Journal of Development and Agricultural Economics Vol. 2(1) pp. 001-006, January, 2010
Available online at http://www.academicjournals.org/JDAE
ISSN 2006- 9774© 2010 Academic Journals

Full Length Research Paper

Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship

Henry de-Graft Acquah

Department of Agricultural Economics and Extension, University of Cape Coast, Cape Coast, Ghana. E-mail: henrydegraftacquah@yahoo.com. Tel: 00233245543956.

Accepted 1 December, 2009

Information criteria provide an attractive basis for model selection. However, little is understood about their relative performance in asymmetric price transmission modelling framework. To explore this issue, this research evaluated the performance of the two commonly used model selection criteria, Akaike information criteria (AIC) and Bayesian information criteria (BIC) in discriminating between asymmetric price transmission models under various conditions. Monte Carlo experimentation indicated that the performance of the different model selection criteria are affected by the size of the data, the level of asymmetry and the amount of noise in the model used in the application. The Bayesian information criterion is consistent and outperforms AIC in selecting the suitable asymmetric price relationship in large samples.

Key words: Model selection, Akaike's information criteria (AIC), Bayesian information criteria (BIC), asymmetry, Monte Carlo.

INTRODUCTION

Alternative methods detect asymmetry at different rates or culminate in different inferences and conclusions (Meyer and von Cramon-Taubadel, 2004; Capps and Sherwell, 2007). However, asymmetric price transmission modelling aims to select a single model from a set of competing models that best captures the underlying asymmetric data generating process for derivation of policy conclusions. This simulates interest in model selection methods. Consequently, information-theoretic criteria such as Akaike's Information Criteria (AIC) (Akaike, 1973) and Bayesian Information Criteria (BIC) (Schwarz, 1978) are increasingly being used to address model selection problems. However, very little is understood about relative performance of AIC and BIC in an asymmetric price transmission modelling context.

Essentially, the two penalized criteria are based on two different model selection approaches. AIC is aimed at finding the best approximating model to the unknown data generating process whilst BIC is designed to identify the true model. AIC does not depend directly on sample size. Bozdogan (1987) noted that because of this, AIC lacks certain properties of asymptotic consistency. Although BIC takes a similar form like AIC, it is derived within a Bayesian framework, reflects sample size and have properties of asymptotic consistency. For reasonable sample sizes, BIC apply a larger penalty than AIC, thus other factors being equal it tend to select simple models than does AIC. From a Bayesian view point this motivates the adoption of the Bayesian information criteria.

Using these two concepts, numerous model selection criteria have been developed, extending on the basic structures of AIC and BIC such as Consistent Akaike Information Criteria (CAIC) (Burhnam and Anderson, 1998) and Draper's Information Criteria (DIC) (Draper,1995). Unlike most of their analytical extensions, AIC and BIC can be readily computed in most standard software and are extensively used in empirical analysis. Subsequently, most previous studies comparing information criteria have focussed on their use with generalized linear models and have focussed on comparison between AIC and BIC. These investigations have generally demonstrated that BIC is consistent whilst in contrast AIC is not (Bickel and Zhang, 1992; Zhang, 1993). Although AIC and BIC have been compared theoretically and empirically (Weakliem, 1999; Kuha, 2004) and examined empirically with respect to selection of stock-recruitment relationships (Wang and Liu, 2006) there has been no empirical comparison for their relative performance in asymmetric price transmission modelling context.

The main purpose of this article is to empirically evaluate and compare the performance of the two commonly used model selection criteria, AIC and BIC in discriminating between alternative methods of testing for asymmetry. A comparison of AIC and BIC will thus contribute to understanding information criteria modelling generally and of their empirical performance in price transmission analysis.

The true data generating process is known in all experiments and the Monte Carlo simulations are essential in deriving the model recovery rates of the true model.

METHODOLOGY

Asymmetric price transmission models

The Granger and Lee asymmetric Error Correction Model data generating process (DGP) can be written as follows:

$$\Delta P_{A,t} = \beta_1 \Delta P_{B,t} + \beta_2^{+} ECT^{+}_{t-1} + \beta_2^{-} ECT^{-}_{t-1} + \varepsilon$$
$$\varepsilon \sim N(0, \sigma_\varepsilon^2) \qquad (1)$$

Using various sample sizes, P_A and P_B are generated as I (1) non stationary variables that are cointegrated. There exist an equilibrium relationship between P_A and P_B which produces I (0) stationary series. This equilibrium equation is estimated by least squares and the lagged deviation from this regression denoted by the Error Correction Term (ECT_{t-1}). The ECT is decomposed into positive and negative deviations using Wolffram segmentation (Granger and Lee, 1989) and plugged into the asymmetric error correction model specified in equation (1).

Where $ECT = P_{At} - \beta_1 P_{Bt}$ and $ECT^{+}_{t-1} = ECT_{t-1}$ if $ECT_{t-1} > 0$ and 0 otherwise and $ECT^{-}_{t-1} = ECT_{t-1}$ if $ECT_{t-1} < 0$ and 0 otherwise.

Asymmetry is introduced by allowing the speed of adjustment to differ for the positive and negative components of the Error Correction Term (ECT) since the long run relationship captured by the ECT was implicitly symmetric (see Cook et al., 2000 ; Holly et al., 2003, Cook et al., 1999). Symmetry in equation (1) is tested by determining whether the coefficients (β_2^{+} and β_2^{-}) are identical (that is $H_0 : \beta_2^{+} = \beta_2^{-}$).

Granger and Lee (1989) proposed a model to test for asymmetries in the adjustments to the equilibrium level. Alternatively, von Cramon-Taubadel and Loy (1996) proposed a model in which asymmetries specified affects the direct impact of price increases and decreases as well as adjustments to the equilibrium level.

The von Cramon-Taubadel and Loy (1996) asymmetric Error Correction Model can be written as follows:

$$\Delta P_{A,t} = \beta_1^{+} \Delta P^{+}_{B,t} + \beta_1^{-} \Delta P^{-}_{B,t} + \beta_2^{+} ECT^{+}_{t-1} + \beta_2^{-} ECT^{-}_{t-1} + e$$
$$e \sim N(0, \sigma_e^2) \qquad (2)$$

Where $\Delta P^{+}_{B,t}$ and $\Delta P^{-}_{B,t}$ are the positive and negative changes in $P_{B,t}$ and the remaining variables are defined as in equation (1).

Von Cramon-Taubadel and Loy (1996) suggested that the $\Delta P_{B,t}$ in equation (1) can also be split into positive and negative components to allow for more complex dynamics and applied equation (2) to study spatial asymmetric price transmission on world wheat markets. The remaining model variables were defined as in equation (1) and formal test of the asymmetry hypothesis using equation (2) is: $H_0 : \beta_1^{+} = \beta_1^{-}$ and $\beta_2^{+} = \beta_2^{-}$.

Noticeably, since equation (2) involves a linear combination of coefficients, a joint F-test can be used to determine symmetry or asymmetry of the price transmission process.

In contrast to von Cramon-Taubadel and Loy (1996) and Houck (1977), proposed a model in which asymmetries specified affects the direct impact of price increases and decreases and does not take into account adjustments to the equilibrium level. The Houck approach can be specified as follows:

$$\Delta P_{A,t} = \beta_1^{+} \Delta P^{+}_{B,t} + \beta_1^{-} \Delta P^{-}_{B,t} + \nu_t$$
$$\nu_t \sim N(0, \sigma_\nu^2) \qquad (3)$$

Model variables are defined as in equation (2). Symmetry is tested by determining whether the coefficients (β_1^{+} and β_1^{-}) are identical (that is $H_0 : \beta_1^{+} = \beta_1^{-}$).

Model selection using information criteria

Model selection refers to the problem of using the data to select one model from the list of competing models. Essentially, it involves the use of a model selection criterion to find the best fitting model to the data (Wasserman, 2000). Model selection using information criteria has been developed to summarize data evidence in favor of a model. Specifically, information criteria techniques emphasize minimizing the amount of information required to express the data and model. This results in selection of models that are efficient representation of the data.

Akaike's Information Criteria (AIC)

One of the most commonly used information criteria is AIC. The idea of AIC (Akaike, 1973) is to select the model that minimises the negative likelihood penalised by the number of parameters as specified in the equation (4).

$$AIC = -2 \log p(L) + 2p \qquad (4)$$

Where L refers to the likelihood under the fitted model and p is the number of parameters in the model.

Specifically, AIC is aimed at finding the best approximating model to the unknown true data generating process and its applications

Table 1. Relative performance of the model selection methods across sample size.

Experiment criterion	Model fitted			
	Methods	CECM (%)	HKD (%)	SECM (DGP) (%)
$n = 50$ $\sigma = 1$	AIC	17.0	4.8	78.2
	BIC	6.3	11.9	81.8
$n = 150$ $\sigma = 1$	AIC	17.5	0.0	82.5
	BIC	3.0	0.1	96.9
$n = 500$ $\sigma = 1$	AIC	16.8	0.0	83.2
	BIC	1.6	0.0	98.4

Note: Recovery rates based on 1000 replications.

draws from (Akaike, 1973; Bozdogan, 1987; Zucchini, 2000).

Bayesian information criteria (BIC)

Another widely used information criteria is the BIC. Unlike Akaike Information Criteria, BIC is derived within a Bayesian framework as an estimate of the Bayes factor for two competing models (Schwarz, 1978; Kass and Rafftery, 1995). BIC is defined as:

$$BIC = -2\log p(L) + p\log(n) \quad (5)$$

Superficially, BIC differs from AIC only in the second term which now depends on sample size n. Models that minimize the Bayesian Information Criteria are selected. From a Bayesian perspective, BIC is designed to find the most probable model given the data.

Performance of the model selection criteria in selecting good models for the observed data is examined using simulation studies. Such a comparison is not straight forward and even its relevance could be questioned, given that the two criteria are based on different theoretical motivations and objectives. However, for application purpose, the Akaike Information Criteria and the Bayesian Information Criteria do have the same aim of identifying good models even if they differ in their exact definition of a “good model”. Comparing them is thus justified, at least to examine how each criterion performs according to recovery of the correct model or how they behave when both should prefer the same model.

A simulation study

The objective of the simulation study is to see whether the model selection methods are capable of identifying the true model. Following the experimental designs of Holly et al. (2003) among others the value of β_1 is set to 0.5 and $(\beta_2^+, \beta_2^-) \in (-0.25, -0.75)$ are considered for the coefficients of the asymmetric error correction terms in the true model. The competing models are fitted to the simulated data and their ability to recover the true model was measured. The recovery rates were derived using 1000 Monte Carlo simulations. The data generation process is specified in equation (1) and the data is simulated from the standard ECM as follows:

$$\Delta P_{A,t} = 0.5\Delta P_{B,t} - 0.25ECT^{+}_{t-1} - 0.75ECT^{-}_{t-1} + \varepsilon \quad (6)$$

P_A and P_B are generated as I (1) non stationary variable that are cointegrated. The ECTs denotes the positive and negative deviations from the equilibrium relationship between P_A and P_B.

In order to examine the effect of the increase in difference of asymmetric adjustment parameters on model recovery the study simulated data of sample size 150 with an error size of 1 from the standard asymmetric price transmission model specified in equation 6 and asymmetry values $(\beta_2^+, \beta_2^-) \in (-0.25, -0.50)$ or $(-0.25, -0.75)$ are considered for the coefficients of the asymmetric error correction terms.

RESULTS AND DISCUSSION

Model recovery rates of the different model selection criteria

This section compares the performance of AIC and BIC in recovering the true data generating process (DGP) by simulating the effect of sample size, noise levels and the level of asymmetry on model selection. The relative performance of the two methods are compared in terms of their ability to recover the true data generating process (DGP) across various sample size conditions (that is Model Recovery Rates) as illustrated in Table 1. In the foregoing discussion, the standard asymmetric error correction model, the complex asymmetric error correction model and the Houck’s model in first differences are denoted by SECM, CECM and HKD respectively.

For each method the model recovery rate defines the percentages of samples in which each competing model provides a better model fit than the other competing models. The model selection methods performed reasonably well in identifying the true model, though their ability to recover the true asymmetric data generating process (DGP) increases with increase in sample size. In small samples (upper part of Table 1), the model selection methods recovered at most 81.8%. When the sample size was large (Lower part of Table 1), the model selection methods recovered at most 98.4%. AIC

Table 2. Relative performance of the selection methods across error size.

Experiment criterion	Model fitted			
	Methods	CECM (%)	HKD (%)	SECM (DGP) (%)
$n = 150$ $\sigma = 3$	AIC	12.3	22.4	65.3
	BIC	1.2	52.1	46.7
$n = 150$ $\sigma = 2$	AIC	17.3	15.1	77.7
	BIC	1.8	18.7	79.5
$n = 150$ $\sigma = 1$	AIC	18.3	0.0	81.7
	BIC	2.4	0.1	97.5

Note: Recovery rates percentages based on 1000 replications.

Table 3. Effects of sample size and stochastic variance on model recovery.

Experiment criterion	Model fitted			
	Methods	CECM (%)	HKD (%)	SECM (DGP) (%)
$\sigma = 2$ $n = 50$	AIC	9.9	35.4	54.7
	BIC	2.8	55.9	41.3
$n = 150$ $\sigma = 0.5$	AIC	18.3	0.0	81.7
	BIC	2.5	0.0	97.5

Note: Recovery rates based on 1000 replications.

performs well in small samples, but is inconsistent and does not improve in performance in large samples whilst BIC in contrast is consistent and improves in performance in large sample size.

Generally, model selection performance improved as sample sizes increased. Two distinct patterns can be noted with regards to the recovery rates of the true model (DGP) in Table 1. First, recovery rates of the Bayesian criteria varied strongly as a function of sample size. Second, although AIC performed well in the small samples, it did not make substantial gains in recovery rates as the sample size increased.

The observed patterns are consistent with previous studies on model selection in other applications. Ichikawa (1998)'s simulation results in a factor analysis indicated that the ability of AIC to select a true model rapidly increased with sample size but at larger sample sizes it continued to exhibit a slight tendency to select complex models. Similarly, Markon and Krueger (2004) reviewed existing work on factor analysis and noted that AIC performs relatively well in small samples, but is inconsistent and does not improve in performance in large samples whilst BIC in contrast appears to perform relatively poorly in small samples, but is consistent and improves in performance with larger sample size.

In order to simulate the effects of noise level on model selection, this study considers three error sizes (σ) ranging relatively from small to large and corresponding to 1.0, 2.0 and 3.0. Using 1000 Monte Carlo simulations, data is generated from equation (6) with the different error sizes and a sample size of 150. The data fitting abilities of alternative models are compared in relation to the true model as the error in the data generating process was increased systematically.

The performance of the model selection algorithms analysed deteriorates with increasing amount of noise in the true asymmetric price transmission data generating process (SECM) as illustrated in Table 2. BIC outperforms AIC in recovering the true data generating process at lower noise levels (σ = 1 - 2) but at higher noise levels (σ = 3), AIC outperforms BIC.

Simulating the effects of sample size and stochastic variance concurrently affirms that a small error and large sample improves recovery of the true asymmetric data generating process and vice versa. With a small sample of 50 and an error size of 2.0 the true data generating process was recovered at least 41.3% of the time by the model selection criteria as illustrated in Table 3.

On the other hand, with a relatively large sample of 150 and error size of 0.5 at least 81.7% of the correct model was recovered across all the model selection methods as indicated in the Table 3.

The model recovery rates of the model selection methods are derived under combined conditions of a small sample size of 50 and large error size of 2 (that is Unstable conditions) and a relatively large sample size of 150 and a small error size of 0.5 (that is Stable conditions). Under stable conditions, model selection performance or recovery rates improves for the true model.

Table 4 illustrates how the different model selection methods exhibit different relative performance in

Table 4. Effects of the level of asymmetry on model recovery.

Experiment criterion	Model Fitted			
	Methods	CECM (%)	HKD (%)	SECM (DGP) (%)
$\beta_2^+ - \beta_2^- = 0.25$	AIC	16.35	0.22	83.43
	BIC	2.71	1.72	95.57
$\beta_2^+ - \beta_2^- = 0.50$	AIC	16.5	0.00	83.5
	BIC	2.83	0.03	97.14

Note: Recovery rates based on 1000 replications.

recovering the true model at different levels of asymmetry. An increase in the difference between the asymmetric adjustments parameters from 0.25 to 0.50 led to improvement in the model recovery rates of the true asymmetric data generating process by the model selection methods. Generally, recovery rates of the Bayesian criteria responds more strongly to increases in the difference between the asymmetric adjustments parameters for the true model.

In effect, another factor which may influence model selection or the recovery of the true data generating process is the difference in asymmetric adjustment parameters as illustrated.

An important feature of the current study is that they generally echo existing theoretical and empirical work on the performance of model selection methods in other applications. First, the results of the Monte Carlo experimentation establish that AIC and BIC does identify the true asymmetric data generating process. Similarly, Myung (2000) demonstrated via Monte Carlo experimentation that BIC and AIC clearly identify the true data generating process in a cognitive psychology modeling framework.

With regards to the effect of noise levels on model selection, the results of the current study suggest that the amount of noise in the asymmetric data generating process is important for the purposes of model selection. Results obtained are generally consistent with trends suggested by previous studies (Gheissari and Bab-Hadiashar, 2003; Yang, 2003), in that the performance of AIC and BIC deteriorates with increases in the amount of noise in the data generating model. Yang (2003) finds that the recovery rates of the true data generating process decreases with increasing noise levels in linear regression models. Within the asymmetric price transmission modeling framework, AIC outperforms BIC when there is a high amount of noise in the true model. Similarly, Chen et al. (2007) note the tendency of BIC to perform worse than AIC at high noise levels in a factorial analysis.

The findings of the current study reinforce the importance of design informativeness in conducting asymmetric price transmission analysis. Simulation results suggested that AIC performs relatively well in small samples but is inconsistent and does not improve performance in large samples whilst BIC in contrast appears to perform relatively poorly in small samples but is consistent and improves in performance with sample size in the price transmission modeling context. This is consistent with previous studies which demonstrated that BIC is consistent (that is tends to choose the true model with a probability equal to 1 in large samples (Bickel and Zhang 1992). Overall, the current results suggest that generally AIC should be preferred in smaller samples whilst BIC should be preferred in larger samples in the price transmission modeling context.

On the level of asymmetry, the results indicated that the performance of the model selection methods in recovering the true data generating process depends on the difference in asymmetric adjustments parameters or speeds. Similarly, without regards to information criteria, Cook et al. (2003) observed that the difference in asymmetric adjustment parameters from 0.25 to 0.50 has positive effect on the test of asymmetry. On the basis of the recovery of the true model, BIC should be preferred to AIC in applications in which the data has strong levels of asymmetry.

Within the asymmetric price transmission modeling framework, this study has not only shed light empirically on the relative performance of the model selection algorithms of which no studies has been undertaken, but has also established that the Bayesian criteria correctly identifies the true asymmetric data generating process.

Conclusion

The model selection criteria examined clearly identify the correct asymmetric model out of alternative competing models. Fundamentally, the results reinforce the importance of design characteristics in conducting asymmetric price transmission studies. The results of the Monte Carlo simulation indicate that the samples sizes, level of asymmetry and noise levels are important in the selection of the true asymmetric model. Larger sample sizes or lower noise levels improve the ability of the model selection methods to point to the correct asymmetric price transmission models. Under unstable conditions such as small sample and large noise levels AIC outperforms BIC. An increase in the difference between the asymmetric adjustments parameters improves model recovery rates of the true asymmetric data generating process by the model selection methods. The comparison provided contributes to knowledge and understanding of the relative performance of the Akaike's Information Criteria and the Bayesian Information Criteria in an asymmetric price transmission modeling framework which has remained understudied. The validity of AIC and BIC in selecting the correct model in the price transmission modeling framework in the current studies suggest that other AIC and BIC based estimators hold promise as model selection criteria.

Notes

All simulations were performed using R programming language version 2.9.2 with the random number procedure used to generate pseudo i.i.d. N (0, 1) random.

ACKNOWLEDGEMENTS

The author is grateful to Prof. Dr. Stephan von Cramon-Taubadel and three other anonymous reviewers for valuable comments and support.

REFERENCES

Akaike H (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In: B.N. Petrov and F. Csaki (eds.) 2nd International Symposium on Information Theory: 267-81 Budapest: Akademiai Kiado.

Bickel P, Zhang P (1992). Variable selection in nonparametric regression with categorical covariates. J. Am. Stat. Assoc. 87: 90–97.

Bozdogan H (1987). Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions. Psychometrika (52)3: 345-370.

Burhnam KP, Anderson DR (1998). Model selection and inference: a practical information theoretic approach. New York: Springer.

Capps O, Sherwell P (2007). Alternative Approaches in Detecting Asymmetry in farm-retail prices transmission of fluid milk. J. Agribusiness (23)3: 313-331.

Cook S, Holly S, Turner P (2000). The Power of Tests for Non-linearity: The Escribano-Pfann Model. Computational Economics 15: 223-226.

Cook S, Holly S,Turner P (1999). The Power of Tests for Non-linearity: the case of Granger–Lee Asymmetry. Econ. Lett. 62: 155–59.

Chen L, Giannakouros P, Yang Y (2007). Model Combining in Factorial Data Analysis. J. Stat. Plann. Inference (137)9: 2920-2934.

Draper D (1995). Assessment and Propagation of Model Uncertainty. J. Roy. Stat. Soc. Series B (Methodological) 57: 45-97.

Gheissari N, Bab-Hadiasher A (2004). Effect of Noise on Model Selection Criteria in Visual Applications. Pattern Recognition 2: 23-26 229-232.

Granger CWJ, Lee TH (1989). Investigation of Production, Sales and Inventory Relationships using Multicointegration and non-symmetric Error Correction Models. J. Appl. Econ. 4: 135-159.

Holly S, Turner P, Weeks M (2003). Asymmetric Adjustment and Bias in Estimation of an Equilibrium Relationship from a Co-integrating Regression. Comput. Econ. 21: 195-202.

Houck JP (1977). An Approach to specifying and estimating nonreversible Functions. American J. Agric. Econ. 59: 570-572.

Ichikawa M (1988). Empirical Assessments of AIC Procedure for Model Selection in Factor Analysis. Behaviormetrika 24: 33–40.

Kass RE, Raftery A (1995). Bayes Factors. J. Am. Stat. Assoc. 90: 773-795.

Kuha J (2004). AIC and BIC: Comparisons of Assumptions and Performance. Sociol. Methods Res. (33)2: 188-229.

Markon KE, Krueger RF (2004). An Empirical Comparison of Information- Theoretic Selection Criteria for Multivariate Behavior Genetic Models. Behav. Genet. (34)6: 593- 609.

Meyer J, von Cramon-Taubadel S (2004). Asymmetric Price Transmission: A survey. J. Agric. Econ. (55)3: 581-611.

Myung Jae I (2000). The Importance of Complexity in Model Selection. J. Math. Psychol. 44: 190-204.

Schwarz G (1978). Estimating the Dimension of a Model. Annals of Statistics 6: 461–464.

Von Cramon-Taubadel S, Loy JP (1996). Price Asymmetry in the international Wheat Market: Comment. Canadian J. Agric. Econ. 44: 311-317.

Wang Y, Liu Q (2006). Comparison of Akaike information criteria (AIC) and Bayesian information criteria (BIC) in selection of stock-recruitment relationships. Fish. Res. 77: 220-225.

Wasserman L (2000). Bayesian Model Selection and Model Averaging. J. Math. Psychol. 44: 92-107.

Weakliem LD (2004). Introduction to the Special Issue on Model Selection. Soc. Methods Res. 33: 167-186.

Yang Y (2003). Regression with Multiple Candidate Models: Selecting or Mixing? Statistica Sinica 13: 783-809.

Zhang P (1993). On the convergence of model selection criteria. Comm. Stat.-Theory Meth. 22: 2765-2775.

Zucchini W (2000). An Introduction to Model Selection. J. Math. Psychol. 44: 41-6.