Journal of Agricultural and Applied Economics, 41,1(April 2009):227–240

Alternative Model Selection Using Forecast Error Variance Decompositions in Wholesale Chicken Markets

Andrew M. McKenzie, Harold L. Goodwin, Jr., and Rita I. Carreira

Although Vector Autoregressive models are commonly used to forecast prices, specification of these models remains an issue. Questions that arise include choice of variables and lag length. This article examines the use of Forecast Error Variance Decompositions to guide the econometrician's model specification. Forecasting performance of Variance Autoregressive models, generated from Forecast Error Variance Decompositions, is analyzed within wholesale chicken markets. Results show that the Forecast Error Variance Decomposition approach has the potential to provide superior model selections to traditional Granger Causality tests.

Key Words: broiler markets, DAGs, forecasting, market structure, VAR

JEL Classifications: C53, D4, L1, Q00

Vector Autoregressive (VAR) models have become the main workhorse in the econometrician's stable of techniques for forecasting prices of related commodities that may be naturally modeled as a dynamic system. However, in applied forecasting settings, the econometrician is typically faced with the unenviable task of determining which price variables belong in the system and at what lag length. More often than not, economic theory provides no specific guidelines other than appealing to broad generalizations underpinning some underlying supply and demand model.

The main objective of this article is to empirically evaluate the potential of Forecast Error Variance Decompositions (FEVDs) derived from structural VARs to help inform the econometrician regarding the specification of forecasting models. Specifically, Directed Acyclic Graphs (DAGs) are used to identify structural VARs, and resulting FEVDs are used to determine if target forecast variables are endogenous or exogenous to the system. It is hypothesized that VARs' forecasts will be superior when target variables are endogenous, while simple univariate autoregressive (AR) models will outperform their VAR counterparts when target variables are deemed to be exogenous. Our FEVD forecasting model selections are also compared with forecasting models that would have been chosen using Granger Causality tests.

Wholesale chicken cuts (parts) prices provide us with a fertile data set to explore this particular forecasting issue. Goodwin, McKenzie, and Djunaidi illustrated that significant dynamic price relationships exist between some wholesale chicken cuts and wholesale broilers

Andrew M. McKenzie is associate professor, Harold L. Goodwin, Jr. is professor, and Rita I. Carreira is research associate, Department of Agricultural Economics and Agribusiness, University of Arkansas, Fayetteville, AR.

without giblets (WOG). For example, WOG production shocks may induce price responses in chicken cuts prices, while demand shocks generated in broiler cuts markets could lead to derived demand-induced price responses in WOG markets. Also, demand shocks in broiler cuts markets and the WOG market could cause price substitution effects that would be transmitted across all wholesale markets (Goodwin et al.). In particular, in this article, we investigate whether VAR forecasting models which supposedly exploit such relationships outperform simple univariate AR forecasting models.

From a practical standpoint, the analysis of wholesale chicken cut prices is of interest because price forecasting is of prime importance in the U.S. poultry industry. According to the National Chicken Council (NCC), the form in which broilers are marketed has changed markedly over the period 1985–2006. In 1985, the percentages of broilers marketed as whole, parts, and further processed were 29.1%, 53.4%, and 17.5%, respectively. By 2006, these percentages had changed to 9%, 44%, and 47% for whole, parts, and further processed, respectively. Similarly, the shares of broilers going to foodservice, retail grocery, and further processing changed from 38%, 44%, and 18% in 1995 to 26%, 20%, and 54% in 2006, respectively (NCC, 2007). Awareness of this market evolution is paramount in developing an understanding of the importance of accurate forecasting of wholesale broiler prices. Although the vast majority of broiler meat is currently sold under contract to foodservice and further processors, the prices for these contracts are benchmarked, to a large extent, off of the wholesale prices, of which the Urner Barry (UB) price is the "gold standard" for many in the industry. Therefore, it is essential for integrator firms (e.g., Tyson Foods, Conagra, Purdue, Pilgrims Pride, etc.) to accurately forecast chicken cut prices from one to nine months ahead. It is from such forecasts that forward contract prices for chicken cuts are established between integrator firms and buyers (retail grocery firms, restaurants, and fast food chains). Inaccurate forecasts lead to either lower volume sales or reduction in potential profit margins.

Data

Urner Barry's Price-Current report publishes daily wholesale chicken and chicken part prices, based on sales, bids, or offers for cash terms, and collected from buyers, sellers, and brokers. The UB quotes are used by the poultry industry as standard base prices from which formula prices are derived. The data set used in this study consists of monthly averages of UB's daily prices of boneless skinless chicken breast tender out (BSBTO), broiler wings (WING), bulk leg quarters (LQ), and whole broilers without giblets 2 1/4 lbs (WOG) for the period starting January 1989 and ending June 2007. Descriptive statistics of these price series measured in cents per pound (¢/lb) are reported in Table 1; Figure 1 contains a plot of the data series. BSBTO exhibited the most volatility followed by WING, while WOG and LQ were relatively more stable over the sample period. Augmented Dickey Fuller (ADF) unit root tests were employed to test if each of the price series were stationary; the results are presented in Table 2. The evidence, on balance, suggested that the four price series were stationary at the 5% level. Hence, the specification of VAR models, rather than Vector Error Correction models (VEC) was deemed appropriate.

The data were split into two subsamples: (1) an in-sample-period (January 1989–December 1999) yielding 132 monthly observations is used to estimate FEVDs from structural VARs and select the appropriate forecasting model (VAR or AR). In addition, in-sample Granger Causality tests were also used to select between VAR or AR forecasting models; and (2) an out-

Table 1. Descriptive Statistics of Monthly Average Prices (¢/lb.) of WOGs and Selected Broiler Parts for the Northeast U.S. Market (January 1989–June 2007)

	Mean	Variance	Minimum	Maximum
BSBTO	167.02	1118.05	105.00	288.64
WING	66.72	534.06	34.05	136.00
LQ	27.53	52.60	13.76	48.24
WOG	60.17	74.02	46.59	90.00

Source: Urner Barry Publications, Inc.

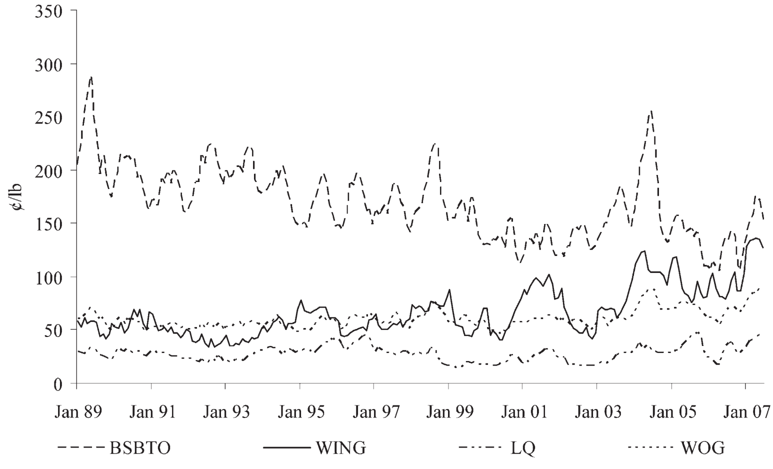

Figure 1. Plot of Monthly Average Prices (¢/lb) of Boneless Skinless Chicken Breast Tender Out (BSBTO), Broiler Wings (WING), Bulk Leg Quarters (LQ), and Whole Broilers without Giblets 2 1/4 lbs (WOG) Between January 1989 and June 2007

of-sample period (January 2000–June 2007) yielding 90 monthly observations is used to evaluate the model forecasting performance.

Modeling Approach

Granger Causality tests are frequently used to determine which variables should be included in a VAR for forecasting purposes. Although previous research has explored alternative multivariate approaches to select subset VAR forecasting models, where variables may either be excluded or enter the model with different time lags (just own lags enter into exogenous series and lags of other variables enter endogenous series) (Hsiao 1979, 1982; Kling and Bessler), traditional Granger Causality tests provide the econometrician with a quick and easy approach to select VAR forecast models. Therefore in this article we compare our FEVD generated forecasting models with models determined by Granger tests.

Given a choice of R variables, X_r such that $r = 1, \ldots, m, \ldots, R$, to include in the VAR, a general test of Granger Causality of the mth variable on, say, the first variable can be formally specified by setting up a full model/reduced model hypothesis test framework. The goal of the test is to determine whether including the mth variable in the model improves the estimation of the dependent variable, in this case the first variable.

Assume that our data set contains S useable observations of the R variables and note that $t = 1, \ldots, S$. The full model for the above test is specified as

Table 2. Results of Augmented Dickey-Fuller Test of Null Hypothesis of Existence of a Unit-Root Given by RATS

Variable	Model Specification Selected	Test Statistic	Critical Value ($\alpha = 1\%$)	Critical Value ($\alpha = 5\%$)	AIC Lag	Test Conclusion
BSBTO	With constant & linear trend	−4.437	−3.99	−3.43	18	Stationary at 1%
WING	With constant & linear trend	−5.025	−3.99	−3.43	12	Stationary at 1%
LQ	With constant only	−3.175	−3.46	−2.88	3	Stationary at 5%
WOG	With constant & linear trend	−4.306	−3.99	−3.43	1	Stationary at 1%

Note: Model specifications were determined by the Akaike Information Criterion and by testing the statistical significance of the constant and trend terms.

$$X_{1t} = p_{10} + \sum_{u=1}^{U} p_{1u} X_{1t-u} + \sum_{r=2}^{R} \sum_{v=1}^{V_r} q_{rv} X_{rt-v} + \upsilon_{1t} \tag{1}$$

where p_{10} is this model's intercept and p_{1u} refers to the parameter of the uth lag of the first variable; q_{rv} refers to the parameter of the vth lag of the rth variable; and υ_{1t} is the white noise residual.

The reduced model, that is, the model that does not include the mth variable, is then specified as

$$X_{1t} = p_{20} + \sum_{u=1}^{U} p_{2u} X_{1t-u} + \sum_{r=2}^{m-1} \sum_{v=1}^{V_r} q_{rv} X_{rt-v} + \sum_{r=m+1}^{R} \sum_{v=1}^{V_r} q_{rv} X_{rt-v} + \upsilon_{2t} \tag{2}$$

with the parameters and error term defined similarly as above. Note that V_r refers to the maximum lag length of the rth variable included in the model. The test statistic and its distribution are

$$F^* = \frac{SSE_2 - SSE_1}{V_m} \Bigg/ \frac{SSE_1}{S - U - \sum_{r=1}^{R} V_r - 1} \overset{H_0}{\sim} F_{V_m, S-U-\sum_{r=1}^{R} V_r - 1} \tag{3}$$

where the null hypothesis is that the mth variable does not have any explanatory power over the dependent variable, that is, the mth variable does not Granger-cause the first variable (formally, the null hypothesis is denoted as $H_0{:}p_{mv} = 0, \forall v = 1, \ldots, V_m$; the alternative hypothesis is that at least one of the parameters is different from zero). SSE1 and SSE2 are the sum of squared errors of the full and reduced models, respectively.

Traditional Granger Causality tests when applied to a large number of candidate variables, with a large lag structure, require a great number of observations (Carnot et al.). Alternatively, when considering a VAR, FEVDs, which can account for contemporaneous as well as lagged relationships between variables, can better characterize the dynamic relationships between variables and implied causality and exogeneity. FEVDs provide valuable information in the model selection process. For example, a variable that explains a very low fraction of the target variable is a good candidate for omission or replacement (Estima, p. 371).

Perhaps the most important advantage of the FEVD approach over Granger Causality tests is the information it provides the econometrician with respect to system dynamics over extended forecast horizons (Bessler). For example, FEVDs illustrate the potential contribution of system variables in helping to forecast a target variable over a range of forecast horizons. In practical applications this can be particularly useful in the model selection process if forecasts need to be projected for different forecast horizons and FEVDs indicate that variable influence changes over those same horizons. In contrast, Granger Causality tests select variables—with predictive power—to be included in a forecasting model irrespective of the required forecast horizon. However, if FEVDs are to be used as guides in building a VAR model useful for forecasting, accurate estimation of FEVDs is obviously critical. Given that innovation-accounting estimation is highly dependent upon specification of contemporaneous shocks, it is important to correctly determine contemporaneous correlation and causality across variables.

In this paper, as an alternative to Granger Causality forecast model selection, we turn to DAG, first described in 1993 by Spirtes, Glymour, and Scheines in their first edition of *Causation, Prediction, and Search*. DAG sheds light on contemporaneous causality across chicken cuts, and we use this information to impose a Sims-Bernanke decomposition on the system (Bernanke; Sims). The advantage of using DAG is that it is a data-driven method that is not subjective and can be used to analyze three or more variables (Awokuse and Bessler). The rationale and methodology behind DAG and its relevance to VAR analysis is clearly described in Awokuse and Bessler. Subsequent FEVD estimates are then used to determine optimal specifications of forecast models. For example if FEVD estimates suggest target forecast variables are endogenous with respect to other variables, then a VAR model is preferred. Conversely, if target forecast variables are deemed to be exogenous to the system, then a univariate AR model is preferred. In the latter case, the inclusion of superfluous variables (with no predictive power) may result in less accurate forecasts. Finally, out-of-sample forecasting performance of our FEVD generated VAR and AR models are evaluated using Diebold Mariano tests.

First, a standard four variable VAR model is specified as:

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{T} + \sum_{k=1}^{K} \begin{bmatrix} b_{11}(k) & . & . & . & b_{1n}(k) \\ . & & & & \\ . & & & & \\ . & & & & \\ b_{n1}(k) & . & . & . & b_{nn}(k) \end{bmatrix} \times Y_{t-k} + \sum_{l=1}^{11} a_l D_l + \varepsilon_t, \tag{4}$$

where $\mathbf{Y}_t$ represents a 4 × 1 vector containing the variables BSBTO, WOG, WING, and LQ in period t ($t = 1, \ldots, T$), $\mathbf{c}$ represents a 4 × 1 vector of constant terms, $\mathbf{T}$ is a 4 × 1 vector of trend terms, k indicates the lag order of the system, $b_{ij}(k)$ are the parameters we wish to estimate with n = 4, D_l are seasonal dummy variables ($l = 1, \ldots, 11$), a_l are the parameters of the seasonal dummy variables to be estimated, and finally ε_t is a 4 × 1 vector of serially uncorrelated random errors (innovations) with constant variance.

Each equation in the VAR system is estimated using ordinary least squares (OLS). Model selection in terms of the number of lags (k) to include in the system is determined using the Schwartz Bayesian Information Criterion (SBIC) and likelihood ratio tests in conjunction with residual diagnostic tests.

Following Bessler and Akleman and McKenzie, contemporaneous causal structures are determined using DAG theory with respect to the VAR innovation vector ε_t. Directed graph theory assigns causal flows between variables based upon partial correlations.[1] The DAG analysis is implemented using TETRAD II-Version 3.1 software (Spirtes et al. 1999). A Sims-Bernanke decomposition of ε_t is then imposed, based upon the directed graph results of contemporaneous causal flow between the system variables to identify a structural VAR model, and standard innovation accounting procedures are implemented to obtain FEVDs. This approach avoids having to make unrealistic assumptions about the contemporaneous relationship between variables to achieve system identification. As noted by Bessler and Akleman, early work applied Choleski factorization, which is a recursive ordering identification. However, the world might not be recursive, and incorrectly imposing it would lead to invalid innovation-accounting results. It is well known that FEVD estimates are highly sensitive to variable ordering when a Choleski factorization is used and variables are correlated (Estima, pp. 371–72).

The structural VAR FEVDs are computed from conditional within-sample forecasts for each of the variables in the system over one to nine month forecast horizons. The FEVD explains the relative proportion of the movements in a sequence due to its own shocks versus shocks to other variables. If own shocks explain all of the forecast error variance (FEV) of a variable, the price series in question may be considered exogenous to the other variables within the system. However, if a large proportion of the FEV associated with the sequence of a particular variable is explained by shocks to one or more of the other variables, then the price series in question would be considered endogenous to the system. The approach also allows one to draw inferences as to the relative importance in terms of the magnitude and sequence of influence among the system's variables, and hence determine the final specification of models useful for forecasting chicken cuts and WOG prices.

Finally, VAR and univariate AR models are used to estimate one to nine month out-of-sample forecasts for each of the target forecast variables (BSBTO, WOG, WING, and LQ). These models are formally specified as

$$Y_{it} = c_i + T_i + \sum_{k=1}^{K} [b_i(k)]\ Y_{it-k} + \sum_{l=1}^{11} a_l D_l + \varepsilon_{it}, \tag{5}$$

where Y_{it} represents each of the i = 1 through 4 variables and K is the optimal lag order identified with the SBIC. Thus each AR(K) model,

[1] For an in depth discussion of DAG theory see: Spirtes, Glymour, and Scheines. Also, Bessler and Akleman provide a good treatment of DAG and the identification of structural VARs with respect to farm and retail prices for pork and beef.

which is nested by the VAR(K) model, contains only K lags of the dependent (target forecast) variable along with a constant term, trend term and 11 seasonal dummies. In essence the AR(K) model is simply a restricted version of the VAR(K) model, and both types of model can be estimated using OLS.

Forecasting performance is evaluated using Diebold Mariano tests (DM) based on both Mean Square Error (MSE) and Mean Absolute Error (MAE) loss functions. Under the null hypothesis of the DM test, the forecasts from a VAR model specification are not different from those from an AR specification, that is, H_0: VAR forecasts $\approx$ AR forecasts. Two alternative hypotheses are considered: (1) VAR forecasts are preferred to AR forecasts (H_{A1}: VAR forecasts $\succ$ AR forecasts), and (2) the converse, that is, HA2: VAR forecasts $\prec$ AR forecasts. Forecast performance is also documented with respect to Mean Error (ME), MAE, Root Mean Square Error (RMSE), and Theil's U Statistic (Theil's U). Theil's U is a unit free measurement ranging from zero to infinity, with unit value being equivalent to a random walk forecast. Forecasting accuracy increases with lower ME, MAE, RMSE, and Theil's U values. Out-of-sample forecasting performance results are compared with in-sample exogeneity results implied by the structural VAR FEVD's. This comparison allows us to evaluate the usefulness of the FEVD approach in selecting forecasting models. If the optimal model specification selected by FEVD's is a VAR (AR), and this model provides superior out-of-sample forecasts to the AR (VAR) model, then we may conclude that the FEVD approach is a useful tool for choosing forecasting models. Similarly if our FEVD model selections provide superior out-of-sample forecasts to Granger Causality model selections, then we may conclude that the FEVD approach is preferred to the traditional Granger Causality approach.

Model Selection Using DAG-Generated Sims-Bernanke FEVD Results

Preliminary model estimations were performed on VAR systems incorporating from 1 to 12 lags for each variable. The SBIC and likelihood ratio test statistics indicated that a parsimonious VAR system with a lag order of three months was optimal, thus, in Equation (2), $K = 3$. In addition, Ljung-Box Q-statistics indicated that serial correlation was not a serious problem for the three lag VAR specification.

Prior to performing standard innovation accounting procedures to obtain FEVDs, contemporaneous causality between system variables was determined using the DAG analysis described in the previous section of this article. Specifically, the DAG analysis is based upon the contemporaneous correlation coefficients of the VAR system residuals presented in Table 3. Figure 2 presents the causal flow of contemporaneous shocks based upon a 1% significance level. The results indicate that system variables are contemporaneously linked, and the chain of causality is clearly identified with BSBTO, WING, and LQ shocks each directly affecting WOG price movements. This is consistent with the premise that demand shocks generated in broiler cut markets could theoretically lead to derived demand-induced price responses in the WOG market (Goodwin, McKenzie and Djunaidi).

Although the DAG results provide interesting implications with respect to the relationship between markets in contemporaneous time, FEVDs are required to analyze the system dynamics and provide insights into forecast model selection. Thus a second VAR is estimated by imposing a Sims-Bernanke decomposition of ε_t based upon the DAG results of contemporaneous causal flow between the system variables. Table 4 reports Sims-Bernanke FEVDs and standard errors for in-sample forecasts for periods of one through nine months ahead. In all cases, as the forecast

Table 3. Contemporaneous Correlation Coefficients of VAR System Residuals (January 1989–December 1999)

	BSBTO	WING	LQ	WOG
BSBTO	1.000			
WING	0.184	1.000		
LQ	0.134	0.203	1.000	
WOG	0.542	0.382	0.393	1.000

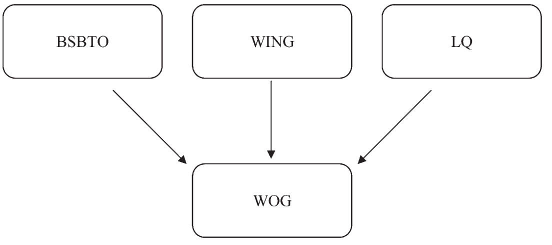

Figure 2. Directed Acyclical Graph of the Variables

horizon increases, the standard errors grow, as would be expected, but tend to level off, implying the system is stationary. Also, standard errors are much larger for BSBTO and WING forecasts, which are consistent with the summary statistics presented in Table 1, and indicate BSBTO and WING price series are more volatile than WOG and LQ. Results indicate that BSBTO, WING, and LQ are in large part exogenous, although WOG innovations appear to have some influence on LQ for forecast horizons beyond four months. Interestingly, WOG forecasts are influenced by all three of the other variables, which is consistent with DAG results. In particular, BSBTO has a large impact on WOG forecasts followed by LQ and to a lesser extent WING, over all forecast horizons. It should be noted that interpreting FEVDs to explain whether a variable is endogenous or exogenous is somewhat subjective. In applied research it is typical for a variable to explain almost all of its FEV at short horizons and smaller proportions at longer horizons (Enders, p. 280). Given that other variables explain 40% of WOG FEV for the first two months ahead and 50% or more thereafter we believe it appropriate to consider WOG to be endogenous to the system. Sims-Bernanke FEVDs for longer forecast horizons (not presented in the paper) show that WOG has some influence on LQ and WING beyond one year out. WOG shocks account for 20% of LQ FEV and 10% of WING FEV from one year onwards.

The Sims-Bernanke FEVDs suggest that simple univariate AR models will outperform VAR models for BSBTO, WING, and LQ forecasts, for one to nine months ahead. On the other hand, we would expect one to nine month ahead WOG forecasts would be better modeled using a VAR that can take advantage of dynamic interactions between WOG and the other system variables. The Sims-Bernanke FEVD generated model selections are statistically tested in the next section using 90 months of out-of-sample data. Formally, AR(3) forecasting models are compared with VAR(3) models using DM tests and other criteria (e.g., ME, MAE, RMSE, and Theil's U values). The AR(3) models were specified as in Equation 2 of the modeling approach section.

In contrast to our findings, results presented in Goodwin, McKenzie, and Djunaidi suggest a greater level of variable interaction. While we used FEVDs generated from DAG and a Sims-Bernanke decomposition, Goodwin, McKenzie, and Djunaidi used a Choleski decomposition to orthogonalize the VAR innovations. Their variable ordering (BSBTO, WOG, WING, and LQ) was based on economic theory and prior industry knowledge, and reflected the fact that white meat is the most valuable part, in dollar terms, of a broiler in the U.S. (Goodwin et al., pp. 486–87). To better compare our results and approaches we replicated Goodwin, McKenzie, and Djunaidi's approach using UB prices over our in-sample period. FEVD results from this replication are presented in Table 5. With respect to BSBTO, WING, and LQ, results from the replication are similar to our Sims-Bernanke FEVDs. BSBTO, WING, and LQ are in large part exogenous, although WOG innovations appear to have some influence on WING and LQ for one to three month ahead forecast horizons. However, unlike our Sims-Bernanke results, replication results for WOG forecasts indicate only BSBTO and WOG itself influence WOG prices.

In sum, with respect to chicken cuts markets, our replication FEVD results are more in line with our Sims-Bernanke FEVD results than the FEVD results presented in Goodwin, McKenzie, and Djunaidi. It would appear that chicken cuts markets have become more exogenous, irrespective of methodological approach. It should be emphasized that in spite of our similar FEVD findings with respect to chicken cuts markets, we in no way advocate

Table 4. In-Sample Forecast Error Variance Decomposition Attributed to Innovations in Respective Series Using a Sims-Bernanke Decomposition (January 1989–December 1999)

Variable	Months Ahead	Standard Error	BSBTO	WING	LQ	WOG
BSBTO	1	8.734	100.000	0.000	0.000	0.000
	2	10.857	98.992	0.002	0.962	0.044
	3	12.897	99.146	0.063	0.682	0.109
	4	13.666	98.919	0.139	0.766	0.176
	5	14.128	98.266	0.364	1.054	0.317
	6	14.324	97.390	0.777	1.249	0.584
	7	14.462	96.313	1.328	1.421	0.937
	8	14.561	95.247	1.870	1.578	1.305
	9	14.643	94.274	2.332	1.734	1.660
WING	1	4.208	0.000	100.000	0.000	0.000
	2	5.677	0.072	99.552	0.145	0.231
	3	6.613	0.129	99.170	0.119	0.582
	4	7.322	0.106	98.080	0.119	1.695
	5	7.827	0.104	97.276	0.169	2.451
	6	8.187	0.143	96.575	0.239	3.043
	7	8.459	0.226	95.752	0.324	3.698
	8	8.672	0.351	94.832	0.445	4.372
	9	8.839	0.495	93.913	0.600	4.992
LQ	1	2.096	0.000	0.000	100.000	0.000
	2	3.436	0.832	1.864	97.240	0.063
	3	4.040	1.294	2.135	94.778	1.793
	4	4.379	1.243	1.981	91.584	5.193
	5	4.659	1.105	1.906	88.873	8.115
	6	4.921	1.118	2.029	86.703	10.150
	7	5.161	1.279	2.277	84.647	11.797
	8	5.374	1.539	2.534	82.660	13.267
	9	5.561	1.852	2.770	80.860	14.518
WOG	1	2.526	23.625	6.402	8.893	61.080
	2	3.470	27.208	3.733	10.072	58.986
	3	3.809	35.225	3.485	9.769	51.521
	4	3.968	38.542	3.860	9.847	47.751
	5	4.067	40.264	4.016	9.827	45.892
	6	4.112	40.950	4.043	9.806	45.201
	7	4.130	41.239	4.040	9.835	44.885
	8	4.137	41.291	4.032	9.926	44.751
	9	4.140	41.237	4.025	10.021	44.680

the use of Choleski decompositions to choose forecasting model specifications. As noted earlier, in general, FEVD estimates are highly sensitive to variable ordering when a Choleski factorization is used and variables are correlated. Also, the use of such an approach presupposes that economic theory may be used to select variable ordering and that variable influence occurs in a recursive manner. This is unlikely to be the case in most situations, and FEVDs generated from DAG-inspired structural VARs will be preferred to their Choleski counterparts.

Model Selection Using Granger Causality Generated Results

Granger Causality tests presented in Table 6—derived from the VAR(3) model and estimated using the in-sample-data—suggest that WING and BSBTO are exogenous. Thus, Granger Causality tests, like our Sims-Bernanke

Table 5. In-Sample Forecast Error Variance Decomposition Attributed to Innovations in Respective Series Using a Choleski Decomposition as Done in Godwin, McKenzie, and Djunaidi (January 1989–December 1999)

Variable	Months Ahead	Std. Error	BSBTO	WOG	WING	LQ
BSBTO	1	8.734	100.000	0.000	0.000	0.000
	2	10.933	99.040	0.290	0.003	0.667
	3	12.928	99.175	0.231	0.104	0.490
	4	13.651	98.904	0.207	0.236	0.653
	5	14.065	98.152	0.197	0.589	1.063
	6	14.236	97.169	0.196	1.214	1.420
	7	14.359	95.978	0.205	2.037	1.780
	8	14.454	94.807	0.221	2.853	2.119
	9	14.537	93.740	0.239	3.570	2.451
WOG	1	2.668	29.418	70.582	0.000	0.000
	2	3.637	32.272	66.752	0.885	0.091
	3	4.002	39.928	59.001	0.745	0.326
	4	4.185	43.261	55.020	1.034	0.685
	5	4.297	44.954	53.164	1.099	0.782
	6	4.347	45.615	52.486	1.097	0.803
	7	4.366	45.894	52.172	1.096	0.838
	8	4.373	45.951	52.047	1.095	0.907
	9	4.377	45.940	51.990	1.093	0.977
WING	1	4.209	3.393	11.248	85.359	0.000
	2	5.726	4.196	13.829	81.943	0.032
	3	6.680	4.555	11.525	83.794	0.126
	4	7.393	4.356	9.629	85.647	0.369
	5	7.901	4.111	8.561	86.681	0.648
	6	8.261	3.867	7.892	87.321	0.920
	7	8.531	3.652	7.406	87.714	1.228
	8	8.742	3.479	7.054	87.869	1.598
	9	8.908	3.353	6.800	87.844	2.004
LQ	1	2.096	1.799	14.509	0.295	83.397
	2	3.387	3.664	10.869	0.708	84.759
	3	3.979	4.482	8.058	0.530	86.931
	4	4.317	4.378	7.175	0.630	87.817
	5	4.588	3.924	6.872	0.714	88.490
	6	4.835	3.546	6.660	0.671	89.124
	7	5.056	3.341	6.619	0.614	89.425
	8	5.252	3.294	6.749	0.569	89.388
	9	5.425	3.355	6.947	0.534	89.164

Note: Goodwin, McKenzie, and Djunaidi assumed that the causality order of the variables was BSBTO, WOG, WING, and LQ.

FEVD approach, suggest that simple univariate AR models should outperform VAR models for WING and BSBTO forecasts. Turning to WOG forecasts, Granger Causality test results indicate that only BSBTO and WOG itself have predictive power in explaining WOG price movements. In contrast, recall that Sims-Bernanke FEVDs imply that WOG forecasts are influenced by all three of the other variables in addition to WOG itself. Therefore, the Sims-Bernanke FEVD approach in this case lends stronger support to the selection of a fully inclusive VAR model. Finally, regarding LQ forecasts, Granger Causality tests endorse the notion that a VAR forecasting model would be preferable. Results show that LQ, WING, and WOG all have power in predicting LQ. Again, recall that Sims-Bernanke FEVDs suggest LQ

Table 6. Granger Causality Test Results Given by RATS

Dependent Variable	BSBTO *F*-Statistic (significance)	WING *F*-Statistic (significance)	LQ *F*-Statistic (significance)	WOG *F*-Statistic (significance)
BSBTO	31.407 (0.000)	0.535 (0.659)	1.127 (0.342)	0.093 (0.964)
WING	0.285 (0.837)	123.792 (0.000)	0.410 (0.746)	1.326 (0.270)
LQ	1.119 (0.345)	2.207 (0.092)	241.854 (0.000)	2.311 (0.080)
WOG	2.457 (0.067)	1.339 (0.266)	0.315 (0.814)	19.413 (0.000)

is exogenous and hence support an AR model selection. In sum, Granger Causality tests and Sims-Bernanke FEVDs lead to the same conclusions with respect to model selection for WING and BSBTO forecasts.

Out-of-Sample Forecast Results

Out-of-sample forecast summary statistics presented in Table 7 clearly indicate that, as expected, the VAR(3) model produces superior WOG forecasts compared with the AR(3) model. ME, MAE, RMSE, and Theil U values are lower for the VAR(3) model across all forecast horizons. Most importantly, DM test results presented in the last five columns of Table 7 (and *p* values associated with the MSE loss function presented in column 11) show that the VAR(3) model forecasts are statistically more accurate than the AR(3) forecasts, again across all forecast horizons. These results translate into cost reductions in terms of WOG prices as measured by MAE of between 0.145¢/lb and 0.327¢/lb for lag periods 1 and 3, respectively, if VAR(3) is utilized as the forecasting model for contract prices of chicken rather than the AR(3) model. To put this into perspective, according to NCC data, in 2005, U.S. broiler production was estimated at 34,986 million lbs, of which 9% were sold as WOG; thus, our results would roughly translate into savings between $4.57 million and $10.3 million.

Turning to LQ forecasts, results are once again consistent with our *a priori* expectations based upon Sims-Bernanke FEVDs. In this case AR(3) is the preferred model, generating forecasts with lower ME, MAE, RMSE, and Theil U values across all forecast horizons. Similarly DM test results using MSE criteria demonstrate a statistically significant preference for the AR(3) model over one, six, seven, eight, and nine month ahead forecast horizons. Although forecasting models cannot be differentiated at conventional significance levels for two, three, four, and five month horizons, the AR(3) model is preferred at 0.15, 0.18, 0.16, and 0.11 levels, respectively, using the MSE criteria. In this case, using the AR(3) model to forecast prices for LQ for contracting purposes results in cost reductions of between 0.145 and 0.372 ¢/lb in lag periods 1 and 7/9, respectively, as measured by MAE.

Results are mixed with respect to BSBTO forecasts. For shorter forecast horizons (ranging from one to five months ahead) DM test results reveal the AR(3) model outperforms the VAR(3) model by between 0.402¢/lb and 1.371¢/lb in lag periods 1 and 4, respectively, as measured by MAE, which is consistent with Sims-Bernanke FEVDs. However, for the longest forecast horizons (eight and nine months ahead) VAR(3) is the preferred model by about 1.40¢/lb as measured by MAE based upon DM tests. In terms of ME, AR(3) forecasts provide superior forecasts over all horizons, but with respect to other criteria (MAE, RMSE, and Theil U) there is little difference between models.

As regards WING forecasts, it is impossible to distinguish between the two forecasting models irrespective of evaluation criteria. Recall our Sims-Bernanke FEVD results showed that WING is exogenous to other system variables, suggesting that an AR(3) forecasting model should outperform a VAR(3) model. Although this is not the case, clearly taking account of dynamic interactions between system variables by using a VAR(3)

Table 7. Diebold-Mariano Test of VAR(3) vs. AR(3) Out-of-Sample Forecasts (January 2000–June 2007)

										D-M p-Value of Test of H_0: VAR(3) = AR(3)				
		AR(3) Estimation				VAR(3) Estimation				H_a: $VAR \succ AR$		H_a: $VAR \prec AR$		Concl.: Data
Step	N	ME.	MAE.	RMSE.	Theil U	ME.	MAE.	RMSE.	Theil U	MSE	MAE	MSE	MAE	Supports
Forecast Statistics for Series BSBTO														
1	90	0.230	9.236	11.868	0.907	0.517	9.638	12.289	0.939	0.972	0.953	0.028	0.047	$VAR \prec AR^{**}$
2	89	0.671	13.656	18.805	0.855	1.519	14.335	19.393	0.881	0.979	0.968	0.021	0.032	$VAR \prec AR^{**}$
3	88	1.287	17.259	23.994	0.852	2.723	18.308	24.853	0.882	0.992	0.993	0.009	0.007	$VAR \prec AR^{***}$
4	87	1.752	20.186	27.515	0.840	3.623	21.557	28.458	0.869	0.984	0.997	0.016	0.003	$VAR \prec AR^{***}$
5	86	2.320	21.952	29.628	0.835	4.528	22.358	30.327	0.855	0.928	0.779	0.072	0.220	$VAR \prec AR^{*}$
6	85	2.921	22.565	30.507	0.835	5.376	22.416	30.796	0.843	0.719	0.396	0.281	0.604	FTR H_0
7	84	3.582	23.079	30.659	0.844	6.209	22.385	30.515	0.840	0.379	0.119	0.621	0.881	FTR H_0
8	83	4.343	22.557	30.370	0.867	7.072	21.154	29.720	0.849	0.053	0.007	0.947	0.993	$VAR \succ AR^{***}$
9	82	4.836	22.532	30.933	0.909	7.612	21.148	30.040	0.883	0.008	0.003	0.992	0.997	$VAR \succ AR^{***}$
Forecast Statistics for Series WING														
1	90	0.879	6.340	7.717	0.873	0.841	6.304	7.659	0.866	0.354	0.399	0.646	0.601	FTR H_0
2	89	2.022	10.477	12.980	0.918	1.998	10.452	12.896	0.912	0.354	0.459	0.646	0.541	FTR H_0
3	88	3.272	13.342	16.074	0.930	3.363	13.405	16.108	0.932	0.551	0.590	0.449	0.410	FTR H_0
4	87	4.312	14.758	18.069	0.918	4.525	15.026	18.132	0.921	0.582	0.800	0.418	0.199	FTR H_0
5	86	5.287	15.957	19.795	0.893	5.628	16.154	19.838	0.895	0.554	0.709	0.446	0.291	FTR H_0
6	85	6.201	17.142	21.398	0.874	6.638	17.317	21.462	0.876	0.585	0.686	0.415	0.314	FTR H_0
7	84	6.958	18.635	22.758	0.863	7.503	18.736	22.886	0.867	0.651	0.599	0.349	0.401	FTR H_0
8	83	7.563	19.760	23.926	0.857	8.196	19.955	24.088	0.863	0.674	0.679	0.326	0.321	FTR H_0
9	82	8.077	20.285	24.888	0.845	8.779	20.651	25.066	0.851	0.682	0.811	0.318	0.189	FTR H_0
Forecast Statistics for Series LQ														
1	90	0.424	1.944	2.652	0.861	0.464	2.089	2.828	0.919	0.946	0.939	0.054	0.061	$VAR \prec AR^{*}$
2	89	0.951	3.713	4.769	0.912	1.053	3.822	5.010	0.958	0.856	0.730	0.144	0.269	FTR H_0
3	88	1.425	5.129	6.291	0.923	1.611	5.145	6.540	0.959	0.819	0.531	0.181	0.469	FTR H_0
4	87	1.807	6.021	7.415	0.917	2.076	6.136	7.687	0.951	0.837	0.674	0.163	0.326	FTR H_0
5	86	2.145	6.595	8.203	0.905	2.507	6.797	8.532	0.941	0.886	0.759	0.114	0.240	FTR H_0
6	85	2.422	6.976	8.684	0.898	2.881	7.303	9.072	0.938	0.926	0.872	0.074	0.128	$VAR \prec AR^{*}$

Table 7. Continued.

		AR(3) Estimation				VAR(3) Estimation				D-M p-Value of Test of H_0: VAR(3) = AR(3)				
										H_a: $VAR \succ AR$		H_a: $VAR \prec AR$		Concl.: Data
Step	N	ME.	MAE.	RMSE.	Theil U	ME.	MAE.	RMSE.	Theil U	MSE	MAE	MSE	MAE	Supports
7	84	2.675	7.315	9.008	0.902	3.221	7.687	9.398	0.941	0.927	0.893	0.073	0.107	$VAR \prec AR^{*}$
8	83	2.891	7.468	9.142	0.911	3.523	7.826	9.567	0.953	0.921	0.865	0.079	0.135	$VAR \prec AR^{*}$
9	82	3.093	7.484	9.134	0.916	3.797	7.856	9.662	0.969	0.932	0.859	0.068	0.140	$VAR \prec AR^{*}$
Forecast Statistics for Series WOG														
1	90	0.392	2.555	3.352	1.001	0.224	2.410	3.163	0.945	0.085	0.123	0.915	0.877	$VAR \succ AR^{*}$
2	89	0.968	4.546	5.855	1.045	0.618	4.289	5.307	0.947	0.026	0.124	0.974	0.876	$VAR \succ AR^{**}$
3	88	1.480	5.665	7.358	1.010	1.092	5.338	6.740	0.925	0.037	0.099	0.963	0.901	$VAR \succ AR^{**}$
4	87	1.914	6.291	8.320	0.979	1.543	6.022	7.766	0.914	0.059	0.144	0.941	0.856	$VAR \succ AR^{*}$
5	86	2.294	6.592	8.926	0.945	1.971	6.358	8.510	0.901	0.092	0.172	0.908	0.828	$VAR \succ AR^{*}$
6	85	2.694	6.604	9.177	0.907	2.404	6.381	8.835	0.873	0.097	0.164	0.903	0.836	$VAR \succ AR^{*}$
7	84	3.024	6.678	9.417	0.882	2.772	6.374	9.070	0.849	0.078	0.051	0.922	0.949	$VAR \succ AR^{*}$
8	83	3.344	6.648	9.474	0.861	3.122	6.364	9.048	0.822	0.025	0.076	0.975	0.924	$VAR \succ AR^{**}$
9	82	3.656	6.593	9.457	0.842	3.453	6.336	8.985	0.800	0.005	0.102	0.995	0.898	$VAR \succ AR^{**}$

Note: * indicates statistical significance at α = 10%; ** indicates statistical significance at α = 5%; and *** indicates statistical significance at α = 1%.

would not lead to superior out-of-sample forecasts.

On a final note, it is constructive to compare our Sims-Bernanke FEVD forecasting model selections with forecasting models that would have been chosen using Granger Causality tests. In other words, does the Sims-Bernanke FEVD approach contain information beyond that provided by Granger Causality tests that would aid the econometrician? Recall that both Granger Causality tests and Sims-Bernanke FEVDs lead to the same conclusions with respect to model selection for WING and BSBTO forecasts. However, the two approaches provide different modeling selections with respect to WOG and LQ forecasts. Importantly, out-of-sample forecasting results for WOG and LQ are consistent with Sims-Bernanke FEVD model selections, and hence lend credence to the assumption that Sims-Bernanke FEVDs may well choose better forecasting models than Granger Causality tests.

Conclusions

Selection of appropriate time-series forecasting models, although of paramount importance to industry econometricians because of the role accurate prices play in determining a firm's profit margins for various products, is fraught with problems. Typically, without the aid of any underlying economic model, the econometrician has relied upon Granger Causality tests to choose which variables should be included in a forecasting model. However, these tests require many observations to choose between large numbers of candidate variables, and perhaps more crucially, select variables independent of forecast horizon. In this paper we promote an alternative forecasting model selection tool—Sims-Bernanke FEVDs generated from DAG-inspired structural VARs. On balance, out-of-sample forecast results for chicken cuts and WOG prices indicate that this approach has much promise and in the case of WOG, if the poultry industry were to improve price forecasts with this approach, savings of between $4.57 million and $10.3 million could be attained. Out-of-sample forecast results are fairly consistent with in-sample Sims-Bernanke FEVD model selections, and indeed the approach appears to provide superior model selections to traditional Granger Causality tests. This is not an insignificant finding, particularly because of the primary role accurately estimated wholesale prices have in determining eventual contract prices for chicken sold to foodservice and further processing markets. Obviously this enthusiasm should be tempered by the fact that our approach has only been tested in one market over a specific sample period. However, results presented in this paper suggest that further research using the Sims-Bernanke FEVD forecasting model selection approach may prove fruitful in other market settings. Another interesting avenue of future research would be to compare forecasting performance of Sims-Bernanke FEVD selected models with Kling and Bessler type subset VARs

[Received November 2007; Accepted May 2008.]

References

Awokuse, T.O., and D.A. Bessler. "VAR, Policy Analysis, and Directed Graphs: An Application to the US Economy." *Journal of Applied Econometrics* 5(2003):1–24.

Bernanke, B. "Alternative Explanations of the Money-Income *Correlation." Carnegie-Rochester Conference Series on Public Policy* 25(1986): 49–100.

Bessler, D.A. "An Analysis of Dynamic Economic Relationships: An Application to the U.S. Hog Market." *Canadian Journal of Agricultural Economics* 32(1984):109–24.

Bessler, D.A., and D.G. Akleman. "Farm Prices, Retail Prices, and Directed Graphs: Results for Pork and Beef." *American Journal of Agricultural Economics* 80(1998):1144–49.

Carnot, N., V. Koen, and B. Tissot. *Economic Forecasting*. New York: Palgrave MacMillan, Ltd., 2005.

Diebold, F.X., and R.S. Mariano. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 13(1995):253–63.

Enders, W. *Applied Econometric Time Series*, 2nd ed. Hoboken, NJ: John Wiley & Sons, Inc., 2004.

Estima. RATS Version 6 User's Guide and Reference Manual. Evanston, IL: Estima, 2004.

Goodwin, H.L., A.M. McKenzie, and H. Djunaidi. "Which Broiler Part is the Best Part?" *Journal of Agricultural and Applied Economics* 35(2003):483–95.

Hsiao, C. "Causality Tests in Econometrics." *Journal of Economic Dynamics & Control* 1(1979):321–46.

———. "Autoregressive Modeling and Causal Ordering of Economic Variables." *Journal of Economic Dynamics & Control* 4(1982):243–59.

Kling, J.L., and D.A. Bessler. "A Comparison of Multivariate Forecasting Procedures for Economic Time Series." *International Journal of Forecasting* 1(1985):5–24.

McKenzie, A.M. "The Effects of Barge Shocks on Soybean Basis Levels in Arkansas: A Study of Market Integration." *agribusiness International Journal (Toronto, Ont.)* 21(2005):37–52.

National Chicken Council. *NCC Statistics and Research*. Internet Site: http://www.nationalchickencouncil.com/statistics/. (Accessed October 2007).

Sims, C.A. "Are Forecasting Models Usable for Policy Analysis?" *Federal Reserve Bank of Minneapolis Quarterly Review* 10(1986):2–16.

Spirtes, P., C. Glymour, and R. Scheines. *Causation, Prediction, and Search*, 2nd ed. New York: The MIT Press, 2001.

Spirtes, P., R. Scheines, C. Meek, T. Richardson, C. Glymour, A. Boomsam, and H. Hoijtink. *TETRAD II Version 3.1 Tools for Causal Modeling*. Pittsburgh, PA: Carnegie Mellon University, 1999.

Urner Barry Publications, Inc. *Urner Barry's Price-Current*. Toms River, N.J. Various Issues, January 1989–June 2007.

U.S. Department of Agriculture/Economic Research Service. *Background Statistics on U.S. Broiler Industry*. Internet site: http://www.ers.usda.gov/news/broilercoverage.htm (Accessed July 2007).