Optimal Hedge Ratio Estimation

by

Robert J. Myers and Stanley R. Thompson

OPTIMAL HEDGE RATIO ESTIMATION

Robert J. Myers and Stanley R. Thompson*

A major problem faced by participants in spot and futures markets is to choose the proportion of spot positions that are covered by opposite positions on the futures market. This is the problem of optimal hedge ratio estimation (Johnson, Stein, Heifner). A commonly proposed solution is for individuals who are long in the spot market to choose a hedge ratio equal to the ratio of the covariance between spot and futures prices to the variance of the futures price (Benninga, Eldor, and Zilcha; Kahl, 1983). The optimal hedge ratio is then estimated as the slope coefficient in a simple regression of spot price on futures price (Ederington), or from a simple regression of spot price changes on futures price changes (Brown, 1985; Carter and Loyns). The question of whether price levels or price changes should be used has become a controversial issue (Brown, 1986; Kahl, 1986; Hill and Schneeweis; Bond, Thompson, and Lee).

In this paper we examine whether prices should enter the simple regression model in the form of levels or changes when estimating optimal hedge ratios. It is found that the debate so far has been too narrowly focused in that a simple regression model of any form is usually inadequate. A more general time series approach is required for optimal hedge ratio estimation. A model selection strategy for the time series approach is provided and a set of statistical tests for evaluating the adequacy of the simple regression approach is developed. In an example of optimal hedge ratio estimation for soybean and corn storage, it is shown that both simple regression approaches may lead to errors in the estimation of optimal hedge ratios.

Derivation of the Optimal Hedging Rule

In this section, the optimal hedging rule is derived using the mean-variance framework of Kahl (1983). The derivation is for a storage hedge, but the extension to a production hedge is straight forward when the production process is linear.

The profit from an individual's spot and futures transactions is denoted

$$\pi_t = (p_t - p_{t-1} - c_{t-1})q_{t-1} + (f_{t-1} - f_t)b_{t-1}$$

where π is profit, p is the spot price, c is the (linear) cost of storage, q is quantity stored, f is the futures price, b is the quantity of futures contracts sold (purchased if negative), and the subscripts indicate time. Individuals choose spot and futures positions in period t-1 to maximize a linear function of the mean and variance of profits in period t:

---

* Assistant Professor and Professor, Department of Agricultural Economics, Michigan State University, East Lansing, Michigan 48824.

$$\max_{q_{t-1}, b_{t-1}} \quad E(\pi_t/\Phi_{t-1}) - \alpha Var(\pi_t/\Phi_{t-1})$$

where E is the expectations operator, Φ_{t-1} is an information set available at t-1, α is a measure of the individual's risk aversion and Var is the variance operator. Notice that the mean and variance of profit are <u>conditional</u> moments that depend on information available at time t-1.

Kahl and others have shown that the optimal hedge ratio for this case is

$$(1) \qquad r_{t-1} = \frac{\mu_{ft}\sigma_p^2 - \mu_{pt}\sigma_{pf}}{\mu_{ft}\sigma_{pf} - \mu_{pt}\sigma_f^2}$$

where $r_{t-1} = b_{t-1}/q_{t-1}$ is the hedge ratio, $\mu_{ft} = E(f_t/\Phi_{t-1}) - f_{t-1}$ is the expected return from holding futures, $\mu_{pt} = E(p_t/\Phi_{t-1}) - p_{t-1} - c_{t-1}$ is the expected return from holding the commodity, $\sigma_f^2 = Var(f_t/\Phi_{t-1})$ is the conditional variance of the futures price, $\sigma_p^2 = Var(p_t/\Phi_{t-1})$ is the conditional variance of the spot price, and $\sigma_{pf} = Cov(p_t, f_t/\Phi_{t-1})$ is the conditional covariance between spot and futures prices. Although this was not made explicit by Kahl, all of the moments are conditional on information available at time t-1, and conditional variances and covariances (but not conditional means) are assumed constant through time.

To determine the optimal hedge ratio, the first two conditional moments of spot and futures prices must be estimated and substituted into (1). Before discussing how this might be done, however, consider the following well known proposition:

<u>Proposition 1</u> If the expected return from holding a futures contract is zero, then the optimal hedging rule reduces to:

$$(2) \qquad r_{t-1} = \frac{\sigma_{pf}}{\sigma_f^2}$$

<u>Proof</u> Set μ_{ft} equal to zero in equation (1).

This proposition is crucial to optimal hedge ratio estimation. It shows that if futures market returns satisfy the martingale difference property (futures markets are "efficient") then the simplified hedging rule, (2), can be used.

If the hedge ratio is chosen to minimize the conditional variance of profit rather than to maximize a linear function of mean and variance, then the expected returns from holding futures would be irrelevant. Thus a corollary to Proposition 1 is:

<u>Corollary</u> If the objective is to minimize the variance of profits then equation (2) is the optimal hedge ratio, irrespective of whether futures prices satisfy $\mu_{ft} = 0$.

The minimum variance objective is convenient because it leads to the simple hedging rule (2) without requiring the assumption that futures market returns satisfy the martingale difference property. However, variance minimization is not consistent with expected utility maximization (except at extremely high levels of risk aversion) and so the corollary is of limited usefulness. In practice, we need to test for the martingale difference property in futures market returns before equation (2) can be used to estimate optimal hedge ratios.

A Time Series Approach to Estimation

The model in the previous section focuses on the behavior of an individual hedger. Thus, it says nothing about how to estimate the conditional moments and implement the optimal hedging rule. To estimate these conditional moments, an equilibrium model of the spot and futures markets is required. A general specification of such a model in its structural form might be

$$By_t = A_o + A(L)y_{t-1} + u_t$$

where y_t is a (kx1) vector of endogenous variables, B is a (kxk) matrix of structural parameters, A_o is a (kx1) vector of constant terms, A(L) is (kxk) matrix of finite order polynomials in the lag operator L, and u_t is a (kx1) vector of serially uncorrelated error terms with $E(u_t) = 0$ and $E(u_t u_t') = \Omega$ for all t.

In this structure, all variables are specified endogenous but the possibility that some variables are exogenous is not ruled out.[1] Candidates for the components of y_t include futures prices, spot prices, quantities produced, quantities in storage, the interest rate, the exchange rate and any other variable that affects market equilibrium in the commodity under investigation.

The reduced form of this structural model is

$$(3) \qquad y_t = C_o + C(L)y_{t-1} + v_t$$

where $C_o=B^{-1}A_o$, $C(L)=B^{-1}A(L)$, and $v_t=B^{-1}u_t$. Notice that $E(v_t)=0$ and $E(v_t v'_t)=B^{-1}\Omega B^{-1'}$. This covariance matrix for v_t is denoted Γ.

Knowledge of the structural form is not necessary for optimal hedge ratio estimation because the conditional moments of spot and futures prices can be estimated straight from the reduced form. Thus, the reduced form provides all the information necessary to implement the optimal hedging rule. More formally, suppose the information set available

---

1 Defining some variables to be exogenous is equivalent to imposing a set of restrictions on B and A(L).

to hedgers is $\Phi_{t-1} = \{Y_{t-s}:s\geq 1\}$. Then $E(y_t/\Phi_{t-1}) = C_o+C(L)y_{t-1}$ and $Cov(y_t/\Phi_{t-1})=\Gamma$ so that ordinary least squares (OLS) estimates of C_o and the parameters in $C(L)$ provide estimates of the first and second conditional moments of y_t.[2] Substituting the relevant conditional moments in equation (1) gives an estimate of the optimal hedge ratio. If the conditions of Proposition 1 or its corollary are met then the simplified equation (2) can be used.

This reduced form approach requires that the optimal hedge ratio be estimated <u>indirectly</u> by first estimating the conditional moments of spot and futures prices and then using equation (1) or (2). An alternative is to first determine whether the conditions of Proposition 1 or its corollary are met. If so, then equation (2) is the optimal hedging rule. Equation (2) can then be estimated <u>directly</u> by <u>augmenting</u> the reduced form equation for spot price through the addition of the current futures price, f_t, as an additional regressor. The OLS estimate of the coefficient on f_t in the augmented reduced form equation is then exactly equal to the estimated ratio of the conditional covariance between spot and futures prices to the conditional variance of the futures price, as would be obtained through the reduced form approach. Thus, the OLS estimate of the coefficient on f_t in the augmented reduced form equation gives a direct estimate of equation (2). This is shown in the following proposition.

<u>Proposition 2</u>: Let two reduced form equations from the model be:

$$y_{it} = C_{io} + C_i(L)y_{t-1} + v_{it}$$

and

$$y_{jt} = C_{jo} + C_j(L)y_{t-1} + v_{jt}$$

where the i and j subscripts indicate the ith and jth row of a vector or matrix. Furthermore, define the <u>augmented</u> reduced form equation

$$(4) \quad y_{it} = C_{io} + \delta y_{jt} + C_i(L)y_{t-1} + w_{it}$$

where $w_{it}= v_{it}-\delta y_{jt}$. Then the OLS estimate of δ is equal to the ratio of $Cov(y_{it},y_{jt}/\Phi_{t-1})$ to $Var(y_{jt}/\Phi_{t-1})$.

---

2 An estimate of the required covariance matrix is obtained by cross multiplying vectors of OLS residuals from the estimated reduced form. In principle, a seemingly unrelated regression (SURE) approach that allows for contemporaneous correlation among the reduced form residuals should be used. But since the reduced form equations all have the same regressors, SURE reduces to OLS and using the OLS residuals provides a fully efficient estimate of the required covariance matrix.

Proof: See the appendix.

It is important to note that the augmented reduced form (4) has no structural interpretation. Its estimation is merely a mechanism for direct estimation of a ratio of conditional covariance to conditional variance. Thus, simultaneous equations bias from including the endogenous y_{jt} as a regressor in the augmented reduced form equation is not an issue. Nor are any of the other "estimation problems" stemming from the properties of the error term in the augmented reduced form equation. Inclusion of the y_{jt} variable is just a simple way of estimating the required conditional covariance to conditional variance ratio. Examination of two special cases will illustrate the approach.

Simple Regression Models as a Special Case

The standard approach to optimal hedge ratio estimation is to regress spot price levels or changes on futures price levels or changes and interpret the resulting slope coefficient as the optimal hedge ratio, presumably under the assumption that the conditions of Proposition 1 or its corollary are satisfied. In this section, it is shown that this is just a special case of the general time series approach described above.

First, consider the simple regression using price levels:

$$(5) \qquad p_t = c_i + \delta f_t + w_t$$

where the OLS estimate of δ is taken as the optimal hedge ratio. This is an augmented reduced form equation for spot price from the simple reduced form system:

$$p_t = c_i + v_{it}$$

$$f_t = c_j + v_{jt}.$$

If p_t and f_t are equal to a constant plus a serially uncorrelated error then OLS estimation of δ in (5) will give an estimate of the ratio of $Cov(p_t, f_t/\Phi_{t-1})$ to $Var(f_t/\Phi_{t-1})$. This is obviously a very restrictive assumption that is unlikely to be satisfied in most circumstances.[3] And even if spot and futures prices do satisfy this property, then the returns on the futures market cannot satisfy the martingale difference property so Proposition 1 cannot be applied. Thus, while OLS estimation of δ would give an estimate of the relevant covariance to variance ratio, this covariance to variance ratio is not the optimal hedge ratio, unless the objective is variance minimization (see the corollary). Therefore, the

---

3 The assumption implies that high (low) prices are likely to be followed by lower (higher) ones thus creating significant arbitrage opportunities.

simple regression of spot price levels on futures price levels is generally an incorrect approach to optimal hedge ratio estimation.

Second, consider the simple regression using price changes:

$$(6)\ p_t - p_{t-1} = c_i + \delta(f_t - f_{t-1}) + w_t$$

where again the OLS estimate of δ is taken as the optimal hedge ratio. This is an augmented reduced form equation for spot price from the reduced form system

$$p_t - p_{t-1} = c_i + v_{it}$$

$$f_t - f_{t-1} = c_j + v_{jt}.$$

If spot and futures price changes are equal to a constant plus a serially uncorrelated error, then OLS estimation of δ in (6) will give an estimate of the ratio of $Cov(p_t, f_t/\Phi_{t-1})$ to $Var(f_t/\Phi_{t-1})$.[4] Notice that futures price changes may satisfy the martingale difference property in this case, provided the hypothesis that c_j equals zero is accepted. This reduced form is also more consistent with knowledge of the time series properties of spot and futures prices than the reduced form implicit in the price levels equation. Thus, the use of price changes has some distinct advantages over the use of price levels if one is going to use a simple regression approach. Nevertheless, the simple regression using price changes is still a very special case. In terms of equation (4), the restriction is that all of the parameters of $C_i(L)$ are zero (no available information is helpful in predicting price changes).

A Strategy for Model Selection

The reduced form equations (4) are a system of stochastic difference equations. Thus, a strategy for estimation is to begin with a general model that is deliberately overfitted and then test down to a more specific model. This approach has two distinct advantages. First, it allows for testing (rather than simply assuming as in previous studies of optimal hedge ratio estimation) the critical hypothesis that futures market returns satisfy the Martingale difference property. Second, it allows for testing the adequacy of the simple regression approaches to optimal hedge ratio estimation versus the more general time series approach.

The general model would be characterized by a y_t vector of high dimension and a C(L) matrix of high order polynomials. The decisions on exactly which variables to include and what lag lengths to use will be

---

[4] Since p_{t-1} and f_{t-1} are known when the hedging decision is made, the conditional variances and covariance of price changes are equal to the conditional variances and covariance of the price levels themselves.

determined in part by economic theory and in part by the length of available data series.

After deciding which variables to include in the model, it is important to determine whether these variables are stationary or nonstationary. Dickey and Fuller (1979, 1981) have developed tests for unit roots which involve estimating models of the form

$$Y_{it} - Y_{it-1} = C_i + \beta Y_{it-1} + d(L)\,(Y_{it-1} - Y_{it-2}) + e_{it}$$

and testing whether $\beta=0$. If $\beta=0$ then $\{Y_{it}\}$ is nonstationary and if $\beta<0$ then $\{Y_{it}\}$ is stationary.

If evidence of unit roots is found, the usual procedure would be to specify the model in first difference form (or higher order differences if necessary). However, the theory of co-integration has highlighted some difficulties with this approach (Engle and Granger). In particular, if the variables in Y_t are co-integrated of order (1,1) no finite order autoregressive representation exists for the $\{Y_t\}$ process modeled in first difference form. A simple test for co-integration is provided by the Durbin-Watson statistic from the OLS regression

$$Y_{it} = \beta_0 + \beta_1 Y_{jt} + e_{it}$$

where Y_{it} and Y_{jt} are two variables hypothesized to be co-integrated.

Once the degree of differencing has been determined, a number of hypothesis tests can be undertaken. One important hypothesis is whether futures market returns satisfy the martingale difference property. If this hypothesis is rejected, then conditional means are important and equation (1) must be used to estimate the optimal hedge ratio. In this case, the optimal hedge ratio will vary over time as conditional means change. If the martingale hypothesis is accepted, then the simplified equation (2) can be used. In this case, the optimal hedge ratio can be estimated as the coefficient on f_t in the augmented reduced form equation for p_t. But even in this case, there is still the question of whether lagged values of y_t belong in the reduced form equation. By estimating the general model and testing whether the parameters in $C_i(L)$ are zero, it can be determined whether (5) or (6) are misspecified by the exclusion of relevant regressors.

An Example of Storage Hedging in Michigan

We estimated optimal hedge ratios for corn and soybean storage in Michigan to illustrate the time series method. Taking the simplest possible approach, we assumed that y_t for each commodity contains only spot and futures prices for that commodity.[5] Data were collected, relevant hypothesis tests were undertaken, and optimal hedge ratios were estimated using three different approaches.

Weekly observations on spot prices were taken as the mid-week (Wednesday) closing price in Saginaw. The appropriate futures price series depends on the hedging strategy. Many previous studies have used data on a nearby futures contract and rolled forward to the next as that contract approached maturity. This is appropriate for hedgers who follow a similar strategy of hedging in nearby contracts and then rolling the hedge forward if the spot position is not liquidated before the contract matures. An alternative which would save transaction costs is to hedge in a more distant contract (but not the new crop future) and simply liquidate the futures position at whatever time (prior to the next harvest) the spot position is liquidated. For this strategy, the appropriate price data run from the beginning of the crop year until the beginning of the month in which the chosen futures contract matures.

The second alternative was applied in this study using the July futures contract for both corn and soybeans. Weekly prices were taken as daily closing July futures price on the Wednesday of that week at the Chicago Board of Trade. Notice that this approach leads to a set of time series data with missing observations between each July and the beginning of the next crop year. Futures price data were collected for the crop years 1978 through 1984 from various issues of the Statistical Annual, Chicago Board of Trade. For the identical mid-week point, corresponding spot price quotes in Saginaw were obtained from Mid-States Terminals, Toledo, Ohio.

Tests for nonstationarity provided t statistics of -0.22 for corn spot prices, -0.44 for corn futures prices, -0.65 for soybean spot prices, and -2.20 for soybean futures prices. From table 8.5.2 in Fuller, the null hypothesis of nonstationarity can only be rejected at high significance levels (greater than 10%). Similar tests on the first difference model revealed it to be stationary. The co-integrating regression provided a Durbin-Watson statistic of 0.14 for corn and 0.2 for soybeans both of which indicate that spot and futures prices are not co-integrated. Thus, the time series model was specified in first difference form.

A bivariate vector autoregression (VAR) on price changes, with one through fifteen week lags, was specified as the general model. We then tested four special cases using the likelihood ratio statistic and chi-square distribution. The first hypothesis was the simple price change model in which spot and futures price changes are equal to a constant plus a serially uncorrelated error. This implies that $C(L) = 0$ and OLS estimation of the augmented reduced form (6) would give the correct

---

5 Thus we show that even if one restricts attention to the information contained in past spot and futures prices, without including other sources of information, th time series method remains a more general approach to optimal hedge ratio estimation.

conditional covariance to variance ratio. The second hypothesis was that spot and futures price changes have zero expectation. This is the joint martingale model and implies that C(L) = 0 and C_0 = 0 so that OLS estimation of (6), excluding the constant term, would give the correct optimal hedge ratio estimate. The third and fourth hypotheses are separate tests of the martingale difference property for spot price changes and futures price changes, respectively. The martingale test for futures price changes is crucial to whether the simplified equation (2) can be used to estimate the optimal hedge ratio.

Significance levels for each of these tests are given in Table 1. There are strong grounds for rejecting all of the null hypotheses except the martingale difference property for corn and soybean futures price changes. Thus the simple price change model (6) will give an incorrect optimal hedge ratio estimate. Furthermore, since price levels are nonstationary and past price information clearly helps predict future (spot) price changes, the simple price level model, (5), is also inappropriate.

TABLE 1

Significance Levels for Test Statistics

	Simple Price Change Model[a]	Joint Martingale Model	Spot Price Martingale Model	Futures Price Martingale Model
Corn	0.0000	0.0000	0.0020	0.0211
Soybeans	0.0000	0.0000	0.0065	0.0200

a The null hypothesis is that spot and futures price changes are both equal to a constant plus a serially uncorrelated error.

If one accepts the hypothesis that corn and soybean futures price changes satisfy the martingale difference property, then the correct optimal hedge ratio estimate can be obtained from the OLS estimate of an augmented reduced form equation (4). The equation should include lagged (in this case one through fifteen weeks) values of spot and futures price changes as well as the current futures price change.

We have discussed three approaches to optimal hedge ratio estimation using augmented reduced forms: the simple regression using price levels; the simple regression using price changes; and the more general time series approach. Hypothesis tests showed that the simple models are inconsistent with the data so that a general time series approach is required. Results from optimal hedge ratio estimation using the three approaches are provided in Table 2. Results from the simple price level regression are very different from the general time series approach but the simple price change regression gives results that are somewhat consistent with the time series method. Nevertheless, the general time series approach does lead to slightly different estimates and its use is supported by the test results in Table 1.

TABLE 2

Estimated Optimal Hedge Ratios

	Simple Price Level Regression	Simple Price Change Regression	General Time Series Regression
Corn	0.93	0.88	0.86
Soybeans	0.86	1.02	1.00

Conclusion

In this paper we investigate alternative hedge ratio estimation procedures. Estimation of a time series model is shown to be the preferred procedure and simple regression models are special cases of the time series approach under restrictions on the spot and futures market equilibria.

A set of statistical tests for determining the adequacy of the simple regression approach are provided. For the case of storage hedging of corn and soybeans in Michigan, the tests rejected the simple regression approach in favor of the more general time series approach. Nevertheless, the optimal hedge ratio estimates for the simple regression model using price changes are reasonably close to those obtained using the time series method.

Optimal hedge ratio estimation depends on the time series properties of spot and futures prices. Whether futures price changes satisfy the martingale difference property, and whether available information can be used to improve optimal hedge ratio estimation, are empirical questions that must be answered on a case by case basis through estimation of general reduced forms and testing of nested models. This paper provides a framework that allows these procedures to be carried out.

REFERENCES

Benninga, S., Eldor, R., and I. Zilcha. "The Optimal Hedge Ratio in Unbiased Futures Markets." J. Futures Markets. 4(1984):155-59.

Brown, S. L. "A Reformulation of the Portfolio Model of Hedging." Amer. J. Agr. Econ. 67(1985):508-12.

Brown, S. L. "A Reformulation of the Portfolio Model of Hedging: Reply." Amer. J. Agr. Econ. 68(1986):1010-12.

Bond, G. E., S. R. Thompson, and B. M. .S. Lee. "Application of a Simplified Hedging Rule." J. Futures Markets 7(1987):65-72.

Carter, C. A., and R. M. A. Loyns. "Hedging Feedlot Cattle: A Canadian Perspective." Amer. J. Agr. Econ. 67(1985):32-39.

Dickey, D. A. and W. A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." J. Amer. Stat. Assoc. 74(1979):427-31.

Dickey, D. A. and W. A. Fuller. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." Econometrica 49(1981):1057-72.

Ederington, L. H. "The Hedging Performance of the New Futures Markets." J. Finance 34(1979):157-70.

Engle, R. F. and C. W. J. Granger. "Co-Integration and Error Correction: Representation, Estimation and Testing." Econometrica 55(1987):251-76.

Fuller, W. A. Introduction to Statistical Time Series. New York: John Wiley and Sons, 1976.

Heifner, R. G. "Optimal Hedging Levels and Hedging Effectiveness in Cattle Feeding." Agr. Econ. Res. 24 (1972):25-36.

Hill, J. and T. Schneeweis. "A Note on the Hedging Effectiveness of Foreign Currency Futures." J. Futures Markets 1(1981):659-64.

Johnson, L. L. "The Theory of Hedging and Speculation in Commodity Futures." Rev. Econ. Stud. 27(1960):139-51.

Kahl, K. H. "Determination of the Recommended Hedging Ratio." Amer. J. Agr. Econ. 65(1983):603-5.

Kahl, K. H. "A Reformulation of the Portfolio Model of Hedging: Comment." Amer. J. Agr. Econ. 68(1986):1007-9.

Stein, J. L. "The Simultaneous Determination of Spot and Futures Prices." Amer. Econ. Rev. 51(1961):1012-25.

Theil, H. Principles of Econometrics. New York: John Wiley and Sons, 1971.

APPENDIX

Proof of Proposition 2

A minor change in notation will facilitate the proof. Let y_i, y_j, v_i, and v_j be (Tx1) vectors of realizations on y_{it}, y_{jt}, v_{it}, and v_{jt}. Furthermore, let X be a (Tx(lk+1)) matrix of observations on the constant and one through l period lags of all k variables in y_t. Then the two reduced form equations in the statement of Proposition 2 can be written

$$\text{(A.1)} \qquad \begin{aligned} y_i &= X\beta_i + v_i \\ y_j &= X\beta_j + v_j \end{aligned}$$

where β_i and β_j are vectors of reduced form parameters. We begin with a lemma and then prove the proposition.

Lemma Let equations (A.1) represent two reduced form equations from the model. Then the estimated conditional (on X) covariance between y_i and y_j is:

$$\frac{y_i{}'My_j}{T}$$

where $M=I-X(X'X)^{-1}X'$.

Proof: The maximum likelihood estimate of the conditional covariance between y_i and y_j is

$$\frac{(y_i - X\hat{\beta}_i)'(y_j - X\hat{\beta}_j)}{T}$$

where $\hat{\beta}_i = (X'X)^{-1}X'y_i$ and $\hat{\beta}_j = (X'X)^{-1}X'y_j$ are the OLS estimates of the reduced form parameters. Substituting the expressions for the OLS estimators into the conditional covariance equation gives

$$\frac{y_i{}'M'My_j}{T}$$

But M is an idempotent matrix so M'M=M. This completes the proof.

Proof of Proposition 2: By the lemma, the ratio of the conditional covariance between y_i and y_j to the conditional variance of y_j is

$$\frac{y_i{}'My_j}{y_j{}'My_j}.$$

Now redefine equation (4) as

$$y_i = Z\gamma + v_i$$

where $Z=[y_j \ X]$ and $\gamma=[\delta \ \beta_j']'$. The OLS estimate of γ is

$$\hat{\gamma} = (Z'Z)^{-1}Z'y_i.$$

Using the definition of Z and computing $(Z'Z)^{-1}$ by partitioned inverse (see Theil, p. 17), the first row of γ is

$$\hat{\delta} = \frac{[y_j' - y_j'X(X'X)^{-1}X']y_i}{y_j'My_j}$$

or, using the definition of M

$$\hat{\delta} = \frac{y_j'My_i}{y_j'My_j} .$$

This shows that the OLS estimate of δ gives the required ratio and therefore completes the proof.