-neutral. This qualification can lead to complications in interpreting the "true" message in RNDs. To illustrate this, suppose an RND is estimated with a stock market index as the underlying asset. If the RND has a negative skewness coefficient (i.e., it is skewed to the left), it is not possible to distinguish the extent to which this reflects a large degree of risk aversion on the part of investors or the extent to which the market attaches a high probability to a sharp downward correction of the stock market. However, if one assumes that the degree of risk aversion over shorter time periods is broadly constant, then changes in RNDs over short periods of time can be interpreted as changes in market expectations concerning higher moments of the value of the underlying asset.

This exercise starts out by estimating an RND and its corresponding confidence band on OMX (broad base Swedish stock market index) November 8, 2001. During this day both the ECB and the BoE cut their interest rates by 50 basis points (ECB at 1.45pm and the BoE at 1.00pm, Central European Time). A second RND was estimated when the markets opened the day after on November 9. By computing the confidence interval around the RND estimated before the cuts and investigating whether the RND estimated after the easening of monetary conditions falls within the confidence bands, a sense of whether the subsequent reaction of stock market uncertainty was statistically significant or not can be provided. If the RND estimated after the event were to lie outside the confidence bands this would indicate that the market did significantly change its expectations concerning higher moments of the value of the underlying asset. Although comparisons like these have been made in the literature before, to our knowledge no attempts have been made to quantify whether the changes are statistically significant or not.

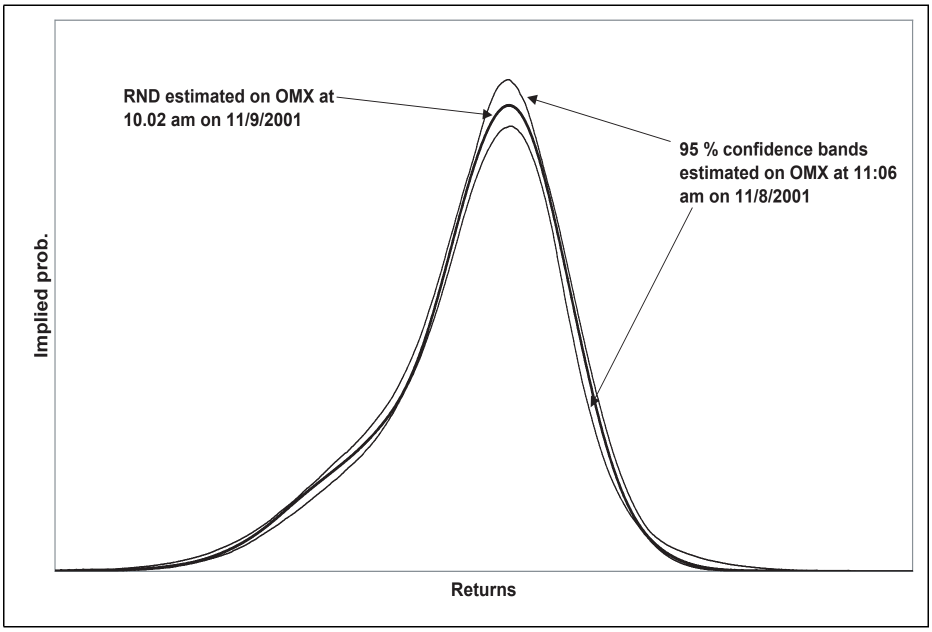
In this exercise the confidence bands were estimated by bootstrapping from current data applying the DLN estimation technique.[14] Figure 5 below shows the RND estimated at 10.02 am the day after the decisions together with the 95 percent confidence bands around the RND estimated at 11.06 am on the day of the interest rate cuts. The RND estimated after the interest rate cuts falls inside the 95 percent confidence bands. This can be interpreted as indicating market participants did not significantly alter their view about the uncertainty of the Swedish stock market outlook following the decisions.

6 Summary and Conclusions

This paper shows that is important to take the non-normal and the heteroscedastic features of the pricing errors into consideration when computing confidence bands for the estimated RND. Two different statistical methods have been applied: bootstrap from historical pricing errors and bootstrap from current pricing errors. Both methods resulted in less narrow, and perhaps more plausible confidence bands, than earlier work have provided. Furthermore, two techniques of estimating RNDs, the SPLINE and the DLN methods, have been implemented in the confidence interval analysis. The SPLINE technique seems to produce more accurate RNDs, since the confidence intervals are more narrow than the DLN counterpart. In addition, the DLN technique produces some computational drawbacks that can be avoided by im-

[14]Hence, the DLN method is chosen even though the SPLINE method is to be preferred generally. The reason is that the DLN method generates wider confidence bands and hence better illustrates the task at hand.

Figure 5: RND and confidence bands



plementing the SPLINE method. The SPLINE method also requires somewhat less computing time than the DLN method. Finally, estimation of the confidence bands provides an opportunity to deliberate upon whether a shift in the shape of the RND over a short period of time can be attributed to increased perceived uncertainty about the development of the underlying asset.

A Data Appendix

The data used in this study consist of equity options on the Swedish stock market (OMX) index. The options are European style, and in this study options with 30 days to maturity are utilized which is favorable from a liquidity point of view. An average of the closing bid and ask quotations is used as the price proxy. The data has, on daily basis, been downloaded from Reuter during the trading day, thereby avoiding the problems arising from using settlement prices, mentioned in Section 2.1.

The OMX is traded at the Swedish Open Market (OM) and the contracts mature the fourth Friday of each month. The OMX dataset stretches from January 1992 until November 2001. The options are traded under a rolling schedule which means that there is one contract with 30 days to maturity once a month. The smallest tick size is 0.01.

The option contract has the index as the underlying asset but the future contract matures at the same time as the option, which means that the future could be used as a proxy for the underlying asset. The reliability of the data can be questioned from an arbitrage point of view, especially when perturbing the theoretical price in the confidence interval estimations. Computing implied call and put prices from the put/call parity indicates that the parity clearly does not hold for some of the observations. However, in most cases, the deviations are small enough to be explained by transaction costs.

References

Bahra, B. (1997), "Implied Risk-Neutral Probability Density Functions from Option Prices," *Bank of England Working Paper* No. 66.

Bates, D. (1991), "The Crash of '87—Was It Expected? The Evidence from Options Markets," *Journal of Finance* 46(3), 1009–1044.

Björk, T. (1998), "Arbitrage Theory in Continuous Time," *Oxford University Press.*

Bliss, R.R. and Panigirtzoglou (2001a), "Testing the Stability of Implied Probability Density Functions,"*Journal of Banking and Finance*, forthcoming.

Bliss, R.R. and Panigirtzoglou (2001b), "Recovering Risk aversion from Options," Unpublished Working Paper.

Breeden D.T. and Litzenberger R.H. (1978), "Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business* 51(4), 621–51.

Craven, P. and Wahba, G. (1979), "Smoothing Noisy Data With Spline Functions," *Numerische Matematik* 31, 377–403.

Malz, A. M. (1997), "Estimating the Probability Distribution of the Future Exchange Rate from Option Prices," *Journal of Derivatives* 5, 18–36.

Melick, W.R. and Thomas C.P. (1997), "Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil During the Gulf Crisis," *Journal of Financial and Quantitative Analysis* 32(1), 91–115.

Melick, W.R. and Thomas C.P. (1998), "Confidence Intervals and Constant Maturity Series for Probability Measures Extracted from Option Prices," Conference paper, Bank of Canada.

Ritchey. R.J. (1990), "Call Option Valuation for Discrete Normal Mixtures," *Journal of Financial Research* 13, 285–296.

Schimko, D.C. (1993), "Bounds of probability," *Risk*, 6(4), 33–37.

Söderlind, P. (2000), "Market Expectations in the UK Before and After the ERM Crisis," *Economica* 67, 1–18.

Söderlind, P. and Svensson, L.E.O. (1997), "New Techniques to Extract Market Expectations from Financial Instruments," *Journal of Monetary Economics* 40, 383–429.

Earlier Working Papers:

Sveriges Riksbank
Visiting address: Brunkebergs torg 11
Mail address: se-103 37 Stockholm

Website: www.riksbank.se
Telephone: +46 8 787 00 00, Fax: +46 8 21 05 31
E-mail: registratorn@riksbank.se

ISSN 1402-9103 Riksbank Printing Office Stockholm 2003