

The Hidden Dangers of Historical Simulation

Matthew Pritsker*

June 19, 2001

Abstract

Many large financial institutions compute the Value-at-Risk (VaR) of their trading portfolios using historical simulation based methods, but the methods' properties are not well understood. This paper theoretically and empirically examines the historical simulation method, a variant of historical simulation introduced by Boudoukh, Richardson and Whitelaw (1998) (BRW), and the Filtered Historical Simulation method (FHS) of Barone-Adesi, Giannopoulos, and Vosper (1999). The Historical Simulation and BRW methods are both under-responsive to changes in conditional risk; and respond to changes in risk in an asymmetric fashion: measured risk increases when the portfolio experiences large losses, but not when it earns large gains. The FHS method appears promising, but requires additional refinement to account for time-varying correlations; and to choose the appropriate length of historical sample period. Preliminary analysis suggests that 2 years of daily data may not contain enough extreme outliers to accurately compute 1% VaR at a 10-day horizon using the FHS method.

*Board of Governors of the Federal Reserve System, and University of California at Berkeley. Address correspondence to Matt Pritsker, The Federal Reserve Board, Mail Stop 91, Washington DC 20551. Alternatively, Matt Pritsker can be reached by telephone at (202) 452-3534, or (510) 642-0829 or Fax: (202) 452-3819, or by email at mpritsker@frb.gov.

1 Introduction

The growth of the OTC derivatives market has created a need to measure and manage the risk of portfolios whose value fluctuates in a nonlinear way with changes in the risk factors. One of the most widely used of the new risk measures is Value-at-Risk, or VaR.¹ A portfolio's VaR is the most that the portfolio is likely to lose over a given time horizon except in a small percentage of circumstances. This percentage is commonly referred to as the VaR confidence level. For example, if a portfolio is expected to lose no more than \$10,000,000 over the next day, except in 1% of circumstances, then its VaR at the 1% confidence level, over a one-day VaR horizon is \$10,000,000. Alternatively, a portfolio's VaR at the k % confidence level is the k 'th percentile of the distribution of the change in the portfolio's value over the VaR time horizon.

The main advantage of VaR as a risk measure is that it is very simple: it can be used to summarize the risk of individual positions, or of large multinational financial institutions, such as the large dealer-banks in the OTC derivatives markets. Because of VaR's simplicity, it has been adopted for regulatory purposes. More specifically, the 1996 Market Risk Amendment to the Basle Accord stipulates that banks' and broker-dealers' minimum capital requirements for market risk should be set based on the ten-day 1-percent VaR of their trading portfolios. The amendment allows ten-day 1-percent VaR to be measured as a multiple of one-day 1-percent VaR.

Although VaR is a conceptually simple measure of risk, computing VaR in practice can be very difficult because VaR depends on the joint distribution of all of the instruments in the portfolio. For large financial firms which have tens of thousands of instruments in their portfolios, simplifying steps are usually employed as part of the VaR computation. Three steps are commonly used. First the dimension of the problem is reduced by modeling the change in the value of the instruments in the portfolio as depending on a smaller (but still large) set of risk factors f . Second the relationship between f and the value of instruments which are nonlinear functions of f is approximated where necessary.² Finally, an assumption about the distribution of f is required.

The errors in VaR estimation depend on the reasonableness of the simplifying assumptions. One of the most important assumptions is the choice of distribution for the risk factors. Many large banks currently use or plan to use a method known as historical simulation to model the distribution of their risk factors. The distinguishing feature of the historical simulation method and its variants is that they make minimal parametric assumptions about

¹For a review of the early literature on VaR, see Duffie and Pan (1997).

²For instruments that require large amounts of time to value, it will typically be necessary to approximate how the value of these instruments change with f in order to compute VaR in a reasonable amount of time.

the distribution of f , beyond assuming that the distribution of changes in value of today's portfolio can be simulated by making draws from the historical time series of past changes in f .

The purpose of this paper is to conduct an in-depth examination of the properties of historical simulation based methods for computing VaR. Because of the increasing use of these methods among large banks, it is very important that market practitioners, and regulators understand the properties of these methods and ways that they can be improved. The empirical performance of these methods has been examined by Hendricks (1996), and Beder (1995) among others. In this paper I study the performance of the historical simulation method, as well as two variants: the BRW method [Boudoukh, Richardson, and Whitelaw (1998)] and the Filtered Historical Simulation Method (Barone-Adesi, Giannopoulos, and Vosper, 1999). In a related paper, Hull and White (1998) empirically examine the performance of the historical simulation method, the BRW method, and the Hull-White method, which is a variant of the Filtered Historical Simulation method.³ The analysis here departs from related work on the empirical properties of the VaR methods in two ways. First, I analyze the historical simulation based estimators of VaR from a theoretical as well as empirical perspective. The theoretical insights aid in understanding the deficiencies of the historical simulation and BRW methods. Second, the earlier empirical analysis of these methods was based on how the method performed with real data. A disadvantage of using real data to examine the methods is that since true VaR is not known, the quality of the VaR methods, as measured by how well they track true VaR, can only be measured indirectly. As a result it is very difficult to quantify the errors associated with a particular method of measuring VaR when using real data. In my empirical analysis, I analyze the properties of the historical simulation method's estimates of VaR with artificial data. The artificial data are generated based on empirical time series models that were fit to real data. The advantage of working with the artificial data is that true VaR is known. This makes it possible to much more closely examine the properties of the errors made when estimating VaR using historical simulation.

Because my main focus in this paper is on the distributional assumptions used in historical simulation methods, in all of my analysis, I abstract from other sources of error in VaR estimates. More specifically, I only examine VaR for simple spot positions in underlying stock indices or exchange rates. For all of these positions, there is no possibility of choosing incorrect risk factors, and there is no possibility of approximating the nonlinear relationship between instrument prices and the factors incorrectly. The only sources of error in the VaR

³The Filtered Historical Simulation and Hull-White methods are the same when the time horizon over which VaR is calculated is one period.

estimates is the error associated with the distributional assumptions.

Before presenting my results on historical simulation based methods, it is useful to illustrate the problems with the distributional assumptions associated with historical simulation. The distributional assumptions used in VaR, as well as the other assumptions used in a VaR measurement methodology, are judged in practice by whether the VaR measures provide the correct conditional and unconditional coverage for risk [Christofferson (1998), Diebold, Gunther, and Tay (1998), Berkowitz (1999)]. A VaR measure achieves the correct unconditional coverage if the portfolio's losses exceed the k percent VaR measures $k\%$ percent of the time in very large samples. Because losses are predicted to exceed k -percent VaR k -percent of the time, a VaR measure which achieves correct unconditional coverage is correct on-average. A more stringent criterion is that the VaR measure provides the correct conditional coverage. This means that if the risk, and hence the VaR of the portfolio changes from day to day, then the VaR estimate needs to adjust so that it provides the correct VaR on every day, and not just on average.

It is probably unrealistic to expect that a VaR measure will provide exactly correct conditional coverage. But, one would at least hope that the VaR estimate would increase when risk appears to increase. In this regard, it is useful to examine an event where risk seems to have clearly increased, and then examine how different measures of VaR respond. The simplest event to focus on is the stock market crash of October 19, 1987. The crash itself seemed indicative of a general increase in the riskiness of stocks, and this should be reflected in VaR estimates.

Figure 1 provides information on how three historical simulation based VaR methods performed during the period of the crash for a portfolio which is long the S&P 500. All three VaR measures use a one-day holding period and a one-percent confidence level.

The first VaR measure uses the historical simulation method. This method involves computing a simulated time series of the daily P & L that today's portfolio would have earned if it was held on each of N days in the recent past. VaR is then computed from the empirical CDF of the historically simulated portfolio returns.

The principle advantage of the historical simulation method is that it is in some sense nonparametric because it does not make any assumptions about the shape of the distribution of the risk factors that affect the portfolio's value. Because the distribution of risk factors, such as asset returns, is often fat-tailed, historical simulation might be an improvement over other VaR methods which assume that the risk factors are normally distributed.

The principle disadvantage of historical simulation method is that it computes the empirical CDF of the portfolios returns by assigning an equal probability weight of $1/N$ to each day's return. This is equivalent to assuming that the risk factors, and hence the historically

simulated returns are independently and identically distributed (i.i.d.) through time. This assumption is unrealistic because it is known that the volatility of asset returns tends to change through time, and that periods of high and low volatility tend to cluster together [Bollerslev (1986)].

When returns are not i.i.d., it might be reasonable to believe that simulated returns from the recent past better represent today portfolio's risk than returns from the distant past. Boudoukh, Richardson, and Whitelaw (1998), BRW hereafter, used this idea to introduce a generalization of the historical simulation method in a way that assigns a relatively high amount of probability weight to returns from the recent past. More specifically, BRW assigned probability weights that sum to 1, but decay exponentially. For example, if λ , a number between zero and 1, is the exponential decay factor, and $w(1)$ is the probability weight of the most recent historical return of the portfolio, then the next most recent return receives probability weight $w(2) = \lambda * w(1)$, and the next most recent receives weight $\lambda^2 * w(1)$, and so on. After the probability weights are assigned, VaR is calculated based on the empirical CDF of returns with the modified probability weights. The historical simulation method is a special case of the BRW method in which λ is set equal to 1.

The analysis in figure 1 provides results for the historical simulation method when VaR is computed using the most recent 250 days of returns. The figure also presents results for the BRW method when the most recent 250 days of returns are used to compute VaR and the exponential decay factors are either $\lambda = 0.99$, or $\lambda = 0.97$. The size of the sample of returns and the weighting functions are the same as those used by BRW. The VaR estimates in the figure are presented as negative numbers because they represent amounts of loss in portfolio value. A larger VaR amount means that the amount of loss associated with the VaR estimate has increased.

The main focus of attention is how the VaR measures respond to the crash on October 19th. The answer is that for the historical simulation method the VaR estimate has almost no response to the crash at all (Figure 1 panel A). More specifically, on October 20th, the VaR measure is at essentially the same level as it was on the day of the crash. To understand why, recall that the historical simulation method assigns equal probability weight of $1/250$ to each observation. This means that the historical simulation estimate of VaR at the 1% confidence level corresponds to the 3rd lowest return in the 250 day rolling sample. Because the crash is the lowest return in the 250 day sample, the third lowest return after the crash turns out to be the second lowest return before the crash. Because the second and third lowest returns happen to be very close in magnitude, the crash actually has almost no impact on the historical simulation estimate of VaR for the long portfolio.

The BRW method involves a simple modification of the historical simulation method.

However, the modification makes a large difference. On the day after the crash, the VaR estimates for both BRW methods increase very substantially, in fact, VaR rises in magnitude to the size of the crash itself (Figure 1, panels B and C). The reason that this occurs is simple. The most recent P & L change in the BRW methods receive probability weights of just over 1% for $\lambda = 0.99$ and of just over 3% for $\lambda = 0.97$. In both cases, this means that if the most recent observation is the worst loss of the 250 days, then it will be the VaR estimate at the 1% confidence level. Hence, the BRW methods appear to remedy the main problems with the historical simulation methods because very large losses are immediately reflected in VaR.

Unfortunately, the BRW method does not behave nearly as well as the example suggests. To see the problem, instead of considering a portfolio which is long the S&P 500, consider a portfolio which is short the S&P 500. Because the long and short equity positions both involve a “naked” equity exposure, the risk of the two positions should be similar, and should respond similarly to events like a crash. Instead, the crash has very different effects on the BRW estimates of VaR: following the crash the estimated risk of the long portfolio increases very significantly (Figure 1, panels B and C), but the estimated VaR of the short portfolio does not increase at all (Figure 2, panels B and C). The estimated risk of the short portfolio did not increase until the short portfolio experienced significant losses in response to the markets partial recovery in the two days following the crash.⁴

The reason that the BRW method fails to “see” the short portfolio’s increase in risk after the crash is that the BRW method and the historical simulation method are both completely focused on nonparametrically estimating the lower tail of the $P&L$ distribution. Both methods implicitly assume that whatever happens in the upper tail of the distribution, such as a large increase in $P&L$, contains no information on the lower tail of $P&L$. This means that large profits are never associated with an increase in the perceived dispersion of returns using either method. In the case of the crash, the short portfolio happened to make a huge amount of money on the day of the crash. As a consequence, the VaR estimates using the BRW and historical simulation methods did not increase.

The BRW method’s inability to associate increases in P&L with increases in risk is disturbing because large positive returns and large negative returns are both potentially indicative of an increase in overall portfolio riskiness. That said, the GARCH literature suggests that the relationship between conditional volatility and equity index returns is asymmetric: conditional volatility increases more when index returns fall than when they rise. Because the BRW method updates risk based on movement in the portfolio’s $P&L$, and

⁴The short portfolio’s losses on October 20 exceeded the VaR estimate for that day. As a result, the VaR figure for October 21 was increased. This new VaR figure was exceeded on October 21, hence the VaR figure was increased again to its level on October 22.

not on the price of the assets, it can respond to this asymmetry in precisely the wrong way. For example, the short portfolio registers larger increases in risk when prices rise, than when they fall. This is just the opposite of the relationship suggested by the GARCH literature.

The sluggish adjustment of the BRW and historical simulation methods to changes in risk at the 1% level are much worse at the 5% level; and in this case the BRW method with $\lambda = 0.97$ and $\lambda = 0.99$ provide very little improvement above and beyond that of the historical simulation method. The strongest evidence for the problem is the number of days in October where losses exceed the 5% VaR limits. For example, for the long portfolio losses exceed the VaR limits on 7 of 21 days in October using historical simulation or BRW with $\lambda = 0.99$, and losses exceed the VaR limits on 5 days using the BRW method with $\lambda = 0.97$ (Figure 3). Losses for the short-portfolio exceed their limits as well, but the total number of times is fewer (Figure 4).

Sections 2 and 3 explore the properties of the historical simulation and BRW methods from a theoretical and empirical viewpoint. Section 4 examines a promising variant of the historical simulation method introduced by Barone-Adesi, Giannopoulos, and Vosper. Section 5 concludes.

2 Theoretical Properties of Historical Simulation Methods

The goal of this section is to derive the properties of historical simulation methods from a theoretical perspective. Because historical simulation is a special case of BRW's approach, all of the results here are derived for the BRW method; and hence they generalize to the historical simulation approach.

The simplest way to implement BRW's approach *without using their precise method* is to construct a history of N hypothetical returns that the portfolio would have earned if held for each of the previous N days, r_{t-1}, \dots, r_{t-N} , and then assign exponentially declining probability weights w_{t-1}, \dots, w_{t-N} to the return series.⁵ Given the probability weights, VaR at the C percent confidence level can be approximated from $G(.; t, N)$, the empirical cumulative distribution function of r based on return observations r_{t-1}, \dots, r_{t-N} .

⁵The weights sum to 1 and are exponentially declining at rate λ ($0 < \lambda \leq 1$):

$$\sum_{i=1}^N w_{t-i} = 1$$

$$w_{t-i-1} = \lambda w_{t-i}$$

$$G(x; t, N) = \sum_{i=1}^N 1_{\{r_{t-i} \leq x\}} w_{t-i}$$

Because the empirical cumulative distribution function (unless smoothed) is discrete, the solution for VaR at the C percent confidence level will typically not correspond to a particular return from the return history. Instead, the BRW solution for VaR at the C percent confidence level will typically be sandwiched between a return which has a cumulative distribution which is slightly less than C , and one which has a cumulative distribution that is slightly more than C . These returns can be used as estimates of the BRW method's VaR estimates at confidence level C . The estimate which slightly understates the BRW estimate of VaR at the C percent confidence level is given by:

$$BRW^u(t|\lambda, N, C) = \inf(r \in \{r_{t-1}, \dots, r_{t-N}\} | G(r; t, N) \geq C),$$

and the estimator which tends to slightly overstate losses is given by:

$$BRW^o(t|\lambda, N, C) = \sup(r \in \{r_{t-1}, \dots, r_{t-N}\} | G(r; t, N) \leq C).$$

where λ is the exponential weight factor, N is the length of the history of returns used to compute VaR, and C is the VaR confidence level.

In words, $BRW^u(t|\lambda, N, C)$ is the lowest return of the N observations whose empirical cumulative probability is greater than C , and $BRW^o(t|\lambda, N, C)$ is the highest return whose empirical cumulative probability is less than C .

The $BRW^u(t|\lambda, N, C)$ method is not precisely identical to BRW's method. The main difference is that BRW smooths the discrete distribution in the above approaches to create a continuous probability distribution. VaR is then computed using the continuous distribution. For expositional purposes, the main analytical results will be proven for the $BRW^u(t|\lambda, N, C)$ estimator of value at risk. The properties of this estimator are essentially the same as those for the estimator used by BRW, but it is much easier to prove results for this estimator.

The main issue that I examine in this section is the extent to which estimates of VaR based on the BRW method respond to changes in the underlying riskiness of the environment. In this regard, it is important to know under what circumstances risk estimates increase (i.e. reflect more risk) when using the $BRW^u(t|\lambda, N, C)$ estimator. The result is provided in the following proposition:

Proposition 1 *If $r_t > BRW^u(t, \lambda, N)$ then $BRW^u(t + 1, \lambda, N) \geq BRW^u(t, \lambda, N)$.*

Proof: See the appendix.

The proposition basically verifies my main claim in the introduction to the paper. Specifically, the proposition shows that when losses at time t are bounded below by the BRW VaR estimate at time t , then the BRW VaR estimate for time $t + 1$ will indicate that risk at time $t + 1$ is no greater than it was at time t . The example of a portfolio which was short the S&P 500 at the time of the crash is simply an extreme example of this general result.

To get a feel for the importance of this proposition, suppose that today's VaR estimate for tomorrow's return is conditionally correct, but that risk changes with returns, so that tomorrow's return will influence risk for the day after tomorrow. Under these circumstances, one might ask what is the probability that a VaR estimate which is correct today will increase tomorrow. The answer provided by the proposition is that tomorrow's VaR estimate will not increase with probability $1 - c$. So, for example, if c is equal to 1%, then a VaR estimate which is correct today, will not increase tomorrow with probability 99%.

The question is how often should the VaR estimate increase the next day. The answer depends on the true process which is determining both returns and volatility. The easiest case to consider is when returns follow a GARCH(1,1). This is a useful case to consider for two reasons. First, it is a reasonable first approximation to the pattern of conditional heteroskedasticity in a number of financial time series. Second, it is very tractable.⁶ I will assume that returns are normally distributed, have mean 0, and follow a GARCH(1,1) process:

$$r_t = h_t^{.5} u_t \tag{1}$$

$$h_t = a_0 + a_1 r_{t-1}^2 + b_1 h_{t-1} \tag{2}$$

where, u_t is distributed standard normal for all t ; a_0 , a_1 , and b_1 are all greater than zero; and $a_1 + b_1 < 1$.

Under these conditions, it is straightforward to work out the probability that a VaR estimate should increase tomorrow given that it is conditionally correct today. The answer turns out to have a very simple form when h_t is at its long run mean. The probability that the VaR estimate should increase tomorrow given that h_t is at its long run mean is given in the follow proposition.

⁶Although deriving analytical results may be difficult, all of the simulation analysis that I perform when the data is generated by GARCH(1,1) models could be performed for generalizations of simple GARCH models that are better optimized to fit the data. For example I could instead use a Skewed Student Asymmetric Power ARCH (Skewed Student APARCH) specification to model the conditional heteroskedasticity of exchange rates (Mittnik and Paoletta, 2000) or equity indices (Giot and Laurent, 2001).

Proposition 2 When returns follow a GARCH(1,1) process as in equations (1) and (2) and h_t is at its long run mean, then

$$\text{Prob}(\text{VaR}_{t+1} > \text{VaR}_t) = 2 * \Phi(-1) \approx .3173$$

where $\Phi(x)$ is the probability that a standard normal random variable is less than x .

Proof: See the appendix.

Propositions 1 and 2 taken together suggest that roughly speaking, when a VaR estimate is near the long-run average value of VaR using the BRW methods, then VaR should increase about 32 percent of the time when in fact it will only increase about C percent of the time, i.e. at the 1% confidence level, 31% of the time VaR should have increased, but didn't, or at the 5% confidence level, 27% of the time VaR should have increased, but did not.

The quantitative importance of the historical simulation and BRW methods not responding to certain increases in VaR depend on how much VaR is likely to have increased over a single time period (such as a day) without being detected. This is simple to work out if returns follow a GARCH(1,1) process.

Proposition 3 When returns follow a GARCH(1,1) process as in equations (1) and (2), then when h_t is at its long run mean, and $y(c, t)$, the VaR estimate for confidence level c , at time t , using VaR method y , is correct, and y is either the BRW or historical simulation methods, then the probability that VaR at time $t + 1$ is at least $x\%$ greater than at time t , but the increase is not detected at time $t + 1$ using the historical simulation or BRW methods is given by:

$$\text{Prob}(\Delta \text{VaR} > x \%, \text{ no detect}) = \begin{cases} 2 \Phi \left(-\sqrt{1 + \frac{x^2+2x}{a_1}} \right) - c & 0 < x < k(a_1, c) \\ \Phi \left(-\sqrt{1 + \frac{x^2+2x}{a_1}} \right) & x \geq k(a_1, c), \end{cases} \quad (3)$$

where $k(a_1, c) = -1 + \sqrt{1 - a_1 + a_1[\Phi^{-1}(c)]^2}$.

Proof: See the appendix.

To get a feel for how much of a change in VaR might actually be missed, I considered VaR for 10 different spot foreign exchange positions. Each involves selling U.S. currency and purchasing the foreign currency of a different country. To evaluate the VaR for these positions and to study historical simulation based estimates of VaR, I fit GARCH(1,1) models to the log daily returns of the exchange rates of 10 currencies versus the U.S. dollar. The

data was for the time period from 1973 through 1997.⁷ The results of the estimation are presented in Table 1. The restrictions of proposition 2 are satisfied for most, but not all of the exchange rates. The parameter estimates for the French franc and Italian lira, do not satisfy the restriction that $a_1 + b_1 < 1$. Instead their parameter estimates indicate that their variances are explosive and hence their variances do not have a long-run mean. As a consequence, some of the theoretical results are not strictly correct for these two exchange rates, but they are correct for processes with slightly smaller values of b_1 .

When the variance of exchange rate returns has a long-run mean, equation (3) shows that when variance is near its long run mean, then of the three parameters of the GARCH model, only a_1 determines how much of the increase in true VaR is not detected. For the 10 exchange rates that I consider, a_1 ranges from a low of about 0.05 for the yen, to about 0.20 for the lira. When VaR is computed at the 1% confidence level using the historical simulation or BRW methods, the probability that VaR could increase by at least $x\%$ without being detected is presented in figure 5 for the low, high, and average values of a_1 .⁸ The figure shows that there is a substantial probability (about 31 percent) that increases in VaR will go undetected. Many of the increases in VaR that go undetected are modest. However, there is a 4% probability that fairly “large” increases in VaR will also go undetected. For example, for the largest value of a_1 , with 4% probability (i.e. 4% of the time) VaR could increase by 25% or more, but not be detected using the historical simulation or BRW methods. For the average value of a_1 , there is 4% chance VaR could increase by 15% with being detected, and for the low value of a_1 , there is a 4% chance that a 7% increase in VaR would go undetected. A slightly different view of these results is provided in Table 2. Unlike the figure, which presents probabilities that VaR will actually increase, the table computes the expected size of the increase in VaR conditional on it increasing, but not being detected. For example, the results for the British pound show that conditional on VaR increasing but not being detected (an event that occurs with 31% probability), the expected increase in VaR is about 5-1/2 percent with a standard deviation of about the same amount. Taken as a whole, the table and figure suggests that conditional on VaR being understated for these currencies, the expected understatement will probably be about 7 percent, but because the conditional distribution is skewed right, there is a nontrivial chance that the actual increase in VaR could be much higher.

It is important to emphasize that proposition 3, table 2, and figure 5 quantify the probability that a VaR increase of a given size will not be detected on the day that it occurs. It is

⁷The precise dates for the returns are Jan 2, 1973, through November 6, 1997. The currencies are the British pound, the Belgian franc, the Canadian dollar, the French franc, the Deutschmark, the Yen, the Dutch guilder, the Swedish kronor, the Swiss franc, and the Italian lira.

⁸The average value of a_1 is 0.1184.

possible that VaR could increase for many days in a row without being detected. This allows VaR errors to accumulate through time and occasionally become large. But, the proposition does not quantify how large the VaR errors typically become. Only simulation can answer that question. This is done in the next section.⁹

3 Simulated Performance of Historical Simulation Methods

3.1 Simulation Design

This section examines the performance of the BRW method using simulation in order to provide a more complete description of how the method performs. Results for simulation of the BRW and historical simulation methods are presented in Tables 3 and 5. For purposes of comparison, analogous results are presented in Tables 4 and 6 for when VAR is computed using a Variance-Covariance method in which the variance-covariance matrix of returns is estimated using an exponentially declining weighted sum of past squared returns.¹⁰

All simulation results were computed by generating 200 years of daily data for each exchange rate when the process followed by the exchange rates is the same as those used to generate the theoretical results in Table 2. The simulation results are analyzed by examining how well each of the VaR estimation methods perform along each 200 year sample path. Simulation results are not presented for the Italian lira because for its estimated GARCH parameters, its conditional volatility process was explosive.

3.2 Simulation Results

The main difference between the simulations and the theory is that the simulations compute how the methods perform on average over time. The theoretical results, by contrast, condition on volatility starting from its long run mean. Because of this difference, one would expect the simulated results to differ from the theoretical results. In fact, the theoretically predicted probability that VaR increases will not be detected, and the theoretically predicted conditional distribution of the nondetected VaR increases (Table 2) appear to closely match the results from simulation. In this respect, table 3 provides no new information

⁹An additional reason to perform simulations is that the analytical results on VaR increasing are derived under the special circumstances that the variance of returns are at their long-run mean, and the VaR estimate using the BRW or historical simulation method at this value is correct.

¹⁰The variance covariance matrix for Riskmetrics is estimated using a similar procedure.

beyond knowledge that the predictions from the relatively restrictive theory are surprisingly accurate in the special case of the GARCH(1,1) model.

The more interesting simulation results are presented in Table 5. The table shows that the correlation of the VaR estimates with true VaR is fairly high for the BRW methods, and somewhat lower, for the Historical Simulation methods. This confirms that the methods move with true VaR in the long run. However, the correlations of changes in the VaR estimates with changes in true VaR are quite low. This shows that the VaR methods are slow to respond changes in risk. As a result, the VaR estimates are not very accurate: The average Root Mean Squared Error (RMSE) across the different currencies is approximately 25% of true VaR (Table 5, panels A and B.). The errors as a percent of true VaR turn out not to be symmetrically distributed, but instead are positively skewed. For example in the case of Historical Simulation estimates of 1-day 1-percent VaR for the British pound, VaR is slightly more likely to be overstated than understated; and the errors when VaR is overstated are much larger than when it is understated (Figure 6). On this basis, it appears that the BRW and historical simulation methods are conservative. However, the risks when VaR is understated are substantial: for example, there is a 10% probability that VaR estimates for a spot position in the British pound/dollar exchange rate will be understate true VaR by more than 25% (Figure 6); the same error, expressed as a percent of the value of the spot position is about 1/2 % (Figure 7).

A more powerful method for illustrating the poor performance of the methods involves directly examining how the VaR estimates track true VaR. For the sake of brevity, this is only examined for the British pound over a period of 2 years. The figures for the British pound tell a consistent story: true VaR and VaR estimated using historical simulation or the two BRW methods tend to move together over the long-run, but true VaR changes more often than the estimates, and all three VaR methods respond slowly to the changes (Figures 8, 9, and 10). The result is that true VaR can sometimes exceed estimated VaR by large amounts and for long periods of time. For example, over the two-year period depicted in Figure 11, there is a 0.2 year episode during which VaR estimated using the historical simulation method understates true VaR by amounts that range from a low of 40% to a high of 100%. Over the same 2 years, even with the best BRW method ($\lambda = 0.97$) there are four different episodes which last at least 0.1 years during which VaR is understated by 20% or more; and for one of these episodes, VaR builds up over the period until true VaR exceeds estimated VaR by 70% or more before the VaR estimate adjusts (Figure 12).

The problems with the BRW and historical simulation methods are striking when one compares true VaR against the VaR estimates. In particular, the errors seem to persist for long periods, and sometimes build up to become quite large. Given this poor performance,

it is important that the methods that regulators and risk practitioners use to detect errors in VaR methods are capable of detecting these errors. These detection methods are briefly examined in the next subsection.

3.3 Can Back-testing Detect The Problems with Historical Simulation?

VaR methods are often evaluated by backtesting to determine whether the VaR methods provide correct unconditional coverage, and to examine whether they provide correct conditional coverage. The standard test of unconditional coverage is whether losses exceed VaR at the k percent confidence level more frequently than k percent of the time. A finding that they do would be interpreted as evidence that the VaR procedure understates VaR unconditionally.

Based on standard tests, both BRW methods and the historical simulation method appear to perform well when measured by the percentage of times that losses are worse than predicted by the VaR estimate. Losses exceed VaR 1.5% of the time. This is only slightly more than is predicted. Given that the VaR estimates are actually persistently poor, my results here reconfirm earlier results that unconditional coverage tests have very low power to detect poor VaR methodologies (Kupiec, 1995).

The second way to examine the quality of VaR estimates is to test whether they are conditionally correct. If the VaR estimates are conditionally correct, then the fact that losses exceeded today's VaR estimate should have no predictive power for whether losses will exceed VaR in the future. If we denote a VaR exceedance by the event that losses exceeded VaR, then correct conditional coverage is often tested by examining whether the time series of VaR exceedances is autocorrelated. To provide a sort of baseline feel for the power of this approach, for the 200 years of simulated data for the British pound, I computed the autocorrelation of the actual VaR errors, and of the series of VaR exceedances. Results are presented for 1-day 1% VaR, and 1-day 5% VaR for both the BRW method and for the historical simulation method.

The autocorrelation of the true VaR errors reinforces my earlier results that these VaR methods are slow to adjust to changes in risk. The autocorrelation at a 1-day lag is about 0.95 for all three methods. For the best of the three methods, the autocorrelation of the VaR errors dies off very slowly: it remains about 0.1 after 50 days (Figure 13). The errors of the historical simulation method die off much more slowly. The 50th order autocorrelation of the errors of the 1-day 1% VaR historical simulation estimates is about 0.5 (Figure 14)!

Given the high autocorrelations of the actual VaR errors, it is useful to examine the autocorrelations of the exceedances. Unfortunately, the autocorrelation of the exceedances is generally much smaller than the autocorrelation of the VaR errors. For example, in the case of the BRW method with $\lambda = 0.97$, the autocorrelation of the VaR exceedances for 1% VaR is only about 0.015 for autocorrelations 1 - 6, and it drops towards 0 after that.¹¹ For the historical simulation method, the first six autocorrelations are 0.02 - 0.03 for 1% VaR, and 0.05 for 5% VaR. Because all of the autocorrelations of the exceedances are generally very small, the power of tests for correct conditional coverage, when based on exceedances is very low.¹²

The low power of tests based on exceedances suggests that alternative approaches for examining the performance of VaR measures are needed.¹³ The alternative that I advocate is the one I use here: evaluate a VaR method by comparing its estimates of VaR against true VaR in situations where true VaR is known or knowable.

3.4 Comparison with VaR estimates based on variance-covariance methods

To put the results on the BRW and historical simulation methods in perspective, it is useful to contrast the results with a variance-covariance method with equally weighted observations and with variance-covariance methods which use exponentially declining weights. The performance of the variance-covariance method with equal weighting is about as good as the historical simulation methods. Neither method does a good job of capturing conditional volatility; and this shows up in the performance of the methods. The variance-covariance

¹¹The first six autocorrelations of the exceedances for $\lambda = 0.99$ are about 0.02 for the 1% VaR estimates about 0.02 - 0.03 for the 5% VaR estimates.

¹²An informal illustration of the power of the tests involves calculating the number of time-series observations that would be necessary to generate a rejection of the null if the correlations that were measured for the test are the true correlations. Let ρ_i represent the i 'th autocorrelation. Consider a test based on the first six autocorrelations of the exceedances. Under the null that all autocorrelations are zero,

$$N \sum_{i=1}^6 \rho_i^2 \sim \chi^2(6).$$

If instead all six measured autocorrelations are about 0.05, then about 839 observations (3.36 years of daily data) are required for the test statistic to reject the null of no autocorrelation at the 0.05 percent confidence level. If instead all six measured autocorrelations are about 0.015, then 37.3 years of daily data are required to reject the null using this test.

¹³Despite the low power of tests based on VaR exceedances, in Berkowitz and O'Brien's (2001) study of VaR estimates at 6 commercial banks, they found that VaR exceedances for two of the 6 banks they examined had VaR exceedances whose first order autocorrelations were statistically different from zero. The first order autocorrelation for the two banks were 0.158, and 0.330, both of which are much larger than the autocorrelation of the VaR exceedances for the cases considered here.

methods with exponentially declining weights are unambiguously better than historical simulation, and also perform better than the BRW methods: the probability that increases in VaR are not detected is with one exception, less than 10%, the mean and standard deviation of undetected increases in VaR is generally low (Table 4), and the correlation of these measures with true VaR and with changes in VaR is high. There are two reasons why these methods perform better than the BRW method in the simulations. The first is that the variance-covariance methods recognize changes in conditional risk whether the portfolio makes or loses money; the BRW method only recognizes changes in risk when the portfolio experiences a loss. The second reason is that computing variance-covariance matrices using exponential weighting is similar to updating estimates of variance in a GARCH(1,1) model. This similarity helps the variance-covariance method capture changes in conditional volatility when the true model is GARCH(1,1). Moreover, the same exponential weighting methods perform well for all of the GARCH(1,1) parameterizations.

Given that these simulations suggest that the exponential weighting method of computing VaR appears to be better than the BRW method with the same weights, the empirical results in Boudoukh, Richardson, and Whitelaw (1997) are puzzling because they show that their method appears to perform better when using real data. The reason for the difference is almost surely that returns in the real world are both heteroskedastic and leptokurtic but the exponential smoothing variance-covariance methods ignore leptokurtosis and instead assume that returns are normally distributed. It turns out that the normality assumption is a first-order important error; it is this error which makes the BRW and historical simulation methods appear to perform well by comparison.

Although the BRW method appears to be better than exponential smoothing when using real data, it is far from an ideal distributional assumption. The BRW method's inability to associate large profits with risk, and its inability to respond to changes in conditional volatility are disturbing. More importantly, there is not a strong theoretical basis for using the BRW method. More specifically, except for the case of $\lambda = 1$, one cannot point to any process for asset returns and say to compute VaR for that process, the BRW method is the theoretically correct approach. Because of the disturbing features of the BRW and historical simulation methods, it is desirable to pursue other approaches for modeling the distribution of the risk factors. Ideally, the methodology which is adopted should model conditional heteroskedasticity and non-normality in a theoretically coherent fashion. There are many possible ways that this could be done. A relatively new VaR methodology introduced by Barone-Adesi, Giannopoulos, and Vosper combines historical simulation with conditional volatility models in a way which has the potential to achieve this objective. This new methodology is called Filtered Historical Simulation (FHS). The advantages and pitfalls of

the filtered historical simulation method are discussed in the next section.

4 Filtered Historical Simulation

In a recent paper Barone-Adesi, Giannopoulos, and Vosper, introduced a variant of the historical simulation methodology which they refer to as filtered historical simulation (FHS). The motivation behind using their method is that the two standard approaches for computing VaR make tradeoffs over whether to capture the conditional heteroskedasticity or the non-normality of the distribution of the risk factors. Most implementations of Variance-Covariance methods attempt to capture conditional heteroskedasticity of the risk factors, but they also assume multivariate normality; by contrast most implementations of the historical simulation method are nonparametric in their assumptions about the distribution of the risk factors, but they typically do not capture conditional heteroskedasticity.

The innovation of the filtered historical simulation methodology is that it captures both the conditional heteroskedasticity and non-normality of the risk factors. Because it captures both, it has the potential to very significantly improve on the variance-covariance and historical simulation methods that are currently in use.¹⁴

4.1 Method details

Filtered historical simulation is a Monte Carlo based approach which is very similar to computing VaR using fully parametric Monte Carlo. The best way to illustrate the method is to illustrate its use as a substitute for the fully parametric Monte Carlo. I do this first for the case of a single-factor GARCH(1,1) model (equations (1) and (2)), and then discuss the general case.

Single Risk Factor

To begin, suppose that the time-series process for the risk factor r_t is described by the GARCH(1,1) model in equations (1) and (2), and that the conditional volatility of returns tomorrow is h_{t+1} . Given, these conditions, VaR at a 10-day horizon (the horizon required by the 1996 Market Risk Amendment to the Basle Accord) can be computed by simulating 10-day return paths using fully parametric monte carlo. Generating a single path involves

¹⁴Engle and Manganelli (1999) propose an alternative approach in which the quantiles of portfolio value based on past data follow an autoregressive process. This approach has two main disadvantages. First, every time the portfolio changes the parameters of the autoregression need to be reestimated. Second, when risk increases using the Engle and Manganelli approach, the source of the increase in risk will not be apparent because the approach models the behavior of the P&L of the portfolio, but not the behavior of the individual risk factors.

drawing the innovation ϵ_{t+1} from its distribution (which is $\mathcal{N}(0, 1)$). Applying this innovation in equation (1) generates r_{t+1} . Given h_{t+1} and r_{t+1} , equation (2) is then used to generate h_{t+2} . Given h_{t+2} , the rest of the 10-day path can be generated similarly. Repeating 10-day path generation thousands of times provides a simulated distribution of 10-day returns conditional on h_t .

The difference between the above methodology and FHS is that the innovations are drawn from a different distribution. Like the monte-carlo method, the FHS method assumes that the distribution of ϵ_t has mean 0, variance 1, and is i.i.d., but it relaxes the assumption of normality in favor of the much weaker assumption that the distribution of ϵ_t is such that the parameters of the GARCH(1,1) model can be consistently estimated. For the moment, suppose that the parameters can be consistently estimated, and in fact have been estimated correctly. If they are correct, then the estimates of h_t at each point in time are correct. This means that since r_t is observable, equation (1) can be used to identify the past realizations of ϵ_t in the data. Barone-Adesi, Giannopoulos, and Vosper (1999) refer to the series of ϵ_t that is identified as the time series of filtered shocks. Because these past realizations are i.i.d., one can make draws from their empirical distribution to generate paths of r_t .

The main insight of the FHS method is that it is possible to capture conditional heteroskedasticity in the data and still be somewhat unrestrictive about the shape of the distribution of the factors returns. Thus the method appears to combine the best elements of conditional volatility models with the best elements of the historical simulation method.

Multiple Risk Factors

There are many ways the methodology can be extended to multiple risk factors. The simplest extension is to assume that there are N risk factors which each follow a GARCH process in which each factor's conditional volatility is a function of lags of the factor and of lags of the factor's conditional volatility.¹⁵ To complete this simplest extension, an assumption about the distribution of ϵ_t , the N -vector of the innovations, is needed. The simplest assumption is that ϵ_t is distributed i.i.d. through time. Under this assumption, the implementation of FHS in the multifactor case is a simple extension of the method in the single factor case. As in the single-factor case, the elements of the vector ϵ_t are identified by estimating the GARCH models for each risk factor. Draws from the empirical distribution of ϵ_t are made by randomly drawing a date and using the realization of ϵ for that date.

This simple multivariate extension is the main focus of Barone-Adesi and Giannopoulos (1999). This extension has two convenient properties. First, the volatility models are very

¹⁵The specification in equations (1) and (2) is a special case of a more general specification which has these features.

simple. One does not need to estimate a multivariate GARCH model to implement them. The second advantage is that the method does not involve estimation of the correlation matrix of the factors. Instead, the correlation of the factors is implicitly modelled through the assumption that ϵ_t is i.i.d.

Although the simplest multivariate extension of FHS is convenient, the assumptions that it uses are not necessarily innocuous. The assumption that volatility depends only on a risk factor's own past lags, and its own past lagged volatility can be unrealistic whether there is a single risk factor, or many. For example, if the risk factors are the returns of the FTSE Index and that of the S & P 500, then if the S & P 500 is highly volatile today, then it may influence the volatility of the FTSE tomorrow. A separate issue is the assumption that ϵ_t is i.i.d. This assumption implies that the conditional correlation of the risk factors is fixed through time. This assumption is also likely to be violated in practice.

Although the assumptions of the simplest extension of FHS may be violated in practice, these problems can be fixed by complicating the modelling where necessary. For example, the volatility modelling may be improved by conditioning on lagged values for other assets. Similarly, time-varying correlations can be modelled within the framework of multivariate GARCH models.

To show the potential for improving upon simple implementations of the FHS method, suppose that the conditional mean and variance-covariance matrix of the factors depends on the past history of the factors.¹⁶ More specifically, let r_t be the factors at time t ; let $h_{r_{t-}}$ be the history of the risk factors prior to time t ; let θ be parameters of the data generating process; let $\mu(h_{r_{t-}}, \theta)$ be the mean of the factors at time t conditional on this history and θ , and let $\Sigma(h_{r_{t-}}, \theta)$ be the variance-covariance matrix of r_t conditional on this history, and θ .

Given this notation, suppose that r_t is generated according to

$$r_t = \mu(h_{r_{t-}}, \theta) + \Sigma(h_{r_{t-}}, \theta)^{.5} \epsilon_t \quad (4)$$

where θ are the parameters of the conditional mean and volatility model, and ϵ_t is i.i.d. through time with mean 0 and variance I .¹⁷ If equation (4) is the data generating process, then under appropriate regularity conditions (Bollerslev and Wooldridge (1992)), the θ parameters can be estimated by quasi-maximum likelihood.¹⁸ Therefore, ϵ_t can be identified,

¹⁶GARCH models are a special case of the general formulation.

¹⁷Since ϵ_t is i.i.d., assuming its variance is I is without loss of generality since this assumption simply normalizes $\Sigma(h_{r_{t-}}, \theta)$.

¹⁸In quasi-maximum likelihood estimation (QMLE), the parameters, θ , are estimated by maximum likelihood with a Gaussian distribution function for ϵ_t . Under appropriate regularity conditions, the parameter estimates of θ are consistent and asymptotically normal even if ϵ_t is not normally distributed.

and the FHS method can be implemented in this more general case.¹⁹

It may be possible to improve on this general case even further. One of the issues in implementing the FHS method, or any historical simulation method is whether the filtered data series contain a sufficient number of shocks to fill the probability space of what could happen; i.e. are there important shocks, or combinations of shocks that are under-represented in the historical data series. If there are, then one possible fix is to create a set of filtered shocks which are uncorrelated. This could be done by using a multivariate GARCH model with time-varying covariances, as above. The marginal distribution of the filtered shocks could then be fit using a semi-parametric or nonparametric method. These marginal distributions could then be used in the same way as draws of random days in the FHS method. In this case, VaR estimation would proceed by making i.i.d. draws from the marginal distributions of the shocks, and then these would be applied to the GARCH model in the same way that they are applied with Filtered Historical Simulation.

4.2 Preliminary Analysis of Filtered Historical Simulation

In this section, I conduct some preliminary analysis of the simplest Filtered Historical Simulation Methodology. To analyze the method, I estimated GARCH(1,1) models for the same exchange rates as in Table 1, but to reserve data for out of sample analysis, I only used data from January 1973 through June 1986 in the estimation.

As noted above, an important assumption in the simple FHS approach is that the correlation of the filtered data sets are constant through time. To investigate whether this is satisfied, I split my estimation sample into two subsamples—one for the first half of the estimation sample—and one for the second. I then used the distribution of Fisher’s z transformation of the correlation coefficient to individually test whether each correlation coefficient is different in the two time periods.²⁰ The tests for each correlation coefficient are not independent, but the null hypothesis is overwhelmingly rejected for 86 out of 90 of the correlation coefficients. This suggests that the changes in the correlations across these

¹⁹Let $\hat{\theta}$ of θ be a consistent estimate of θ . Then since $h_{r_{t-}}$ is observable, from (4), a consistent estimate of ϵ_t is

$$\epsilon_t = \Sigma(h_{r_{t-}}, \theta)^{-.5} [r_t - \mu(h_{r_{t-}}, \theta)]$$

²⁰Formally, I tested the null hypotheses that the correlation coefficients in the two time periods were the same. Following Kendall and Stewart (1963), let n be the sample size and let r and ρ be the estimated and true correlation coefficients; and define z and ξ by $z = .5 \log \frac{1+r}{1-r}$, and let $\xi = .5 \log \frac{1+\rho}{1-\rho}$. Kendall and Stewart show that $z - \xi$ is approximately normal with approximate mean $\frac{\rho}{2(n-1)}$ and approximate variance $\frac{1}{n-3}$. Let z_1 and z_2 be the estimates of z in the two subsamples and assume that both subsamples have the same number of observations, denoted n . Assuming that the estimates in the subsamples are independent, it follows that under the null of constant correlation, $.5(n-3)(z_1 - z_2)^2$ is asymptotically $\chi^2(1)$.

periods are statistically significant. Examination of the differences in the coefficients in the two periods suggests that they are economically significant as well (Table 7).

Perhaps, the fact that the correlations appear to have changed over a 13 year period of time is not surprising. But the fact that they have shows that using historical simulation with simple filtering could be problematic. The difficulty is that one may be making draws from time periods where the correlation of the shocks is different than it is today. This could have the effect of making a risky position appear hedged, or a hedged position appear risky.

Even if correlations appear to have changed over a period of 13 years, if they appear to remain fixed over reasonably long periods of times, then a simple solution to the time-varying correlation problem is to only use recent data when doing filtered historical simulation. To investigate this possibility, I examined the stability of the correlations over a much shorter period of time—adjacent one-year intervals. Over this shortened period, there is less evidence that correlations change through time, but there still appear to be economically and statistically significant changes in many correlations (Table 8). One reason that the correlations vary through time is that the univariate GARCH(1,1) models that are used here may be improperly specified. However, it appears that better-specified univariate GARCH models will not fix the problem because there is substantial evidence in the GARCH literature which suggests that foreign exchange rates, do not have constant conditional correlations. The only other quick fix to the problem with the correlations is to shorten the sample that is used for filtered historical simulation by even more, but I am hesitant to do this for fear of losing important nonparametric information about the tails of the risk factors.

The second exercise that I conducted examines the ability of the FHS method to accurately estimate VaR at a 1% confidence level at a 10-day horizon. The accuracy of these VaR estimates is of particular interest because the BIS capital requirements for market risk are based on VaR at this horizon and confidence level. If the VaR estimates are excellent at this horizon, then a strong case could be made for moving toward using the FHS method to compute capital for market risk.

Barone-Adesi, Giannopoulos, and Vosper, have written a number of papers in which they examine the performance of the FHS method using real data. Because the FHS method makes strong assumptions about how the data is generated, it is necessary to examine the method using real data. But, examining the method using real data is not sufficient to understand how the method performs. There are two weaknesses with only relying on real data when analyzing a method. The first is that the typical methods for analyzing model performance, such as backtesting, have low power to detect poor models because detection is based on losses exceeding VaR estimates of loss, and this is a rare event even if the VaR estimates are poor. The second is that when a VaR model is examined using real data, the

ability to understand the properties of the VaR model is obfuscated by the simultaneous occurrence of other types of model errors including errors in pricing, errors in GARCH models, and other potential flaws in the VaR methodology.

To focus exclusively on potential difficulties with the FHS approach, while abstracting from other sources of model error, the simulations in the second exercise are conducted under "ideal" conditions for using the FHS methodology. In particular, I assume that the parameters of the GARCH processes are estimated exactly, the filtered innovations that are used in the simulation are the true filtered innovations, and that today's conditional volatility is known, and that all pricing models are correct. Because these conditions eliminate most forms of model misspecification, the simulations help to examine how well the FHS method works under nearly perfect conditions.

To examine the performance of FHS at a 10-day horizon, for each exchange rate's estimated GARCH process, I generated 2 years of random data, and then used the simulated data with the true GARCH parameters to estimate VAR at a 10-day horizon using filtered historical simulation. Each filtered historical simulation estimate of VaR was computed using 10,000 sample paths. The VAR estimates using the FHS method were then evaluated by comparing them with VAR estimates based on a full Monte-Carlo simulation (which used 100,000 sample paths). The errors of any particular FHS estimate will depend on the initial conditional volatilities that are used to generate the sample paths and on any idiosyncracies in a particular sample of filtered innovations. To date, I have only examined the behavior of the method from initial conditional volatilities that were selected to be the same as those at two different points in time. Therefore, my results should be viewed as conditional on a particular set of beginning volatilities. That said, to keep the results from depending on any particular set of filtered innovations, for each set of beginning conditional volatilities the results that I report for each exchange rate are based on 100 independent simulations of the FHS method.²¹ Additional details on these simulations are provided in the appendix.

My assessment of the FHS method is designed to answer two questions. The first is whether, for a fixed span of historical data, VaR estimates using filtered historical simulation do a good job of approximating VaR at a 10-day horizon. The second question is if filtered historical simulation does not do well, then what is the source of the errors? Two sources are examined. The first is that the number of sample paths in the filtered historical simulation may have been too small. To address this issue, in the appendix I derive confidence bounds for how big the errors in the VaR estimates would be if for a given set of historical data, VaR was computed using an infinite number of sample paths. If the errors with an infinite number

²¹Although the filtered innovations in each FHS estimate are independent, to economize on computation I use the same monte-carlo simulation to assess the results of each estimate.

of sample paths remain large, then the number of sample paths is not the source of the error. The second potential source of error is that the size of the historical sample that is used for the bootstrap is too small. In other words, 2 years of historical data may not have enough extreme observations to generate good estimates of 10-day VaR at the 1% confidence level. Logically, this second source of error is the only other potential source because virtually all other sources of error have been eliminated from the simulation by design. Below, I provide intuition for why I believe this second source of error is important.

4.3 Results of Analysis of FHS

The analysis here is still preliminary. The first issue I address is whether the VaR estimates using filtered historical are downward biased; i.e. do they appear to understate risk. To address this question, I use order-statistics from the monte carlo simulation of VaR to create 95% confidence intervals for the percentage errors in the filtered historical simulation VaR estimates.²² When the confidence interval for the percentage errors did not contain 0 percent error, the VaR estimate was classified as under- or over- stating true VaR based on whether the confidence interval was bounded above or below by 0. The amount of under- or overstatement was measured as the shortest distance from the edge of the confidence interval to 0. This is probably a conservative estimate of the true amount of under- or overstatement. My main finding is that the filtered historical simulation method is biased towards understating risk. At a 10-day horizon and 1% confidence interval the method was found to understate risk about 2/3rds of the time, and when VaR was understated the average amount of understatement was about 10% (Table 9). VaR was overstated about 28 percent of the time; and conditional on VaR being overstated, the average amount of overstatement was about 9%.

To investigate the source of the downward bias in the VaR estimates I first investigated whether additional monte-carlo draws using filtered historical simulation would produce better results. I performed this analysis by creating 90% confidence intervals for the percentage difference between the infinite sample (population) estimate of VaR using filtered historical simulation and the infinite sample (population) estimate of VaR using pseudo monte carlo. The difference between this analysis, and my analysis of the VaR percentage errors above is that this analysis uses statistics to test whether the FHS VaR estimates could be improved by increasing the number of sample paths to infinity. These confidence intervals are not as precise as those used to analyze whether the VaR estimates were over or under stated, but

²²Details on this method are contained in Pritsker (1997).

they are suggestive, and seem to indicate that even in infinite samples the VaR estimates using filtered historical simulation in this context would be downward biased (Table 10).

The most likely explanation for the downward bias is a lack of extreme observations in a short filtered data set. For example, when making draws from a 500 observation historical sample of filtered data, the extremes of the filtered sample are the lowest and highest observation in the data, which corresponds to roughly the 0.2'd and 99.8th percentile of the probability distribution. Draws greater than or less than these percentiles cannot happen in the filtered historical simulation. The question is how likely are draws outside of these extremes when simulating returns over a 10-day period. The answer is that on a 10 day sample path, the probability of at least one draw outside of these extremes is $1 - (.996)^{10}$, which is about 4%. Since these sample paths are also those where GARCH volatilities will be higher (through GARCH-style amplification of shocks to volatility) these 4% of sample paths are probably very likely to be important for risk at the 1% confidence level, but they cannot be accounted for by the filtered historical simulation method when only a two year sample of daily historical data is used in the simulation. A longer span of historical data may improve the filtered historical simulation method by allowing for more extreme observations, but longer data spans mean it is necessary to address the time-varying correlation issue.

Before closing, it is useful to contrast my results on filtered historical simulation with those of Barone-Adesi, Giannopoulos, and Vosper (2000). When working with real data, they found that at long horizons (their longest horizon was 10 days) the FHS method tended to overstate the risk of interest rate swap positions, and of portfolios which consist of interest rate swaps, futures, and options. The finding that risk is overstated at these long horizons stands in contrast to my results. If both sets of results are correct (remember my results are still preliminary), they suggest that several sources of bias are present. To better understand the properties of filtered historical simulation in the real data, it would be useful to attempt to separate the effects of model risk from errors due to the filtered historical simulation procedure. This remains a task for future research.

5 Conclusions

Historical simulation based methods for computing VaR are becoming popular because these methods are easy to implement. However, the properties of the methods are not well understood. In this paper, I explore the properties of these methods and show that a number of the methods are under-responsive to changes in conditional volatility. Despite the under-responsiveness of these methods to changes in risk, there is strong reason to believe that backtesting methods have little power to detect the problems with historical simulation

methods. I also investigated the properties of the filtered historical simulation method that was recently introduced by Barone-Adesi, Giannopoulos, and Vosper (1999). The advantage of the method is it allows for time-varying volatility but the distribution of the risk factor innovations is modeled nonparametrically. While I think this new method has a great deal of promise, the paper illustrates two areas in which the method needs further development. The first is modeling time-varying correlations. The evidence in this paper suggests that time-varying correlations may be important. Filtered historical simulation methods need to be improved in ways that account for these correlations. The second area which needs further development is establishing the number of years of historical data which are needed to produce accurate VaR estimates. Results in this paper suggest that 500 days of daily data may not be enough to accurately compute VaR at a 10-day horizon because a sample period this short may not contain enough extreme observations. Methods for choosing the length of the historical data series, or methods to augment historical data with additional extreme observations (perhaps with a parametric model fit to the historical data) are topics for further research.

Appendix

A Proofs

A.1 Proof of Proposition 1

Proposition 1 *If $r_t > BRW^u(t, \lambda, N)$ then $BRW^u(t + 1, \lambda, N) \geq BRW^u(t, \lambda, N)$.*

Proof:

When the VaR estimate using the BRW method is estimated for returns during time period $t + 1$, the return at time $t - N$ is dropped from the sample, the return at time t receives weight $\frac{1-\lambda^N}{1-\lambda}$, and the weight on all other returns are λ times their earlier values.

Define,

$$r(C) = \{r_{t-i}, i = 1, \dots, N | G(r_{t-1}; t, N) \leq C\}.$$

To verify the proposition, it suffices to examine much probability weight the VaR estimate at time $t + 1$ places below $BRW^u(t, \lambda, N)$. There are two cases to consider:

Case 1: $r_{t-N} \notin r(C)$. In this case, since $r_t \notin r(C)$ by assumption, then $G(BRW^u(t, \lambda, N); t + 1, \lambda, N) = \lambda G(BRW^u(t, \lambda, N))$. Therefore,

$$BRW^u(t + 1, \lambda, N) = \inf(r \in \{r_t, \dots, r_{t-1-N}\} | G(r; t + 1, \lambda, N) \geq C) \geq BRW^u(t, \lambda, N).$$

Case 2: $r_{t-N} \in r(C)$. In this case, since $r_t \notin r(C)$ by assumption, then

$$G(BRW^u(t, \lambda, N); t + 1, \lambda, N) < \lambda G(BRW^u(t, \lambda, N)).$$

Therefore,

$$BRW^u(t + 1, \lambda, N) = \inf(r \in \{r_t, \dots, r_{t-1-N}\} | G(r; t + 1, \lambda, N) \geq C) \geq BRW^u(t, \lambda, N). \square$$

A.2 Proof of Proposition 2

Proposition 2 *When returns follow a GARCH(1,1) process as in equations (1) and (2), then when h_t is at its long run mean, then*

$$Prob(VaR_{t+1} > VaR_t) = 2 * \Phi(-1) \approx .3173$$

where $\Phi(x)$ is the probability that a standard normal random variable is less than x .

Proof:

The long run mean of h_t , denoted \bar{h} , is equal to $\frac{a_0}{1-a_1-b_1}$. Therefore, when $h_t = \bar{h}$, $h_{t+1} > h_t$ if and only if

$$\begin{aligned} a_0 + a_1 r_t^2 + b_1 h_t &> h_t \\ a_0 + a_1 h_t u_t^2 + b_1 h_t &> h_t \\ \frac{-a_0}{a_1 h_t} + \frac{1-b_1}{a_1} &< u_t^2 \\ \frac{a_1 + b_1 - 1}{a_1} + \frac{1-b_1}{a_1} &< u_t^2 \\ 1 &< u_t^2. \end{aligned}$$

Finally, the result follows because

$$\text{Prob}(1 < u_t^2) = \text{Prob}(u_t < -1) + \text{Prob}(u_t > 1) = 2\Phi(-1)\square$$

A.3 Proof of Proposition 3

Proposition 3 When returns follow a GARCH(1,1) process as in equations (1) and (2), then when h_t is at its long run mean, and $y(c, t)$, the VaR estimate for confidence level c , at time t , using VaR method y , is correct, and y is either the BRW or historical simulation methods, then the probability that VaR at time $t + 1$ is at least $x\%$ greater than at time t , but the increase goes undetected by the historical simulation or BRW methods is given by:

$$\text{Prob}(\Delta \text{VaR} > x \%, \text{ no detect}) = \begin{cases} 2 \Phi \left(-\sqrt{1 + \frac{x^2+2x}{a_1}} \right) - c & 0 < x < k(a_1, c) \\ \Phi \left(-\sqrt{1 + \frac{x^2+2x}{a_1}} \right) & x \geq k(a_1, c), \end{cases} \quad (5)$$

where $k(a_1, c) = -1 + \sqrt{1 - a_1 + a_1[\Phi^{-1}(c)]^2}$.

Proof:

Let $A(x)$ denote the event that $\frac{\text{VaR}_{t+1}}{\text{VaR}_t} > 1 + x$.

Let B denote the event that VaR increases and is not detected.

Let $C(x) = A(x) \cap B$.

It suffices to compute $C(x)$ for all $x > 0$ to complete the proof.

$$\text{Prob}[A(x)] = \text{Prob} \left(\sqrt{\frac{a_0 + a_1 r_t^2 + b_1 h_t}{h_t}} > 1 + x \right) \quad (6)$$

$$= \text{Prob} \left(\sqrt{\frac{a_0 + a_1 h_t u_t^2 + b_1 h_t}{h_t}} > 1 + x \right) \quad (7)$$

Imposing the condition that h_t at time t is at its long run mean, and that VaR is correctly estimated at time t , equation (7) simplifies to become:

$$\text{Prob}[A(x)] = \text{Prob} \left(\sqrt{1 + a_1 (u_t^2 - 1)} > 1 + x \right) \quad (8)$$

$$= \text{Prob} \left(|u_t| \geq \sqrt{1 + \frac{x^2 + 2x}{a_1}} \right) \quad (9)$$

From the proof of propositions 1 and 2, we know that VaR will increase, but not be detected (the event B) for $u_t \in \{[\phi^{-1}(c), -1] \cup [1, -\phi^{-1}(c)] \cup [-\phi^{-1}(c), \infty]\}$.

Partition B into the union of the disjoint events B_1 and B_2 where:

$$B_1 = u_t \in \left\{ [\Phi^{-1}(c), -1] \cup [1, -\Phi^{-1}(c)] \right\}$$

$$B_2 = u_t \in [-\Phi^{-1}(c), \infty]$$

Algebra shows, the event B_2 corresponds to $x > k(a_1, c)$. Therefore, for $x > k(a_1, c)$,

$$\text{Prob}(C(x)) = \text{Prob} \left(u_t \geq \sqrt{1 + \frac{x^2 + 2x}{a_1}} \right) \quad (10)$$

$$= \Phi \left(-\sqrt{1 + \frac{x^2 + 2x}{a_1}} \right) \quad (11)$$

For $0 \leq x \leq k(a_1, c)$

$$\text{Prob}(C(x)) = \text{Prob}(B_2) + \text{Prob}\left(\sqrt{1 + \frac{x^2 + 2x}{a_1}} \leq |u_t| \leq -\Phi^{-1}(c)\right) \quad (12)$$

$$= c + 2 \left(\Phi[-\Phi^{-1}(c)] - \Phi\left[\sqrt{1 + \frac{x^2 + 2x}{a_1}}\right] \right) \quad (13)$$

$$= c + 2 \left(1 - c - \Phi\left[\sqrt{1 + \frac{x^2 + 2x}{a_1}}\right] \right) \quad (14)$$

$$= -c + 2\Phi\left[-\sqrt{1 + \frac{x^2 + 2x}{a_1}}\right] \square \quad (15)$$

A.4 Error Bounds for Filtered Historical Simulation Quantiles

Let $S = \{S_i, i = 1, \dots, N_S\}$ be a sample of draws from continuous distribution function F . Let $G(\cdot)$ be a function which maps $S_i \rightarrow R$, and let $\hat{\xi}_p$ be the p 'th quantile of $G(\cdot)$ when making i.i.d. draws with replacement from S , and let ξ_p be the p 'th quantile of $G(\cdot)$ when making i.i.d. draws from F . The error associated with sampling from S instead of F is $\hat{\xi}_p - \xi_p$.²³ Our goal is to compute a confidence interval for this error as a percent of ξ_p .

To begin, suppose that there exist confidence bounds \hat{L} and \hat{H} , and L^* and H^* such that

$$\text{Prob}(0 < \hat{L} \leq \hat{\xi}_p \leq \hat{H}) = q_1,$$

and

$$\text{Prob}(L^* \leq \xi_p \leq H^*) = q_2$$

The only loss of generality in this assumption is that that the true p 'th quantile of the distribution must involve losses in portfolio value. If it does not, then the possibility that true VaR is equal to 0 makes it meaningless to compute errors as a percent of VaR. The assumption that confidence bounds for the quantiles exist is with no loss of generality because if one draws large random samples of size N_S and N_F for S and F respectively, then order statistics from those samples can be used to construct nonparametric confidence bounds for ξ_p and $\hat{\xi}_p$, and as N_S and N_F go to infinity the upper and lower bounds of the confidence bounds converge to the true quantiles (see for example David (1991)).

²³The reason for the error is that infinitely repeated sampling from a finite subsample is not equivalent to infinite sampling from the true population.

Let A and B be the events:

$$A = \{0 < \hat{L} \leq \hat{\xi}_p \leq \hat{H}\}, \quad (16)$$

$$B = \{L^* \leq \xi_p \leq H^*\}, \quad (17)$$

and let $C = A \cap B$, then because A and B are independent $\text{Prob}(C) = q_1 q_2$.²⁴

Suppose that C is true, then, multiplying equation (17) by -1 and rearranging it follows that:

$$-H^* \leq -\xi_p \leq -L^* \quad (18)$$

Adding equations (18) and (16) and dividing by ξ_p it then follows that:

$$\frac{\hat{L} - H^*}{\xi_p} \leq \frac{\hat{\xi}_p - \xi_p}{\xi_p} \leq \frac{\hat{H} - L^*}{\xi_p} \quad (19)$$

Because B is true, it is simple to verify that (see Pritsker 1997)

$$\bar{L} = \text{Inf} \left(\frac{\hat{L} - H^*}{L^*}, \frac{\hat{L} - H^*}{H^*} \right) \leq \frac{\hat{L} - H^*}{\xi_p}, \quad (20)$$

and,

$$\bar{H} = \text{Sup} \left(\frac{\hat{H} - L^*}{L^*}, \frac{\hat{H} - L^*}{H^*} \right) \geq \frac{\hat{H} - L^*}{\xi_p} \quad (21)$$

Substituting equations (21) and (20) in equation (19) produces the result that when C is true, then

$$\bar{L} \leq \frac{\hat{\xi}_p - \xi_p}{\xi_p} \leq \bar{H}. \quad (22)$$

Because C is true with probability $q_1 q_2$, equation (22) provides a $q_1 q_2$ percent confidence interval for the percentage difference between the population and sampling quantile.

By choosing q_1 and q_2 as close to 1 as desired, and by allowing N_S and N_F to become large, this confidence interval can approximate the true percentage error with with an arbitrary

²⁴ A and B are independent because the Monte-Carlo draws to create the confidence bounds are independent.

degree of accuracy provided that the quantiles being considered involve losses in portfolio value.

B Details on How Data was Generated for Tests of Filtered Historical Simulation

To generate the data I assumed that the 10×1 vector of filtered shocks for the exchange rate series at each time t is a transformation of a set of primitive shocks e_t :

$$s_t = (\rho^5)e_t \tag{23}$$

where e_t is a 10×1 vector of independent $t(6)$ random variables that have been normalized to have a standard deviation of 1. ρ in the above expression is the estimated unconditional correlation of the exchange rates during the period January 1973 until June 1986. I chose to only use data from this period in order to allow for future out of sample data analysis.

The resulting set of filtered shocks s_t will be i.i.d. with mean 0, and variance-covariance matrix ρ ; but because the shocks are linear combinations of student t random variables their distributions will be leptokurtic.

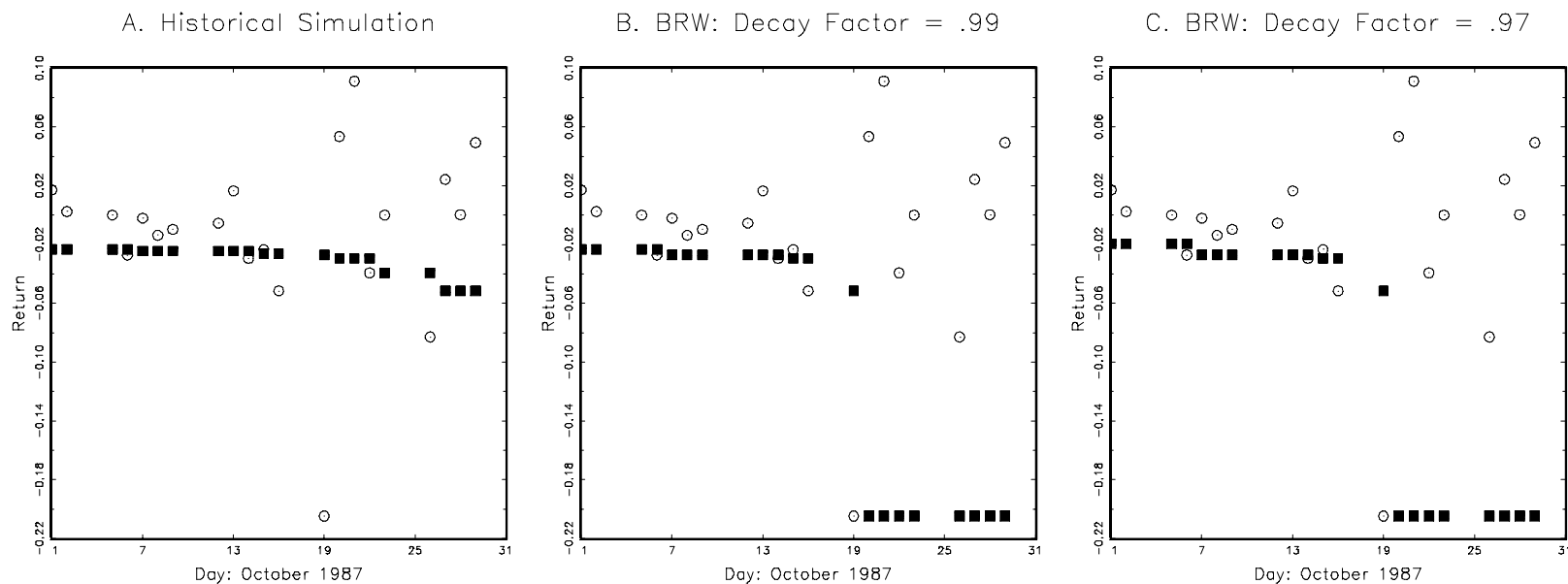
The data was generated by choosing a set of beginning conditional volatilities for each exchange rate by picking days between January 1973 and June 1986 and starting the process with that day's conditional volatilities. From this starting point the future returns were generated by using the filtered shocks on each day t as the innovations in GARCH(1,1) models for each exchange rate. Because the parameters for these GARCH(1,1) models were estimated using the data from January 1973 to June 1986, they differ somewhat from the estimates in Table 1. The GARCH parameters and correlation estimates that were used in the simulation are provided in table 11.

BIBLIOGRAPHY

- Barone-Adesi, G., Giannopoulos, K., and L. Vosper, 1999, "VaR without Correlations for Nonlinear Portfolios," *Journal of Futures Markets*, 19 (April), 583 - 602.
- Barone-Adesi, G., Giannopoulos, K., and L. Vosper, 2000 "Filtering Historical Simulation. Backtest Analysis," Working Paper, University of Westminster, March.
- Barone-Adesi, G., and K. Giannopoulos, November 2000 "Non-parametric VaR Techniques. Myths and Realities," Working Paper, University of Westminster, November.
- Beder, T.S., 1995, "VAR: Seductive but Dangerous," *Financial Analysts Journal*, September, 12-24.
- Berkowitz, J., 1999, "Testing the Accuracy of Density Forecasts," manuscript, University of California, Irvine.
- Berkowitz, J. and James O'Brien, 2001, "How Accurate are Value at Risk Models at Commercial Banks?," manuscript, University of California, Irvine.
- Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T., Wooldridge, J.M., 1992, "Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances," *Econometric Reviews*, 11 iss. 2, 143 - 72.
- Boudoukh, J., Richardson, M., and R. Whitelaw, 1998. "The Best of Both Worlds," *Risk* 11 (May), 64-67.
- Christofferson, P. 1998, "Evaluating Interval Forecasts," *International Economic Review*, 39 , 841-862.
- David, H.A., *Order Statistics, 2nd Edition*, John Wiley and Sons, inc., New York, 1981.
- Diebold, F.X., Gunther, T.A., and Tay, A.S., 1998, "Evaluating Density Forecasts," *International Economic Review*, 39, 863-883.
- Duffie, D., and J. Pan, 1997, "An Overview of Value at Risk," *Journal of Derivatives*, 4, Spring, 7 - 49.

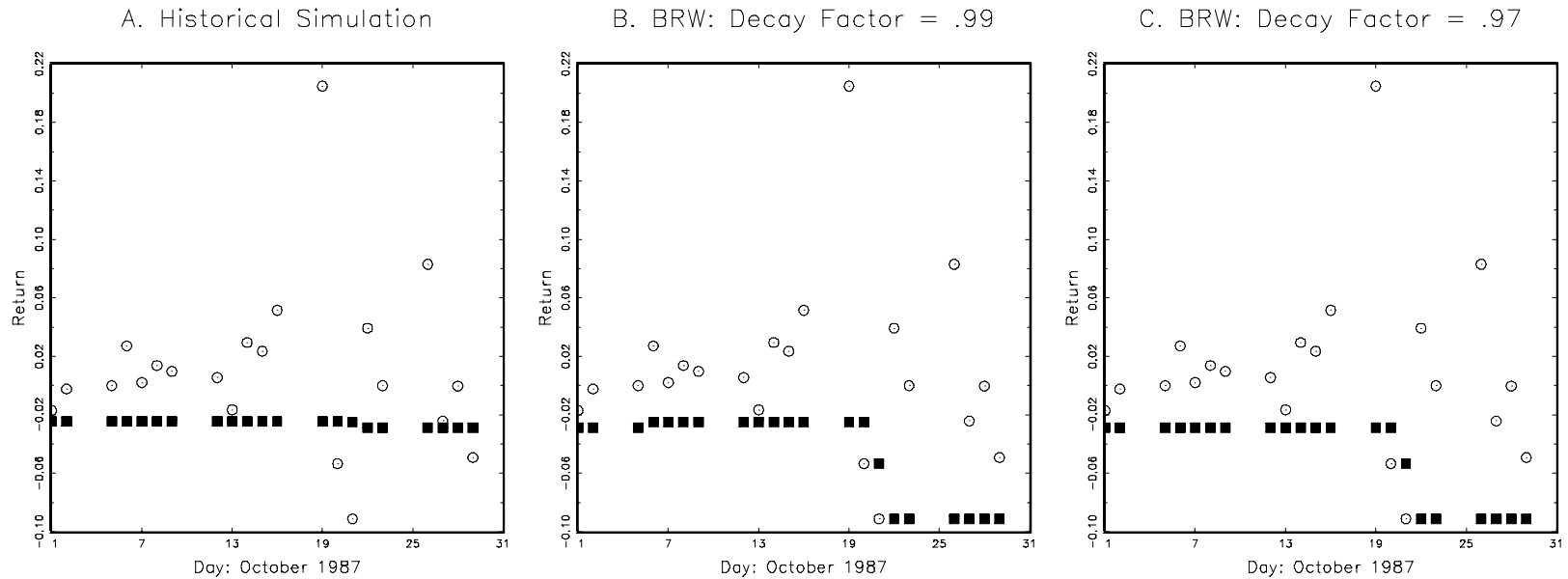
- Engle, R.F., and S. Manganelli, 1999, "CaViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," Working Paper, University of California, San Diego, July.
- Giot, P., and S. Laurent, 2001, "Value at Risk for Long and Short Trading Positions," Mimeo, Maastricht University.
- Hull, J., and A. White, 1998, "Incorporating volatility updating for value-at-risk," *Journal of Risk*, 1 no. 1, Fall (1998), 5 - 19.
- Hendricks, D., 1996, "Evaluation of Value at Risk Models Using Historical Data," *Economic Policy Review*, Federal Reserve Bank of New York, April, 36-69.
- Kendall, M.G., and A. Stuart, 1963, *The Advanced Theory of Statistics*, Hafner Publishing Company, New York.
- Kupiec, Paul, 1995, "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives* 3 (Winter), 73-84.
- Mittnik, S., and M. Paoletta, 2000, "Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates," *Journal of Forecasting*, 19, 313 - 333.
- Pritsker, M., 1997 "Evaluating Value at Risk Methodologies: Accuracy versus Computational Time," *Journal of Financial Services Research* (Oct/Dec), 201 - 241.

Figure 1: One-Percent VaR Measures for Long Equity Portfolio in October 1987



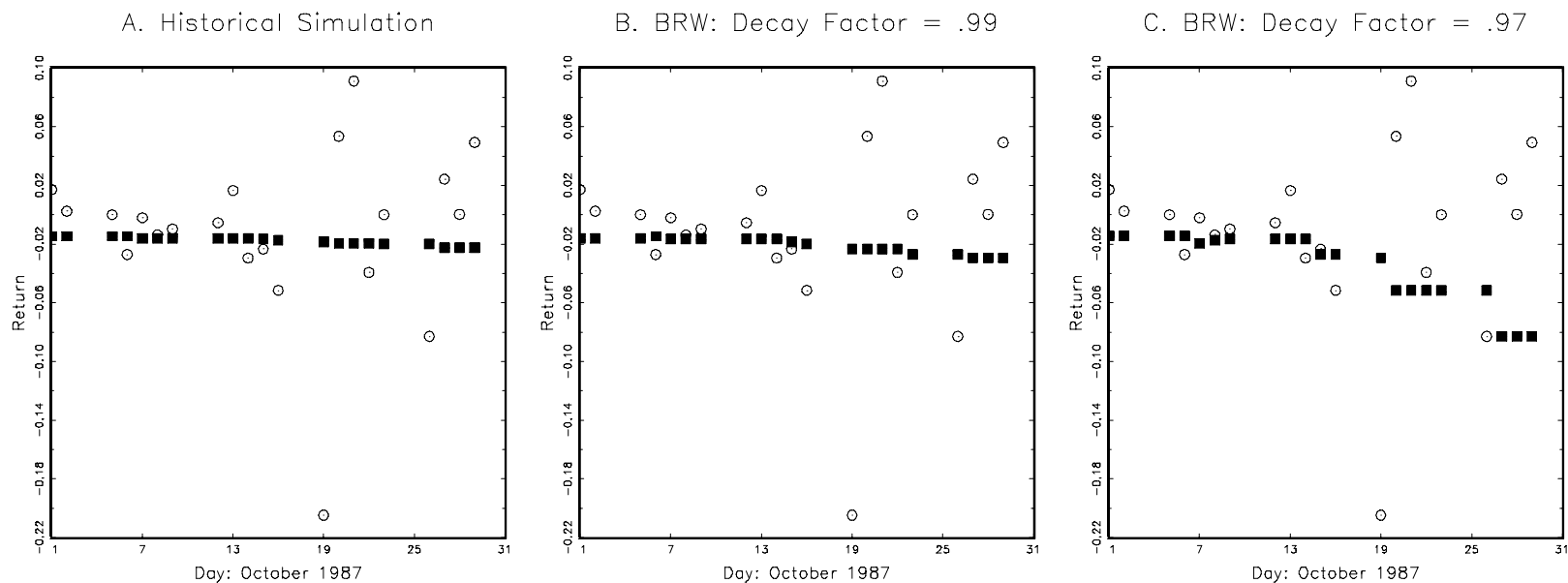
Notes: The figure examines the behavior of VaR estimates for a portfolio which is long the S&P 500 in the period surrounding the October 19, 1987 market crash. Each panel plots the portfolios daily return (clear circles), and a one-percent VaR estimate for each days return (solid boxes). The VaR methods are “Historical Simulation”, and “BRW” with decay factors 0.99 and 0.97. Details on the VaR methodologies are provided in the text.

Figure 2: One-Percent VaR Measures for Short Equity Portfolio in October 1987



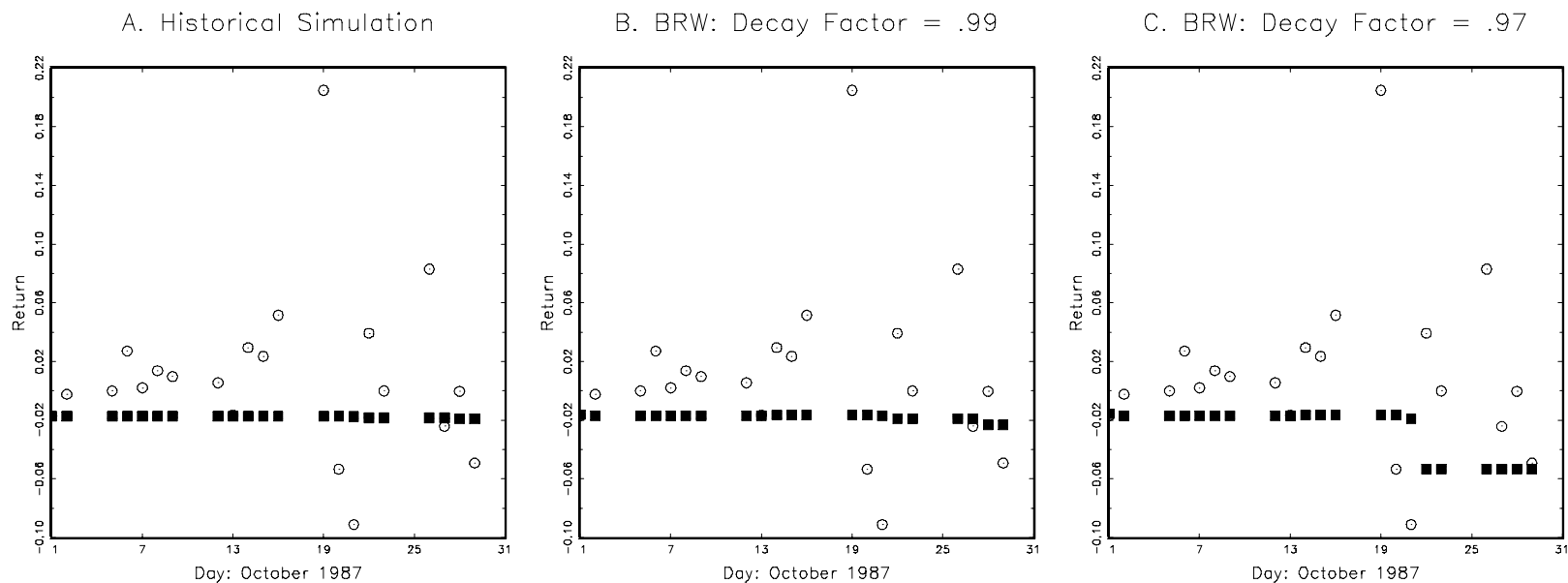
Notes: The figure examines the behavior of VaR estimates for a portfolio which is short the S&P 500 in the period surrounding the October 19, 1987 market crash. Each panel plots the portfolios daily return (clear circles), and a one-percent VaR estimate for each days return (solid boxes). The VaR methods are “Historical Simulation”, and “BRW” with decay factors 0.99 and 0.97. Details on the VaR methodologies are provided in the text.

Figure 3: Five-Percent VaR Measures for Long Equity Portfolio in October 1987



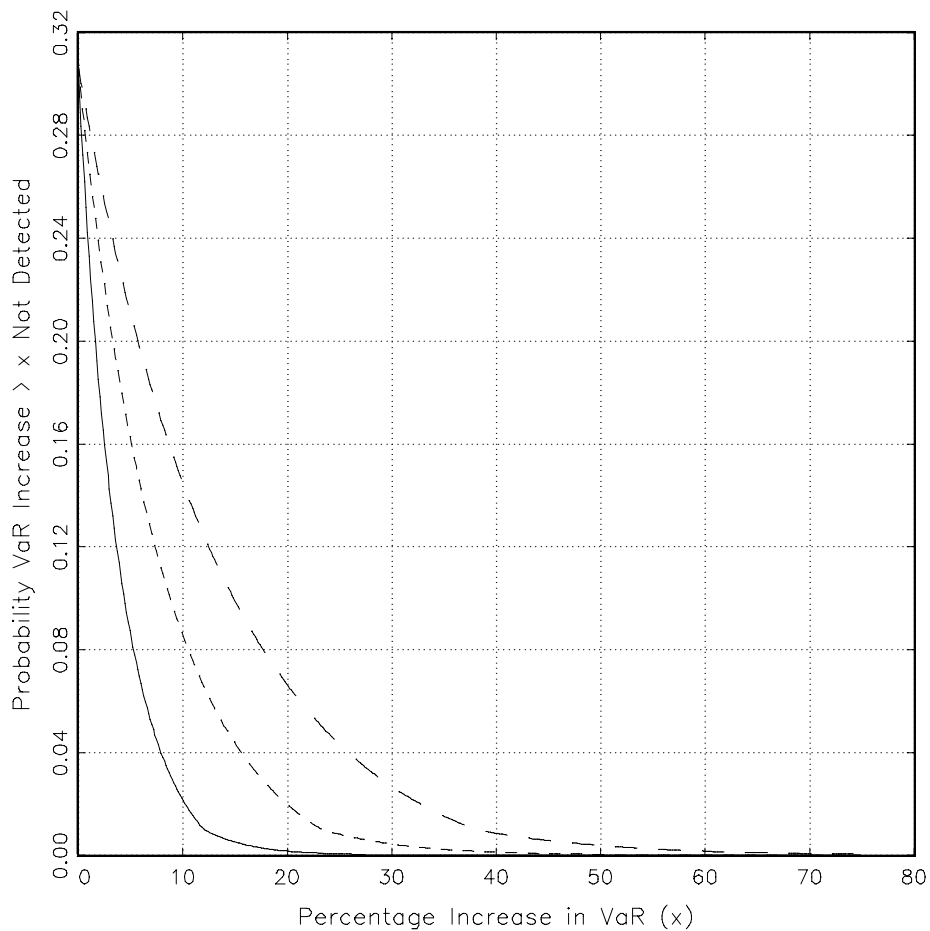
Notes: The figure examines the behavior of VaR estimates for a portfolio which is long the S&P 500 in the period surrounding the October 19, 1987 market crash. Each panel plots the portfolios daily return (clear circles), and a five-percent VaR estimate for each days return (solid boxes). The VaR methods are “Historical Simulation”, and “BRW” with decay factors 0.99 and 0.97. Details on the VaR methodologies are provided in the text.

Figure 4: Five-Percent VaR Measures for Short Equity Portfolio in October 1987



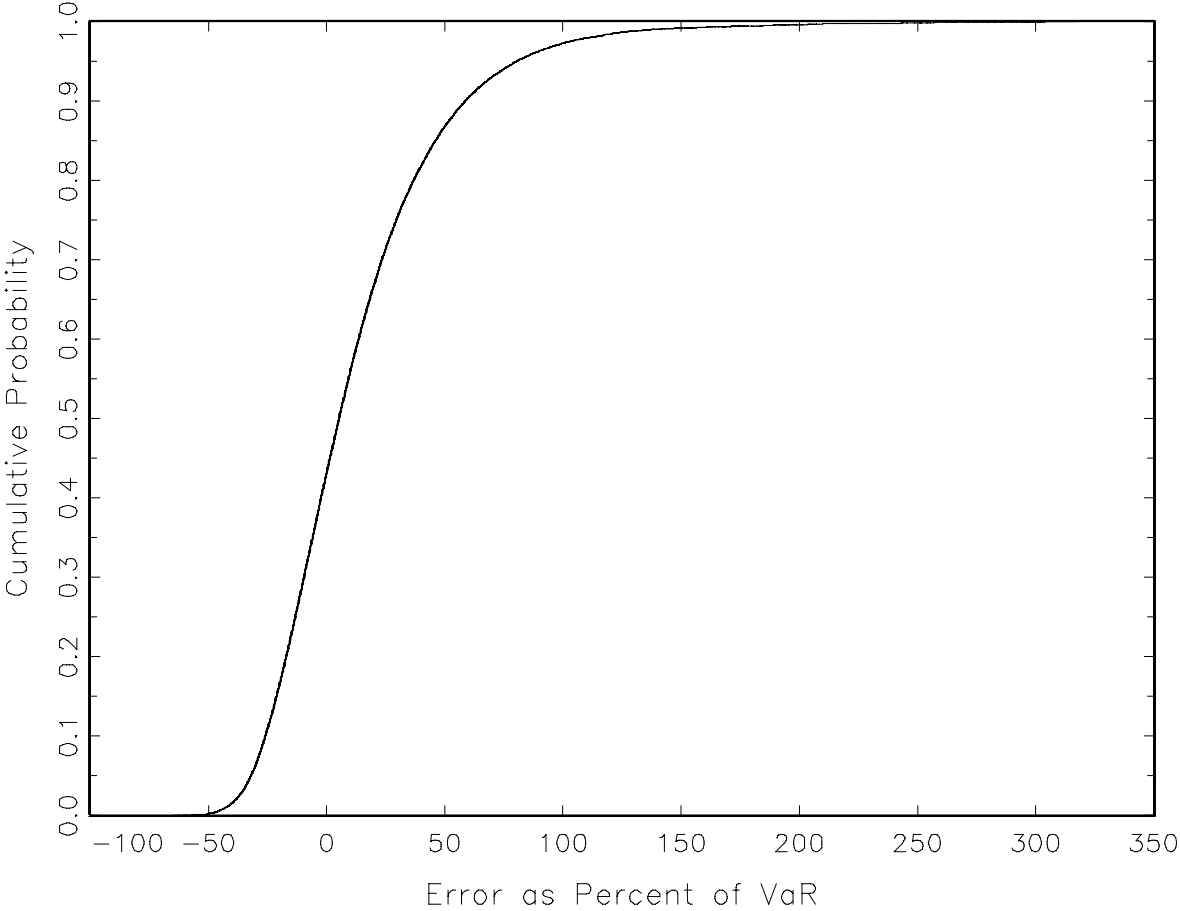
Notes: The figure examines the behavior of VaR estimates for a portfolio which is short the S&P 500 in the period surrounding the October 19, 1987 market crash. Each panel plots the portfolios daily return (clear circles), and a five-percent VaR estimate for each days return (solid boxes). The VaR methods are “Historical Simulation”, and “BRW” with decay factors 0.99 and 0.97. Details on the VaR methodologies are provided in the text.

Figure 5: Probability VaR Increases will Go Undetected



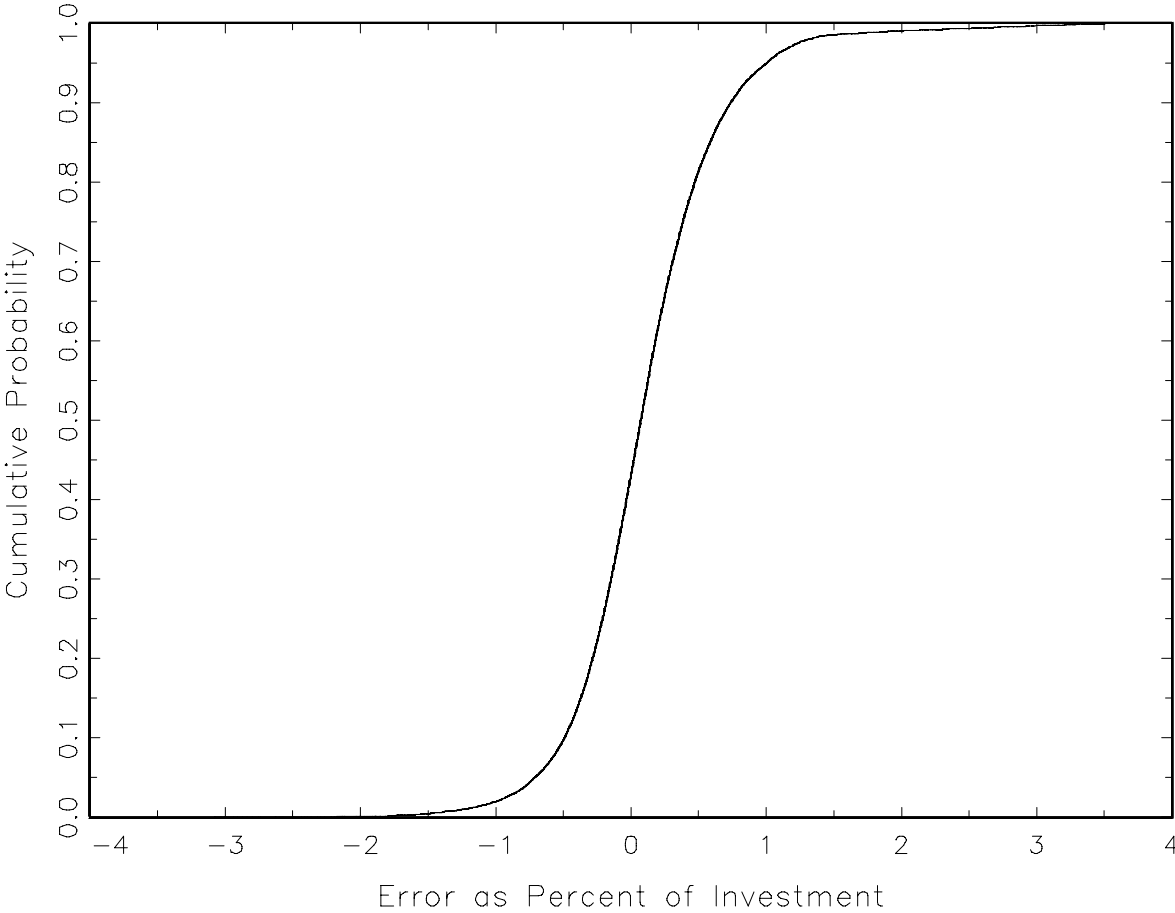
Notes: The figure plots the probability that various percentage increases in 1-day 1-percent Value at Risk will not be detected using the historical simulation methodology when exchange rate returns follow GARCH(1,1) processes with the low, high, and mean values of a_1 from Table 1. Results are presented for $a_1 = 0.05810$ (short dashes), 0.1184 (solid), and 0.2057 (long dashes). All curves were computed conditional on the event that before true VaR changes, the VaR estimates are correct, and the variance of returns are at their long run averages.

**Figure 6: Distribution of Errors as Percent of VaR for Historical Simulation
1% VaR Estimates for British Pound**



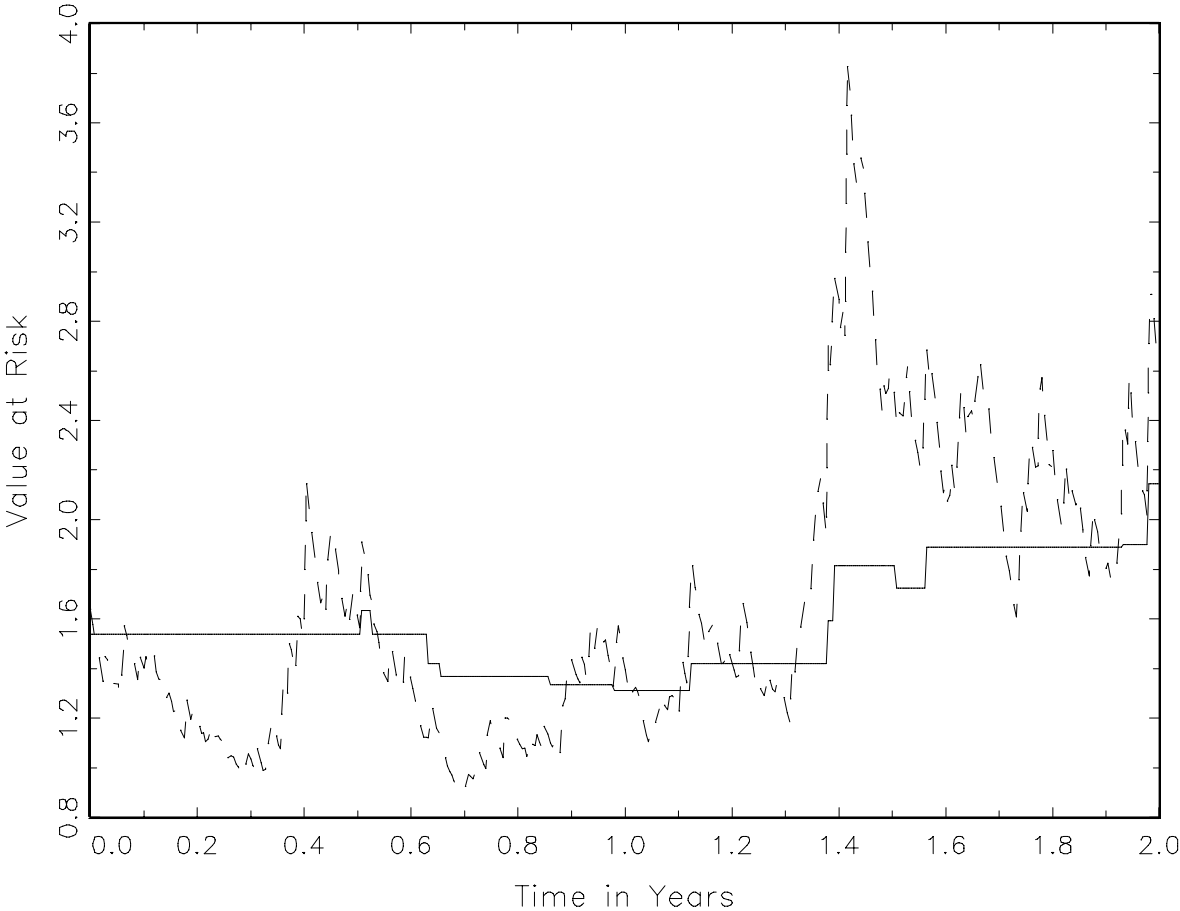
Notes: When the exchange rate for the British pound is simulated using the GARCH model in Table 1, the figure presents the cumulative distribution function for the percentage errors associated with computing VaR by historical simulation with 250 observations. Errors are reported as a percent of true VaR.

Figure 7: Distribution of Errors as Percent of Investment for Historical Simulation 1% VaR Estimates for British Pound



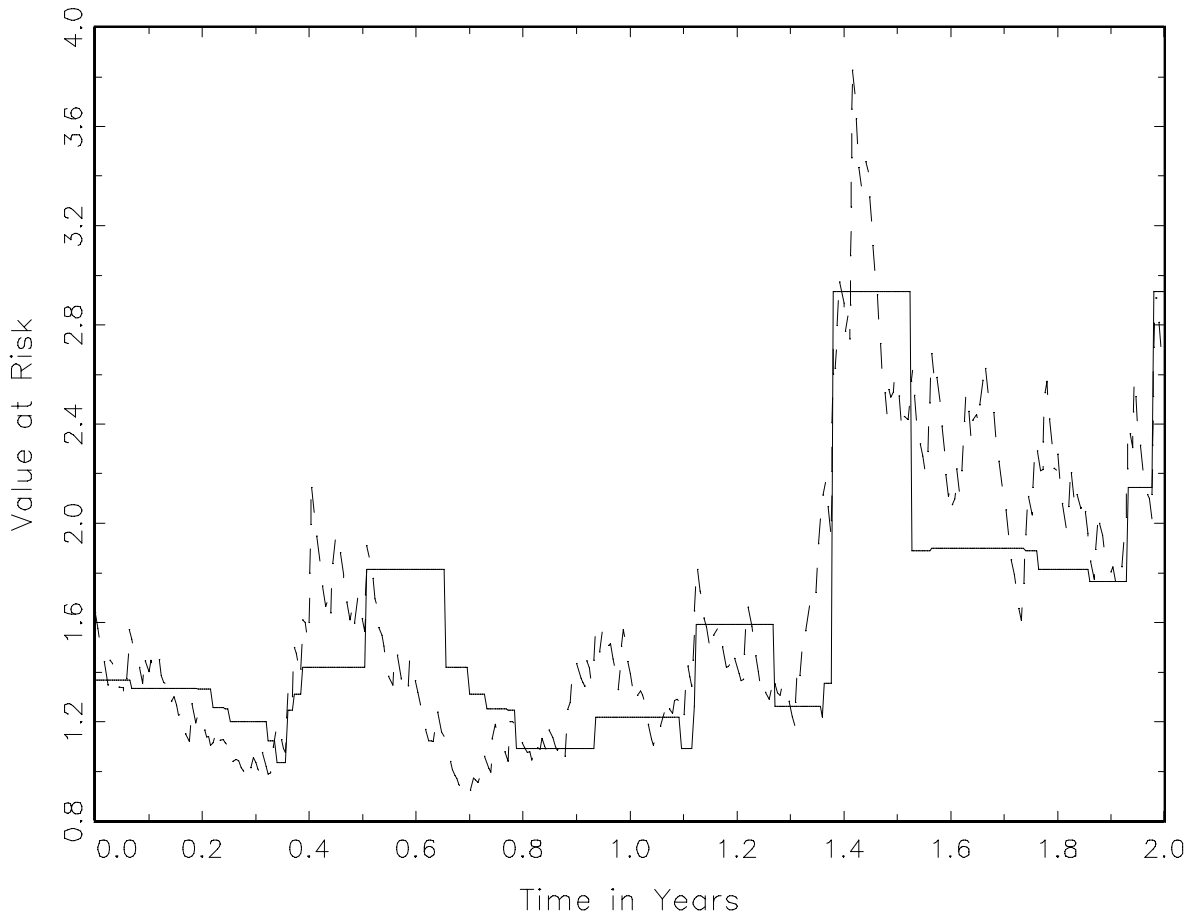
Notes: When the exchange rate for the British pound is simulated using the GARCH model in Table 1, the figure presents the cumulative distribution function for the percentage errors associated with computing 1% VaR for a 1-day horizon by historical simulation with 250 observations. Errors are reported as a percent of the size of a position in the dollar/sterling exchange rate.

Figure 8: True and Historical Simulation Estimates of 1% VaR for British Pound



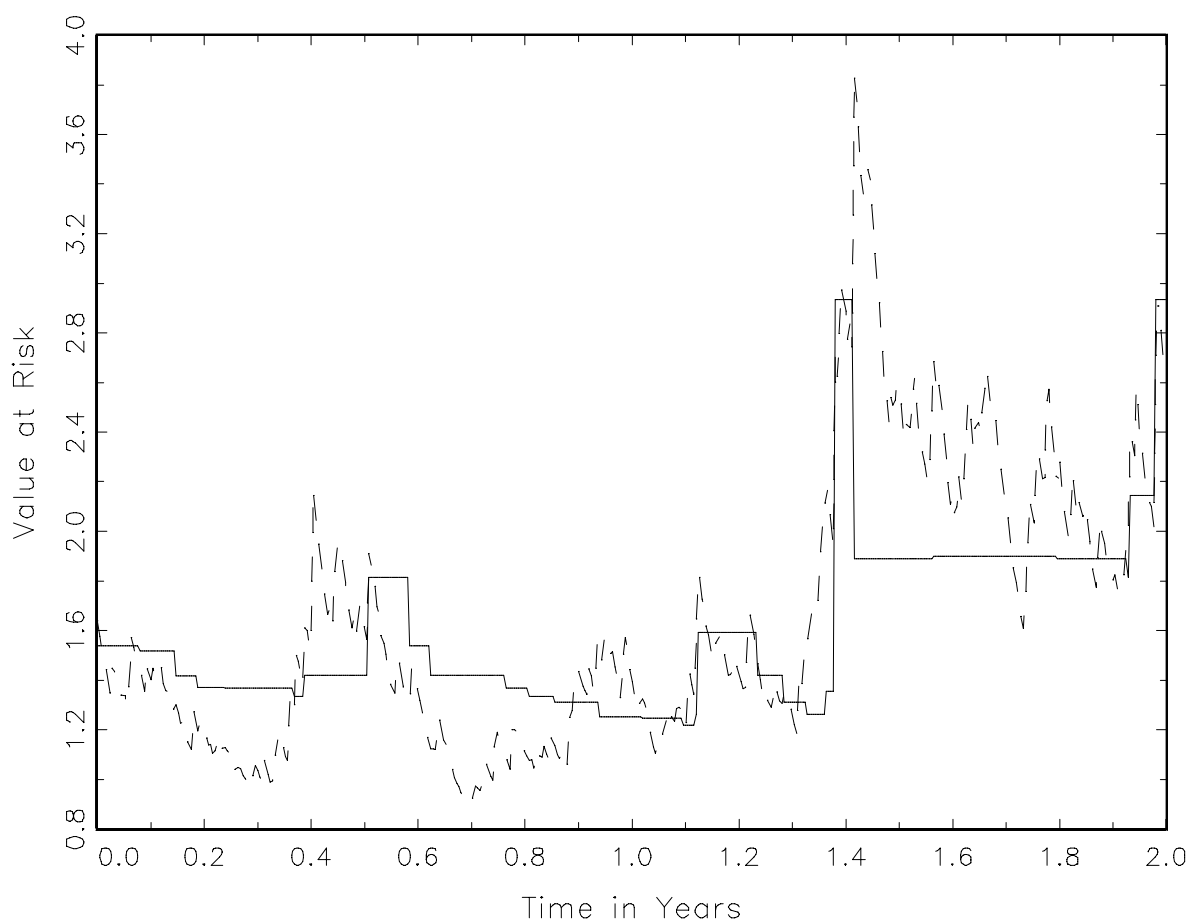
Notes: When the exchange rate for the British pound is simulated using the GARCH model in Table 1, the figure plots 2 years of the simulated time series of true (dashed line) and estimated (solid line) 1% VaR for a one-day time horizon. VaR is estimated using the historical simulation method with 250 observations of daily data.

Figure 9: True and BRW ($\lambda = .97$) Estimates of 1% VaR for British Pound



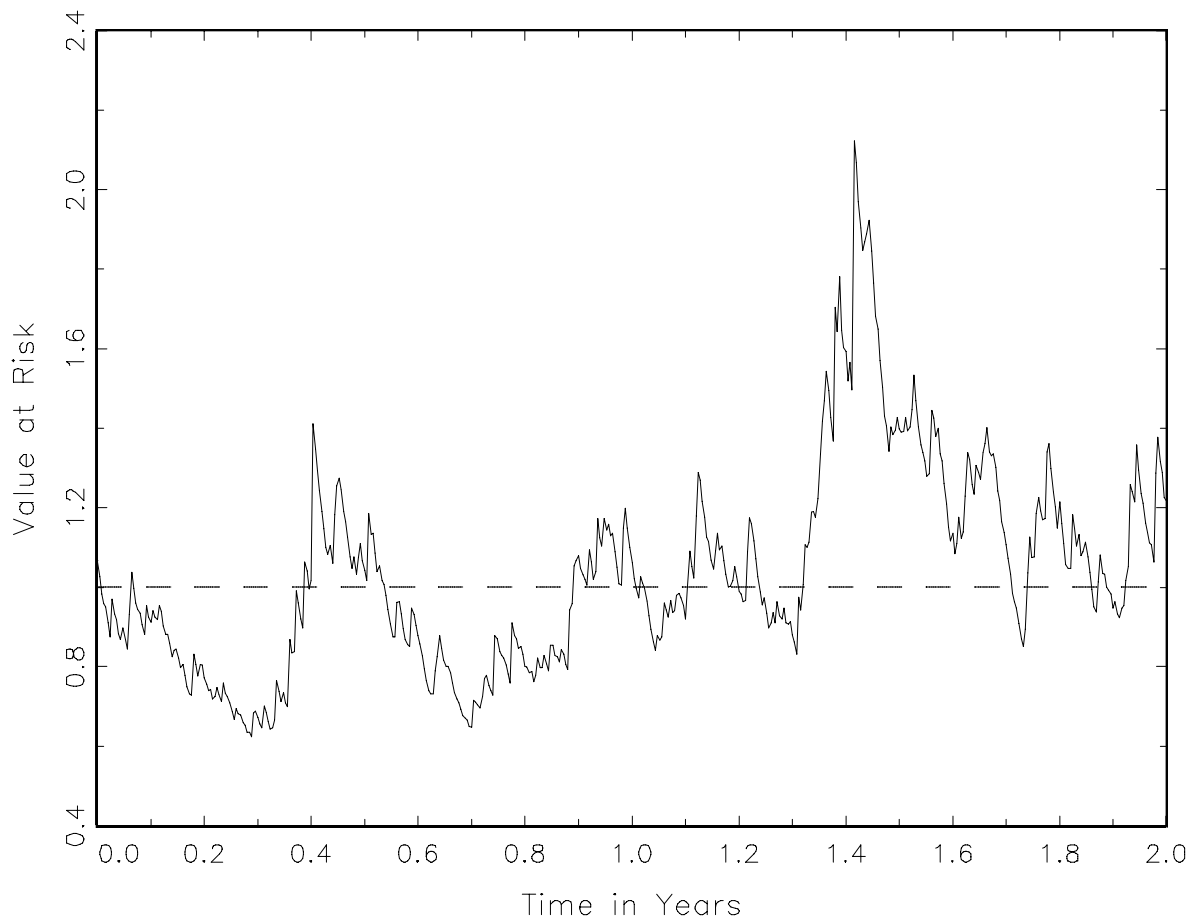
Notes: When the exchange rate for the British pound is simulated using the GARCH model in Table 1, the figure plots 2 years of the time series of true (dashed line) and estimated (solid line) 1% VaR for a one-day time horizon. VaR is estimated using the BRW method with 250 observations of daily data, and a weighting factor of 0.97.

Figure 10: True and BRW ($\lambda = .99$) Estimates of 1% VaR for British Pound



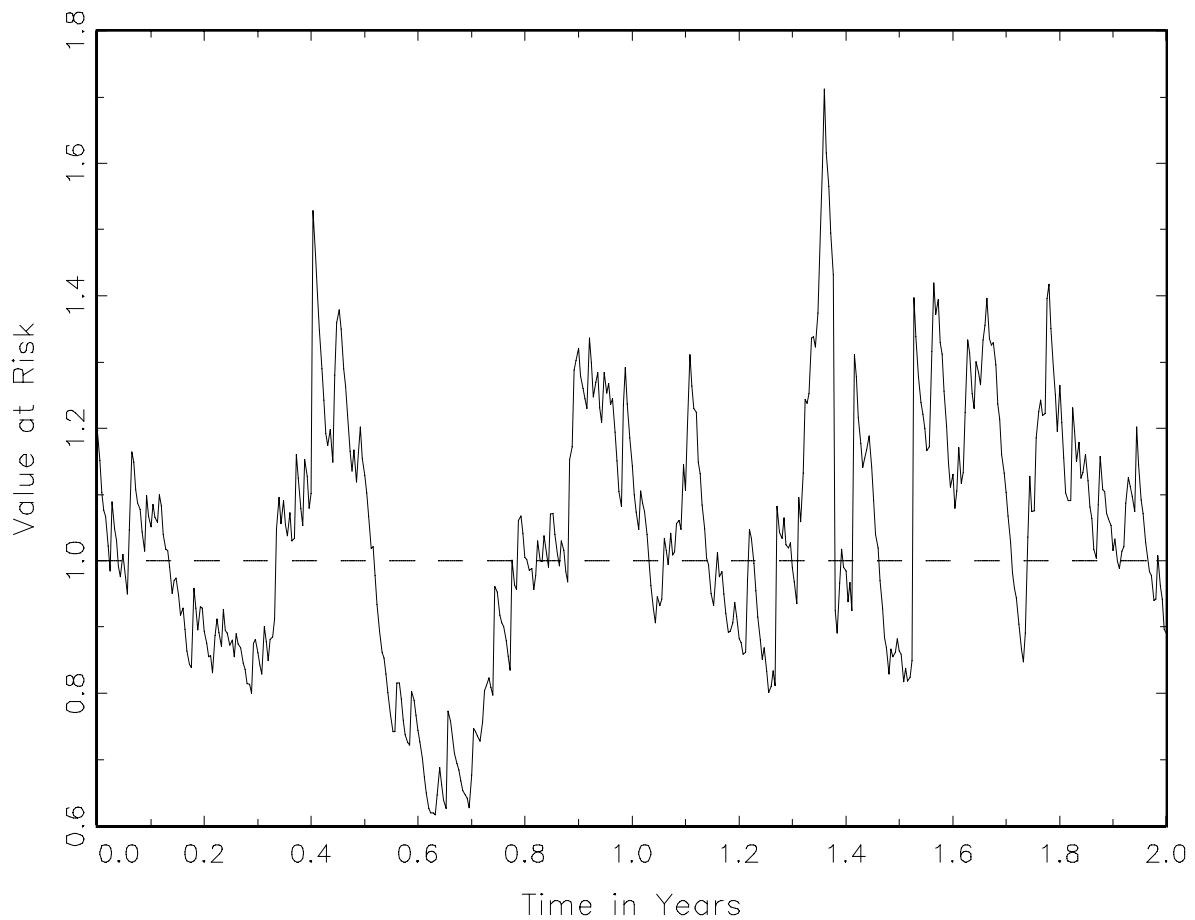
Notes: When the exchange rate for the British pound is simulated using the GARCH model in Table 1, the figure plots 2 years of the time series of true (dashed line) and estimated (solid line) 1% VaR for a one-day time horizon. VaR is estimated using the BRW method with 250 observations of daily data, and a weighting factor of 0.99.

Figure 11: Ratio of True and Historical Simulation Estimates of 1% VaR for British Pound



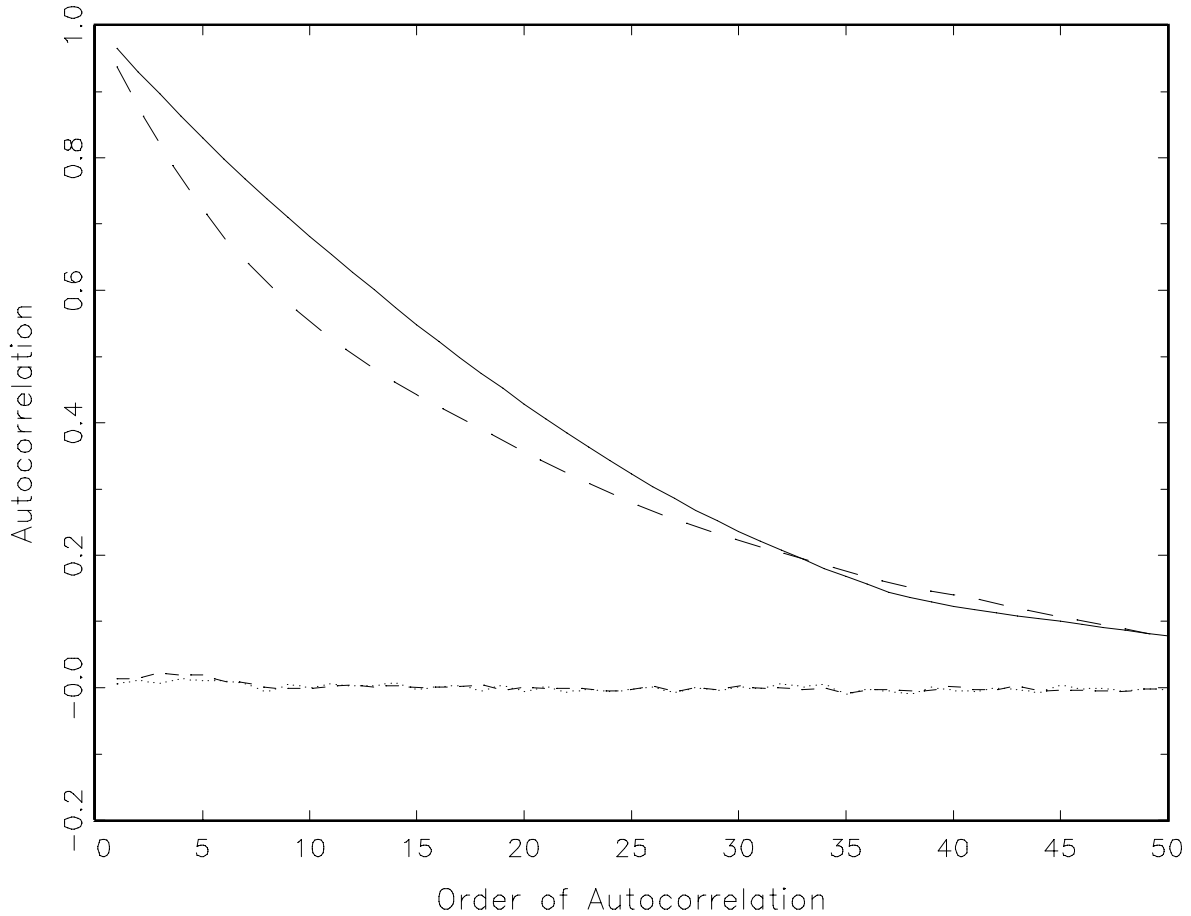
Notes: For a two-year period, the figure presents the ratio of true Value at Risk at the 1% confidence level for a one day holding period to estimated value at risk based on historical simulation. Details on the methods are provided in the text.

Figure 12: Ratio of True and BRW ($\lambda = 0.97$) Estimates of 1% VaR for British Pound



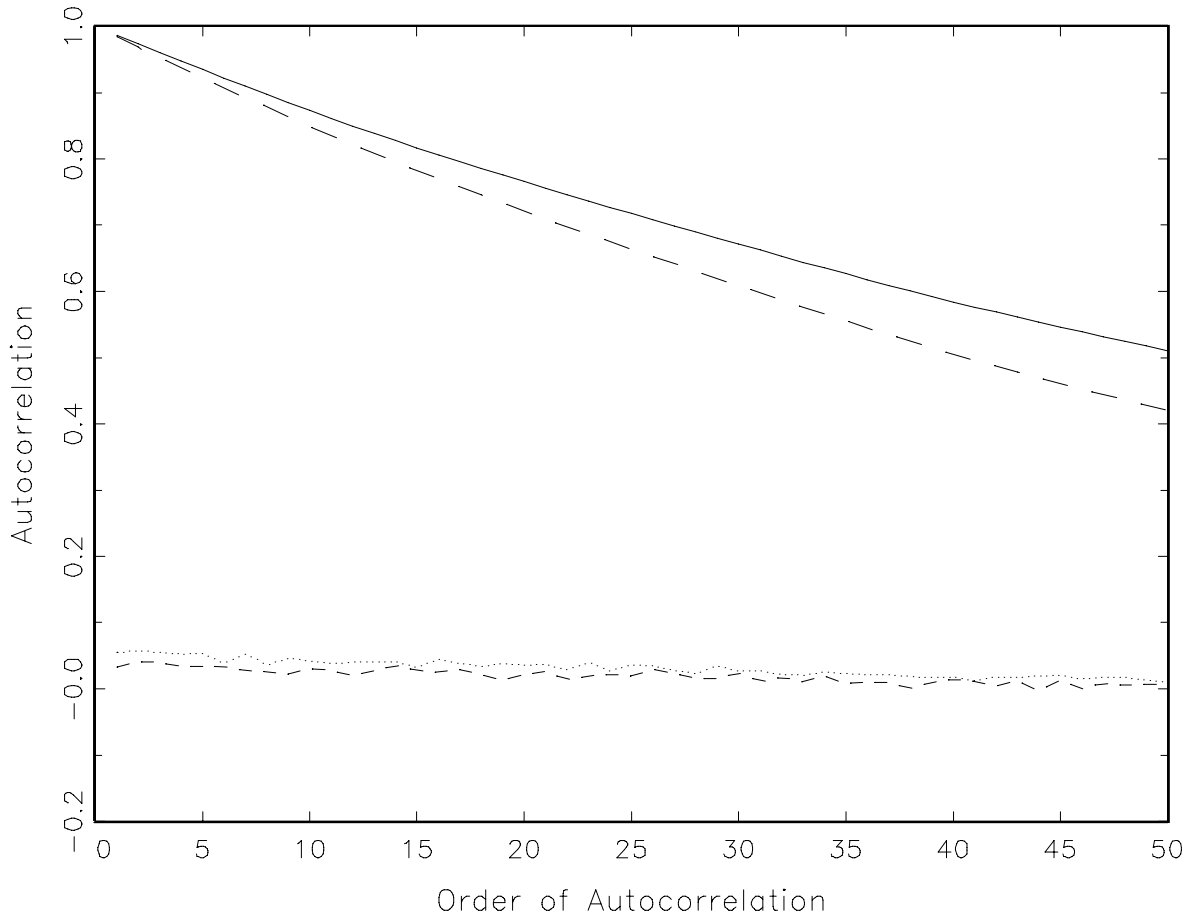
Notes: For a two-year period, the figure presents the ratio of true Value at Risk at the 1% confidence level for a one day holding period to estimated value at risk based on the BRW method with $\lambda = 0.97$. Details on the methods are provided in the text.

Figure 13: Autocorrelation of VaR Errors and VaR Exceedances for BRW ($\lambda = 0.97$) Estimates of 1% VaR for British Pound



Notes: The figure presents daily autocorrelations for the errors in BRW ($\lambda = 0.97$) VaR estimates, and for the VaR exceedances associated with these VaR estimates. All computations were performed for a one-day holding period. Results are presented for 1% VaR errors (solid), 5% VaR errors (dashed), 1% VaR exceedances (short dashes), and 5% VaR exceedances (dots). A VaR exceedance occurs on date t if the portfolio's losses for that date exceed the VaR estimate for that date.

Figure 14: Autocorrelation of VaR Errors and VaR Exceedances for Historical Simulation Estimates of 1% VaR for British Pound



Notes: The figure presents daily autocorrelations for the errors in historical simulation VaR estimates, and for the VaR exceedances associated with these VaR estimates. All computations were performed for a one-day holding period. Results are presented for 1% VaR errors (solid), 5% VaR errors (dashed), 1% VaR exceedances (short dashes), and 5% VaR exceedances (dots). A VaR exceedance occurs on date t if the portfolio's losses for that date exceed the VaR estimate for that date.

Table 1: GARCH(1,1) Estimates for U.S. Dollar Denominated Exchange Rates

Currency	a_0	a_1	b_1
British pound	7.059×10^{-7} (3.511×10^{-8})	0.08428 (0.004302)	0.9010 (0.004434)
Belgian franc	1.177×10^{-6} (1.109×10^{-7})	0.1141 (0.005891)	0.8690 (0.006798)
Canadian dollar	1.088×10^{-7} (1.169×10^{-8})	0.1232 (0.005398)	0.8697 (0.005337)
French franc	6.746×10^{-7} (4.720×10^{-8})	0.1446 (0.005239)	0.8586 (0.004694)
Deutschemark	1.115×10^{-6} (9.750×10^{-8})	0.09343 (0.004589)	0.8856 (0.006038)
Japanese yen	2.147×10^{-6} (1.032×10^{-7})	0.05810 (0.002247)	0.9165 (0.002931)
Netherlands guilder	6.888×10^{-7} (4.996×10^{-8})	0.1043 (0.003866)	0.8865 (0.003323)
Swedish kronor	2.470×10^{-6} (7.781×10^{-8})	0.1588 (0.006095)	0.8031 (0.006042)
Swiss franc	1.261×10^{-6} (1.231×10^{-7})	0.09751 (0.005329)	0.8857 (0.006428)
Italian lira	2.618×10^{-8} (6.486×10^{-9})	0.2057 (0.003426)	0.8428 (0.002422)

Notes: The table provides estimates of the parameters of a GARCH(1,1) model for the exchange rates of the listed currencies versus the U.S. dollar. Standard errors are provided in parenthesis. The estimated model has the form:

$$\begin{aligned}
 r_t &= \sqrt{h_t}u_t, \\
 h_t &= a_0 + a_1r_{t-1}^2 + b_1h_{t-1},
 \end{aligned}$$

where r_t is the natural log of the exchange rate return at time t , h_t is the volatility of r_t conditional on the history of returns through time $t - 1$, and u_t is standard normal and i.i.d. . Further details are provided in the text.

Table 2: Distribution of Undetected Percentage Increases In VaR

Currency	Mean	Std	Skewness
British Pound	5.49244	5.34002	1.80512
Belgian franc	7.31864	7.02216	1.73131
Canadian dollar	7.8657	7.51898	1.71149
French franc	9.13466	8.65992	1.66888
Deutschemark	6.05846	5.86547	1.78085
Japanese yen	3.84307	3.78585	1.88463
Netherlands guilder	6.72428	6.4788	1.75396
Swedish kronor	9.96375	9.39721	1.64327
Swiss franc	6.30919	6.09703	1.77052
Italian lira	12.6349	11.7329	1.57052

Notes: For spot foreign exchange positions in the listed currencies against the U.S. dollar, the table presents theoretical results on the mean, standard deviation, and skewness²⁵ of percentage increases in VaR over a one-day period conditional on the increases not being detected when using the historical simulation or BRW methods. All of the figures were calculated under the assumption exchange rate returns follow the GARCH(1,1) processes in Table 1. Furthermore, all calculations are conditioned on the event that before true VaR changes, the VaR estimates are correct, and the variance of returns are at their long run averages. All figures are for VaR at the 1% confidence level.

Table 3: Simulated Distribution of Undetected Percentage Increases in VaR: Historical Simulation Methods

A. BRW with $\lambda = 0.97$				
Currency	Prob Not Detected	Mean	Std	Skewness
British Pound	0.317996	5.39	5.46	2.0755
Belgian franc	0.323725	7.20	7.24	1.9993
Canadian dollar	0.323363	7.80	7.82	2.0009
French franc	0.327022	9.10	9.12	1.9574
Deutschemark	0.320650	5.95	6.01	2.0460
Japanese yen	0.314639	3.75	3.83	2.1612
Netherlands guilder	0.321212	6.64	6.69	2.0201
Swedish kronor	0.334559	9.82	9.76	1.9017
Swiss franc	0.321032	6.20	6.25	2.0319
B. BRW with $\lambda = 0.99$				
Currency	Prob Not Detected	Mean	Std	Skewness
British Pound	0.323464	5.54	5.60	2.0171
Belgian franc	0.329193	7.40	7.43	1.9595
Canadian dollar	0.328670	8.01	8.03	1.9611
French franc	0.331585	9.31	9.33	1.9171
Deutschemark	0.326318	6.10	6.14	1.9738
Japanese yen	0.320027	3.85	3.91	2.0772
Netherlands guilder	0.326318	6.83	6.89	2.0022
Swedish kronor	0.340469	10.07	9.97	1.8454
Swiss franc	0.326459	6.37	6.42	1.9837
C. Historical Simulation				
Currency	Prob Not Detected	Mean	Std	Skewness
British Pound	0.322238	5.58	5.69	2.0448
Belgian franc	0.328167	7.46	7.56	1.9845
Canadian dollar	0.326539	8.06	8.14	1.9721
French franc	0.328007	9.37	9.49	1.9702
Deutschemark	0.325675	6.16	6.26	2.0104
Japanese yen	0.319966	3.87	3.95	2.0813
Netherlands guilder	0.324027	6.85	6.98	2.0226
Swedish kronor	0.339625	10.16	10.12	1.8644
Swiss franc	0.325172	6.41	6.51	2.0068

Notes: For three VaR methods, when foreign exchange returns are generated as in Table 1 the table presents the simulated empirical frequency with which VaR increases are not detected (Prob Not Detected), and conditional on a VaR increase not being detected, the table presents the mean increase in VaR, the standard deviation of the increase, and the skewness of the increase. Results for the Italian lira are not presented because its simulated return series was explosive for its estimated GARCH parameter values. Additional details on the VaR methods and simulations are contained in the text.

**Table 4: Simulated Distribution of Undetected Percentage Increases in VaR:
Exp Weighting Methods**

A. Exp Weighting VCOV with $\lambda = 0.97$				
Currency	Prob Not Detected	Mean	Std	Skewness
British Pound	0.039961	0.96	0.93	2.1390
Belgian franc	0.057127	1.84	1.83	2.1654
Canadian dollar	0.060644	2.22	2.26	2.3066
French franc	0.074454	3.32	3.71	2.8011
Deutschemark	0.045589	1.17	1.10	1.8345
Japanese yen	0.027096	0.43	0.38	1.7437
Netherlands guilder	0.050674	1.56	1.54	2.1416
Swedish kronor	0.079278	3.15	3.06	1.9064
Swiss franc	0.047720	1.31	1.26	1.8231

B. Exp Weighting VCOV with $\lambda = 0.99$				
Currency	Prob Not Detected	Mean	Std	Skewness
British Pound	0.066494	1.70	1.82	3.0408
Belgian franc	0.085469	2.80	3.01	2.8323
Canadian dollar	0.096786	3.72	4.19	2.7383
French franc	0.117249	5.41	5.94	2.3978
Deutschemark	0.070775	1.87	1.90	2.7070
Japanese yen	0.046674	0.78	0.74	2.1200
Netherlands guilder	0.083318	2.70	3.02	2.8891
Swedish kronor	0.104464	4.04	4.01	2.2063
Swiss franc	0.074956	2.15	2.29	2.9606

C. Exp Weighting VCOV with $\lambda = 1$				
Currency	Prob Not Detected	Mean	Std	Skewness
British Pound	0.068685	3.54	3.77	2.2413
Belgian franc	0.073670	4.91	5.27	2.3930
Canadian dollar	0.080122	5.70	6.10	2.4022
French franc	0.089690	7.28	7.71	2.2046
Deutschemark	0.068725	3.86	4.04	2.0396
Japanese yen	0.060946	2.24	2.35	2.0615
Netherlands guilder	0.075097	4.63	4.96	2.3503
Swedish kronor	0.078172	6.49	6.73	2.0673
Swiss franc	0.070675	4.11	4.40	2.3131

Table 4 Continued...

Notes: For three VaR methods, when foreign exchange returns are generated as in Table 1 the table presents the simulated empirical frequency with which VaR increases are not detected (Prob Not Detected), and conditional on a VaR increase not being detected, the table presents the mean increase in VaR, the standard deviation of the increase, and the skewness of the increase. Results for the Italian lira are not presented because its simulated return series was explosive for its estimated GARCH parameter values. Additional details on the VaR methods and simulations are contained in the text.

Table 5: Simulated Properties of Generalized Historical Simulation VaR

Estimators					
A. BRW with $\lambda = 0.97$					
Currency	% Violations	RMSE	Percent RMSE	Corr w/VaR	Corr w/VaR Changes
British Pound	1.9276	0.0037	23.9760	0.8096	0.3292
Belgian franc	1.9678	0.0054	27.4308	0.7891	0.3254
Canadian dollar	1.9819	0.0031	29.7743	0.8762	0.3047
French franc	2.0060	0.0398	34.1141	0.9120	0.3076
Deutschemark	1.9317	0.0040	24.6741	0.7494	0.3294
Japanese yen	1.8653	0.0041	21.0454	0.6796	0.3362
Netherlands guilder	1.9176	0.0057	26.9853	0.8685	0.3160
Swedish kronor	1.9839	0.0063	30.4198	0.6248	0.3204
Swiss franc	1.9236	0.0050	25.4269	0.7911	0.3286
B. BRW with $\lambda = 0.99$					
Currency	% Violations	RMSE	Percent RMSE	Corr w/VaR	Corr w/VaR Changes
British Pound	1.3809	0.0045	23.9027	0.6970	0.3137
Belgian franc	1.4211	0.0070	28.3879	0.6609	0.3124
Canadian dollar	1.4492	0.0047	32.8127	0.7526	0.3173
French franc	1.5477	0.0638	39.4913	0.8156	0.3037
Deutschemark	1.3648	0.0048	24.1135	0.6296	0.3105
Japanese yen	1.3266	0.0039	17.5141	0.5780	0.3138
Netherlands guilder	1.4070	0.0082	29.0913	0.7507	0.3167
Swedish kronor	1.3889	0.0073	30.2564	0.4967	0.3015
Swiss franc	1.3809	0.0063	25.7077	0.6702	0.3121
C. Historical Simulation					
Currency	% Violations	RMSE	Percent RMSE	Corr w/VaR	Corr w/VaR Changes
British Pound	1.5196	0.0057	28.6479	0.4990	0.2271
Belgian franc	1.5497	0.0087	33.8256	0.4600	0.2184
Canadian dollar	1.6844	0.0060	39.4759	0.6063	0.2175
French franc	1.9518	0.0840	48.5647	0.6910	0.2490
Deutschemark	1.4492	0.0058	28.4983	0.4102	0.2227
Japanese yen	1.3447	0.0046	19.8334	0.3448	0.2270
Netherlands guilder	1.6563	0.0105	35.0124	0.5945	0.2202
Swedish kronor	1.4995	0.0086	35.1230	0.2740	0.1952
Swiss franc	1.5337	0.0078	30.6758	0.4654	0.2230

Table 6: Simulated Properties of Exp Weighted VCOV VaR EstimatorsA. Exp weighting VCOV with $\lambda = 0.97$

Currency	% Violations	RMSE	Percent RMSE	Corr w/VaR	Corr w/VaR Changes
British Pound	1.1658	0.0022	12.2719	0.9233	0.9706
Belgian franc	1.2583	0.0038	16.8230	0.8898	0.9458
Canadian dollar	1.2844	0.0023	19.2503	0.9345	0.9379
French franc	1.4613	0.0298	23.9478	0.9501	0.9139
Deutschemark	1.1819	0.0025	13.4012	0.8870	0.9624
Japanese yen	1.1317	0.0018	8.1686	0.9096	0.9850
Netherlands guilder	1.2362	0.0039	15.9886	0.9371	0.9548
Swedish kronor	1.3789	0.0045	20.7734	0.7681	0.9045
Swiss franc	1.2060	0.0032	14.2889	0.9029	0.9598

B. Exp weighting VCOV with $\lambda = 0.99$

Currency	% Violations	RMSE	Percent RMSE	Corr w/VaR	Corr w/VaR Changes
British Pound	1.3447	0.0038	20.4414	0.7458	0.9120
Belgian franc	1.4271	0.0059	25.7861	0.6979	0.8718
Canadian dollar	1.6221	0.0039	30.7626	0.7940	0.8439
French franc	1.8814	0.0533	38.8553	0.8389	0.8013
Deutschemark	1.3246	0.0040	20.8050	0.6790	0.9045
Japanese yen	1.1276	0.0029	12.7292	0.6870	0.9529
Netherlands guilder	1.4714	0.0069	26.3906	0.7924	0.8737
Swedish kronor	1.4714	0.0060	27.9627	0.5324	0.8298
Swiss franc	1.3950	0.0053	22.6497	0.7130	0.8949

C. Exp weighting VCOV with $\lambda = 1$

Currency	% Violations	RMSE	Percent RMSE	Corr w/VaR	Corr w/VaR Changes
British Pound	1.6060	0.0051	27.5902	0.5149	0.6022
Belgian franc	1.8090	0.0077	33.7106	0.4648	0.5683
Canadian dollar	2.1085	0.0053	41.4323	0.6127	0.5287
French franc	2.6291	0.0734	53.6503	0.6922	0.5022
Deutschemark	1.5899	0.0051	27.1936	0.4266	0.6022
Japanese yen	1.2764	0.0038	16.9711	0.4067	0.6539
Netherlands guilder	1.8975	0.0094	35.7822	0.6012	0.5570
Swedish kronor	1.7508	0.0071	34.2068	0.2900	0.5486
Swiss franc	1.6784	0.0070	29.9581	0.4761	0.5886

Table 7: Correlation Stability of Filtered Innovations: Full Sample

A. Correlations: First Half of Sample.

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	0.4135								
CAD	-0.0252	0.0334							
FFR	0.4305	0.6677	0.0237						
DEM	0.4297	0.8085	0.0447	0.6968					
JPY	0.2220	0.3618	-0.0742	0.3326	0.3901				
NLG	0.4060	0.8199	0.0299	0.6944	0.8565	0.3810			
SKR	0.2973	0.6047	0.0331	0.4954	0.6264	0.2681	0.6027		
CHF	0.3650	0.7042	0.0485	0.6269	0.8089	0.3465	0.7483	0.5490	
ITL	0.3599	0.5445	0.0296	0.5551	0.5279	0.2824	0.5259	0.3789	0.4939

B. Correlations: Second Half of Sample.

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	0.6304								
CAD	0.3905	0.3800							
FFR	0.6517	0.8368	0.4014						
DEM	0.7017	0.8750	0.4355	0.9098					
JPY	0.4994	0.5817	0.3220	0.6060	0.6683				
NLG	0.6800	0.8527	0.4296	0.8891	0.9635	0.6385			
SKR	0.2514	0.2918	0.1824	0.3141	0.3407	0.2199	0.3248		
CHF	0.6739	0.8080	0.4329	0.8363	0.9039	0.6829	0.8772	0.3192	
ITL	0.5218	0.6857	0.3406	0.7187	0.7439	0.4825	0.7252	0.4493	0.6814

C. Correlation Difference: A-B

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	-0.2169								
CAD	-0.4157	-0.3466							
FFR	-0.2211	-0.1692	-0.3777						
DEM	-0.2721	-0.0664	-0.3908	-0.2130					
JPY	-0.2774	-0.2199	-0.3962	-0.2735	-0.2782				
NLG	-0.2740	-0.0328	-0.3997	-0.1947	-0.1069	-0.2575			
SKR	0.0459	0.3128	-0.1493	0.1813	0.2857	0.0482	0.2779		
CHF	-0.3088	-0.1039	-0.3845	-0.2094	-0.0950	-0.3364	-0.1288	0.2298	
ITL	-0.1619	-0.1412	-0.3109	-0.1636	-0.2160	-0.2002	-0.1993	-0.0704	-0.1875

Table 7: Correlation Stability of Filtered Innovations: Full Sample Contd.

D. P-Value for Correlation Difference = 0

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	0.0000								
CAD	0.0000	0.0000							
FFR	0.0000	0.0000	0.0000						
DEM	0.0000	0.0000	0.0000	0.0000					
JPY	0.0000	0.0000	0.0000	0.0000	0.0000				
NLG	0.0000	0.0019	0.0000	0.0000	0.0000	0.0000			
SKR	0.1579	0.0000	0.0000	0.0000	0.0000	0.1446	0.0000		
CHF	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
ITL	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0156	0.0000

Notes: The table tests whether the normalized innovations used in filtered historical simulation have a correlation structure which is stable over time. The data used for this analysis are the daily exchange rates of specific currencies against the U.S. dollar over the period from January 1973 through June 1986. The currencies are the British pound (GBP), the Belgian franc (BFR), the Canadian dollar (CAD), the French franc (FFR), the Deutschemark (DEM), the Japanese yen (JPY), the Netherlands guilder (NLG), the Swedish kronor (SKR), the Swiss franc (CHF), and the Italian lira (ITL). Details on how the filtered historical simulation was implemented are contained in the text.

Table 8: Correlation Stability of Filtered Innovations: Two-Year Sample

A. Correlations: First Year of Sample.

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	0.8442								
CAD	0.5352	0.5771							
FFR	0.8567	0.9639	0.6185						
DEM	0.8591	0.9712	0.5956	0.9910					
JPY	0.7261	0.7539	0.5243	0.7846	0.7943				
NLG	0.8655	0.9675	0.6034	0.9868	0.9899	0.7821			
SKR	0.0062	-0.0049	0.1056	0.0055	-0.0049	-0.0086	0.0020		
CHF	0.8179	0.9149	0.5551	0.9370	0.9448	0.8082	0.9349	-0.0196	
ITL	0.7994	0.9239	0.5948	0.9455	0.9441	0.7485	0.9401	0.0469	0.8840

B. Correlations: Second Year of Sample.

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	0.7641								
CAD	0.3142	0.2596							
FFR	0.7581	0.9635	0.2490						
DEM	0.7734	0.9565	0.2741	0.9758					
JPY	0.5748	0.7100	0.2318	0.7461	0.7623				
NLG	0.7682	0.9511	0.2732	0.9703	0.9940	0.7529			
SKR	-0.0531	-0.1014	-0.0239	-0.0735	-0.0378	0.0125	-0.0482		
CHF	0.7478	0.9037	0.3124	0.9254	0.9465	0.7844	0.9424	-0.0351	
ITL	0.0153	0.0204	-0.0007	0.0408	0.0673	0.0107	0.0620	0.8077	0.0714

C. Correlation Difference: A-B

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	0.0801								
CAD	0.2209	0.3176							
FFR	0.0986	0.0004	0.3695						
DEM	0.0857	0.0147	0.3215	0.0152					
JPY	0.1512	0.0438	0.2924	0.0386	0.0321				
NLG	0.0973	0.0165	0.3302	0.0165	-0.0041	0.0291			
SKR	0.0593	0.0965	0.1294	0.0790	0.0328	-0.0210	0.0502		
CHF	0.0701	0.0112	0.2427	0.0116	-0.0017	0.0238	-0.0074	0.0155	
ITL	0.7841	0.9034	0.5955	0.9047	0.8768	0.7378	0.8780	-0.7608	0.8126

**Table 8: Correlation Stability of Filtered Innovations: Two-Year Sample
Contd.**

D. P-Value for Correlation Difference = 0

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	0.0107								
CAD	0.0025	0.0000							
FFR	0.0013	0.9561	0.0000						
DEM	0.0037	0.0199	0.0000	0.0000					
JPY	0.0032	0.2932	0.0001	0.2999	0.3654				
NLG	0.0009	0.0199	0.0000	0.0000	0.0037	0.4303			
SKR	0.5095	0.2816	0.1491	0.3791	0.7152	0.8151	0.5768		
CHF	0.0425	0.4729	0.0008	0.3293	0.8605	0.4690	0.4862	0.8631	
ITL	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: The table tests whether the normalized innovations used in filtered historical simulation have a correlation structure which is stable over time. The data used for this analysis are the daily exchange rates of specific currencies against the U.S. dollar over the period from January 1973 through June 1986. The GARCH(1,1) estimates to perform the filtering were computed using all of this data. However, the correlations of the filtered innovations were examined using only the last 2-years of the filtered innovations. The currencies are the British pound (GBP), the Belgian franc (BFR), the Canadian dollar (CAD), the French franc (FFR), the Deutschemark (DEM), the Japanese yen (JPY), the Netherlands guilder (NLG), the Swedish kronor (SKR), the Swiss franc (CHF), and the Italian lira (ITL). Details on how the filtered historical simulation was implemented are contained in the text.

Table 9: Filtered Historical Simulation: VaR Errors

Currency	RMSE	Understatement			Overstatement		
		Freq.	Avg.	Std.	Freq.	Avg.	Std.
British pound	11.64	67	10.07	5.81	30	8.05	9.22
Belgian franc	12.18	63	10.90	6.80	34	7.79	8.35
Canadian dollar	13.07	62	11.51	7.27	31	10.27	8.63
French franc	12.38	64	10.40	6.86	33	10.15	7.77
Deutschemark	10.56	65	9.20	5.93	28	8.07	7.38
Japanese yen	14.04	74	11.87	8.37	23	10.34	8.41
Netherlands guilder	10.89	70	9.68	6.15	26	7.06	7.26
Swedish kronor	13.20	60	12.50	8.15	35	7.92	7.30
Swiss franc	10.94	68	9.38	6.33	27	8.19	7.33
Italian lira	14.70	72	11.61	7.71	24	12.73	12.45

Notes: The table presents simulation results on the errors from computing VaR using filtered historical simulation for Spot Foreign Exchange Positions. All results are presented as a percent of true VaR. RMSE is the root mean squared error of the VaR estimates. VaR estimates were labelled as understated when a 95% confidence interval for the VaR percentage error is bounded above by 0. VaR estimates were labelled as overstated when a 95% confidence interval for the VaR percentage error is bounded below by 0. “Freq.” is the percentage of simulations for which VaR was understated; “Avg.” is the average understatement conditional on VaR being understated; and Std. is the standard error of understatement conditional on VaR being understated. Similar labels apply when VaR is overstated.

Table 10: Results on Bias of FHS with Infinite Monte-Carlo Draws

Results from random start date 1				
Currency	Understatement		Overstatement	
	Frequency	Average	Frequency	Average
British pound	37	7.72	32	7.67
Belgian franc	36	8.27	29	8.69
Canadian dollar	37	9.28	26	9.25
French franc	38	7.11	24	7.03
Deutschemark	36	6.26	26	7.63
Japanese yen	45	9.23	26	12.68
Netherlands guilder	47	6.11	28	6.58
Swedish kronor	39	8.83	26	9.96
Swiss franc	42	6.08	29	8.71
Italian lira	43	7.87	29	10.50
Results from random start date 2				
Currency	Understatement		Overstatement	
	Frequency	Average	Frequency	Average
British pound	38	6.23	30	8.05
Belgian franc	38	6.79	34	7.79
Canadian dollar	35	7.78	31	10.27
French franc	32	7.61	33	10.15
Deutschemark	36	5.84	28	8.07
Japanese yen	40	9.69	23	10.34
Netherlands guilder	38	6.47	26	7.06
Swedish kronor	36	8.56	35	7.92
Swiss franc	40	5.59	27	8.19
Italian lira	38	7.51	24	12.73

Notes: The table reports results on the difference between the true 1% quantile of the distribution of ten-day portfolio returns and the 1% quantile of the distribution if it was estimated by filtered historical simulation with an infinite number of draws. Historical simulation was classified as understating (overstating) true VaR if a 90% confidence interval for the difference between the true and filtered historical simulation 1% quantiles was bounded above (below) by zero. The table reports the frequency with which the FHS estimates were classified as understating or overstating VaR. All results are based on 100 simulations of the FHS method using two different random start dates. For those FHS estimates which were classified as over- or under- stated, the table also reports the frequency of under and overstatement. Details on the simulation are contained in the text and in the appendix.

Table 11: Parameters for FHS Simulations

A. Garch Parameters

Currency	$a_0 \times 10^6$	a_1	a_2
British Pound	0.8900 (0.0043)	0.1363 (0.0069)	0.8520 (0.0061)
Belgian franc	1.0341 (0.0078)	0.1768 (0.0082)	0.8253 (0.0071)
Canadian dollar	0.1333 (0.0017)	0.1664 (0.0097)	0.8212 (0.0099)
French franc	0.3520 (0.0037)	0.2229 (0.0077)	0.8198 (0.0054)
Deutschemark	1.2367 (0.0114)	0.1275 (0.0074)	0.8530 (0.0086)
Japanese yen	0.7025 (0.0031)	0.1948 (0.0035)	0.8213 (0.0032)
Netherlands guilder	0.4953 (0.0042)	0.1148 (0.0047)	0.8840 (0.0033)
Swedish kronor	3.2522 (0.0124)	0.3077 (0.0117)	0.6833 (0.0110)
Swiss franc	1.1766 (0.0119)	0.1397 (0.0088)	0.8511 (0.0088)
Italian lira	0.0317 (0.0006)	0.3054 (0.0066)	0.7866 (0.0042)

Table 11: Parameters for FHS Simulations Contd...

B. Conditional Correlations.

	GBP	BFR	CAD	FFR	DEM	JPY	NLG	SKR	CHF
BFR	0.5320								
CAD	0.1962	0.2200							
FFR	0.5464	0.7572	0.2190						
DEM	0.5746	0.8444	0.2500	0.8072					
JPY	0.3731	0.4826	0.1390	0.4767	0.5391				
NLG	0.5508	0.8374	0.2383	0.7947	0.9124	0.5178			
SKR	0.2742	0.4453	0.1070	0.4080	0.4853	0.2428	0.4668		
CHF	0.5284	0.7600	0.2491	0.7350	0.8588	0.5255	0.8156	0.4367	
ITL	0.4405	0.6141	0.1838	0.6343	0.6338	0.3829	0.6231	0.4120	0.5854

Notes: The table presents the parameters that were used to simulate data to study the properties of the filtered historical simulation VaR methodology. All parameters were estimated using data from January 1973 through June 1986. The GARCH(1,1) estimates that were used to generate the data are contained in panel A (with standard errors in parentheses). The conditional correlation matrix of the innovations is contained in panel B. All simulations are for exchange rates of the listed currencies against the U.S. dollar. The currencies are the British pound (GBP), the Belgian franc (BFR), the Canadian dollar (CAD), the French franc (FFR), the Deutschemark (DEM), the Japanese yen (JPY), the Netherlands guilder (NLG), the Swedish kronor (SKR), the Swiss franc (CHF), and the Italian lira (ITL). Details on how the simulations were conducted are contained in the text and the appendix.