MODEL AVERAGING IN ECONOMETRIC MODELING: REVIEW OF RECENT DEVELOPMENTS

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research
(CEAR)
University of Ibadan

26th July, 2018

OUTLINE

- ⋆ INTRODUCTION
- ⋆ BACKGROUND
- ⋆ PROBLEM STATEMENT
- ⋆ MODEL AVERAGING TECHNIQUES
- ⋆ REVIEW OF LITERATURE
- ⋆ EMERGING ISSUES
- ⋆ PLAUSIBLE RESEARCH QUESTIONS
- ⋆ JUSTIFICATION
- ⋆ AIM AND OBJECTIVES
- ⋆ WHAT NEXT
- ⋆ REFERENCES

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

INTRODUCTION

- ⋆ Modelling: A General and an Econometric Outlook
  - ⋆ Developing models that strike reasonable balance between realism and manageability
- ⋆ Characterization of real world phenomena: Complex and highly Uncertain
- ⋆ Uncertainty: The difference between Truth and Fact
  - ⋆ Emanates from the complexity of the observed/perceived relationship(s) between/among economic variables
  - ⋆ Often trivialized by researchers who fail to capture it in their model specification
  - ⋆ Recent researches have attempted to account for same in a view to provide answers to hinging on its relevance.

INTRODUCTION CONT'D

Rational Expectation Models

⋆ Is a widely adopted procedure in econometric analyses

⋆ Restricts variables composition in a choice model on the basis of perceived relationship between/among variables and available data

⋆ Proponents often disregard the possibility of mis-specification and neglect the inherent uncertainty by choosing a single model as "best" model through pre-testing

⋆ A Model Selection technique

INTRODUCTION CONT'D

- ⋆ Inferences and predictions are subsequently based on the single selected model
- ⋆ Limits the predictor space to an incomplete subset of the unknown whole, usually constrained one's level of exposure and/or experience
- ⋆ Although restrictive, it still provides useful information for understanding real world phenomena
- ⋆ Arguments opposed to the adoption of a single model to fully explain the complexities of the real world phenomena still holds strongly, premised on the fact that models are mere approximations of the truth and not the truth in themselves

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

INTRODUCTION CONT'D

- Methods of averaging as plausible alternatives (Moral-Benito, 2015; Steel, 2017)

- Extant model averaging techniques can be broadly classified into two paradigms
  - Bayesian (see Hoeting et al., 1999; Fragoso and Neto, 2015)
  - Classicalist/Frequentist (see Buckland et al., 1997; Wang et al., 2009)

- Although, both paradigms attempt to account for associated uncertainty, weights (as a measure of uncertainty) generation and assignment differ markedly one from the other

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

BACKGROUND

- ⋆ Major research challenges, especially with respect to modelling, revolve round the appropriateness of researcher's preference or choice of model
- ⋆ An immediate cause being the unwieldy size of the model space cum the uncertainty about each consisting model
- ⋆ Researchers often have numerous plausible models that could be used to represent the relationship between variables of interest in a particular study, in a view to answer specified research question(s)
- ⋆ Plausible models in the model space vary with respect to satisfying one or more structural assumptions
  - ⋆ functional form, included/excluded variable(s) [that is, variable selection] and statistical distribution of the residual

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

BACKGROUND CONT'D

- ⋆ Appropriate handling of uncertainty has been a bothersome task for most econometric researchers
- ⋆ Conventionally, many researchers opt for estimation procedures involving a single equation
  - ⋆ chosen on the bases of pre-specified selection criteria
  - ⋆ restrictive and often misleading as no single equation model could be all-encompassing, to completely explain the inherent relationship between variables
  - ⋆ observed to understate the real uncertainty [Draper, 1995], since the chosen model is usually perceived and used as the true model [Moral-Benito, 2012]
- ⋆ Models generally contain aspects (structure and/or parameter) that are not known with certainty and the mere search for the "best" model implies unassessed structural uncertainty [Draper, 1995]

BACKGROUND CONT'D

- Researchers' uncertainty could be classified into two
  - uncertainty of parameter estimates conditional on the chosen model
  - uncertainty of model selection from the unwieldy model space
- Extant methodologies that attempt to tackle inherent uncertainty can be sub-grouped into three
  - Model Selection (Magnus, 2002; Leeb and Potscher, 2003; 2006; Berk et al., 2013; among others)
  - Model Averaging (Leamer, 1978; Draper, 1995; Buckland et. al., 1997; Hoeting, et. al., 1999; Moral-Benito, 2015; Liu et al., 2016; Forte et al., 2018; among others)
  - Shrinkage & Penalized Methods (Tibshirani, 1996; Knight and Fu, 2000; Hansen, 2016; Schneider, 2016; among others).

PROBLEM STATEMENT

- Model selection processes do not explicitly incorporate model uncertainty despite the probabilistic manner in which the final model was chosen through the search exercise
- Increased chances of mis-specification resulting from omitted variable bias
- Conflicting results from competing models given the available data and sample periods considered
- Model selection process under-reports the variance
  - could be mis-leading with respect to policy decisions
  - is too conservative, especially, as it relates to the true nature of most noisy (financial and macro-econometric) series
  - A large but truthfully reported variance would be preferred over a small but misleading variance [Magnus et al., 2013]

MODEL AVERAGING TECHNIQUES

- Traced back to Bates and Granger (1969) study
  - advocated for averaging over the estimates from several single regression models to improve predictive performance
- Aim to produce best possible estimates rather than best possible model [Magnus et al., 2010]
  - The latter being a model selection objective
- Reduce estimation variance, while controlling for omitted variable bias that characterizes the model selection approach

MODEL AVERAGING TECHNIQUES CONT'D

- Achieve this in two simple steps
  - Estimation of all the plausible candidate models
  - Computation of the weighted averages of corresponding parameter estimates
- Take cognizance of predictor variables' contribution in each candidate model in explaining the dependent variable and weights each accordingly
- Reflect the uncertainty associated with each model in the model space in a more formal manner
  - by averaging inferences that are not model-specific across all plausible models in the constructed model space [Forte et al., 2018]

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

MODEL AVERAGING TECHNIQUES CONT'D

- ⋆ Model averaging techniques can be viewed from two perspectives
- ⋆ Objectivism (classicalists/frequentists)
  Frequentist Model Averaging (FMA) techniques (Buckland et. al., 1997; Hjort and Claeskens, 2003; Poeter and Anderson, 2005; Hansen, 2007; among others)
- ⋆ Subjectivism (Bayesians)
  Bayesian Model Averaging (BMA) techniques (Draper, 1995; Hoeting, et. al., 1999; Doppelhofer and Weeks, 2009; Amini and Parmeter, 2012; among others)

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

MODEL AVERAGING TECHNIQUES CONT'D

- Both techniques account for uncertainty
- BMAs offers a more natural approach, using prior probabilities to fully accounts for model uncertainty [Steel, 2017]
- Setting priors for candidate models and dealing with conflicting results from various specified priors is perceived to be problematic, hence the stance of the frequentist [Hjort and Claeskens, 2003]
- FMAs do not require setting of priors for estimating the parameters of the model
  - the corresponding estimators are entirely data-determined

REVIEW OF LITERATURE

Leamer, 1978

- ⋆ Laid the foundation for development of Bayesian Model Averaging
- ⋆ Argued for the incorporation of Bayesian known prior probability distribution on the parameter values in data analysis using the Bayes' rule
- ⋆ Both estimation and inference come naturally together from the posterior distribution
- ⋆ Computation complexities that hindered it wide usage
- ⋆ Prior elicitation to improve the optimality of the technique
- ⋆ Sensitivity of the technique when errors are characterized by higher moments (skewness and Kurtosis)

REVIEW OF LITERATURE CONT'D

Buckland et al., 1997

- Model averaging under the frequentist paradigm
- Recognized the identification of the "best" model as an estimation problem and assigned weights to each model in proportion to its relative penalized likelihood
- Based on information theoretics - AIC or BIC
- Inference was based on bootstrap (Efron, 1979) samples
- Assume that the true model is nested
- Involves sample splitting, which is considered inefficient
- Excludes the concept of heteroscedasticity, which limits its applicability

REVIEW OF LITERATURE CONT'D

Hansen, 2007

- Frequentist Model Averaging Mallows' Model Average (MMA)
- Based on Mallows' C_p criterion (an estimate of the average squared error from the model average fit)
- Asymptotically optimal in the class of discrete model average estimators
- Displayed decreasing relative risk, thus outperforming contending estimators (Mallows, AIC and S-AIC) except S-BIC (at small n and R^2 when MMA showed increasing relative risk)
- Selection of models to be estimated and averaged over is often computationally burdensome even for moderately sized model spaces
- Optimality of MMA fails under the heteroscedastic condition

REVIEW OF LITERATURE CONT'D

Magnus et al., 2010

- ⋆ Fusion of Frequentist (parameter estimation) and Bayeisan (weighting) - Weighted-Average Least Squares (WALS)
- ⋆ Hinges on two key merits (theoretically and practically) above extant Bayesian Model Averaging technique
  - ⋆ Triviality of the computational burden and the transparent definition of prior ignorance
  - ⋆ Reduced the $2^k - 1$ plausible models in the conventional BMA to k models, given a k predictor space
- ⋆ Treats researchers' ignorance about the priors differently, obtaining better risk profile and avoiding unbounded risk (replace normal prior with Laplace prior)
- ⋆ Places no prior distributions on the model space
  - ⋆ Adopts orthogonalization to enable the selection of a single variable and subsequent linear additions to the model space without placing importance on the order
  - ⋆ The posterior inclusion probability of parameters cannot be determined

REVIEW OF LITERATURE CONT'D

Raftery et al., 2010

- Proposed the dynamic model averaging (DMA) technique (originally engineering oriented)
  - Extended BMA beyond the assumption of constancy of regression coefficients across models
  - Incorporates time variation into the regression coefficients
  - Allows for the interpretation of the PIP of each model
  - Allows the relevant model set to change with time through the forgetting factor
- The time varying coefficients regression models are formed from all possible combinations of available predictors
- The state space and Markov chain models are both specified in terms of forgetting, leading to a highly parsimonious representation
- Shown to outperform the best model among the list of physically motivated models

REVIEW OF LITERATURE CONT'D

Hansen and Racine, 2012

- Developed another Frequentist model averaging technique - Jackknife Model Averaging (JMA)
- Obtained appropriate weights for averaging M approximate (mis-specified) models for improved estimation of an unknown conditional mean in the face of non-nested model uncertainty in heteroscedastic error settings
- Selects weight by minimizing a delete-one cross validation criterion (quadratic in weights, so that computation is a simple application of quadratic programming)
- Applicable to both nested and non-nested cases
- Revealed equivalence of MMA and JMA when errors are homoscedastic, but JMA outperformance under heteroscedastic errors

REVIEW OF LITERATURE CONT'D

Magnus et al., 2013

- ⋆ Proposed the weighted-average least squares (WALS) prediction procedure
- ⋆ Allowed for explicit correlation among random errors, while accounting for both model and error uncertainty and also proposed an appropriate estimate for the variance of the WALS estimator
- ⋆ Measured the accuracy of the estimator in a simulation study, while emphasizing on the importance of a truthfully reported variance of the predictors rather than a small variance
- ⋆ Comparatively showed the outperformance of WALS other contending estimators in a large Monte Carlo simulation experiment, while controlling for sample size, parameter values and variance specification
  - unrestricted maximum likelihood, pretesting, ridge regression and MMA

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

REVIEW OF LITERATURE CONT'D

Onorante and Raftery, 2016

- ⋆ Proposed a way of implementing the DMA in large model space - the Dynamic Occam's Window (DOW)
- ⋆ The strategy considers a subset of models (not the whole model space) and dynamically optimizes the choice of models at each point in time
- ⋆ This is in a verge to deal with problem that may arise when the number of models to be combined exceeds 3000
- ⋆ The authors have to assume that models do not change too fast over time, which is not an ideal assumption when dealing with financial and in some cases monthly economic data

EMERGING ISSUES

- ⋆ Model selection estimates improvement to include uncertainty and determining the optimality of model averaging estimators (schomaker and Heumann, 2018)
- ⋆ Optimization of weighting structure to improve weight choice and adoption of appropriate model screening schemes to eliminate poor models before averaging to reduce computational burden (Wan et al., 2014)
- ⋆ Improving model predictability, ascertaining the sensitivity of model average techniques to nested and non-nested models and adaptation to financial and macro-econometric series (Steel, 2017)
- ⋆ Modelling with Big data (Koop, 2017)
- ⋆ Modelling with endogenous regressors (Clarke, 2017)
- ⋆ Model averaging estimators for generalized linear models (Dardanoni et al., 2015; De Luca et al., 2017), among others

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

PLAUSIBLE RESEARCH QUESTIONS

1. How do the contending model averaging techniques perform when dataset is characterized by certain statistical features

- ⋆ Persistence
- ⋆ Endogeneity
- ⋆ Time varying parameters

2. Following from (1), does the choice of model averaging technique matter?

AIM AND OBJECTIVES

- To investigate the influence of incorporating endogeneity, persistence and time varying parameters in extant model average techniques

- Attempt to explicitly incorporate endogeneity, persistence and time-varying parameters into a chosen model averaging technique

- Empirically examine the performance of contending model averaging techniques when series are characterized by some salient features
  - Persistence
  - Endogeneity
  - Combination of Persistence and Endogeneity
  - Time varying Parameters

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

WHAT NEXT?

- Methodological exposition of the contending model averaging techniques to be investigated.
- Description of the Monte Carlo Experiment with Relevant Scenarios
- Data Simulation and Empirical Data Analysis with appropriate interpretation and discussion
- Statement of findings

Dr. Olusanya Elisa Olubusoye
&
Ahamuefula Ephraim Ogbonna

A Concept Note
@
Centre for Econometric and Allied Research (CEAR)
University of Ibadan

REFERENCES

- Amini, S.M. and Parmeter, C.F. (2012) Comparison of model averaging techniques: Assessing growth determinants. Journal of Applied Econometrics 27: 870876.
- Bates, J. M. and Granger, C. M. W., (1969). The combination of forecasts. Operations Research Quarterly 20, 451-468.
- Berk, R., Brown, L. and Buja, A. (2013) Valid Post-Selection Inference. The Annals of Statistics, Vol. 41, No. 2, 802 837 DOI: 10.1214/12-AOS1077
- Buckland, S.T., Burnham, K.P. and Augustin, N.H. (1997) Model selection: An integral part of inference. Biometrics 53: 603618.
- Claeskens, G. and N. Hjort (2003). The focused information criterion. Journal of the American Statistical Association 98, 90016.
- Clarke, J. A. (2017) Model Averaging OLS and 2SLS: An Application of the WALS Procedure. Department of Economics, University of Victoria, Econometrics Working Paper EWP1701
- Dardanoni, V., De-Luca, G., Modica, S. and Peracchi, F. (2015) Model averaging estimation of generalized linear models with imputed covariates. Journal of Econometrics 184, 452 - 463
- De Luca, G., Magnus, J. R. and Peracchi, F. (2018) Weighted-average least squares estimation of generalized linear models. Journal of Econometrics. https://doi.org/10.1016/j.jeconom.2017.12.007
- De Luca, G. and Magnus, J.R. (2011) Bayesian model averaging and weighted average least squares: Equivariance, stability and numerical issues. Stata Journal 11: 518544.
- Doppelhofer, G. and Weeks, M. (2009) Jointness of growth determinants. Journal of Applied Econometrics 24: 209244.
- Draper, D. (1995): Assessment and Propagation of Model Uncertainty (with discussion). Journal of the Royal Statistical Society, Series B (Methodological), 57, 45-97.
- Efron, B. (1979) Bootstrap Methods: Another Look at the Jackknife. The Annal of Statistics Vol. 7 (1), 1 - 26.
- Forte, A., Garcia-Donato, G. and Steel, M. (2018). Methods and Tools for Bayesian Variable Selection and Model Averaging in Normal Linear Regression. International Statistical Review (2018), 0, 0, 122 doi:10.1111/insr.12249
- Fragoso, T. M. and Neto, F. L. (2015) Bayesian model averaging: A systematic review and conceptual classi cation

REFERENCES

- Hansen, B.E. (2007) Least squares model averaging. Econometrica 75:1175 1189.
- Hansen, B.E. (2008) Least-squares forecast averaging. Journal of Econometrics 146:342350.
- Hansen, B.E. (2016) The Risk of JamesStein and Lasso Shrinkage. Econometric Reviews, 35:8-10, 1456-1470, DOI: 10.1080/07474938.2015.1092799
- Hansen, B.E. and Racine, J.S. (2012) Jackknife model averaging. Journal of Econometrics 167: 3846.
- Hjort, N.L. and Claeskens, G. (2003) Frequentist model averaging estimators. Journal of the American Statistical Association 98: 879899.
- Hoeting, J.A., Madigan, D., Raftery, A.E. and Volinsky, C.T. (1999) Bayesian model averaging: A tutorial. Statistical Science 14: 382417.
- Knight and Fu, (2000) Asymptotics for lasso-type estimators. The Annals of Statistics Vol. 28 (5) 1356 - 1378.
- Koop, G. (2017). Bayesian methods for empirical macroeconomics with big data. Review of Economic Analysis 9, 33 56.
- Leamer, E. (1978): Specication searches: Ad hoc inference with no experimental data, Wiley New York.
- Leeb, H. and Ptscher, B.M. (2003) The nite-sample distribution of post-model-selection estimators and uniform versus non-uniform approximations. Econometric Theory 19: 100142.
- Leeb, H. and Ptscher, B.M. (2006) Can one estimate the conditional distribution of post-model-selection estimators? Annals of Statistics 34: 25542591.
- Liu, Q. F., Okui, R. and Yoshimura, A. (2016) Generalized least squares model averaging. Econometric Reviews 35(810):16921752
- Madigan, D., J. Gavrin, and A. Raftery (1995) Eliciting prior information to enhance the predictive perfromance of Bayesian graphical models. Communications in Statistics, Theory and Methods 24, 227192.
- Magnus, J. R., Wang, W. and Zhang, X. (2016) Weighted-Average Least Squares Prediction, Econometric Reviews, 35:6, 1040-1074, DOI:10.1080/07474938.2014.977065
- Magnus, J.R. (2002) Estimation of the mean of a univariate normal distribution with known variance. Econometrics Journal 5: 225236.

REFERENCES

- ⋆ Magnus, J.R. and Wang, W. (2014) Concept-based Bayesian model averaging and growth empirics. Oxford Bulletin of Economics and Statistics doi:10.1111/obes.12068.
- ⋆ Magnus, J.R., Powell, O. and Prufer, P. (2010) A comparison of two averaging techniques with an application to growth empirics. Journal of Econometrics 154: 139153.
- ⋆ Moral-Benito, (2015). Model Averaging in Economics: An overview. Journal of Economic Surveys 29, 46-75
- ⋆ Moral-Benito, E. (2012) Determinants of economic growth: A Bayesian panel data approach. Review of Economics and Statistics 94: 566579.
- ⋆ Onorante and Raftery (2016) Dynamic model averaging in large model spaces using Dynamic Occams' Window. European Econimic Review 81, 214
- ⋆ Poeter, E., and D. Anderson. 2005. Multimodel ranking and inference in ground water modelling. Ground Water 43 : 597 - 605.
- ⋆ Raftery, A.E., Karny, M. and Ettler, P. (2010) Online Prediction under model Uncertainty via dynamic model averaging: Application to a Cold Rolling Mill. Technometrics 52, 52 - 66
- ⋆ Schneider, U. (2016). Confidence sets based on thresholding estimators in high-dimensional Gaussian regression models. Econometric Review 35(810):14121455. DOI:10.1080/07474938.2015.1092798
- ⋆ Schomaker, M. and Heumann, C. (2018) When and when not to use optimal model averaging. arXiv:1802.04589v1 [stat.ME]
- ⋆ Steel, M. F. J. (2017) Model Averaging and its Use in Economics. Online at https://mpra.ub.uni-muenchen.de/81568/ MPRA Paper No. 81568, posted 25 September 2017 13:47 UTC
- ⋆ Tibshirani, R. J., (1996). Regression shrinkage and selection via the LASSO. Journal of the Royal Statistical Society, Series B, 58, 267-288.
- ⋆ Wan, A.T.K., Zhang, X. and Wang, S. (2014). Frequentist model averaging for multinomial and ordered logit models. International Journal of Forecasting 30: 118128.
- ⋆ Wang, H., Zhang, X. and Zou, G. (2009) Frequentist model averaging estimation: A review. Journal of System Science and Complexity 22: 732748.
- ⋆ Yuan, Z. and Yang, Y. (2005) Combining linear regression models: When and how? Journal of the American Statistical Association 100: 12021214.

Thank you for listening.