Calculating the Optimal Hedge Ratio: 
Constant, Time Varying and the Kalman Filter Approach

Abdulnasser Hatemi-Ja and Eduardo Roca

Abstract

A crucial input in the hedging of risk is the optimal hedge ratio – defined by the relationship between the price of the spot instrument and that of the hedging instrument. Since it has been shown that the expected relationship between economic or financial variables may be better captured by a time varying parameter model rather than a fixed coefficient model, the optimal hedge ratio, therefore, can be one that is time varying rather than constant. In this paper, we suggest and demonstrate the use of the Kalman Filter approach for estimating time varying hedge ratio – a procedure that is statistically more efficient and with better forecasting properties.

Running Title: Optimal Hedge Ratio and the Kalman Filter Approach

JEL: G13, C32, G10

Keywords: Hedge Ratio, Kalman Filter, State-Space Approach, Time Varying Parameters
Calculating the Optimal Hedge Ratio:  
Constant, Time Varying and the Kalman Filter Approach

Abstract

A crucial input in the hedging of risk is the optimal hedge ratio – defined by the relationship between the price of the spot instrument and that of the hedging instrument. Since it has been shown that the expected relationship between economic or financial variables may be better captured by a time varying parameter model rather than a fixed coefficient model, the optimal hedge ratio, therefore, can be one that is time varying rather than constant. In this paper, we suggest and demonstrate the use of the Kalman Filter approach for estimating time varying hedge ratio – a procedure that is statistically more efficient and with better forecasting properties.

Running Title: Optimal Hedge Ratio and the Kalman Filter Approach

JEL: G13, C32, G10

Keywords: Hedge Ratio, Kalman Filter, State-Space Approach, Time Varying Parameters

1. Introduction

Given the volatile financial and economic environment, financial risk has taken centre stage in every sphere of economic activity. The hedging of risk has therefore become a very important issue. Instruments, such as forward and futures contracts, options or derivatives exist which can be used for hedging purposes. However, it is crucial that the optimal quantity of hedging instrument(s) to be used is determined. Thus, the calculation of the optimal hedge ratio plays a critical role in the hedging process. As
will be shown later, basically, the optimal hedge ratio is based on the coefficient of the regression between the change in spot prices and the change in prices of the hedging instrument. However, it is well-recognised that regression parameters could be time varying rather than static. In fact, it has been shown that time varying parameter (TVP) models have performed better than static models. Thus, it is important to test whether a static or dynamic parameter model is more appropriate. If it turns out that a dynamic parameter model is necessary, then this calls for the calculation of time varying optimal hedge ratio. In this paper, we propose and demonstrate a procedure based on the Kalman Filter. Compared to other estimation procedures, the Kalman Filter approach generates estimated parameters with better statistical properties in terms of efficiency and forecasting (see, for instance, Harvey, 1997). At present, there are several methods that have been developed in estimating the optimal hedge ratio, both on a static and dynamic basis (see, Chen and Shrestha, 2003 for a summary). However, the method that we are proposing is not one of those dynamic models. In this paper, we discuss this new method and demonstrate its application using a portfolio of Australian equities to be hedged with the use of futures contracts.

The rest of the paper is organised as follows. The next section presents a brief review of the literature while Section 3 discusses the time varying hedge ratio and its estimation using the Kalman Filter approach. Section 4 demonstrates the application of the Kalman Filter approach in estimating the time varying optimal hedge ratio and the last section provides the conclusion for the paper.
2. Literature Review

There is a significant amount of empirical research on the calculation of the optimal hedge ratio (see, for example, Cechetti et. al., 1988; Myers and Thompson, 1989; Baillie and Myers, 1991; Kroner and Sultan, 1991; Lien and Luo, 1993; and Park and Switzer, 1995). Methods for empirically estimating the optimal hedge ratio have been proposed and developed which generally fall into the use of following: (a) Ordinary Least Squares (OLS) Models, (b) Error Correction (ECM) Models, and (c) Autoregressive Conditional Heteroskedasticity (ARCH)-based Models.

The OLS has been criticised for not taking into account time varying distributions, serial correlation, heteroskedasticity and cointegration. It has been pointed out in the literature that by not considering cointegration, it results in model misspecification and downward bias in hedge ratios and consequently underhedging. Hedge ratios based on ECM models have therefore been found to yield better performance over those derived from OLS methods (see Ghosh, 1995; Lien, 1996; Ghosh and Clayton, 1996; Chou, et. al, 1996 and Sim and Zurbruegg, 2001 among others). ARCH-based models are intended to take into account serial correlation and time varying distributions and hence are also expected to provide better performance over the OLS. There is some evidence in support of this claim (see, for instance Baillie and Myers, 1991, Park and Switzer, 1995 and Lien, et. al. 2002).

Inspite of the existence of evidence pointing to the superiority of hedge ratios estimated with the use of ECM and ARCH-based models over those calculated from OLS models, there is still no clear conclusion as to which method is really best. This is the case because there is also conflicting evidence which can be found in the
literature. For instance, Floros and Vougas (2004) estimated hedge ratios, using data on the Greek stock and futures market – 1999 to 2001, based on the OLS, ECM, VECM and BGARCH models. They found the ECM and VECM to be superior over the OLS model. The BGARCH model even produced a better result than the ECM and VECM models. Lim (1996), in a study of the Nikkei 225 futures contracts hedging performance provided evidence supporting the superiority of the ECM method. Rossi and Zucca (2002) provided support for the superiority of the GARCH hedge ratios over the OLS ones in their study of the hedging performance of the LIFFE traded futures contract Eurolira and German Bund Futures on Italian government bonds.

Bystrom (2003), however, in a study of the hedging effectiveness of the electricity futures contracts in Norway from January 1996 to October 1999, found that the OLS hedge ratio performed slightly better. This is also further supported by the study of Butterworth and Holmes (2001), in a study of the FTSE-Mid250 futures contracts, showed that the OLS through the use of the Least Trujmed Squares, can result in more reliable hedge ratio if outliers in the data are not included. Still further support for the OLS has been put forward by Lien, et. al (2002) who found that the OLS hedge strategy outperforms the GARCH strategy, based on an examination of ten spot and futures markets covering currency futures, commodity futures and stock index returns. There is also the study by Holmes (1995) of the FTSE-100 stock index futures contracts which found that the OLS hedge ratio estimated on an ex-ante basis performed better than the perfect foresight and the naïve hedge ratios.
Thus, there is scope for further research in this area. The methods currently available for estimating the optimal hedge ratio are either based on static or time varying parameter models. However, none of these methods are based on the use of the Kalman filter approach and also on the particular specification of the parameter which is used in this paper.

3. The Optimal Hedge Ratio, Time Varying Coefficient Models and the Kalman Filter Approach

3.1 Estimation of the Optimal Hedge Ratio

The optimal hedge ratio may be defined as the quantities of the spot instrument and the hedging instrument that ensure that the total value of the hedged portfolio does not change. This can be more formally expressed as follows:

\[ V_h = Q_s S - Q_F F \]  \hspace{1cm} (1)
\[ \Delta V_h = Q_s \Delta S - Q_F \Delta F \]  \hspace{1cm} (2)

where:
- \( V_h \) \equiv value of the hedged portfolio,
- \( Q_s \) \equiv quantity of spot instrument,
- \( Q_F \) \equiv quantity of hedging instrument,
- \( S \) \equiv price of spot instrument
- \( F \) \equiv price of hedging instrument
If $\Delta V_h = 0$

then: $\frac{Q_F}{Q_s} = \frac{\Delta S}{\Delta F}$

Let: $h \equiv \frac{Q_F}{Q_s}$ so that

$$ h = \frac{\Delta S}{\Delta F} \quad (3) $$

where: $h \equiv$ hedge ratio.

The hedge ratio, therefore, can be represented as the coefficient in a regression of the price of the spot instrument on the price of the hedging instrument\(^1\). This coefficient, however, may be more appropriately represented as time varying rather than static. It has been argued in the literature that the true relationship between economic or financial variables can be better captured by time varying parameter (TVP) models rather than fixed parameter (FP) models. One reason for this is based on the so-called Lucas (1976) critique, which states that agents rationally anticipate policy changes or the effect of unexpected events and therefore change their behaviour correspondingly. Other reasons, as pointed out by Engle and Watson (1987) and Hatemi-J (2002), is that the data generating process could change due to changes in non-observable factors such as expectations. Also, as pointed out by the same authors, certain models may be misspecified partly because the non-whiteness of their error terms is due to the time varying nature of the coefficient. Finally, it has been shown that TVP models have superior forecasting properties (see Phillips, 1995 and Brown et. al. 1997).

\(^1\) Floros and Vougas (2004) has demonstrated that this is the case whether prices or returns are used (see p. 129 for formal proof).
Thus, it is important to test whether the hedge ratio is one that is captured by a FP model and therefore constant, or one that is better estimated by a TVP model and therefore time varying. Within this context, we propose to estimate the following two alternative models as a basis for calculating the optimal hedge ratio:

\begin{align*}
S_t &= a + h F_t + e_t, \quad \text{(4)} \\
S_t &= a + h_t F_t + u_t, \quad \text{(5)} \\
h_t &= h_{t-1} + v_t, \quad t = 1, ..., T
\end{align*}

where $T$ is the final time period. Equation (4) represents a fixed coefficient model where $a$ and $h$ are fixed coefficients to be estimated and $e_t$ is the error term, which is assumed to be white noise with zero mean, constant variance and no autocorrelation. Hence, in Equation (4), the hedge ratio is static. On the other hand, Equation (5) is a TVP model where the hedge ratio is time varying. The first equation in (5) is known as the observation (or measurement) equation and the second as the state (or transition) equation. The state equation describes the dynamics of the coefficient $h$, which is assumed to follow an autoregressive process of the first degree. The error terms $u$ and $v$ are assumed to be independent white noise processes. This state space model can be estimated by applying the Kalman filter. The time path of the estimated hedge ratio can then be traced. We discuss below the derivation of the Kalman Filter.

In order to choose a model that is more suitable to capture the relationship between the variables of interest, we will test the fixed parameter model – Equation (4) as a null hypothesis against the state space model - Equation (5), as the alternative hypothesis by performing a likelihood ratio test.

\footnote{It should be mentioned that a potential measure of the cost of hedging could be obtained by evaluating the difference between the integral of the fixed hedge ratio ($h$) and the time varying hedge ratio ($h_t$).}
3.2 Derivation of the Kalman Filter

We provide a brief discussion of the Kalman Filter approach. In order to make the discussion concise, we utilise matrix algebra. A general specification of a state space model of \( N \times 1 \) vector of \( y_t \) as a function of \( x_t \) (a \( N \times m \) matrix of variables) is given by the following system of equations:

\[
\begin{align*}
    y_t &= \beta_t' x_t + \varepsilon_t, \\
    \beta_t &= A \beta_{t-1} + \eta_t,
\end{align*}
\]

where \( \varepsilon \) and \( \eta \) are error terms that are assumed to be independently distributed as:

\[
\begin{align*}
    \varepsilon &\sim NID(0, \sigma), \\
    \eta &\sim NID(0, Q).
\end{align*}
\]

These conditions make it possible that the likelihood inference is available. In addition we assume that the error terms across the equation are non-correlated, that is:

\[
E(\varepsilon_t, \eta_s) = 0, \quad \forall \ t \text{ and } s. \quad (8)
\]

The aim is to estimate the parameters \( A, P \) and \( Q \) and make inference about the time varying vector \( \beta \), given observations of \( (y_t, x_t) \) for each time period. This can be done by applying the Kalman filter, which is an algorithm for estimating forecasts of the state vector given the initial values and data observed up to date \( t \). The filter consists of the following equations for each period in time:
\[
\begin{align*}
\hat{\beta}_{t|t-1} &= A\hat{\beta}_{t-1} \\
\hat{P}_{t|t-1} &= AP_{t-1}A' + Q \\
\hat{e}_t &= (y_t - \hat{\beta}_t x_{t|t-1}) \\
\hat{f}_t &= x_t'P_{t|t-1}x_t + \sigma \\
\hat{\beta}_t &= \hat{\beta}_{t|t-1} + (P_{t|t-1}x_t)'\left(\frac{\hat{e}_t}{\hat{f}_t}\right) \\
\hat{P}_t &= P_{t|t-1} - P_{t|t-1}x_t(x_t'P_{t|t-1}x_t + \frac{1}{f_t})
\end{align*}
\]

where \( \hat{\beta}_t \) is the optimal estimated value of the state vector \( \beta_t \). \( P_t \) denotes the variance of \( \hat{\beta}_t \). The one step ahead prediction error is denoted by \( \hat{e}_t \) and its variance is \( \hat{f}_t \). The indices \( t|t-1 \) represents the optimal estimator of a parameter at time \( t \) based on all information available at time \( t-1 \). Once the end of the series is reached, optimal prediction of future observations can be made. It is also possible to use backward recursion to estimate optimal estimator of the state vector to be calculated at all points in time using the full sample according to Harvey (1997).

Note that when estimating our state space model (6) the matrix of the transition equation, i.e. \( A \), is reduced to an identity matrix. Hence, the two first equations will be reduced to:

\[
\begin{align*}
\hat{\beta}_{t|t-1} &= \hat{\beta}_{t-1} \\
\hat{P}_{t|t-1} &= P_{t-1} + Q,
\end{align*}
\]

For more details and further analysis of the Kalman filter based on the maximum likelihood approach, the interested reader is referred to Harvey (1990, 1997).
4. Application of the Proposed Procedure – Time Varying Hedge Ratios Based on the Australian Equity and Futures Markets

We demonstrate the application of the above procedures using the example of an investment in the Australian equity index to be hedged by a short position in the Australian futures market. This application is interesting considering that the Australian equity market is one of the significant and better performing markets in the world with a good amount of index investment going on. The Australian futures market is also one that is fairly active and has been in fact the subject of a number of scholarly studies (see Frino and West, 1999 and 2000, for example). In this application, the spot instrument therefore corresponds to the equity market index while the hedging instrument is the futures market index. We make use of the MSCI price index for the Australian equity market and the Share Price Index for the Australian futures market. We collect daily data covering the period January 1, 1988 to September 8, 2001 for a total of 3,586 observations.

In order to avoid spurious regression results, it is necessary to first test whether each variable is stationary or non-stationary and if each one is non-stationary, we then conduct a cointegration test. To test for stationarity of each variable, we perform the Perron (1989) test for unit roots on the data. This test has better power properties it allows for a shift in the mean value as well as a shift in the trend for the underlying variable. The results from this test are reported in Table 1. For the null hypotheses of $I(1)$, i.e. integration of the first order, the estimated test statistics are found to be less than the critical values at the conventional significance levels. Hence, the null hypothesis that each variable is $I(1)$ cannot be rejected. The next step is to investigate whether each variable becomes stationary after taking the first difference.
Nevertheless, the null hypothesis that each variable is $I(2)$ is rejected at the one percent significance level. Hence, we can conclude that each variable contains one unit root.

Insert Table 1 here

Before estimating each model, we therefore conduct a cointegration test between the two variables in order to avoid spurious and false inference. For this purpose, we make use of the Engle-Granger test. For the FP model (Equation 4), we obtained an estimated value of $-7.447$. Since the critical value is $-3.94$ at the one percent level, the null hypothesis of no cointegration is strongly rejected. We therefore proceed to estimate the parameters of the FP model using OLS. The estimated parameters for Equation 4 are shown below:

$$S_t = 16.13 - 0.76 F_t + e_t$$

(11)

$$(0.816) (0.078)$$

We test the TVP model (Equation 5) for cointegration following Hatemi-J (2002). The estimated value is -9.681 and the critical value at the one percent level of significance is $-3.94$. Thus, the null hypothesis of no cointegration is also strongly rejected for this model. We therefore proceed to estimate the TVP model using the Kalman Filter procedure discussed in Section 2. We obtain the results given below with the standard errors shown in parenthesis:

$$S_t = 16.30 - 0.77 F_t + u_t.$$  

(12)

$$(5.631) (0.001)$$
We then test which of the two models perform better. We set up the FP model (Equation 4) as the null hypothesis and the TVP model (Equation 5) as the alternative hypothesis. For this particular purpose, we make use of the following likelihood ratio (LR) test suggested by Hatemi-J (2002):

\[ LR = -2 \ln \left( \frac{\hat{L}_R}{\hat{L}_U} \right), \tag{13} \]

where \( \hat{L}_R \) and \( \hat{L}_U \) are the estimated values of the likelihood function for the restricted (FP) model and unrestricted (TVP) model, respectively. This test is chi-square distributed with one degree of freedom.

The estimated value for the \( LR \) test is 27.24 and the critical value at the one percent significance level is 6.63. This implies that the null hypothesis of constant parameter model is strongly rejected in favour of the alternative hypothesis of a time varying model.

We checked for the performance of each model in terms of ex ante forecasting properties. Each model was used for forecasting during 30 periods in time. The constant parameter model produced an average forecasting error of 6.1 percent while the time varying parameter model had only 1.7 percent average forecasting error. Thus, the time varying parameter model is superior in forecasting over the constant parameter model.

Thus, we calculate the optimal hedge ratio based on the TVP model. It can be seen from Equation (12) that the estimated \( h \) coefficient (for the last period) is significant. This implies in this case, that futures contract, at least deserves consideration as a possible hedging instrument for a portfolio consisting of Australian equities.
However, since the optimal hedge ratio is based on the TVP model, this means that it varies over time. Its time path is shown in Figure 1 below:

Insert Figure 1 here

As can be seen in Figure 1, the hedge ratio is very volatile for about half of the sample period. For the period between (a) and (b), the time varying hedge ratio is above the fixed hedge ratio. For the rest of the period, the time varying hedge ratio is below the fixed hedge ratio. It can also be seen from Fig. 1 that up to observation 1700 which corresponds approximately to the year 1994, the absolute value of the hedge ratio has a decreasing trend. This indicates that the returns in the futures market have been greater than returns in the stock market over this time period. The development of derivatives and the popularity of these instruments as investment avenues coupled with microstructure reforms instituted in the Australian futures market particularly in relation to making trading more efficient (see Frino and Mckenzie, 2002) could have been the significant reasons behind this good performance of the futures market during the said time period.

If this time varying nature of the hedge ratio is a guide of what is going to happen in the future, there is therefore difficulty in terms of hedging because this implies that for the hedge to work, the hedged portfolio must be rebalanced on a period-by-period basis. This may involve huge transaction costs and therefore it may not be worth using this particular instrument (futures contract) for hedging. Thus, the investor may search for a more suitable hedging instrument subjecting this instrument again to the procedure that we have demonstrated.
5. Conclusion

We have demonstrated in this paper that in calculating the optimal hedge ratio, it is important to determine whether a static or dynamic model is the appropriate model to use. If a time varying coefficient model is called for, we propose the estimation of this based on the Kalman Filter as this approach generates estimated parameters which have better statistical properties. The time path of the estimated hedge ratio can then be analysed to determine whether or not there is great volatility in the hedge ratio as this will have implications in terms of the costs of rebalancing the hedged portfolio. Based on this, the suitability of the hedging instrument can therefore be evaluated.

Acknowledgements

We would like to thank an anonymous referee and the co-editor of this journal for the useful comments that resulted in a significant improvement of this paper. The usual disclaimer applies.

References


Tables and Figures

Table 1
Test for Unit Roots Using the Perron Test

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>$H_0$: $I(1)$, $H_1$: $I(0)$</th>
<th>$H_0$: $I(2)$, $H_1$: $I(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share prices</td>
<td>-2.99(1)</td>
<td>-44.51(0)***</td>
</tr>
<tr>
<td>Futures Prices</td>
<td>-2.83(1)</td>
<td>-7.03(0)***</td>
</tr>
</tbody>
</table>

Notes:
(a) The critical value is −4.90 and −4.24 at the 1% and 5% significance level, respectively.
(b) The notation *** implies significance at the one percent significance level.
(c) The numbers in the parentheses indicate the number of lags required to remove potential
    autocorrelation in the Perron regression (equation 1) at the 5% significance level using the LM
    test.

Figure 1: Time Path of the Hedge Ratio

Notes: TVH represents the time-varying parameter hedge ratio and CH represents the constant parameter hedge ratio.