MODEL SELECTION PROCEDURES; COMPARISON AND EVALUATION THROUGH MONTE CARLO EXPERIMENT

By

IMRAN KHAN

PIDE2016FPHDETS11

Supervisor

Dr. SAUD AHMED KHAN

(Assistant Professor PIDE)

Co-supervisor

Dr. ATIQ UR REHMAN

(Director, Kashmir Institute of Economics UAJK)

PIDE School of Economics

Pakistan Institute of Development Economics
Islamabad

2022

Certificate of Approval

This is to certify that the research work presented in this thesis, entitled: **"Model Selection Procedures; Comparison and Evaluation through Monte Carlo Experiment"** was conducted by Mr. Imran Khan under the supervision of Dr. Saud Ahmed Khan and Co-Supervisor Dr. Attiq ur Rahman. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Econometrics from **Pakistan Institute of Development Economics, Islamabad.**

Student Name: Mr. Imran Khan
PIDE2016FPHDETS11 — Signature: ________________

Examination Committee:

a) **External Examiner: Dr. Adnan Haider** — Signature: ________________
Professor/ Chairperson
IBA, Karachi

b) **Internal Examiner: Dr. Ahsan ul Haq** — Signature: ________________
Assistant Professor
PIDE, Islamabad

Supervisor: Dr. Saud Ahmed Khan — Signature: ________________
Assistant Professor
PIDE, Islamabad

Co-Supervisor: Dr. Atiq-ur-Rehman — Signature: ________________
Associate Professor
City Campus,
The University of Azad Jammu and Kashmir, Muzaffarabad

Dr. Shujaat Farooq — Signature: ________________
Head, PIDE School of Economics (PSE)
PIDE, Islamabad

Author's Declaration

I Mr. Imran Khan hereby state that my PhD thesis titled **"Model Selection Procedures; Comparison and Evaluation through Monte Carlo Experiment"** is my own work and has not been submitted previously by me for taking any degree from **Pakistan Institute of Development Economics, Islamabad'** or anywhere else in the country/world.

At any time if my statement is found to be incorrect even after my Graduation the university has the right to withdraw my PhD degree.

Mr. Imran Khan
PIDE2016FPHDETS11

Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled **"Model Selection Procedures; Comparison and Evaluation through Monte Carlo Experiment"** is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the **HEC** and **Pakistan Institute of Development Economics, Islamabad** towards plagiarism. Therefore I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the University reserves the rights to withdraw/revoke my PhD degree and that HEC and the University has the right to publish my name on the HEC/University Website on which names of students are placed who submitted plagiarized thesis.

Students/Author Signature:____________________

Mr. Imran Khan

PIDE2016FPHDETS11

ABSTRACT

Model selection is a major concern for many sciences and there have been a plethora of studies on model selection methods. Experts from different scientific fields have adopted different model selection procedures to select an appropriate model. Identifying the best subset among a large number of variables is the hardest part of model selection. There exist many methods for variable selection, and different methods select a different subset of variables. We can test their relative performance by comparing them. In the current study, we have compared the latest variables selection criteria from different classes of variables selection procedures; Autometrics, Elastic Net, and Extreme Bound Analysis (EBA). We have analyzed different situations (different data generating processes) by changing the combination of relevant variables, irrelevant variables, orthogonal variables, relevant variables are mutually correlated, irrelevant variables are correlated to relevant variables, relevant variables are serially correlated, heteroscedastic error terms and errors are auto correlated. When the variables are mutually correlated then Extreme Bound Analysis performs better than other model selection procedures. When the variance of the error term is heteroscedastic, the Autometrics performs superior to other rival model selection criteria. Similarly, Autometrics performs better in the case of orthogonal variables. For real data analysis, we use the data related to economic growth and its determinants for 32 countries. For in sample comparison, we use root mean square error (RMSE) and mean square prediction error (MSPE). Extreme Bound Analysis presents a superior predictive performance in terms of the lowest RMSE and MPSE that are 1.03 and 0.05 respectively.

Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AIC	Akaike information criterion
AICc	Akaike information criterion corrected (AICc)
BC	Bridge criterion
BIC	Bayesian information criterion
DGP	Data Generating Process
DHSY	Davidson, Hendry, Srba, and Yeo,(DHSY)
EBA	Extreme Bound Analysis
E-Net	Elastic Net
ESS	Estimated sum squares
FPE	Finite Prediction Error
GETS	General to specific
GRMSE	Geometric Root Mean Square Error
GUM	General Unrestricted Model
HQC	Hannan and Quinn criterion
KICvc	Vector Corrected Kullback Information Criterion
LASSO	Least Absolute Shrinkage and Selection Operator
LDGP	Local Data Generating Process
LR	Likelihood Ratio
MSE	Mean square error
MUSIC	Multiple signal classification
OLS	Ordinary Least Squares
PCA	Principal component analysis
PRS	Progressive Research Strategy
RMSE	Root mean square error
RSS	Residuals sum squares
SSE	Sum of squares of error
SVD	Singular value decomposition
TSS	Total sum square
WIC	Weighted Average Information Criterion

CHAPTER 1

INTRODUCTION

What is a model

A model is a simple representation of a complex real-world phenomenon. It is consists of the salient features of this complex phenomenon. For example, the consumption function is a complex real-world phenomenon and by making an econometric model we try to represent it in a simple form. Econometric models are statistical models which include the uncertainty (error term). There may be many economic theories for a single phenomenon or no theory behind this real-world economic phenomenon; in this situation, econometric models are also being constructed.

In general to construct an econometric model the researcher includes all variables in the model which he thinks are relevant to the dependent variable. The rest of the variables are put in the error term, and this is the main difference between economic and econometric modeling. This is also the difference between mathematical and statistical models. The mathematical models are in exact form but statistical models are in stochastic form. These are the steps in regression analysis (econometric modeling). Problem statement for the phenomenon under consideration then the second step is the selection of variables that are relevant to the dependent variable. Collection of data and model specification, then very important step is techniques for model estimation using data. The next step is the model fitting and model valuation. In end, this model is used for real-world problems.

1.1 Model selection and it's different forms

The term model selection procedure can be defined as; this is a set of rules which are used for selecting a statistical model from different candidate models which are based on data observations. The reality of any (economic) phenomenon is a complex event, so the true model is usually unknown, and model selection methods can only approximate the reality (model selection methods approximate true variables, approximate functional form, and approximate structural breaks, etc.) with observed data. Thus, model selection may take many forms; finding the variables carrying the best possible information about the variable of interest i.e. the variable specification, identifying the functional form of the model, finding the order for the autoregressive process i.e.

lag selection, to discover the change points in the models of time series i.e. identification of breaks, and finding the best estimates of parameters.

In other words, model selection is a set of procedures to select the most suitable statistical model that fits the real-world observed measurements. Modeling a complex phenomenon is a continuous search for factors that minimize the noise in the estimates, which produced hundreds of variables. This proliferation of factors has induced many technical complexities in selecting and evaluating the features for model selection. Favorably, due to the recent availability of abundant computing power, computer automation has now become possible for a new generation of model-selection techniques to be applied at a relatively lower cost and with higher speed.

1.2 The properties of a good model

The selected model should have some statistical properties. The seminal research of Hendry and Richard (1983) laid down the basis of an empirical examination for model-selection as:

- The selected model be data interpretable; meaning that the inference from the model should be logically conceivable/possible.
- The selected model should be compatible with theory; meaning that the model must have a strong theoretical background.
- The independent variables of the selected model are uncorrelated with the error term i.e. weakly exogenous independent variables.
- The coefficients of the selected model are stable over time[1].
- The selected model should be data coherent.

1.3 Different procedures for variable selection

Many procedures have been developed for different types of model selection (like finding the variables carrying the best possible information about the variable of interest i.e. the variable specification, identifying the functional form of the model, finding the order for the autoregressive process i.e. lag selection, to discover the change points in the models of time series i.e. identification of breaks, and finding the best estimates of parameters), a researcher may take in model selection, and our focus in this study is variable selection.

[1] In general, breaks in regressors in-sample should not alter the selection probabilities Krolzig and Hendry (2001)

The variables selection[2] is one of the most crucial steps of econometric modeling and it is still an open problem for all experimental researchers in all fields. In statistical model building, it is a crucial question which variables to be included in the model given a large number of variables. In particular, the technique of having better accuracy in the selection of true underlying variables for the model remained always an open research question.

There are a variety of model selection procedures as well for variable selection, so it is very important to find an appropriate criterion to evaluate the performance under different circumstances. Furthermore, the presence of a large number of potential candidate variables often called the curse of dimensionality makes the task of variable selection more difficult to retain the most relevant variables. The presence of correlation amongst the variables and the existence of outliers are the major challenges to work on for any valid inference and prediction.

The mainstream different classes for variables selection criteria as well as for model selection include; the variables selection procedures based on ordinary least square (OLS) residuals; R square (R^2), Adjusted R squared ($\bar{R}^2$) and finite prediction error (FPE), etc. The type of variable selection procedures based on information criteria; these selection criteria include Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Extended Bayesian Information Criterion, Mallow's Cp, Likelihood Ratio test (LR), Hannan and Quinn Criterion (HQC), and Bridge Criterion (BC). Variables selection criteria based on Stepwise Regression; including Forward Stepwise Regression, Backward Stepwise Regression, and Bi-directional Stepwise Regression. Variables selection procedures based on regularized methodology; including Ridge Regression, Least Absolute Shrinkage, and Selection Operator (LASSO), Elastic Net, etc. Automatic procedures for variable selection; PcGets[3] and Autometrics. Variables selection procedures based on consistency of parameters; Leamer's Extreme Bound Analysis (EBA) and Sala-i-Martin Extreme Bound Analysis. In the following few paragraphs, a brief introduction of these variable selection procedures is given.

---

[2] Model selection and variable section are overlapping concepts with the term model selection having broader coverage. The term model selection is popular in literature and previous authors used this term for variables selection. When we talk about model selection (identifying the true variables) is actually is variable selection. In literature lot of studies have used model selection procedure for variables selection (Gagné et al.,2002) (Cetin, and Aydin, E. R. A. R, 2002) (Karlsson, 2017) (Jianqing and Runze 2020).

[3] PcGets is computer program which is based on automatic econometric procedure of model selection and based on general to specific methodology.

Variables selection procedures based on ordinary least square residuals: The quality of the model (as well as variables selection) was first assessed by the error in the fit. If the selected variables are most relevant to the dependent variable then the model fits the data better, and the better it's quality. This can be done by the ordinary least square methodology, and thus became the usage of R^2 for variable selection purposes. R^2 explains the variation in the dependent variable explained by independent variables. R^2 is a non-decreasing function of independent variables. It increases, as an extra independent variable is included in the regression model, whether the extra independent variable is relevant to the dependent variable or not? To overcome this problem, an adjusted $\bar{R}^2$ was developed. The adjusted $\bar{R}^2$ increases only if the new variable is relevant to the dependent variable otherwise it may decrease or even become negative[4].

Alternative to the Least-Square methodology, the other methodology is the maximum log-likelihood estimation. This method estimates parameters of a model using the log-likelihood function, which gives the breeding ground for another popular class of variables selection that is called the information criteria. Akaike (1973) introduced variable selection criteria which are based on Kullback-Leibeler information[5] and are known as the Akaike information criterion (AIC). The AIC minimizes the discrepancy in empirical likelihood and theoretical likelihood. Later on, many criteria were introduced on similar lines on Kullback-Leibeler information, e.g., corrected AIC (AICc). Schwartz (1978) Bayesian information criterion (BIC). Another extension of the information-based criterion is the likelihood ratio (LR) test. Hannan and Quinn Criterion (HQC), HQC performs very well asymptotically. Bridge Criterion (BC) is a new criterion to control the order of an autoregressive model fitted to time series data. It has the advantages of two well-known techniques for variable selection, the Akaike information criterion, and the Bayesian information criterion.

Efroymson (1960) proposed an alternative algorithm for variable selection that recursively picks the explanatory variables for a multiple regression model from a group of candidate variables through a series of automated steps and it is called stepwise regression. It has three different types. Forward stepwise regression starts the variables selection procedure by estimating a model with

---

[4] R square can have a negative value when the model selected does not follow the trend of the data, therefore leading to a worse fit than the horizontal line. It is usually the case when there are constraints on either the intercept or the slope of the linear regression line.

[5] In mathematical statistics, Kullback-Leibeler information shows the difference of one distribution to the second distribution.

no regressor. Adds an extra variable and check its significance, if the variable is significant, it is retained in the model otherwise the variable will be removed. This process continues until no regressor is left. Mostly this is unable to find the better model (Lovell, 1983; Whittingham, Stephens, Bradbury, & Freckleton, 2006). Henceforth, further improvements were presented in the form of PcGets presented by (Hoover and Perez, (1999); Hendry & Krolzig (2001, 2002). Backward stepwise algorithm starts the selection procedure with a model in which all variables are included and remove the variables one by one which is insignificant. Bidirectional stepwise regression uses both forward and backward selection procedures at a time

A different approach for variable selection was introduced on the concept of regularization; to tackle the problem associated with the instability[6] of least square estimators due to correlation among variables (high multicollinearity). To solve this problem, Hoerl (1962) introduced ridge regression which controls the instability of coefficients associated with the least squared method. The idea rotates around reducing the parameter size to reduce the overfit and increase the predictive power of the model by adding a regularization parameter. This regression method penalizes the coefficients of the regression, making them lower in absolute value and as a result, the variance of regression function remains stable in comparison to OLS. Nevertheless, ridge regression does not perform well for feature selection and considers all variables for the final model.

Tibshirani (1996) proposed a variant of regularized regression in which he used the same minimization function as in ridge regression but subject to different constraints, to solve the variable selection problem. Due to computational limitations, the LASSO regression did not become popular until the development of the least-angle regression (LARS)[7] algorithm in 2002 which provided an efficient way to calculate the estimates of LASSO (Tibshirani, 2011). The following years marked many modifications of LASSO: Fused LASSO (Tibshirani et al., 2005), Grouped LASSO (Yuan and Lin, 2006), Adaptive LASSO (Zou, 2006), Graphical LASSO (Yuan and Lin, 2007), Dantzig Selector (Candes et al, 2007), Matrix Completion (Candes and Tao, 2010) and Near Isotonic Regularization (Tibshirani et al., 2011). The most relevant amongst all of them

[6] The estimates of parameters swing wildly based on inclusion of independent variables are in the model. The coefficients become very sensitive to small changes in the model.

[7] Least Angle Regression (LARS) is an algorithm used in regression for high dimensional data (i.e., data with a large number of attributes). Least Angle Regression is somewhat similar to forward stepwise regression.

was the development of Elastic Net by Zou and Hastie (2005), which is a combination of LASSO and ridge regression.

The two methods (forward stepwise and backward stepwise) are recognized as specific-to-general and general-to-specific (GETS) methods. And the mixture of these two methods is known as Autometrics (Doornik & Hendry, 2007; Doornik, 2009). Autometrics is a new and advanced algorithm for automatic model selection (also known as the 'Hendry' or 'London school of economics (LSE)' methodology). It improves on previous implementations through an enhanced search method which essentially makes a presearch simplification, unnecessary. In addition, the algorithm is presented within a likelihood framework, allowing for applications beyond regression models.

The last type of variable selection procedure (in the current study) is based on the consistency of the coefficient of a variable in estimated models, called Extreme Bound Analysis (EBA). It selects the relevant variables by applying the procedure for parameters consistency of the variables under consideration. Two popular procedures that follow extreme bound analysis are Leamer's extreme bound analysis and Sala-i-Martin extreme bound analysis. In Leamer's extreme bound analysis, different groups of doubtful variables are chosen and the coefficient of the core variable is estimated. If the coefficient of the core variable is beyond the extreme bounds[8], then this variable is not included in the final model. If the coefficients of the underlying variables do not vary across the limits when the doubtful variables are included in the model with different combinations, then the underlying variable will be retained in the final model. In comeback to the professed inflexibility[9] of EBA that is presented by Leamer, a substitute method by Sala-i-Martin (1997) considers the entire distribution of coefficient of underlying core variable, not just on its extreme bounds. In Sala-i-Martin EBA the binary label robust (relevant) or fragile (irrelevant) is not given to a variable under consideration but it assigns such a confidence level for robustness to the variable under consideration.

[8] For any underlying variable v, the lower and upper extreme bounds are clear as the minimum and maximum values $\hat{\beta}_j \pm \tau\hat{\sigma}_j$ across the M estimated regression models, where τ is the critical value for a specific level of confidence. The conventional 95-percent confidence level τ will be equal to approximately 1.96.

[9]In leamer's EBA if only one coefficient from "M" number of coefficients lies outside the bounds then the variable will be dropped.

1.4 How to find the best model

The choice of the variable selection procedure is the persistent unresolved issue; because the reality is complex, simultaneously dynamic, non-synchronous, and involves high-dimensional data structure. Social structures may also influence as they also change from time to time, thus making the model-making process more complex. To coup these challenges, over time advanced and updated, variable selection procedures are emerging (Autometrics Elastic Net). To find which one model selection procedure is better than other procedures is only possible by a valid comparison among model selection procedures by using the same data generating process. However, there is a lack of studies comparing the advanced and updated procedures among them and comparison with existing model selection procedures under the DGPs setting which we are going to do in our study. Modern academic work has highlighted the need to expand this previous work to compare the old variable selection procedures by including advanced and updated procedures of model selection. In this study, we try to address this need, we provide a common understanding of the characteristics and practical performance of different variable selection procedures by reviewing their theoretical and practical strengths, behaviors, and relationships. Therefore, this (comparison study) is an attempt to bridge the gap highlighted in the literature. This study aims to compare the variable selection procedures of different classes of variable selection procedures. We have taken three main classes for variable selection. Then take one updated procedure for variable selection from each main class based on theoretical and empirical grounds, we have taken the latest forms of each class of variable selection.

In the current study, we are going to know the performance of extreme bound analysis (variables selection procedure) in comparison with the regularized type of variable selection procedures(Elastic Net), and the performance of extreme bound analysis in comparison with Autometrics. In literature, there is a lack of this type of comparison under the current setting of DGPs.

1.5 The evaluation of statistical techniques

The evaluation of statistical methods usually follows two routes; the first is using real data that has already undergone some statistical procedure. The advantage of this method is that it allows the comparison of results between current and previous studies. But the problem with this is that the true DGP is unknown and one does not know whether the selection procedure has retained the correct variables or not.

The other method for evaluation of statistical procedure is based on Monte Carlo Simulation. This method starts with known DGP and then retains the DGP by different statistical procedures and compares it with true DGP. The statistical procedure which retains the DGP nearest to true DGP (which is generated already) is preferred to other statistical procedures.

This study uses both Monte Carlo Simulation and real data strategies to analyze the characteristics of variable selection procedures of different classes. The main objectives of the study are the following.

1.6 Objectives

1- To find out the optimal variable selection procedure that gives us the best results for a variety of scenarios.

2- To test the robustness of variable selection procedures in the following situations:

   i. To test the robustness of variable selection procedures for different sample sizes.
   ii. To check the performance of the variable selection procedures for the situation of autocorrelation in errors terms.
   iii. To check the performance of variable selection procedures in case of multi-collinearity and serial correlation in regressors.
   iv. To check the robustness of variable selection procedures in case of heteroscedasticity in the variance of errors terms.

3- To compare the variable selection procedures based on ‘in sample forecasting’ using real data.

1.7 Significance of the Study

The choice of the variable selection procedure is the persistent unresolved issue; because the reality is complex, simultaneously dynamic, non-synchronous, and involves high-dimensional data structure. Social structures may also influence as they also change from time to time, thus adding to the complexity of model making process. For example, there are numerous economic theories for the same empirical phenomena with quite different viewpoints that all seem quite plausible. But there is no clear consensus in the literature as to which theoretical procedure to employ or which factors drive a certain phenomenon such as economic growth. Thus, both the model itself and the variable selection procedure are of significant importance to glean the most valid inferences. Explaining the underlying phenomenon is the fundamental objective of any model and the search for factors that determine the dependent variable; has produced a large number of candidate variables. Which factors are more relevant to the dependent variable is the challenge for variable selection procedures.

So, to tackle these difficulties new techniques of variable selection are emerging, to find which one variable selection procedure is better than other procedures is only possible by a valid comparison among these model selection procedures. So, it is considered important to draw a valid comparison amongst those techniques of variable selection. This study will assist researchers to narrow down the best technique for variable selection, given the available information set (data generating process settings or real data). This study would help to choose the appropriate variable selection procedure according to the properties of the data under consideration. Our contribution is geared on the comparison of different advanced techniques of variables selection procedures, which to the best of our knowledge are not compared[10] before in the current setting of DGPs (DGP with correlated variables, DGP with serially correlated variables, DGP with autocorrelated error terms and DGP with heteroscedastic error terms and DGP with multiple[11] problems at a time). Another contribution is based on the extraction of factors that explain the phenomenon of economic growth by employing the most recent and cutting-edge econometrics techniques.

1.8 Organization of the Study

Chapter one describes the introduction of the study. In chapter two, we start variable selection procedures with the traditional methodology and then progress to the advanced techniques for variable selection methods. After this, we put the studies in which these variables selection procedures are compared using Monte Carlo Simulation as well as with real data. We finally then make some grounds for the current study to make the comparison among different variables selection procedures. In chapter three the methodology is described that draws the mechanism of

[10] In the literature we are unable to find the comparison between EBA type of model selection procedures and Autometrics, secondly, we also unable to find the comparison between EBA (sensitivity analysis family of variable selection) and LASSO (shrinkage based methodology of variable selection procedures).

[11] It is possible that there is multi collinearity autocorrelation and heteroscedasticity is present in model at a time.

comparison between the selected model selection procedures. The DGPs with different assumptions like correlated variables, autocorrelated error terms are explained in the chapter of methodology. We provide a valid rationale for each setting of GDP and describe the complete statistical procedure of each DGP. Simulated results are illustrated in chapter 4. Chapter 5 presents the real-life empirical application of the study on the driver of economic growth of 32 different countries. For extraction of the variables of economic growth by selected variables selection procedures are done in chapter 5 and relative frequencies are also calculated in chapter 5. A summary of the simulation results concerning different settings of DGPs is given in chapter 6. Further, the conclusion is based on the recommendations and future research concepts are also in chapter 6.

CHAPTER. 2
LITERATURE

2.1 Background

Usually, the researcher has a large number of candidate variables that can potentially explain the dependent variable. The variable selection procedure selects the best subset from candidate variables which are assumed to provide the most appropriate explanation for the dependent variable. We start with the traditional methodology and then progress to the advanced techniques for variable selection methods. After this, we put the studies in which these variables selection procedures are compared using Monte Carlo Simulation and also with real data. We finally then make some grounds for the current study to make the comparison among different variables selection procedures.

2.1.1 Traditional Econometric Methodology for variable and model selection

The first step of the traditional methodology is to state the theory and then in the next step the mathematical model for the theoretical hypothesis is presented in the form of:

$$Y_t = X\beta + \varepsilon_t \tag{2.1}$$

For econometric analysis one must have; (a) Economic theory. (b) Specification of the mathematical model of the theory. (c) Specification of the statistical or econometric model. (d) Statistical data. (e) Estimation of the parameters of the econometric model. (f) Hypothesis testing. (g) Forecasting or prediction. (h) Using the model for policy.

The model selection relative to (c) to (f) finally leads to the accuracy of (h). The traditional methodology takes the functional form theory i.e. all variables of the model are based on a pre-defined theory. As we have described before, numerous theories are explaining the same phenomenon and resultantly produce a different model with each model estimated without any concern about other models.

The assumptions of the traditional methodology include; *(1) The regression model is linear in parameters. (2) The mean of residuals is zero.*$E(\varepsilon_t) = 0$. *(3) Homoscedasticity of residuals or equal variance* $var(\varepsilon_t) = \sigma^2$. *(4) No autocorrelation in residuals.*$cov(\varepsilon_t, \varepsilon_{t-1}) = 0$. *(5) The X variables and residuals are uncorrelated* $cov(\varepsilon_t, X) = 0$. *(6) The number of observations must be greater than the number of Xs.* $\mathrm{N} > \mathrm{P}$. *(7) The variability in X values is positive. (8) The regression model is correctly specified. (9) No perfect multi-collinearity.*

When all assumptions stated above are fulfilled, the next step in the traditional methodology is the estimation of the parameters after collecting the data. After estimation, check the diagnostic tests (Autocorrelation test, Hetroscdasticity test, Parameters stability test, Specification test) and if the model passes through the diagnostic testing, then variables in the model are considered as most relevant variables to the dependent variable. Eventually, the resultant model is used for designing and feedback of the policy. The steps of the traditional methodology are outlined in figure 2.1 below:

2.1.2 Leamer's Criticism on traditional methodology and Extreme Bound Analysis

Leamer (1978) criticized the traditional methodology for the regression model. He said the "Traditional version" of the regression model can be easily rejected as this is based on the unreliable assumption of the correct specification. Leamer purposed the Extreme Bounds Analysis (EBA) to overcome this (misspecification) issue. EBA selects the relevant variable by applying the procedure for parameters consistency on core variables. Two popular procedures that follow extreme bound analysis are, Leamer's extreme bound analysis and Sala-i-Martin extreme bound analysis. In Leamer's extreme bound analysis, different groups of doubtful variables are selected and the coefficient of the core variable is estimated. If the coefficient of the core variable is beyond the extreme bounds, then this variable is not included in the final model. If the coefficient remains within the limits then it will be retained in the final model.

In comeback to the professed inflexibility of EBA that is presented by Leamer, a substitute method by Sala-i-Martin (1997) was presented by Sala-i- Martin which considered the entire distribution of an underlying core variable, not just on its extreme bounds. In Sala-i-Martin EBA the binary label robust (relevant) or fragile (irrelevant) is not given to a variable but it assigns such a confidence level for robustness to the variable under consideration. A regressor may be significant on 95% level or 90% level.

2.1.3 Tibshirani (1996) criticism on traditional methodology and regularized methodology

The conventional method assumes that the number of observations must be greater than the number of parameters in the regression model. What will happen, if the number of parameters exceeds the number of observations? Then the traditional methodology fails because when the number of the parameter is greater than the number of observations there is no unique solution for parameters in traditional methodology. One solution is a regularized type method. The first popular type of this method is Ridge regression, which regularized the coefficient towards zero, but is

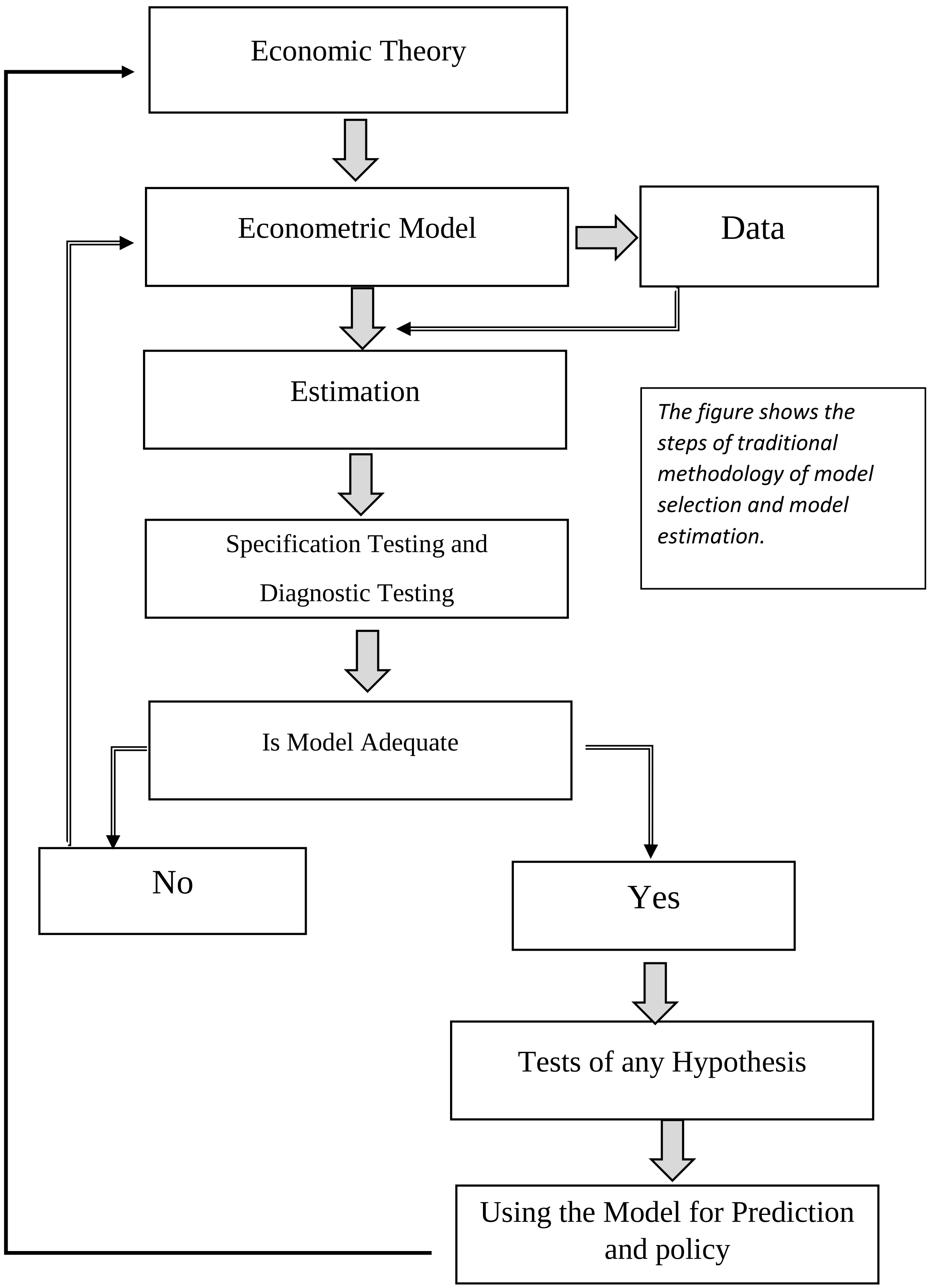


Figure 2.1: Steps of traditional methodology

unable to make them exactly zero then LASSO purposed by Tibshirani (1996) which can shrink some coefficients to zero and can make them exactly zero.

Secondly, the LASSO is a procedure for both estimation of the model and the selection of the variables simultaneously. LASSO can handle the situation in which there are more variables than observations and presents sparse models (Zhao and Yu, 2006, Meinshausen and Yu, 2009). Further regularization path of the LASSO can be calculated efficiently as shown in Efron et al. (2004), or Friedman et al. (2010) lately. Several generalizations and different types of LASSO procedures to overcome many problems are found in Tibshirani (2011). Elastic Net (E-Net), adaptive LASSO (adaLASSO) has received particular attention.

2.2 MODEL SELECTION PROCEDURES

Model selection is an important part of all statistical analysis in the pursuit of science in general. Numerous authors have examined this question extensively, and several tools have been suggested to select the "best model" and to select the most relevant variables in the literature. There are several ways to choose the most relevant variables, some of which are listed below. The main classes of variables selection criteria are described below in figure 2.2.

2.2.1 Residual Based Criteria

2.2.1.1 R Square

R^2 Explains the variation in the dependent variable explained by independent variables. R^2 is a non-decreasing function of independent variables. So it always increases as an extra independent variable is included in the regression model. So, if the R^2 of a model is 0.50, then about half of the observed variation is explained by the model's inputs.

$$R^2 = \frac{ESS}{TSS} \tag{2.2}$$

$$R^2 = 1 - \frac{RSS}{TSS} \tag{2.3}$$

ESS= Estimated sum of the square, RSS= Residual sum of the square, and TSS= Total sum of squares. The significant property of R^2 is that it is a non-decreasing function of the number of explanatory variables or regressors present in the model, except the added variable is perfectly collinear with the other regressors. As the number of regressors increases, R^2 increases and never decreases.

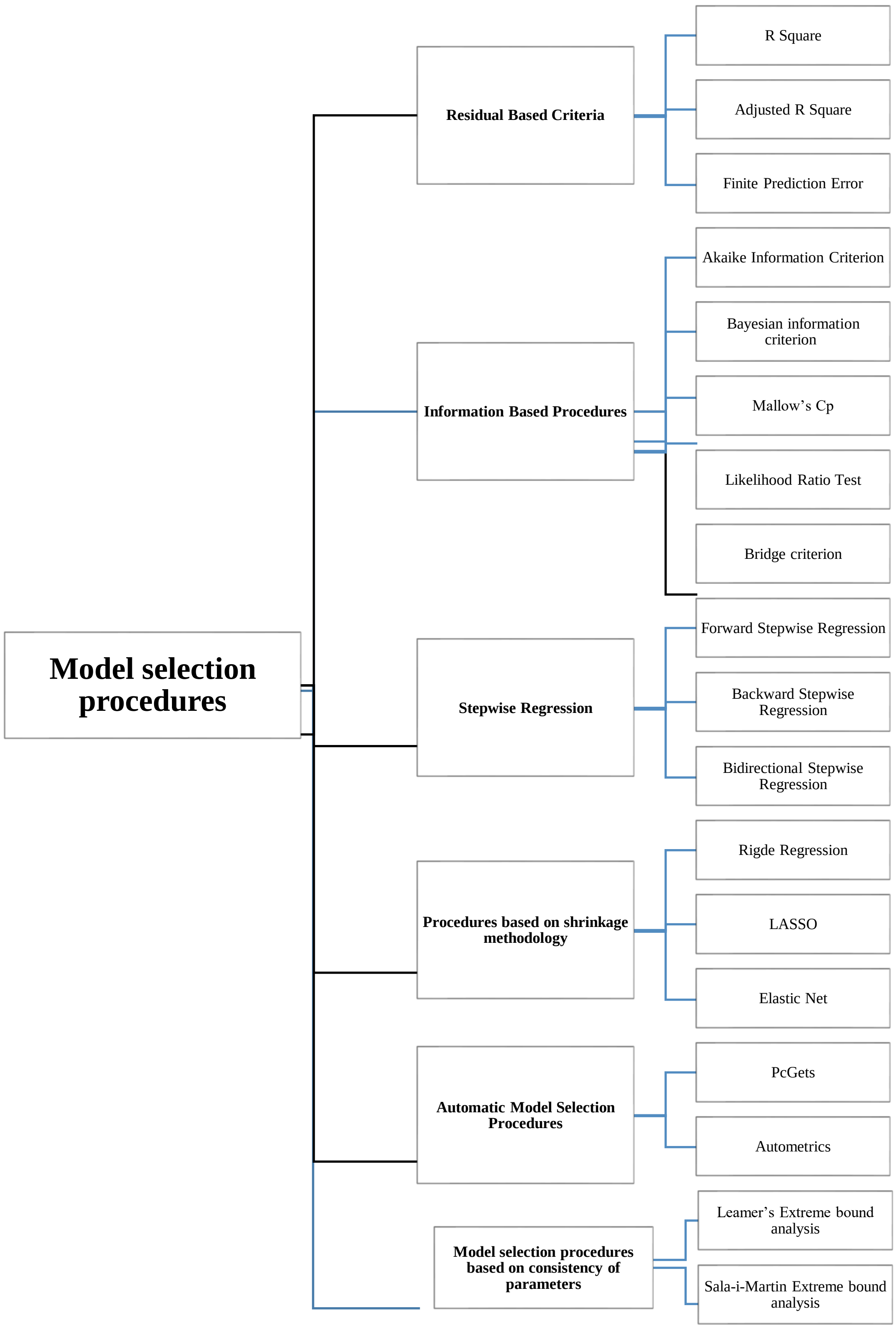


Figure 2.2: Different groups for model selection procedures

2.2.1.2 Adjusted R Square

The problem with R^2 is; it always increasing whether the extra included independent variable is relevant to the dependent variable or not? To overcome this problem, an adjusted $\bar{R}^2$ was introduced. The adjusted $\bar{R}^2$ increases only if the new variable is relevant to the dependent variable, otherwise it may decrease or even become negative. The adjusted $\bar{R}^2$ is a modified version of R^2 that has been familiar with the increasing number of predictors in the model. The adjusted $\bar{R}^2$ increases only if the new regressors improve the model quality. It decreases when a predictor does not improve the model quality.

$$\bar{R}^2 = 1 - \frac{\sum \hat{\mu}_i^2 / (n-k)}{\sum y_i^2 / (n-1)} \tag{2.4}$$

where k = the number of parameters in the model including the intercept term.
$\sum \hat{\mu}_i^2$=residuals sum of the square, $\sum y_i^2$=total sum of square and, n =is the number of observations in the model. The term adjusted means adjusted for the degree of freedom (df) associated with the sums of residuals squares, and has (n – k) degree of freedom in the estimates model with k number of parameters, with intercept term, and $\sum \mathrm{y}_\mathrm{i}^2$ has (n – 1) degree of freedom.

$$\bar{R}^2 = 1 - \frac{\hat{\sigma}^2}{S_Y^2} \tag{2.5}$$

Where $\hat{\sigma}^2$ is the variance of a residual, unbiased estimator of true σ^2 and S_Y^2 is the sample variance of Y. It is easy to see that $\bar{R}^2$ and R^2 are interrelated to each other as.

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k-1} \tag{2.6}$$

It is directly apparent for k > 1, $\bar{\mathrm{R}}^2 < \mathrm{R}^2$ which implies that as the number of X variables increases, the adjusted $\bar{\mathrm{R}}^2$ increases less than the unadjusted R^2; and adjusted $\bar{\mathrm{R}}^2$ can be negative, although R^2 is necessarily nonnegative. Theil notes:

“It is more appropriate to use adjusted $\bar{R}^2$ than R^2 as the R^2 gives an over-optimistic figure for regression fit, mostly in the case where the number of observations is not too large to number of regressors”

2.2.1.3 Finite Prediction Error (FPE)

Akaike's Finite Prediction Error criterion delivers a measure of model quality by simulating the situation where the model is tested on a different data set.

$$FPE = RSS(1 + \frac{2P}{N - P}) \tag{2.7}$$

RSS=Residuals Sum of the square, P=number of regressors and, N=number of observations

We want to choose a model that minimizes the FPE, which represents a balance between the number of parameters and the explained variation in the dependent variable.

2.2.2 Information Based Procedures

The model selection procedures that are based on information theory minimize the loss of information by choosing a different number of variables or group of variables. For those variables where the loss of information is minimum then these variables are declared as the most relevant variables to the dependent variable for the model under consideration.

2.2.2.1 Akaike Information Criterion (AIC)

The Akaike information criterion (AIC) is an estimator for prediction error and the relative quality of statistical models for a given set of data. Given the different models for the data, AIC estimates the quality of each model, relative to the other models. Thus, AIC provides a means for model selection as well as variable selection. As the statistical model is applied to capture the process of data, this will be never exact; something is missing. AIC calculates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.

In the process of calculating the amount of information lost by a model, AIC deals with the trade-off between the goodness of fit and the simplicity of the model.

$$AIC = 2k - 2\ln(\hat{L}) \tag{2.8}$$

K= number of regressors and, $\ln(\hat{L})$= Log-Likelihood of the model

A set of candidate models are given for the data; the preferred model is the model that has the lowest AIC value. Thus, the AIC rewards the goodness of fit (as measured by the likelihood function), but also includes a penalty that is an increasing function of the estimation parameters in the model. Penalties discourage excessive fit, as increasing the number of parameters in the model always improves the goodness of fit.

2.2.2.2 Bayesian information criterion (BIC)

Bayesian information criterion (BIC) is used to select a model among the number of models. BIC is closely related to AIC. In practice, it may increase in likelihood by adding more variables in the model but AIC and BIC can handle this situation by introducing a penalty term. Let $\left(\hat{\mathrm{L}}\right)$ the maximized value of the likelihood function of the model, and K is the number of parameters.

$$BIC = \ln(n)k - 2\ln(\hat{L}) \tag{2.9}$$

BIC compares the different models and selects the model which has the lowest BIC value. With increasing the sample size, mostly the model selection criteria perform better.

2.2.2.3 Mallows Cp

Mallow's C_p solves the problem of over-fitting, C_p statistics are constructed on data sample mean squared prediction error (MSPE).

$$Cp = \frac{RSS_p}{\hat{\sigma}^2} - (n-2p)\hat{\sigma}^2 \tag{2.10}$$

RSSp=Sum of squares of residuals of the model with (p-1) regressors, N=number of observations, and $\hat{\sigma}^2$=standard error of the model.

Mallows suggest the statistic for selecting the model from several alternative models. Under the models that do not have an admirable reduction in fit (bias), the C_p expectation is equal to p.

C_p considers the ratio of the sum of squares of error (SSE) of the model with (p – 1) variables to Mean Square Error (MSE) for the full model; then penalizes for the number of variables:

$$Cp = \frac{SSE_p}{MSE(full)} - (n-2p) \tag{2.11}$$

SSE_p =Sum of squares of errors of the model with (p-1) regressors, MSE(full)= Mean square error and. N= size of the sample.

An estimated model is considered good if $C_p \leq p$. The estimated model may a model with minimum number of regressors. A benefit to Cp is we can use it to select model size getting a good model that contains as few variables as possible. Might also choose to pick the model with the minimum Cp, but it is more about Cp relative to p, and getting a smaller number of variables in the model while still having the same predictive ability.

2.2.2.4 Likelihood Ratio (LR) Test

The likelihood-ratio test assesses the goodness of fit of two competing statistical models based on the ratio of their likelihoods, specifically one likelihood is found by maximization over the entire parameter space and another likelihood is founded after imposing some constraint on the model's parameters. The likelihood ratio is for the comparison of only two models. If the observed data support the constraint (i.e., the null hypothesis) then this indicates the two models have not different from each other. Thus the likelihood-ratio test tests the difference of two likelihoods whether they are different from one another or not.

H_0 : Smaller model is best

H_1 : Larger model is best

$$LR = -2\ln\left(\frac{L(\hat{\theta})}{L(\theta)}\right) \tag{2.12}$$

$L(\hat{\theta})$=log likelihood function of the smaller model and, $L(\theta)$=log likelihood function of the larger model.

The test statistic approach to a chi-square distribution with degree of freedom (the difference of parameters of the two models). If the value of this statistic is very low, then the probability ratio test will reject the null hypothesis. How the test statistic is small, depends on the significance of the test, i.e. type 1 error being considered.

2.2.3 Stepwise Regression

Stepwise regression is a method of fitting a regression model in which explanatory variables are selected automatically. At each stage, a variable is considered to be added or to be subtracted in the model, from a set of explanatory variables based on some predefined criteria. Typically, this takes the form of a sequence of F-tests or t-tests, but other techniques are possible, such as adjusted R^2, Akaike information criterion, Bayesian information criterion, Mallow's C_p. The stepwise regression can be divided into three categories.

2.2.3.1 Forward selection

This method starts with no explanatory variables in the estimated model, then adds each explanatory variable one by one in the model and checks its significance using some predefined criterion. If a variable is significant, it will be included in the final model otherwise drop this explanatory variable from the final model.

2.2.3.2 Backward elimination

This method start will with all explanatory variables in the model and tests the significance of each variable by using some predefined criterion. Delete the most insignificant variable maintaining the overall model fit statistics, then again estimate the model repeat the process of deletion. Continue this process until no more explanatory variables are deleted.

2.2.3.3 Bidirectional elimination

This procedure is a combination of the above procedures (Backward elimination and Forward selection), in which each step is tested to add or remove the explanatory variables based on some predefined criterion.

2.2.4 Model Selection Procedures based on shrinkage methodology

2.2.4.1 Ridge Regression

Ridge regression is like the least square, but the procedure of coefficients estimation is a little bit different. The least-square regression tries to estimate the slope coefficients of the variables by minimizing the RSS (residuals sum squares). While the ridge regression minimizes the below equation for coefficients estimation.

$$\hat{\beta}^{ridge} = \underset{\beta}{agrimin}\left\{\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^2\right\} \qquad (2.13)$$

Where λ ≥ 0 and is a tuning parameter that is calculated separately. The above equation makes the balance between two different criteria. Ridge regression estimates coefficients that fit the data well, by minimizing RSS as like the least square do, and Secondly by a shrinkage penalty, which is minimum as β1, . . . , βp are minimum.

When λ = 0, the value of the penalty is zero and the coefficients of ridge regression are as same as the coefficient of least square regression. When λ→∞ the effect of penalty becomes strong and the coefficients of ridge regression approach to zero (but not exactly zero). In the estimation procedure of least-squares, there is only one set of coefficients but in the case of ridge regression, there are many sets of coefficients for each value of λ. The penalty term is applied only on slopes coefficients β1, . . ., βp, not on intercept βo. The main aim of shrinkage of coefficients is to shrinkage the association between a dependent variable and independent variables but not to shrink the intercept which is the mean value of the dependent variable when all independent variables are zero.

The preference of Ridge regression to least squares is due to the bias-variance trade-off. When the value of λ goes up then there is a decrease in the flexibility of ridge regression fit, further, decrease in variance but bias goes up. With the increasing value of λ, the variance of predictors goes down at the risk of increasing bias.

The least-squares estimates have low bias and may have more variance when the relationship is linear. When the number of observations is equal to or less than to predictors then the least square has no unique solution. But the ridge regression still performs better by trading off in bias-variance.

We can write the ridge regression as.

$$\hat{\beta}^{ridge} = \underset{\beta}{agrimin} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 \tag{2.14}$$

$$subject\ to\ \sum_{j=1}^{p} \beta_j^2 \le t,$$

This provides explicitly the size of the constraint on the parameters. There is a correspondence of lambda and t as one-to-one. When there is high multi-collinearity in many regressors then the coefficients have more variance than usual and are poorly determined. A large positive coefficient cancels the negative coefficient in which case the variables are correlated.

2.2.4.2 LASSO

LASSO is the same as ridge regression but with a minor difference. Before LASSO, the stepwise selection method is more widely used for choosing the regressors in which prediction accuracy is improved in certain cases, mostly the prediction accuracy is worse. Ridge regression recovers prediction error by shrinking the large regression coefficients for reducing the overfit[12] but failed to the selection of covariates. LASSO overcome the above problems by shrinking the coefficient less than from a fixed value and further forcing them to zero. This is the same as in ridge regression but in the ridge, regression coefficients are shrinkage to zero but not exactly zero.

The coefficients of LASSO areas, $\widehat{\mathrm{B}}_{\mathrm{j}}^{\mathrm{L}}$, minimize the quantity

$$\beta\ ^{\wedge \mathrm{lasso}} = \underset{\beta}{\mathrm{agrimin}} \sum_{i=1}^{N} \left(\mathrm{y} - \beta_0 - \sum_{\mathrm{j=1}}^{\mathrm{p}} \mathrm{x_{ij}} \beta_{\mathrm{j}} \right)^2 \tag{2.15}$$

$$\text{subject to } \sum_{\mathrm{j=1}}^{\mathrm{p}} \left| \beta_{\mathrm{j}} \right| \leq \mathrm{t}$$

Same as the ridge regression, LASSO shrinks the coefficient estimates to zero. However, here the L1[13] penalty makes some coefficients exactly zero when the value of λ is sufficiently large. When λ goes up bias increases and as λ decreases, the variance goes up. As $\lambda = 0$, no parameter becomes zero. The coefficients are the same as the least square coefficients. Hence, LASSO makes the variables selected as the best subset selection. Lastly, the results of LASSO are easily interpretable than the results of ridge regression.

We can write the estimation of the parameters as lagrangian form,

$$\hat{\beta}^{lasso} = \underset{\beta}{\mathrm{agrimin}} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(\mathrm{y} - \beta_0 - \sum_{\mathrm{j=1}}^{\mathrm{p}} \mathrm{x_{ij}} \beta_{\mathrm{j}} \right)^2 + \lambda \sum_{j=1}^{p} \left| \beta_j \right| \right\} \tag{2.16}$$

Similar to ridge regression but L_2 the penalty is replaced by L_1 penalty$\sum_{j=1}^{\mathrm{p}} |\beta_{\mathrm{j}}|$. Later the constraint makes the solution nonlinear in the $\mathrm{y_i}$.

[12] An overfit model is a model that is too much complicated (more parameters than justified by the data) for the given data set.

[13] L1 penalty is equivalent to the sum of the absolute value of coefficients $\sum_{j=1}^{p} |\beta_j|$

2.2.4.3 Elastic Net

Although LASSO has shown a lot of success in many ways, it has its limitations. Consider the following three scenarios. Here "p" is a number of variables and "n" is a number of observations.
(a) In the case of p>n, due to the nature of the convection correction problem, LASSO mostly chooses new variables. This appears to be a limited feature of the variable selection procedure. Furthermore, LASSO cannot be well defined unless it is less than a certain value, bound to the L1-norm of the coefficient.
(b) If there is a group of variables that are closely related by a pair, then LASSO selects only one of the groups and does not care which one is chosen.
(c) For the normal situation when n>p and there is a high correlation between independent variables then results of LASSO prediction are dominated by ridge regression (Tibshirani, 1996).
Scenarios (a) and (b) in some cases make LASSO the method of choosing the irrelevant variable. As for as predictive performance is concerned, scenario (c) is no less of a regression issue. Therefore, it is possible to further strengthen LASSO predictive power. Simulation results and real data results demonstrate that Elastic Net often outperforms LASSO to forecasting accuracy.
There are two variants of Elastics Net as follows.

2.2.4.4 Naïve Elastic Net

Let assume there are n observations and p regressors. Let **y**=y1, . . . , yn be the dependent variable and **X** data matrix of independent variables. Location and scale transformation, let us assume that the response is centered and the predictors are standardized.

$$\sum_{i=1}^{n} y_i = 0$$

$$\sum_{i=1}^{n} x_{i,j} = 0$$

For j=1,2,3…p
For any fixed and non-negative λ_1 and λ_2 the elastic net.

$$L(\lambda_1, \lambda_2, \beta) = |Y - X\beta|^2 + \lambda_2 |\beta|^2 + \lambda_1 |\beta| 1 \quad (2.17)$$

Where λ_1 *and* λ_2 are tunning parameters

$$|\beta|^2 = \sum_{j=1}^{p} \beta_j^2 \quad (2.18)$$

$$|\beta|1 = \lambda \sum_{j=1}^{p} |\beta_j| \quad (2.19)$$

The naïve elastic net B estimator is the minimizer of the equation below.

$$\hat{B} =_{b}^{agri\min} \{L(\lambda_1, \lambda_2, \beta)\} \quad (2.20)$$

The procedure can be watched as penalized least square procedure.

Assume= $\lambda_2/(\lambda_1, \lambda_2)$; then explaining B in equation 2.20 is equal to the optimization procedure

$$\hat{B} =_{b}^{agri\min} |Y - X\beta|^2 \quad (2.21$$

Subject to $(1 - \text{a})|\text{B}|1 + \text{a}|\text{B}|^2 \leq \text{t}$ for some t

Function $(1 - \text{a})|\text{B}|1 + \text{a}|\text{B}|^2$ is called the elastic net penalty. That is a mixture of the ridge and LASSO penalty. If a=1 it behaves like ridge regression. When a=0 it is like LASSO and if a<1 this will be elastic net. Below figure 2.3 illustrates this.

When the number of the variable is more than the number of observations then grouped variables are introduced to handle this situation. The naïve elastic net can handle the situation of group variables that are not incorporated in the LASSO. Segal and Conklin (2003) urge to use the regularized regression (Elastic net) procedure to find the grouped genes.

2.2.4.5 Elastic Net

As a method of automatically selecting variables, the naive elastic net removes the limitations of the LASSO in scenarios (a) and (b). However, empirical evidence suggests that the naive elastic net does not perform satisfactorily unless it is very close to ridge regression or LASSO. The naive Elastic Net in regression prediction sequencing, an accurate method of penalization achieves better forecasting performance by trading off the difference in bias-variance. In the estimation process of Naïve elastic, there are two steps: first is each value of λ2, then the calculation of regression coefficients. The second step is the calculation of LASSO type shrinkage coefficients. It is like double shrinkage compared to pure LASSO or the ridge regression and the double shrinkage does not minimize the variations but increases the additional bias.

Given the data on Y and X, and penalty terms λ_1, λ_2 and modified data (y^*, X^*) the naïve elastic net estimate LASSO type coefficients as

$$\hat{B}^* = (y^*, X^*)\hat{B} = \underset{\hat{B}^*}{ágri\min} | y^* - X^*\beta^* |^2 + \frac{\lambda_1}{\sqrt{(1+\lambda_2)}} | \beta^* | 1 \tag{2.22}$$

The elastic net calculation for $\widehat{\mathrm{B}}$ areas.

$$\hat{B}(Elastic - Net) = \sqrt{(1+\lambda_2)}\hat{B}^* \tag{2.23}$$

That $\widehat{\mathrm{B}}$(naïve elastic net) $= \{1/\sqrt{(1+\lambda_2)}\}\hat{B}^*$ thus

$\widehat{\mathrm{B}}$(Elastic net)=$(1+\lambda_2)$ $\widehat{\mathrm{B}}$(naïve elastic net)

Henceforth the coefficients of the Elastic Net have rescaled the coefficients of the naïve Elastic Net. This modification preserves the property of variables selection of naïve Elastic Net and it is a simple method to shrink the coefficients.

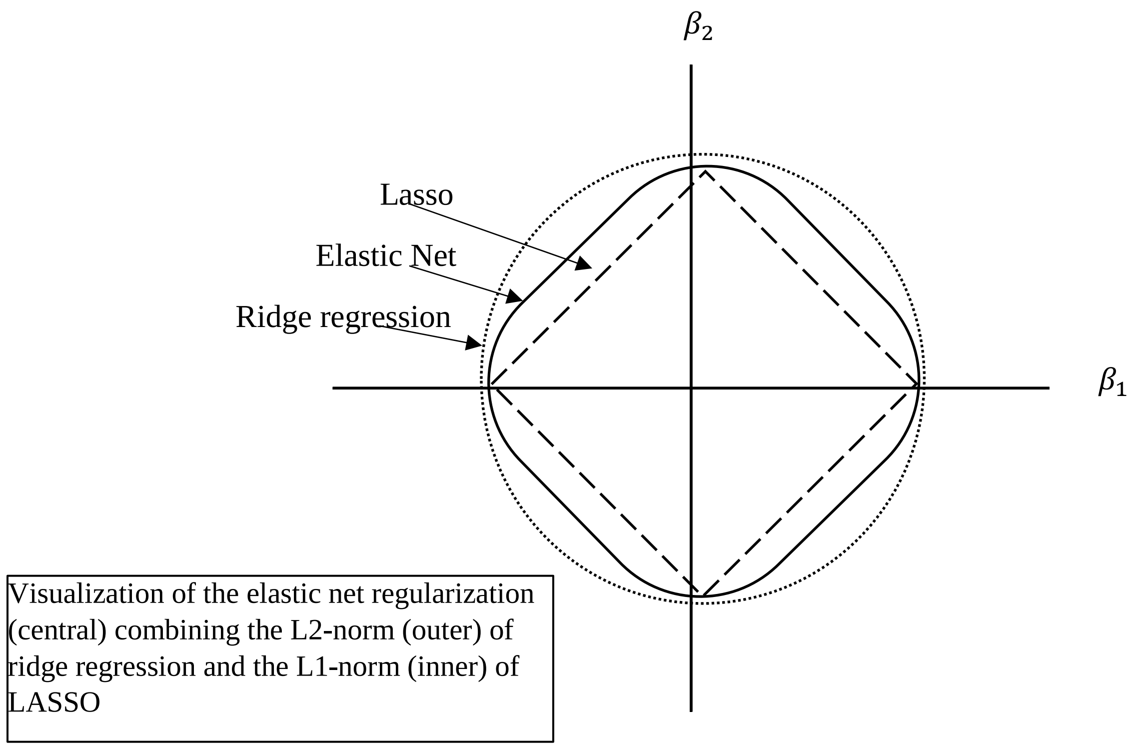


Figure 2.3: Visualization of regularized penalities

Source:Hui Zou and Trevor Hastie (2005)

2.2.4.6 Choosing the tuning parameters *"λs"*

There are two methods, which are used for choosing the tuning parameter λ in regularized type regression models:

One is the traditional method in which information-based criteria like AIC or BIC. Which are used to select the tunning parameters and where AIC or BIC is the smallest that specific λ is chosen. This method for choosing the tuning parameter focuses on the model's fit. The second criterion is the machine learning approach in which cross-validation is done to choose the tuning parameter in the regularized type of regression. This method is focused on the predictive performance of the model.

2.2.5 Automatic Model Selection Procedures (General to specific modeling)

The two methods (forward stepwise and backward stepwise) are recognized as specific-to-general and general-to-specific (GETS) methods for model selection. The mixture of these two methods is known as Autometrics (Doornik & Hendry, 2007; Doornik, 2009). Autometrics is a new and advanced algorithm for automatic model selection tilting more towards a general-to-specific framework (also known as the 'Hendry' or 'LSE' methodology). The algorithm is presented within a likelihood framework, allowing for applications beyond regression models.

Autometrics is an extension of GETS, which incorporates further steps in finding the final model. Autometrics apply the tree search method, which takes up space to overall the model. However, finding all possible models is mathematically ineffective. So many strategies, such as pruning, bunching, and chopping are implemented to remove the irrelevant paths and make the process fast. In general, to specific modeling, the general unrestricted model (GUM) has great importance. The procedure of constricting the GUM is following.

There are six straightforward steps in GETS initial in its formation of the general unrestricted model (GUM) (for more Hendry and Krolzig, 2001).

2.2.5.1 Formulation of General Unrestricted Model (GUM)

1. Formulate the GUM based on theory, official knowledge, historical possibilities, data, and dimension information, to ensure that the resulting sample contains the preceding evidence.

2. Select a set of incorrect specification tests (e.g., residual autocorrelation, etc.), their form (e.g. rth-order), and level of significance, Choose the required information criterion (such as BIC) for the final model selection which are mutually encompassing congruent models. Determine the significance level of all tests to ensure the required frequency of rejection under the null hypothesis. Estimate the GUM properly and check the specification of the estimated model to verify the GUM.

3. Start the search for reductions at 5% significance level (a) Includes adding the smallest absolute "t" values until the correct reduction is originated, (b) Eliminate the regressors of largest t values until the critical value of the rest is reached from above. Then remove the insignificant candidate to reduce the search complexity, and now estimate the GUM for upcoming stages.

4. The search for a reduction in more than one path now begins with every possible early deletion. The accuracy of individual reduction is checked to ensure the congruence of the final sample. If entire reduction and diagnostic tests are acceptable, and all other regressors are statistically significant, then this model becomes the last selection, and the search for the further path begins.

5. The significance of every regressor in the last model is evaluated in two sub-samples to check the selection consistency.

Based on general to specific modeling, we start with the General Unrestricted Model (GUM) and then check whether the GUM captures the essential features of the data then eliminates insignificant regressors in terms of the data, to decrease its complexity. Final Model Candidates with final selection should check the accuracy of the reduction through diagnostic tests to ensure the integration of Progressive Research Strategy (PRS).

Main modern developments in theory and practice of automatic model selection include multi-path searches, encompassing choices, impulse saturation, and non-linearity. Autometrics delivers a powerful model-selection procedure: null rejection frequency close to nominal, power close to starting with Local Data Generating Process (LDGP); near unbiased estimates of fit and standard errors.

Theory of Reduction.

2.2.5.2 Tree Search

To understand the tree search methodology, let us assume a GUM with four variables ABCD; there are 4! possible models. The next model in the tree search method is gained by eliminating A, so the next model will be BCD then eliminate B then the model will be CD then eliminate C and the model will be D. the other models from BCD are also constructed on the same pattern.

Next, the new main node will be started sequentially by eliminating B and the model will be ACD, then eliminate A the model will be CD then eliminate C the model will be AD and the last model will AC where D is removed.

As above the next sub-models will be constructed in the same way by eliminating one-by-one variables. The insignificant variables are illustrated with an open circle and have no label like in most right side of figure 2.4 D is an insignificant variable and it has an open circle and this on the third number on the most right-hand side after D and C.

The main point of Autometrics is the whole range of models constructed by the variables in GUM. Every path of this tree is the main model to be estimated for retention or the removal of variables. Inside the regression, each variable is removed or retained in the model based on the t-test value. Autometrics uses the further steps to make the tree search more efficient as described below:

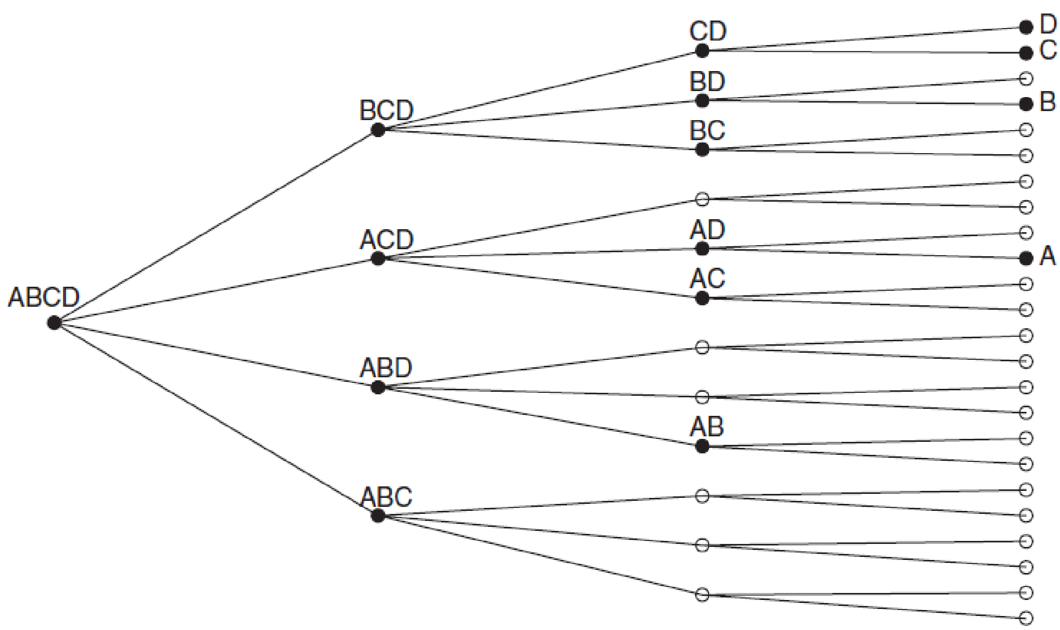


Figure2.4: Searching tree: Entire single models originated from the GUM ABCD.

2.2.5.3 Pruning

By default, in tree search, each reduction removes only one variable. The first rule is that if a reduced model is unable to qualify the diagnostic tests, then sub-nodes from this model will be ignored.

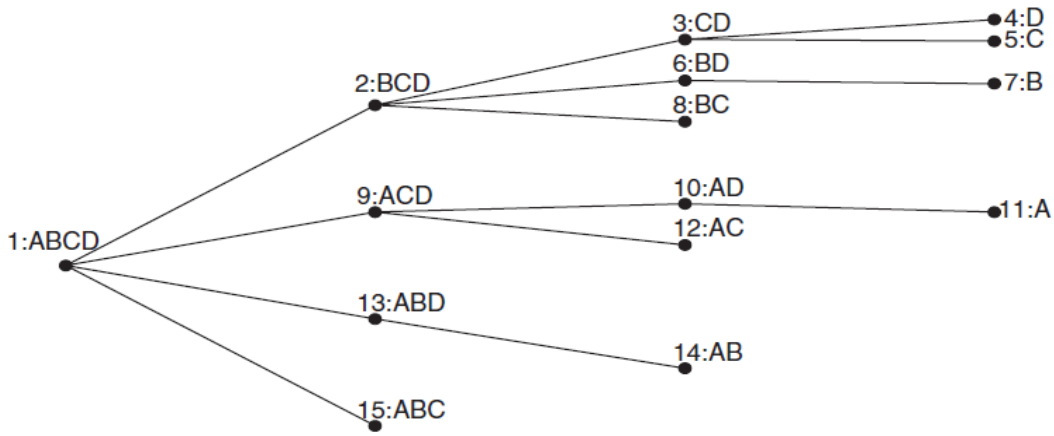


Figure 2.5: The tree Search: All single models originated from GUM ABCD.

If model 3 (CD) in figure 2.5 is unable to qualify the diagnostic tests, then there is no need to consider or test the sub-nodes D and C. Then the next node BD will be considered for retention or removal.

2.2.5.4 Bunching

In bunching, the variables are grouped. Then a group is tested for removal from a process and if a group is deleted then the next group is considered.

2.2.5.5 Chopping

Chopping is a procedure in which a highly insignificant variable is removed permanently and not considered again in the estimation process. And the same procedure of permanent delectation is applied to a group. Chopping saves, computation time and helps the process to fast.

The summary of all the above procedures of Autometrics is given below as a flow chart.

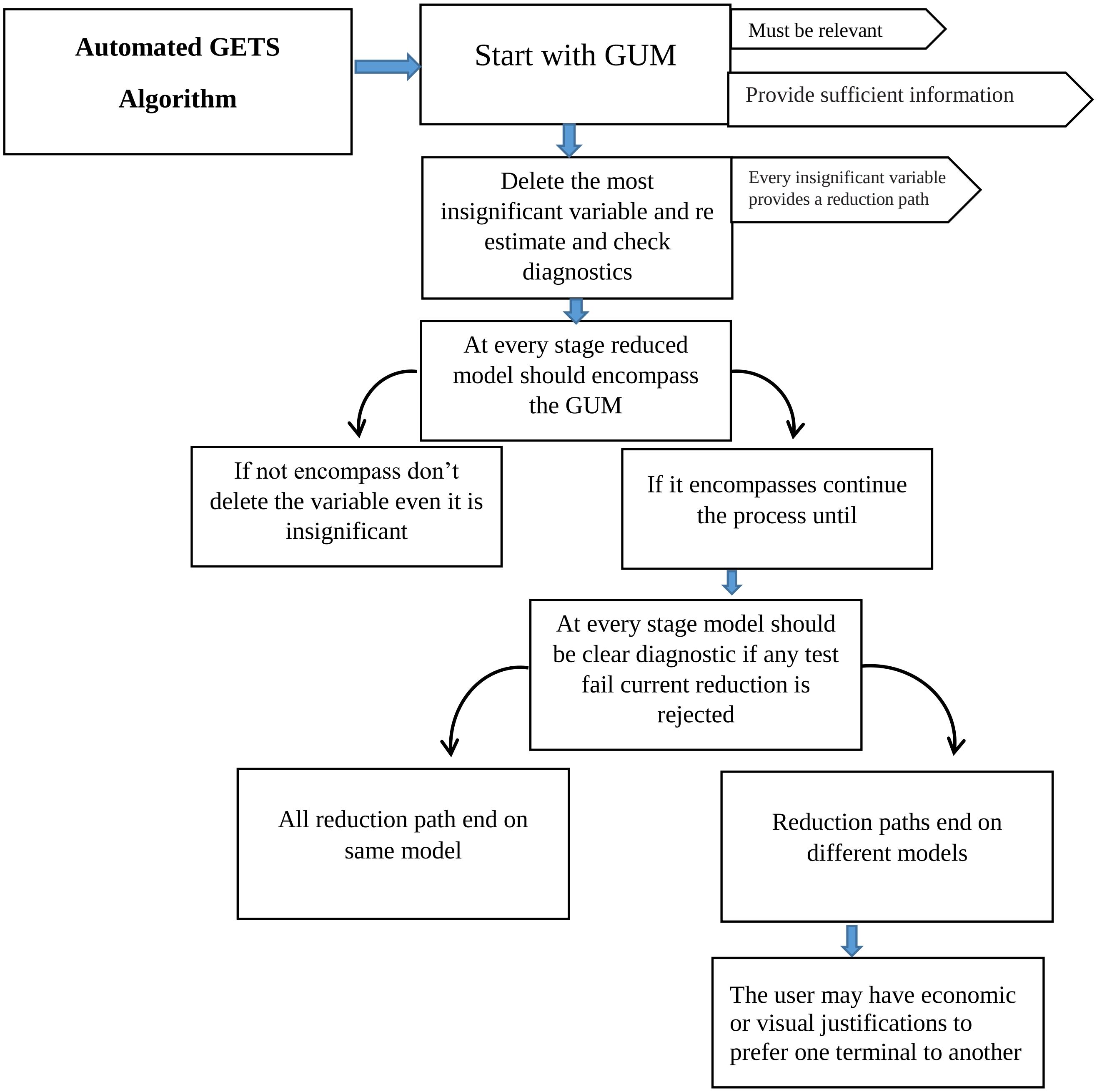


Figure 2.6: Key points of the procedure of Autometrics

The figure shows the procedure of Autometrics for variables selection step by step. The details of every step are given above.

2.2.6 Model selection procedures based on the sensitivity of parameters

We are interested in finding the variables from a set of variables X which are robustly associated with dependent variables Y. For this, we run several regressions in which Y is dependent variables and F is a set of variables that are included in each estimated regression. For each estimated regression another subset D in X is used in the whole process.

X is a set of all variables. Assume F is a set of free variables. Some variables from X are as focused variables to which the researcher is interested to include in the model or removing from the model.

$$y = a_j + \beta_j v + \gamma_j F + \delta_j D_j + \epsilon \tag{2.24}$$

Where j represents the regression model, F is the set of free variables which are used in each estimated model. D_j is the k variables taken from X. The above model is estimated for M number of combinations of $D_j \subset X$. From each estimated model the coefficient $\hat{\beta}_j$ of focused variable v along with $\hat{\sigma}_j$ is estimated. If the coefficient of the focused variable is beyond the extreme bounds, then this variable is not included in the final model. If the coefficient remains within the limits then it will be retained in the final model. This occurs when the focused and doubtful variables are independent. EBA was presented by Edward E. Leamer in 1983, and then enhanced by Clive Granger and Harald Uhlig in 1990.

There are two main types of Extreme Bound Analysis

(i) Leamer's extreme bound analysis

(ii) Sala-i-Martin extreme bound analysis.

2.2.6.1 Leamer's EBA

To identify whether an underlying variable is robust (relevant to the dependent variable) or fragile (irrelevant to the dependent variable). For any underlying variable v, the lower and upper extreme bounds are clear as the minimum and maximum values $\hat{\beta}_j \pm \tau\hat{\sigma}_j$ across the M estimated regression models, where τ is the critical value for a specific level of confidence. The conventional 95-percent confidence level τ will be equal to approximately 1.96. M regressions are run by making different combinations of doubtful variables. If the upper and lower extreme bounds have the same sign, the focus variable v is said to be robust. On the other hand, if the limits have conflicting signs, the variable is declared fragile.

Intuitively Leamer's version of EBA searches numbers of models for lowest to highest values of B_j at some significance level. The Leamer EBA declares a variable either robust or fragile on the same or opposite signs, respectively.

2.2.6.2 Sala-i-Martin's extreme bound analysis

In comeback to the professed inflexibility (declare the variable irrelevant based on only one opposite sign of bound) of EBA that is presented by Leamer, a substitute method by Sala-i-Martin (1997) was introduced, which considered the entire distribution of coefficient of the underlying variable, not just on its extreme bounds. In Sala-i-Martin EBA the binary label, robust or fragile is not given to an underlying variable but it assigns such a confidence level for robustness to the variable under consideration.

Sala-i-Martin introduced two variants of Extreme bound analysis

- Sala-i-Martin EBA that assume a normal distribution
- Sala-i-Martin EBA that does not assume a normal distribution

For the estimation of the normal model, Sala-i-Martin calculates the weighted mean of the coefficient of underlying variables. $\hat{B}_j$ and $\hat{\sigma}_j^2$.

$$\bar{B} = \sum_{j=1}^{M} w_j \hat{\beta}_j \qquad (2.25)$$

$$\bar{\sigma}^2 = \sum_{j=1}^{M} w_j \hat{\sigma}_j \qquad (2.26)$$

Where w_j are the weights that are applied for each estimated model from M models. According to Sala-i-Martin (1997) by using weights a researcher can give more weight to the model that is more likely to be the true model. Assuming the true model is that which has a better fit probability. After calculating the weighted mean of coefficient and its standard error, then Sala-i-Martin calculates CDF (0) on the normal distribution. It is assumed that the regression coefficient follows the normal distribution.

$$\beta \sim N(\bar{\beta}, \bar{\sigma}^2)$$

This will provide a specific value of probability to the variables under consideration.

In the general model, Sala-i-Martin calculates the cumulative density function of coefficients from each regression model separately and combines it into an aggregate CDF(0), then that serves as a measure of the robustness of the variable under consideration. First, the sampling distribution is used of regression coefficient $\hat{B}_j$ to gain the individual coefficient CDF (0), denoted by $\phi_j(0|\hat{\beta}_j, \hat{\sigma}_j^2)$ for every regression model. He then calculates the aggregate CDF (0) for B as the weighted average of all the individual CDF (0)'s:

$$\phi(0) = \sum_{j=1}^{M} w_j \phi_j \, (0|\hat{\beta}_j, \hat{\sigma}_j^2) \tag{2.27}$$

For the normal and the generic model, Sala-i-Martin applies weights that are proportional to the integrated likelihood.

For the combined likelihood to give greater weight to models which gives a better fit:

$$w_j = \frac{L_i}{\sum_{j=1}^{M} L_i} \tag{2.28}$$

The weights could consist of any other portion of the goodness of fit of the model. McFadden's likelihood ratio index (McFadden 1974) is used by Hegre and Sambanis (2006) or can be applied at equal weights to each regression model (Sturm and de Haan 2005; Gassebner et al., 2013).

The table below shows the similarities of the selected variables selection procedure. On these similarities, we make grounds for comparison of selected variables selection procedures.

2.2.7. Similarities among Compared Model Selection Procedures

Similarities among Elastic Net, Autometrics, and Extreme Bound Analysis			
	E-Net	**Autometrics**	**EBA**
Goals	Elastic Net has many goals one of them is Variables Selection	Variables Selection is one of the goals of Autometrics	EBA can be used for Variables Selection
Linearity	Elastic Net can be used for Linear Models	Autometrics can be used for Linear Models	EBA can be used for Linear Models
Base Model	Researchers make GUM usually based on theory. May use the transformation of variables.	GUM: specific by the researcher usually based on theory. May use transforms of the original variables	The base model is specified by the researcher. And it may the general model
Time Series Data	Elastic Net can be applied on Time Series Data	Autometrics can be used for Time Series Data	EBA can be used for Time Series Data
When N<P	Elastic Net can handle when variables are more than observations	Autometrics can handle when variables are more than observations	EBA can handle when variables are more than observations

All variables selection procedures have many goals like feature selection and can be used in time-series data. All of the above-mentioned model selection procedures can be used for variable selection. In our current study, our main objective is variable selection so the property of variable selection of selected procedures; which provides us a common ground for comparison, so we can compare these selected procedures in different situations.

The above model selection procedures can be used for different levels of polynomials as well as for linear models. In the current study, our concern is with linear models and we are using the model selection procedure for linear model variables selection. The base model can be described as the general unrestricted model (GUM). In GUM, all possible variables are put in a single model, and then model selection procedures are applied which select the most relevant variables by using their respective algorithm. Time series data is collected on different units of time, the frequencies are daily, weekly, monthly, and yearly. All of the model selection procedures, which we are comparing, all can be used for time-series data.

Hence, based on the above common grounds of the selected model selection procedures we can make valid comparisons amongst them. Some studies are listed below, in which different model selection procedures are compared by using different DGPs.

2.3 Comparison of model selection criteria

Many studies have been done in different ways based on the comparison of model selection procedures. Most of the studies compared model selection procedures within the same family. Some studies compared model selection procedures across the different classes of model selection procedures.

2.3.1 Comparison of information criteria

Lutkepohl (1985) compared different criteria for selecting intervals (AIC, BIC, FPE, Shibata, etc.) using the Monte Carlo simulation. According to Schwarz's BIC criterion, a small average square in a sample of normally available size leads to a prediction error and it automatically chooses the correct order. Mills and Prasad (1992) compared the selection criteria of models based on information to evaluate their relative performance through Monte Carlo simulation. They examined the relative performance of criteria in several situations, such as distribution of errors, collinearity among regressors, and non-stationary data. They predicted samples for quality assessment and selection of real models. They concluded that the Schwartz Bayesian formation Criteria are consistent, beyond the predictive performance of the sample.

FU (1998) compared different criteria for model selection based on a compression approach. He performed simulation exercises to compare model selection criteria such as Bridge, Ridge, and LASSO. He used the prostate cancer data. According to results, Bridge regression performs better from ridge and LASSO. Kuha (2004) examined the quality of behavior in the selection of good models for observational data, which is based on artificial data and two well-known data sets on social mobility.

Reffalovich et.al. (2008) used the Bayesian Information Criteria (BIC), Adjusted R2, Mallows CP, AIC, AICc, and stepwise regression using Monte Carlo simulation of model selection when true DGPs are known. They found that the ability of these selection methods to add important variables and exclude irrelevant variables increases with the sample size and decreases with the amount of noise in the model. Any model selection procedure does not perform well in small samples, thus, data mining in small samples should be avoided altogether. Instead, the apparent uncertainty in the model's description should be addressed. In large samples, BIC is better than other methods in which they have accurately identified most of the production processes they have developed and performed the same systematically.

2.3.2 Comparison of model selection procedures based on regularized criteria

Wenjiang J. FU, (1998) compared the three model selection procedures; Ridge regression, LASSO, and Bridge regression. He applied all model selection procedures to prostate cancer data. The general finding of this research was the bridge regression performs better than other model selection procedures. The conclusion of the study was for the case of nonlinearity, the bridge model does not always perform the best in estimation and prediction in comparison to other shrinkage models selection procedures the lasso and the ridge.

Meinshausen, (2006) compared the two shrinkage procedures of model selection LASSO and Relaxed LASSO. The results show that the contradicting demands of an efficient computational procedure and fast convergence rates of the 2-loss can be overcome by a two-stage procedure, termed the Relaxed LASSO. For orthogonal designs, the relaxed Lasso provides a continuum of solutions that include both soft- and hard thresholding of estimators. The relaxed Lasso solutions include all regular LASSO solutions and computation of all Relaxed LASSO solutions is often identically expensive as computing all regular LASSO solutions. Theoretical and numerical results demonstrate that the Relaxed LASSO produces sparser models with equal or lower prediction loss than the regular LASSO estimator for high dimensional data.

2.3.3 Comparison across the different classes of variable selection procedures

Pchen et.al (2008) compared model selection criteria (Akaike Information Criterion (AIC), Hannan and Quinn Criterion, BIC, Corrected AIC (AICc), Vector Corrected Kullback Information Criterion (KICvc) and Weighted-Average Information Criterion (WIC).) Using the data number of signals in multiple signal classification (MUSIC) method. They concluded that in a simple MUSIC additive white noise model, for small sample size n, WIC performs nearly as well as AICc and outperforms other criteria, and for moderately large to large n, WIC performs nearly as well as BIC and outperforms other criteria. They also concluded that when the authors are not certain of the relative sample size, WIC may be a practical alternative to any criterion.

Choi and Kurozumi (2008) compared the model selection criteria like Mallows Cp criterion, Akaike AIC, Hurvich and Tsai corrected AIC, and the BIC of Akaike and Schwarz. They run simulations and concluded that the BIC appears to be most successful in reducing mean square error (MSE), and Cp in reducing bias. Another finding was that, in most cases, the selection rules without the restriction that the numbers of the leads and lags be the same have an advantage over those with it.

Wei and Zhou (2010) purposed a modified version of the Akaike information criterion and two modified versions of the Bayesian information criterion. To select the number of principal components and to choose the penalty parameters of penalized splines in a joint model of paired functional data. Numerical results show that, compared with an existing procedure using the cross-validation, the procedure based on the information criteria is computationally much faster while giving similar performance.

Choi and Jeong (2013) analyzed the consistency properties of (AIC), corrected AIC, BIC, and Hannan and Quinn information criteria for factor models. They conducted simulations and concluded that it is difficult to determine which criterion performs better. Ismail et.al (2015) evaluated the forecasting performances of several selected model selection algorithms on air passenger flow data based on Root Mean Square Error (RMSE) and Geometric Root Mean Square Error (GRMSE). The findings of their research show that multiple models selection performed well in one and two-step-ahead forecasts but was outperformed by a single model in three-step ahead forecasts.

2.3.4 Which is the best variable selection procedure

As such, different methods for choosing the relevant set of regressors from a large group of possible determinants have been discussed. After analyzing the literature, we realized that different approaches select distinctive sets of variables. The question arises which is an approach that can produce reliable and consistent results. This is possible by comparing the criteria of variable selection. Several studies in the literature compared different methods of variable selection in different situations. This urges us to evaluate the different variable selection procedures for selecting the variables in different settings of DGPS. Mostly the studies compared the variable selection procedures of the same class and further the comparison is in the narrow sense. Some studies compared the variables selection procedures for orthogonal DGP and some studies compared the variables selection procedure by introducing only one problem in DGP as the correlation among variables, autocorrelation in error terms, or heteroscedasticity in error variance. There are some studies, which compared the Variable selection of different classes in limited situations (limited setting of DGPs as introducing only one problem in DGP) not in a broad sense (introducing many problems in DGPs and introducing multiple problems at a time in DGPs).

2.4 Literature GAP

There are many variable selection procedures, most of which are already discussed in this chapter above. However, our focus is on Autometrics, which takes a general to specific approach, the Elastic Net takes a penalization technique for variable selection, and Extreme bound analysis approaches which select the variables by examining the consistency of their parameters. The performance differs in every scenario and to determine which one performs better is possible only by comparing them in the same settings of DGPs. Therefore, in the current study, we are going to compare the three classes of advanced variables selection procedures[14]. We have taken the latest form of these three classes of variable selection procedures. The EBA, LASSO, and Autometrics all are used in variables selection. However, the procedure for variable selection of all criteria is different from one another. The selection of variables in Autometrics is based on general to specific modeling procedures, LASSO is based on regularized methodology and EBA retains the variable that remains consistent in different experiments. So all the above criteria select the variables with a different approach so it is needed to make a comparison among these procedures in different situations (DGP with correlated variables, DGP with serially correlated variables, DGP with autocorrelated error terms, and DGP with heteroscedastic error terms).

In the literature, we are unable to find the comparison between EBA type of model selection procedures and Autometrics, secondly, we unable to find the comparison between EBA (sensitivity analysis family) and LASSO (shrinkage-based methodology of model selection procedures). We generate different DGPs with correlated variables, autocorrelated error terms, and heteroscedastic error terms and check the performance of variable selection procedures concerning different sample sizes 30, 60, 120,240,480.

[14] Actually these are three main classes of model selection procedures, and we have taken their latest versions.

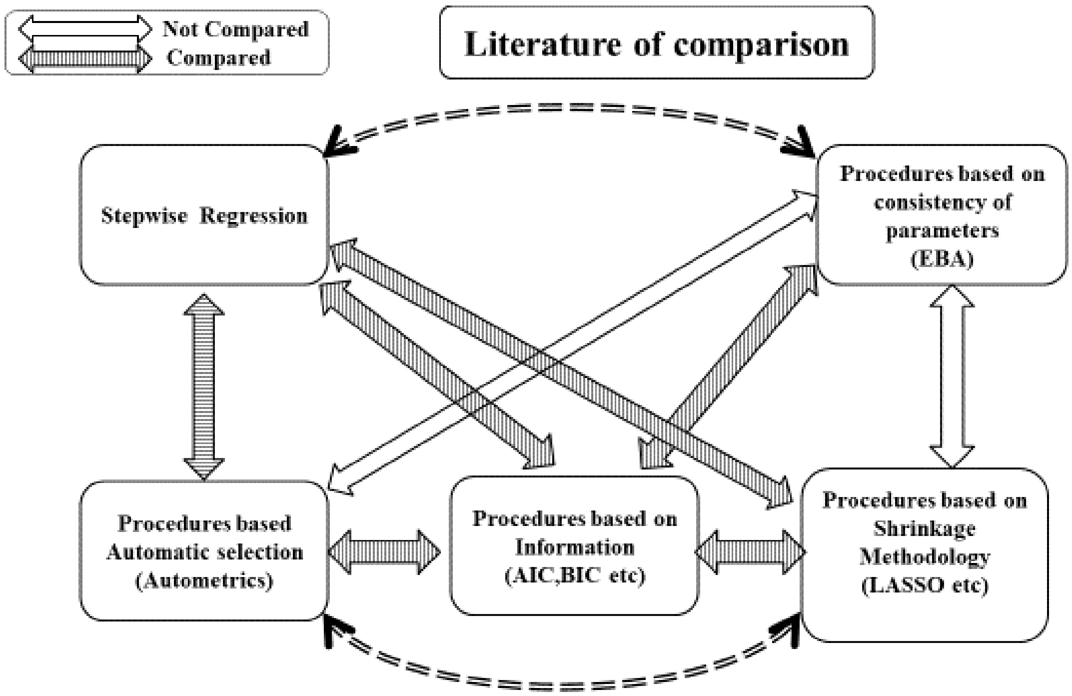


Figure 2.7: The literature gap is outlined below by using the diagram

Figure 2.7 shows the main findings and literature gap for variable selection procedures. In the figure, there are five main classes of variable selection procedures. The dark arrows represent that these classes are already compared in literature but light arrows show that these classes of variables section procedure are not compared in the current settings of DGPs which setting we are going to do in the current study. The figure shows that Extreme bound analysis is not compared with procedures based on regularized based and with the procedures based on automatic methodology.

2.5 Conclusion

Finding the correct specification (relevant variables or true DGP) of the model is a daunting task, but we certainly take a path to find the closest match i.e. the most parsimonious model. Therefore, any method of model selection that can find the closest DGP should be preferred over others in certain situations (different settings of DGPs). There are many methods for variable selection and each method gives a distinct set of variables. To evaluate which set of variables is better, we are resorting to this by comparing the variable selection procedures.

Numerous studies compared the key classes of variable selection procedures. Some studies compare variable selection methods within the same family (like the comparison among the procedures based on information criteria like AIC, BIC) but to the best of our knowledge and based on the literature presented above the family of Extreme Bound Analysis and the family of regularized methods is not compared yet directly. Furthermore, the general to specific methodology (Autometrics) and the methodology of Extreme Bound Analysis are also not compared under changing DGPs (DGP with correlated variables, DGP with serially correlated variables, DGP with autocorrelated error terms, and DGP with heteroscedastic error terms).

So, it is needed to make comparisons among these to fill the gap in the literature of model selection criteria.

CHAPTER 3

METHODOLOGY

We have utilized the Monte Carlo Simulation to find the frequency of correct variables (true DGP) from among the various candidate variables. Monte Carlo Simulation is an essential tool in modern statistical research. Monte Carlo Simulation methods help statisticians to examine the performance of the different statistical procedures in different settings of data generating processes in which mathematical derivations would be difficult. In Monte Carlo Simulation, the distribution is given and GDP is generated according to a given distribution. Then different statistical techniques are applied to already generate GDP and this process is replicated a lot of times like 1000, or 10000 or more than 10000. The statistical procedure that has closets retained DGP to true generated DGP; this will be preferred to other statistical techniques.

The Monte Carlo simulation design consists of two main components 1) Data generating process 2) Simulation design.

3.1 Data Generating Process (DGP)

The question at hand is which method is best at selecting the relevant variables from a large number of candidate variables. The candidate variables may be orthogonal to each other or they may have a specific degree of correlation. The relationship between independent variables and dependent variables may be linear or nonlinear, but here we have taken only a linear relationship. For each scenario, we have DGP matching with assumptions.

Fortunately, we have a model that can generate different scenarios; this model is discussed as follows.

Let

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + ... + \beta_{k1} X_{k1t} + \beta_{(k1+1),1} X_{(k1+1),t} + ... + \beta_k X_{kt} + \varepsilon_t \tag{3.1}$$

$$Y = X\beta + \varepsilon$$

$$
\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_T \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & . & . & X_{(k1),1} & X_{(k1+1),1} & . & . & X_{k1} \\ 1 & X_{12} & X_{22} & . & . & X_{(k1),2} & X_{(k1+1),2} & . & . & X_{k2} \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ 1 & X_{1T} & X_{2T} & . & . & X_{(k1),T} & X_{(k1+1),T} & . & . & X_{kT} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ . \\ \beta_{k1} \\ \beta_{k1+1} \\ . \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_T \end{bmatrix} \tag{3.2}
$$

$$
Y_{(T\times 1)} \qquad X_{(T\times k)} \qquad \boldsymbol{\beta}_{(k\times 1)} \qquad \varepsilon_{(T\times 1)}
$$

We take the linear relationship.

The total number of variables in this DGP is"k". Suppose the first k_1 group of variables in the matrix X is part of DGP of Y, and has non-zero coefficients, and is termed as relevant variables. The other group (k-k_1) variables do not enter into the DGP of Y. Therefore, the corresponding values of $\boldsymbol{\mathcal{B}_i}$ are zero in DGP. X is the matrix of regressors, $\boldsymbol{\mathcal{B}}$ is a set of coefficients from which some are zero. If $\boldsymbol{\mathcal{B}}_i \neq 0$ then the corresponding column of X is determinant of Y and if $\boldsymbol{\mathcal{B}_i = 0}$ then the columns of X are not the determents of Y.

The coefficient $\boldsymbol{\mathcal{B}}_i$ is constructed as.

$$
\boldsymbol{\mathcal{B}}_i = \begin{cases} \frac{2_{k1}}{\sqrt{n}} & if\ i \leq\ k_1 \\ 0 & otherwise \end{cases} \tag{3.3}
$$

$\boldsymbol{\mathcal{B}}$ has two parts first part $\boldsymbol{\mathcal{B}}_1$ and the second part is $\boldsymbol{\mathcal{B}}_2$

$$
\boldsymbol{\mathcal{B}_1} = [\beta_1, \beta_2, \beta_3, \dots, \beta_{k1}] \tag{3.4}
$$

$$
\boldsymbol{\mathcal{B}_2} = [\beta_{k1+1}, \beta_{k1+2}, \beta_{k1+3}, \dots, \beta_k] \tag{3.5}
$$

$$
\boldsymbol{\mathcal{B}} = [\beta_1, \beta_2] \tag{3.6}
$$

The construction of coefficients of relevant variables stated above is now in vector form.

$$
\boldsymbol{\mathcal{B}} = \left[\frac{2k1}{\sqrt{n}}, 0\right] \tag{3.7}
$$

The first part is for relevant variables and zeros for irrelevant variables

The vector of beta coefficients for relevant variables are generated with respect to sample size (as in Krolzig and Hendry, 2001) and beta coefficients for irrelevant variables are a vector of zeros.

If there are k=4 and k1=2, meaning that there are 4 total variables from which 2 are relevant

Then coefficients for all relevant variables:

If N=30:

$2(2)/\sqrt{30}=$ 0.73

$$\boldsymbol{\mathcal{B}} = \left[\frac{2(2)}{\sqrt{30}}, 0\right] \tag{3.8}$$

$$\boldsymbol{\mathcal{B}} = \begin{bmatrix} \beta_0 \\ 0.73 \\ 0.73 \\ 0.00 \\ 0.00 \end{bmatrix} \tag{3.9}$$

β_0 is intercept first, 0.73 and 0.73 are the coefficients of relevant variables, and 0.00 and 0.00 are the coefficients of irrelevant variables.

N=60: $2(2)/\sqrt{60}=$ 0.51

$$\boldsymbol{\mathcal{B}} = \begin{bmatrix} \beta_0 \\ 0.51 \\ 0.51 \\ 0.00 \\ 0.00 \end{bmatrix} \tag{3.10}$$

Error terms are independently and identically distributed with standard normal distribution.

$$\varepsilon = \left(\varepsilon_1, \varepsilon_2, ..., \varepsilon_T\right)' \sim N\left(0, \sigma^2 I_T\right) \tag{3.11}$$

T is for sample size.

X is generated as

The multivariate normal distribution of k dimensional random vector

$$X \sim N_k(\mu, \Sigma)$$

The mean vector of k-dimension

$$\mu = E(X) = (E[X_1], E[X_2], \dots, E[X_k])^T \quad (3.12)$$

Variance covariance matrix

$$\Sigma_{i,j} = E\left[(X_i - \mu_i)\left(X_j - \mu_j\right)\right] = Cov[X_i, X_j] \quad (3.13)$$

The above procedure for generating is a general procedure. Below for each scenario, a separate procedure for DGP is described. First, we take the orthogonal variable case.

3.2 SCENARIO 1:

3.2.1 Orthogonal Regressors

This is the simplest case in which there are no problems in DGP (no autocorrelation in error no heteroscedasticity in error variance and multicollinearity in regressors). After this, we relax the assumptions of DGP one by one and will check the performance of variables selection procedures. In a statistical simulation, orthogonal variables and other no problem (autocorrelation in error terms heteroscedasticity in error variance, no outlier, no structural break, no endogeneity) are an ideal DGP for simulation in case of comparison of different statistical techniques. Therefore, in our simulation work, we have taken the simplest form of DGP for the comparison of variable selection procedures. In this scenario, we find the frequencies of our selected variables selection procedures for retaining the true DGP in several situations. We vary the relevant and irrelevant variables and vary the sample size. For the orthogonal regressors and errors are independent identical distributed. The DGP is known as Krolzig and Hendry (2001).

$$X_{kt} \sim N(\mu, \sigma^2 I)$$

Here "k" is for the number of variables and "t" is time. X is a set of k random variables and μ is a mean vector. Where, Σ is the variance-covariance matrix.

The general form of normal distribution joint density is as,

$$f(x) = (2\pi)^{-2/T} \, |\Sigma|^{-1/2} \, e^{(-1/2)(x-\mu)'\Sigma^{-1}(x-\mu)} \quad (3.14)$$

$$\Sigma = \sigma^2 I_k \quad (3.15)$$

$$R_{ij} = \sigma_{ij} / \left(\sigma_i \sigma_j\right) \tag{3.16}$$

Then,

$$f\left(x\right) = (2\pi)^{-T/2} e^{-\frac{\sigma^2}{2}(x-\mu)^{'}(x-\mu)} \tag{3.17}$$

One can use $\mu = 0$ in which case

$$f\left(x\right) = (2\pi)^{-T/2} e^{-\frac{\sigma^2}{2}x'x} \tag{3.18}$$

$X_1, \ldots, X_k$ are generated through a similar process however the coefficients of the first k_1 group of variables areas.

$$\frac{2_{k1}}{\sqrt{n}},$$

Moreover, the coefficients of the last group ($k - k_1$) variables are zero.

If all variables are orthogonal then

$$\sigma_{ij} = 0 \; for \, i \neq j \tag{3.19}$$

3.3 SCENARIO 2:

3.3.1 Relevant Variables are non-orthogonal

In this scenario, we generate the DGP with correlated variables. In real-world models mostly the regressors are non-orthogonal, either the correlation may weak or it may be strong. So, in simulation studies, it is highly important to incorporate the situation in which the regressors are non-orthogonal. In the current scenario, we analyzed the different levels of correlation among independent variables and have checked the performance of variables selection procedures. If a variable selection procedure remains consistent to retain the true DGP with increasing the level of correlation among regressors then is it superior to other variables selection procedures.

Here we create correlation in relevant variables (as in Khalaf & Shukur, 2006). Same procedure as in 3.1 except.

$$X_{it} \sim N\left[\mu, \begin{pmatrix} \Sigma & 0 \\ 0 & I_{k-k1} \end{pmatrix}\right]$$

Where Σ is the non-zero diagonal variance-covariance matrix. This GDP represents that the relevant variables are non-orthogonal.

3.4 SCENARIO 3:

3.4.1 Relevant Variables are correlated to Irrelevant Variables.

In real-world data, the economic series are correlated. The correlation may be either weak or strong. Statistical analysis of variables selection must analyze the situation in which the regressors are correlated (problem of multi-collinearity). It may be possible that two candidate variables are correlated from which one is relevant and one irrelevant. If there is a weak level of multi-collinearity, then there is no problem. If there is a high level of multi-collinearity, then there may be a problem because the chance of selection of irrelevant variables that are highly correlated with relevant variables increases. So in this situation, the model selection criteria may select non-DGP[15] variables which are correlated to GDP variables[16].

We will vary the collinearity between relevant and irrelevant variables.

In multivariate cases, we can write as

$$\begin{pmatrix} X_t \\ Z_t \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} I_{k1} & \Sigma \\ \Sigma & I_{(k-k1)} \end{pmatrix}\right]$$

Where X variables are relevant and Z are irrelevant variables.

We will assign the values of coefficients of correlation 0 to 0.9 to analyze the different levels of multi-collinearity. As the coefficient goes up, the level of multi-collinearity also goes up and it is expected that the frequency to be selected of correlated irrelevant variables will also increase.

We will check the gauge (selection of irrelevant variables) for different model selection procedures and selection of irrelevant variables, which are correlated with relevant variables. The correlation of irrelevant variables to relevant variables starts from zero-mean they are orthogonal, like 0.1, 0.2, …0.9. In this 0.1 means, a low level of multi-collinearity in the variables, and 0.9 means a high level of multi-collinearity among variables.

[15] Non-DGP variables =irrelevant variables

[16] DGP variables=relevant variables

3.5 SCENARIO 4:

3.5.1 Relevant variables are serially correlated

The economic series in a real-world application depends on their previous history. So it is important to analyze the case in which the regressors are serially correlated. So in this scenario, we discuss this issue. We generate the regressors that are serially correlated and then apply the different model selection criteria and note the performance of different model selection procedures. We vary the covariates and their serial correlation to check the robustness of different model selection procedures under this problem.

3.5.2 The DGP of serially correlated variables

We describe the X matrix for k_1 DGP variables and $(k-k_1)$ are Non-DGP variables.

The relevant variables are;

$$X_{r,t} = \vartheta X_{r,t-1} + \varepsilon_t \tag{3.20}$$

Here "r" is for regressors and "t" is time. r=1,2,3,..,k and $t = 1, 2, 3, \ldots T$

$$\text{where} \qquad \varepsilon_t \sim N\left(0, \sigma^2\right)$$

A variation in the value ϑ tells about serial correlation in regressors. When $\vartheta = 0$ the regressors are serially independent and $\vartheta = .25$ there is a weak serial correlation in the regressors. When $\vartheta = .90$ there is a high serial correlation in the regressors.

3.6 SCENARIO 5:

3.6.1 Performance of Model Selection Procedures in case Moving average in error terms

There is numerous situation in economics when we fail to estimate the model which is correctly specified. In many cases, the model is miss specified. The model's misspecification creates the problems of autocorrelation in residuals, which is problematic. In many cases there exist measurement error in the variables and this creates the problem of autocorrelation in residuals.

The same procedure is discussed in 3.1 besides the following.

The moving average correlates error terms.

$$Y = X\beta + \varepsilon \tag{3.21}$$

$$\varepsilon_t = \emptyset \varepsilon_{t-1} + v_t \tag{3.22}$$

We vary the $\emptyset$ to check the performance of different variables selection procedures for a different level of autocorrelation in error terms.

and

$$v_t \sim N(0,1)$$

3.6.2 The auto correlated error terms

In the usual setting, it is assumed that errors are homoscedastic and serially independent but errors can be serially correlated.

$$\mathrm{E}[\varepsilon'\varepsilon \mid \mathrm{X}] = \sigma^2\Omega \tag{3.23}$$

$\sigma^2\Omega$ *is positive definite with constant*

$$\sigma^2 = \mathrm{Var}\left[\varepsilon_t | \mathrm{X}\right] \text{ on diagnoal} \tag{3.24}$$

Suppose that

$$\Omega_{ts} \text{ is function of} \left| t - s \right|$$

$$\frac{\gamma_s}{\gamma_0} = \rho_s = \rho_{-s} \tag{3.25}$$

$$\rho_{ts} = \frac{\gamma |t-s|}{\gamma_0} \tag{3.26}$$

$$\mathrm{Var}[\varepsilon_t] = \frac{\sigma_u^2}{1-\rho^2} = \sigma_\varepsilon^2 \tag{3.27}$$

$$\sigma^2\Omega = \frac{\sigma_u^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & . & . & \rho^{T-1} \\ \rho & 1 & \rho & \rho^2 & . & . & \rho^{T-2} \\ \rho^2 & \rho & 1 & \rho & . & . & \rho^{T-3} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & \rho \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & . & . & \rho & 1 \end{bmatrix} \tag{3.28}$$

We take only first-order autocorrelation(s=1). We vary the strength of first-order autocorrelation like 0.1, 0.2,...,0.9 and will check the performance of different model selection procedures. With changing the level of autocorrelation, we change the sample size to check the robustness of variable selection procedures in different levels of autocorrelation as well for different sample sizes.

3.7 SCENARIO 6

3.7.1 Heteroscedasticity in error variance

Heteroscedasticity has serious consequences for the OLS estimators. Although the OLS estimators remain unbiased, the estimated standard errors (SE) are wrong. Because of this, confidence intervals and hypotheses tests cannot be relied on. In addition, OLS estimators are no longer the best linear unbiased estimators (BLUE). If the form of heteroscedasticity is known, it can be corrected (via appropriate transformation of the data).

While heteroscedasticity does not cause bias in the coefficient estimates, it does make them less precise. Lower precision increases the likelihood that the coefficient estimates are further far from the correct population parameter value.

Heteroscedasticity tends to produce p-values that are smaller than they should be. This effect occurs because heteroscedasticity increases the variance of the coefficient estimates but the OLS procedure does not detect this increase. Consequently, OLS calculates the t-values and F-values using an underestimated amount of variance. This problem can lead you to conclude that a model term is statistically significant when it is not significant actually. The procedure for DGP is the same as discussed in 3.1 besides the following.

Error terms are generated as heteroscedastic.

Consider the equation

$$Y = X\beta + \varepsilon \tag{3.29}$$

$$\varepsilon_t \sim N\left(0, \sigma_t\right)$$

$$\sigma_t = \begin{cases} \sigma_t = \sigma_1 & \text{if } t < T_o \\ \sigma_2 & \text{otherwise} \end{cases} \tag{3.30}$$

Here T_o is the first part of the sample and for this sample the variance of the error term is σ_1. After T_o the second part of the sample starts and σ_2 is error variance for the second sample.
To change the intensity of heteroscedasticity we vary the variance of error as follows (Castle & Doornik, 2011; Khalaf & Shukur, 2006).

$$\frac{\sigma_1}{\sigma_2} = \mathrm{R} \tag{3.31}$$

$$\mathrm{R} = 0.1, 0.2, \ldots, 0.9$$

Sigma 1 is the first part error variance and Sigma 2 is the second part error variance.

We change the variance of both parts of the error terms and observed the performance of different model selection criteria. After generating the DGP we apply the model selection criteria. The frequencies of retaining the DGP are given in chapter 4.

3.8 SCENARIO 7:

3.8.1 Multiple problems in DGP

In this scenario, we introduced multiple problems in DGP at a time because in real-world models there are more than one problems that can occur at a time. In previous scenarios, we introduced only one problem in DGP at a time but now we introduce more than one problem (DGP with correlated variables, with auto correlated error terms, and with heteroscedastic error terms) in DGP simultaneously.

Simulation Design

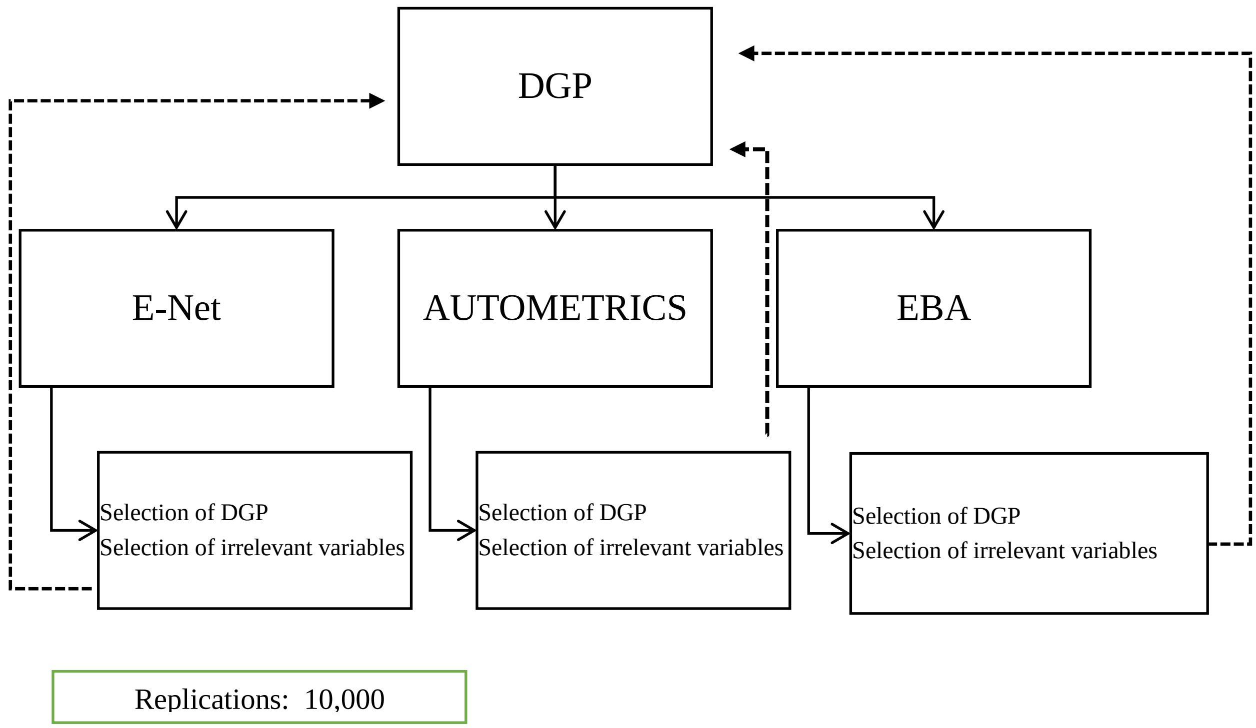


Figure 3.1: The process of simulation

The figure shows the process of simulation design. In simulation design, first, we have DGP then different techniques of variables selection procedure are applied (here we applied three variables selection procedures Elastic Net, Autometrics and Extreme bound analysis). Three variable selection procedures try to retain true DGP. The selection of the DGP (relevant variables) is called the power of the variable selection procedures, and the selection of irrelevant variables is called the size of the variable selection procedures. This process is replicated 10000 times.

3.9 Comparing the performance of variable selection procedures

Suppose

$$Y = X_{k1}\beta_1 + X_{k-k1}\beta_2 + \varepsilon \tag{3.32}$$

Y is the dependent variable X_{k1} is the set of relevant variables and $X_{k-k1}$is the set of irrelevant variables $B_1$is set of coefficients for relevant variables and B_2 is the set of coefficients for irrelevant variables. X_{k1} is a set of relevant variables of DGP and model selection procedure which detects this DGP with high frequency will be superior to other model selection procedures. X_{k-k1} is a set of irrelevant variables and the model which detects all relevant variables and does not detect any irrelevant variable will be superior to other model selection procedures. So the process will be replicated a large number of times and the first criteria used to compare the model selection procedures would be relative frequency to detect the DGP. However, we do not expect all variable selection procedures to detect the DGP without any irrelevant variable with high frequency so the next criteria for comparison of variables selection procedures will be the detection of DGP+ one irrelevant variable, further the criteria will be DGP+ 2 irrelevant variables detection.

We can find the frequencies as; let us find the frequency of x1 variable in percentage.

$$Rf(X) = \left(\frac{n(X)}{N}\right) \times 100 \tag{3.33}$$

Rf=Relative frequency of X, n(X) =number of times the X is retained in the estimated model, N=Total number of replications

10% means that “X” is retained in the estimated model 10 times out of 100 replications. One the same way if we put all relevant variables in place of X this will give us the frequency of true DGP (all relevant variables).

3.10 Frequency of DGP detection with respect to sample size

Model selection procedures may have asymptotic properties so it is interesting to check their performance for different sample sizes. As the size of the sample increases, it is expected that the variable selection procedures increase their DGP detection frequency.

CHAPTER 4

RESULTS AND DISCUSSION

In this chapter, the results of the Monte Carlo Simulation are summarized in each setting of DGP. To gain the first objective (to check the retention frequency true DGP) in the first scenario, we have calculated the relative frequencies of DGPs for different sample sizes for orthogonal and non-orthogonal variables. We have analyzed the case of DGPs with a different number of relevant and irrelevant variables. In the first, we have taken the simple case in which we take six candidate variables in which there are 3 relevant and 3 irrelevant variables. In the next section, we have taken ten candidate variables, five are relevant and five are irrelevant variables. Lastly, we have taken a more general form of DGP in which we take fifteen candidate variables in which five are relevant and ten are irrelevant. Further, we change the number of irrelevant variables in DGP and calculate the probability of irrelevant variables.

In the next scenario, to gain the sub-objective of "to check the performance of variable selection procedures in case of serial correlation in regressors" we introduce the serial correlation in the regressors. Also, change the level of serial correlation in the regressors and calculate the frequencies of DGP for the different sample sizes.

To gain the objective of "to check the size of variables selection (selection frequency of irrelevant variables) of different variables selection procedures in case of irrelevant variables are correlated to relevant variables" we have generated the DGP in which irrelevant variables are correlated with relevant variables. Moreover, apply different variables selection on this DGP and calculate the frequencies of irrelevant variables.

One of the objectives is to check the performance of variables selection procedures in case of auto correlated error terms. For this objective, we generate the DGP with auto correlated error terms, apply the variables section procedures, and calculate their power to retain the DGP. In the next scenario, we introduce heteroscedasticity in the variance of error terms and check the frequency of true DGP. In the last scenario, we introduce multiple problems (autocorrelation in error terms, heteroscedasticity in error variance, and correlation in regressors) at a time and calculate the frequency of true DGP.

A detailed discussion on all the findings are given below:

4.1 SCENARIO 1:

All variables are orthogonal

As described in the objectives, the first objective of the study is to check the power of different variables selection procedures. In this scenario, we take the DGP with orthogonal variables, the variance of error terms is homoscedastic and there is no autocorrelation in error terms. In this scenario, we find the frequency of selecting the correct variables (true DGP) in several situations like increasing the number of relevant and irrelevant variables. Here we vary the relevant variables, irrelevant variables, and sample size to compare the potency (power to find correct variables) of different model selection procedures. The results of six candidate variables (from which three are relevant and three are irrelevant) cases are put in appendix 2. Moreover, the results of 10 and 15 candidate variables are given below.

4.1.1 Relevant Variables=5; Irrelevant Variables=5

Here the total number of candidate variables is 10 of which 5 are relevant and five are irrelevant. Here we put an equal number of relevant and irrelevant variables. In the next section, we increase the candidate variables by increasing the irrelevant variables in the model. A detailed description of DGP is given in section 3.1.

Table 4.1: Relative Frequencies of DGP Orthogonal Variables (%)

	E-Net					Autometrics					EBA				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
DGP	10	13	15	15	16	**40**	**44**	**47**	**48**	**50**	0	10	18	19	19
DGP+1_IRV	17	18	25	26	26	**21**	**23**	**25**	**25**	**29**	4	12	13	17	22
DGP+2_IRVs	24	25	26	27	28	0	1	2	3	5	0	4	6	11	20
DGP+3_IRVs	16	17	17	18	19	0	0	0	0	2	0	0	0	2	5
DGP+4_IRVs	2	3	4	5	6	0	0	0	0	0	0	0	0	0	1
DGP+5_IRVs	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Total Frequency	69	76	88	92	**96**	61	68	74	76	86	4	26	37	49	67

IRV=Irrelevant variable

In the first row, the relative frequencies of true DGP (relevant variables=5 and irrelevant variables=0) are given. The table shows that Autometrics has the highest frequencies of detecting the true model without adding any extra variable. For all sample sizes, the frequencies of retaining the true model are above 40% and it increases with sample size. Similarly, Autometrics has high frequencies of retaining the true model with one extra variable. The frequencies of retaining the true DGP with one extra variable are between 20% to 30%. Combining the above two models (True model with no extra variable and model with one extra variable) the frequencies are between 60% to 80%.

The frequencies of the true model or the model with no extra variable the performance of EBA and Elastic Net is poorer, and frequencies lie between 4% to 26%.

The table also indicates that the frequencies of retaining the true model in the case of all variable selection procedures lie between 0 to 50 percent. In the last row total frequencies are given, here Elastic Net has the highest total frequency for 480 sample size 96%.

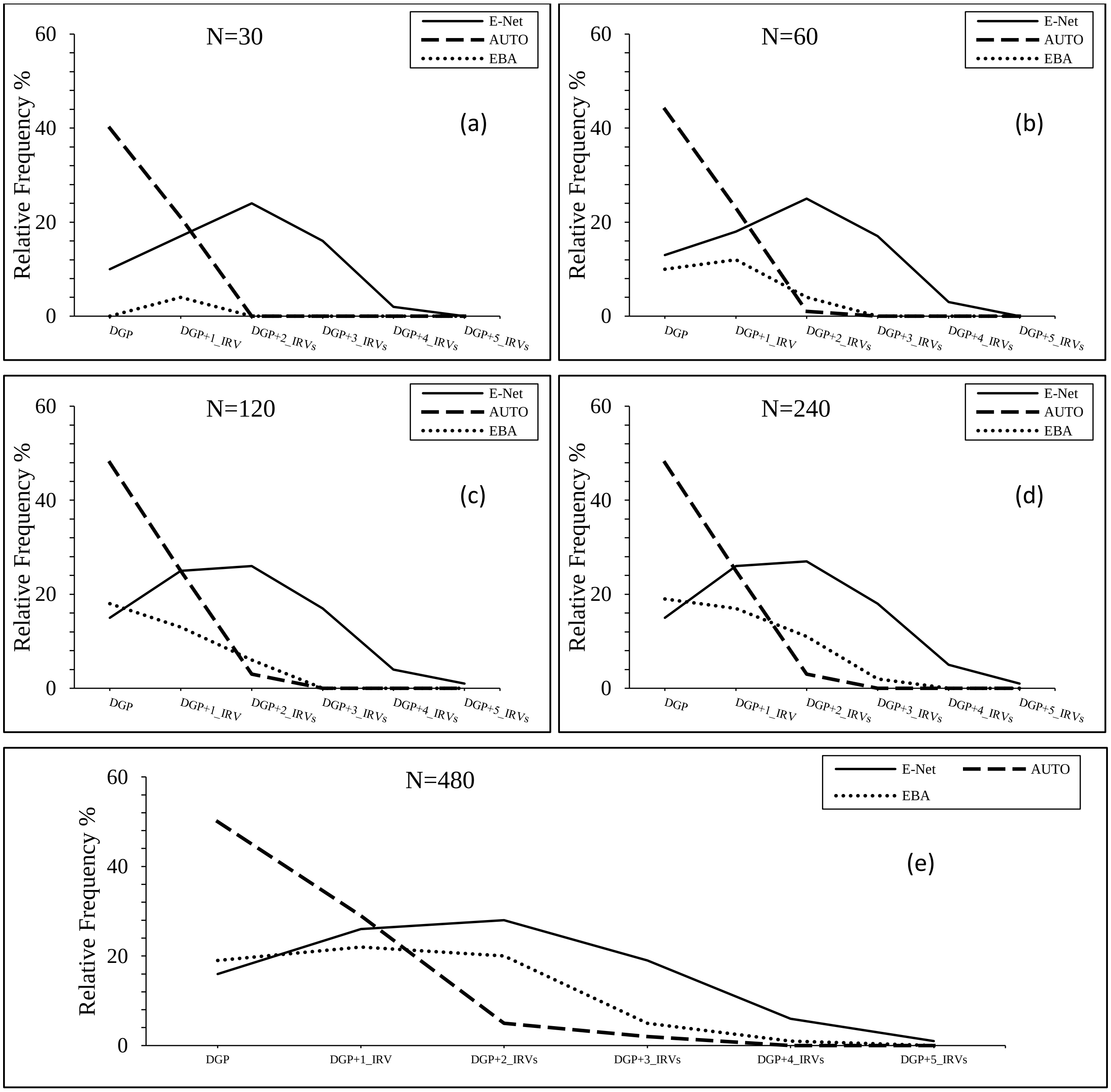


Figure 4.1: Relative Frequencies of DGP Orthogonal Variables (%)

Figure 4.1 represents the frequencies of the true model with different no. of irrelevant variables and for different sample sizes. Panel (a) shows the frequencies of DGP of different model selection procedures in case of 30 sample size, panel (b) shows the frequencies for 60 sample size, panel (c) is for 120 sample size, (d) is for 240 sample size and (e) is 480 sample size. In all sample sizes, Autometrics has higher frequencies for the true DGP and EBA has lower frequencies for the true DGP. The E-Net has a higher frequency in the case of DGP plus one irrelevant variable retained in the model.

4.1.2: Frequencies of Irrelevant Variables

Table 4.2: Relative Frequencies of Irrelevant Orthogonal variables (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
IRVs=1	**71**	**74**	**78**	**80**	**81**	16	20	20	23	36	55	58	60	63	66
IRVs=2	49	50	50	52	53	1	3	4	4	7	13	16	16	23	34
IRVs=3	20	21	21	23	25	0	0	0	1	2	3	4	6	8	13
IRVs=4	2	2	3	5	7	0	0	0	0	0	0	0	0	0	0
IRVs=5	0	1	1	2	3	0	0	0	0	0	0	0	0	0	0

In table 4.2 and figure 4.2 the frequencies of irrelevant variables are given. The results show that Elastic Net has higher frequencies to retain the one irrelevant variable for all sample sizes and frequencies are between 71 to 81 percent. In the case of two irrelevant variables, the Elastic Net has also higher frequencies for all sample sizes. There is an increase in relative frequencies with an increase in sample size and there is a decrease in frequencies with an increase in the number of irrelevant variables retained by different variable selection procedures. As a whole, the E-Net has a higher frequency to retain irrelevant variables and Autometrics has the lowest frequency to retain irrelevant variables.

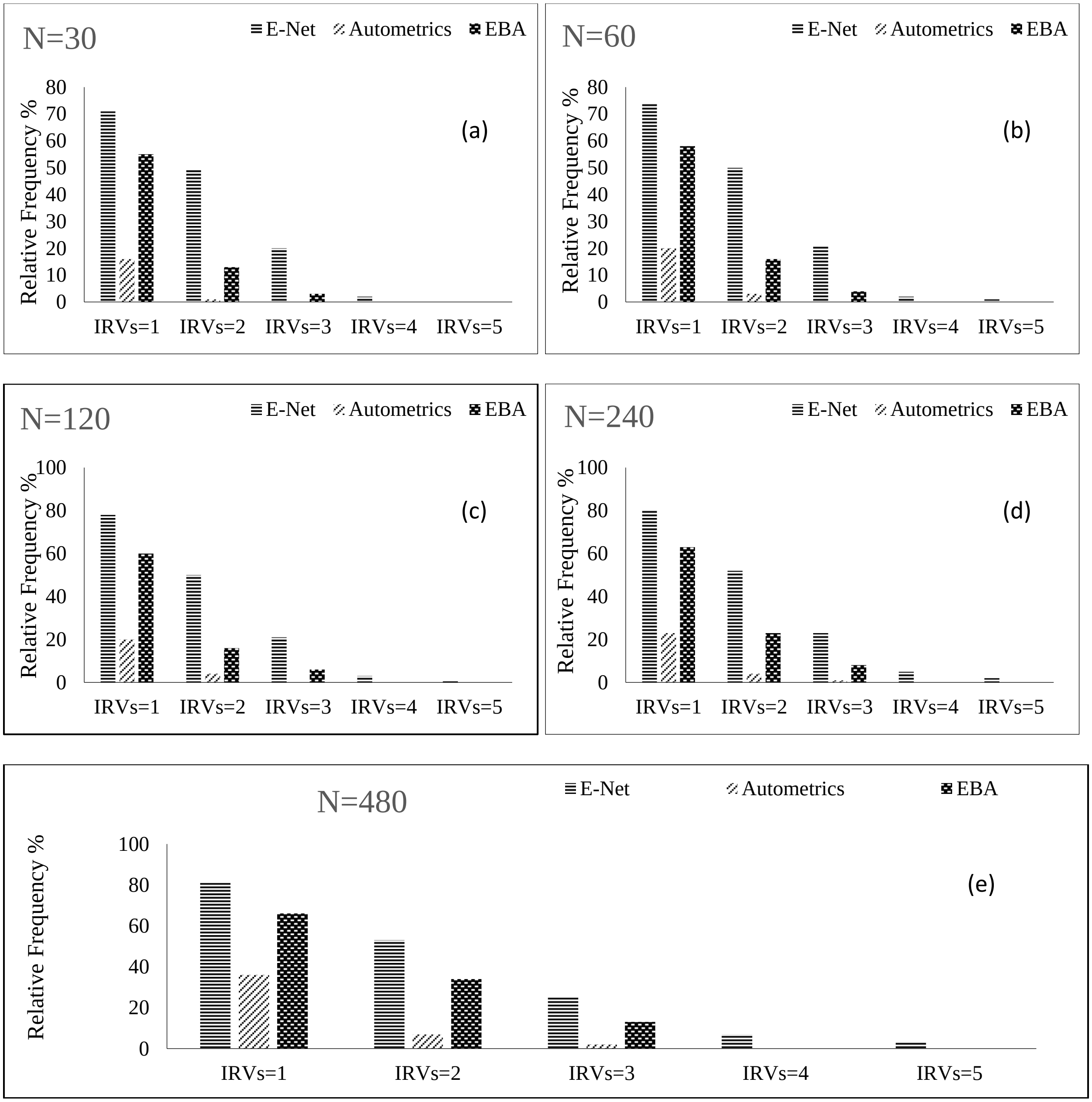


Figure 4.2: Relative Frequencies of Irrelevant Orthogonal variables (%)

Total candidate variables=15

4.1.3 Relevant Variables=5; Irrelevant Variables=10

Here the candidate variables are fifteen and relevant variables are five. All variables are orthogonal to each other. The complete description of DGP is given in section 3.1.

Table 4.3: Relative Frequencies of DGP Orthogonal Variables (%)

	E-Net					Autometrics					EBA				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
DGP	1	1	3	4	6	**30**	**34**	**35**	**40**	**43**	**0**	**0**	**0**	**0**	**0**
DGP+1_IRV	3	3	5	6	7	22	24	25	29	30	0	0	0	0	0
DGP+2_IRVs	19	21	23	24	27	5	8	10	13	14	0	0	0	0	0
DGP+3_IRVs	15	15	17	19	23	2	4	3	7	8	1	1	3	6	9
DGP+4_IRVs	8	10	11	13	15	1	1	2	3	3	8	10	14	15	16
DGP+5_IRVs	4	6	7	8	10	0	0	0	0	0	9	12	17	19	19
DGP+6_IRVs	4	5	5	7	7	0	0	0	0	0	7	11	17	18	19
DGP+7_IRVs	0	0	1	1	2	0	0	0	0	0	2	3	5	8	12
DGP+8_IRVs	0	0	0	0	0	0	0	0	0	0	1	2	4	6	8
DGP+9_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DGP+10_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total Frequency	**54**	**61**	**72**	**82**	**97**	**60**	**71**	**75**	**92**	**98**	**28**	**39**	**60**	**72**	**83**

Here we take 15 candidate variables in the model from which five are relevant and ten are irrelevant. In the first row of table (4.3), the frequencies of true DGP (all relevant variables and zero irrelevant variables) are given for different sample sizes. The results show that Autometrics has higher relative frequencies to retain the true DGP than the other two variables selection procedures. The frequencies of Autometrics lie between 30 to 43 percent. In the next row, the frequencies of DGP with one extra irrelevant variable are given, again the Autometrics has higher frequencies and lies between 22 to 30 percent. By combining the above frequencies these lie between 42 to 73 percent and this is higher than other variable selection procedures.

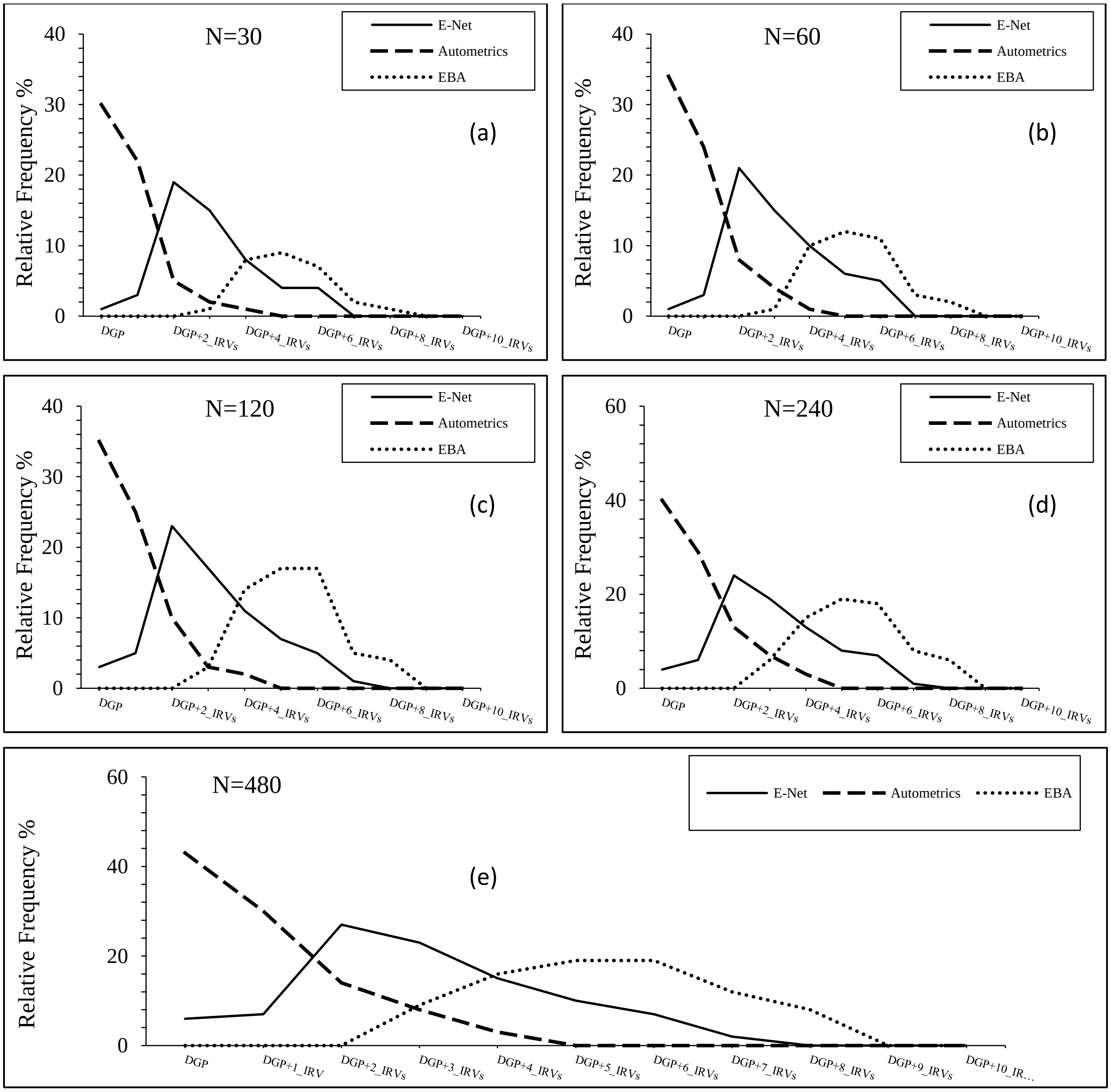


Figure 4.3: Relative Frequencies of DGP Orthogonal Variables (%)

The Elastic Net has a very low relative frequency to retain the true DGP and EBA is unable to retain the true DGP. Then in the next rows, the relative frequencies of DGP with some irrelevant variables are given (like DGP+1 irrelevant variable, DGP+2 irrelevant variables). The results show that Autometrics has higher relative frequencies in the first two rows, and then Elastic Net has higher frequencies in 3rd and fourth row. This shows that Autometrics has more power to retain the DGP than other variables selection procedures. Elastic net retains irrelevant variables in the estimated model. EBA performs poorly in this situation because it is unable to find the DGP but has a low relative frequency to retain the DGP with many irrelevant variables in the estimated model. In the last row aggregate of relative frequencies of all variable selection, procedures are given for all sample sizes. Here Autometrics has the highest total frequency for DGP that 98% for a 480-sample size.

4.2 SCENARIO 2:

Relevant Variables are Non-Orthogonal

In this scenario, we generate the DGP with non-orthogonal (there exist multi-collinearity) variables. In real-world models, the regressors are correlated to each other, either the correlation may weak or it may be strong. So, in simulation studies, it is important to incorporate the situation in which the regressors are correlated. In the current scenario, we analyzed the various levels of correlation among independent variables, and check the performance of selected variables selection procedures to retain the true DGP. If some variables selection procedure has consistent with increasing the level of correlation among regressors then is it superior to other variables selection procedures.

Table 4.4: Relative Frequencies of DGP Non-Orthogonal Variables (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / Σ	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
Σ =0	15	17	20	20	22	**47**	**52**	**56**	**60**	**65**	8	11	17	19	20
Σ =0.1	14	16	25	27	28	**49**	**50**	**54**	**55**	**55**	16	33	42	38	39
Σ =0.25	23	24	25	27	29	39	41	42	45	48	**45**	**47**	**51**	**52**	**53**
Σ =0.5	32	34	35	40	48	31	35	39	40	40	**58**	**60**	**61**	**62**	**64**
Σ =0.75	33	35	36	37	39	27	29	30	30	31	**68**	**70**	**72**	**76**	**76**
Σ =0.9	29	30	31	32	34	14	15	16	16	17	**66**	**67**	**69**	**70**	**74**

Σ=level of correlation

In this scenario, the relevant variables are non-orthogonal to each other and have different levels of correlation among them, which means there exists multi-collinearity of different levels. In the first case, we take zero levels of multi-collinearity, then the only low level of multi-collinearity (Σ =0.1) and then a high level of multi-collinearity (Σ =0.9). In all levels of multi-collinearity, we calculate the frequencies of the true GDP (all relevant variables and zero irrelevant variables).

In the first row, there is no correlation among variables and here the Autometrics has higher frequencies for all sample sizes the frequencies lie between 47 to 65 percent. In the second row, there is a very low level of multi-collinearity and Autometrics has higher frequencies for all sample sizes. The level of multi-collinearity is increased (Σ =0.3) and EBA has higher frequencies for retention of the true DGP, and these lie between 45 to 53 percent.

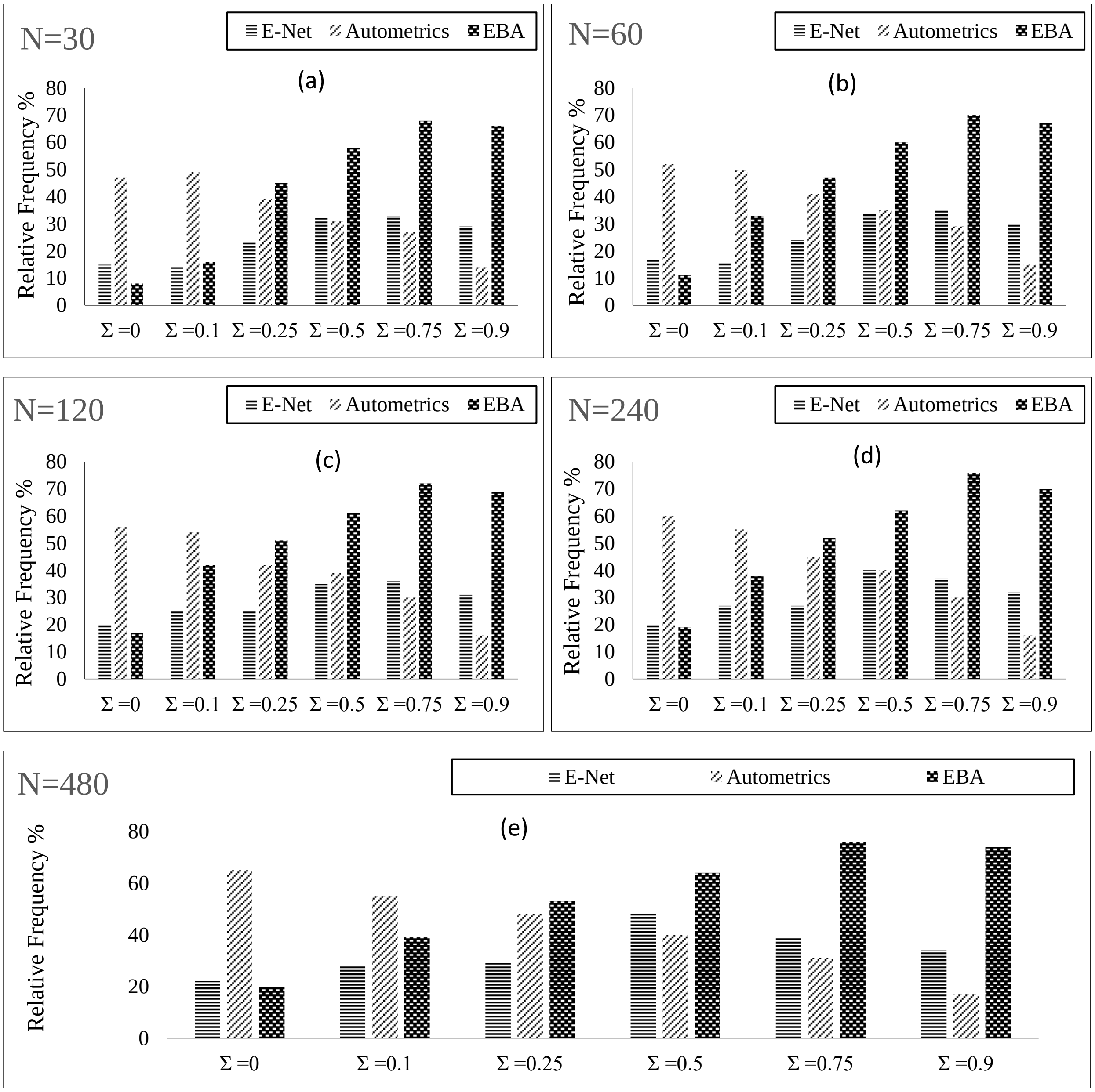


Figure 4.4: Relative Frequencies of DGP Non-Orthogonal Variables (%)

In the following rows we increase the level of multi-collinearity and for all the EBA has higher frequencies than other model selection procedures. In the last row, a high level of multi-collinearity is analyzed and here EBA performs better than other variables selection procedures to retain the true GDP, and frequencies lie between 66 to 74 percent for different sample sizes. It is also noticeable that mostly with increasing the sample size the frequencies are increasing in all cases. For results that are more detailed see appendix 2.

4.3 SCENARIO 3:

Irrelevant Variables are Correlated to Relevant Variable

In real-world data, the economic series are correlated to each other. It may be possible that two variables in the model are correlated, between which one is relevant and one is irrelevant to the dependent variable. The correlation may be either weak or strong. If there is a weak level of multi-collinearity, then there is no problem. If there is a high level of multi-collinearity, then there may be some problems because the chance of selecting irrelevant variables that are highly correlated with relevant variables increases. Statistical analysis of variables selection must analyze the situation in which the regressors are correlated (problem of multicollinearity). Therefore, in this situation, the model selection criteria may select irrelevant variables correlated to relevant variables. The complete description of DGP in which irrelevant variables are correlated with relevant variables is given in section 3.3. The frequencies of irrelevant variables, which are correlated, with relevant variables are given below.

Table 4.5: Frequencies of Irrelevant variables which are correlated to relevant variables (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
								Σ *=0							
IRVs=1	**70**	**79**	**80**	**82**	**83**	24	30	31	32	32	46	50	61	67	69
IRVs=2	21	25	47	79	80	4	4	5	8	9	14	19	22	23	25
IRVs=3	19	20	21	24	25	2	3	5	6	7	1	1	3	2	5
IRVs=4	7	8	9	9	9	0	0	0	0	0	0	0	0	0	0
IRVs=5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
								Σ =0.25							
IRVs=1	**96**	**97**	**98**	**99**	**99**	23	25	25	26	26	60	70	89	96	100
IRVs=2	77	80	83	81	80	2	3	4	19	20	19	40	54	66	80
IRVs=3	36	44	47	51	52	2	0	0	0	0	2	11	14	15	43
IRVs=4	9	9	10	20	21	0	0	0	0	0	0	0	2	7	10
IRVs=5	0	0	0	1	4	0	0	0	0	0	0	0	0	0	1
								Σ =0.5							
IRVs=1	**95**	**96**	**97**	**99**	**99**	16	20	25	26	30	54	76	80	96	99

IRVs=2	75	76	77	80	80	1	4	4	5	5	12	30	38	50	60
IRVs=3	20	29	31	33	35	0	0	0	1	1	2	3	7	19	23
IRVs=4	9	10	11	12	14	0	0	0	0	0	0	0	0	0	4
IRVs=5	0	0	1	1	3	0	0	0	0	0	0	0	0	0	0
Σ =0.75															
IRVs=1	**92**	**93**	**96**	**97**	**97**	20	23	24	25	26	50	80	87	96	97
IRVs=2	60	69	71	73	78	4	4	5	6	6	14	30	37	47	56
IRVs=3	33	36	36	37	40	0	1	1	3	3	3	12	13	14	17
IRVs=4	4	9	11	13	15	0	0	0	0	0	0	0	0	0	2
IRVs=5	0	0	1	2	3	0	0	0	0	0	0	0	0	0	0
Σ =0.9															
IRVs=1	**96**	**97**	**97**	**98**	**98**	27	30	32	32	33	49	68	76	90	93
IRVs=2	75	75	76	77	78	1	1	2	3	6	11	25	30	48	58
IRVs=3	34	35	35	39	41	0	1	2	3	3	1	4	6	12	17
IRVs=4	11	11	12	15	17	0	0	0	0	0	0	0	0	0	0
IRVs=5	0	0	4	4	5	0	0	0	0	0	0	0	0	0	0

Where '*' represents the multicollinearity level in variables

In this scenario, the irrelevant variables are correlated with relevant variables with different strengths, which means there exists multi-collinearity of different levels. In the first case, we take zero levels of multi-collinearity then only a low level of multi-collinearity, and then a high level of multi-collinearity (0.9). For all levels of multi-collinearity, we calculate the frequencies of different no. of irrelevant variables.

The results show that Elastic net and extreme bound analysis have high relative frequencies to retain the irrelevant variables for all levels of correlation of irrelevant variables to relevant variables. However, Autometrics has lower frequencies to retain the different number of irrelevant variables. If we see the relationship of frequencies and sample, it is clear that with increasing the sample size the relative frequencies of all variable selection procedures are increasing behavior.

4.4 SCENARIO 4

Relevant Variables Are Serially Correlated

The economic series in real-world applications depend on their previous history. Therefore, it is important to analyze the case in which the regressors are dependent on their lags. Therefore, in this scenario, we discuss this issue. We generate the data for DGP in which regressors are serially correlated, and then apply the different model selection criteria and note their performance to retain the DGP. We vary the covariates and their serial correlation to check the robustness of different model selection procedures under this problem. We take two cases first with ten candidate variables and second with fifteen candidate variables.

4.4.1 Relevant Variables=5; Irrelevant Variables=5

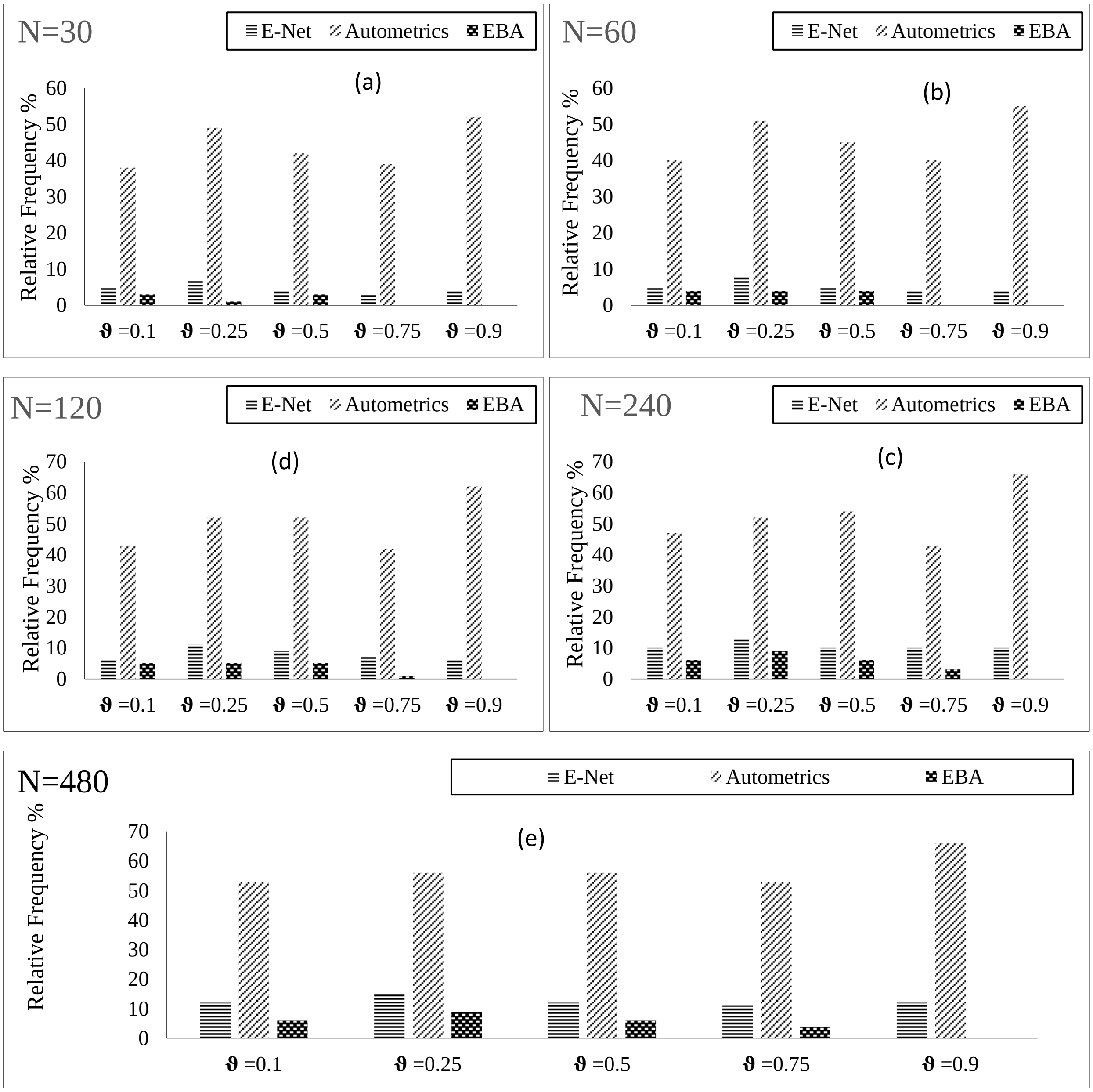
Here we take ten candidate variables for selection in which five are relevant and five are irrelevant variables. The complete description of DGP is given in section 3.4. The results of DGP in the case of serial correlation in regressors are given below.

Table 4.6: Relative Frequencies of DGP when regressors are serially correlated (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / **No.of Vars**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
ϑ =0.1	5	5	6	10	12	**38**	**40**	**43**	**47**	**53**	3	4	5	6	6
ϑ =0.25	7	8	11	13	15	**49**	**51**	**52**	**52**	**56**	1	4	5	9	9
ϑ =0.5	4	5	9	10	12	**42**	**45**	**52**	**54**	**56**	3	4	5	6	6
ϑ =0.75	3	4	7	10	11	**39**	**40**	**42**	**43**	**53**	0	0	1	3	4
ϑ =0.9	4	4	6	10	12	**52**	**55**	**62**	**66**	**66**	0	0	0	0	0

ϑ=level of serial correlation

The results show that Autometrics has higher frequencies to detect the DGP than other variable selection procedures. When the serial correlation is 0.01 the frequencies of Autometrics lie between 38 to 53 percent. And when the strength of the serial correlation is high then the frequencies of Autometrics lie between 52 to 66 percent. The Elastic Net and EBA perform poorer to retain the true DGP in the case of serially correlated variables.



ϑ=level of serial correlation

Figure 4.5: Relative Frequencies of DGP when regressors are serially correlated (%)

When there is a high level of serial correlation the EBA is unable to retain the true DGP and Elastic Net has very low frequencies. With increasing the sample, size the relative frequencies also increase.

4.4.2 Relevant Variables=5; Irrelevant Variables=10

Here we take fifteen candidate variables for selection from which five variables are relevant and ten are irrelevant. The complete description of DGP is given in section 3.4. The results of DGP in the case of serial correlation in regressors are given below.

Table 4.7: Relative Frequencies of DGP when regressors are serially correlated (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
ϑ =0.1	1	2	3	4	4	**25**	**30**	**32**	**39**	**40**	0	0	0	0	0
ϑ =0.25	2	2	3	4	5	**31**	**32**	**33**	**39**	**40**	0	0	0	0	0
ϑ =0.5	0	0	1	1	1	**29**	**30**	**31**	**32**	**33**	0	0	0	0	0
ϑ =0.75	0	0	0	0	0	**37**	**39**	**40**	**42**	**43**	0	0	0	0	0
ϑ =0.9	0	0	0	0	0	**30**	**31**	**32**	**33**	**35**	0	0	0	0	0

The results show that Autometrics has higher frequencies to detect the DGP than other variable selection procedures. When the serial correlation is 0.01 the frequencies of Autometrics lie between 25 to 40 percent. And when the strength of the serial correlation is high (0.9) then the frequencies of Autometrics lie between 30 to 35 percent. The Elastic Net and EBA perform poorer to retain the true DGP in the case of serially correlated variables. The EBA is unable to retain the true DGP and Elastic Net has very low frequencies. With increasing the sample, size the relative frequencies also increase.

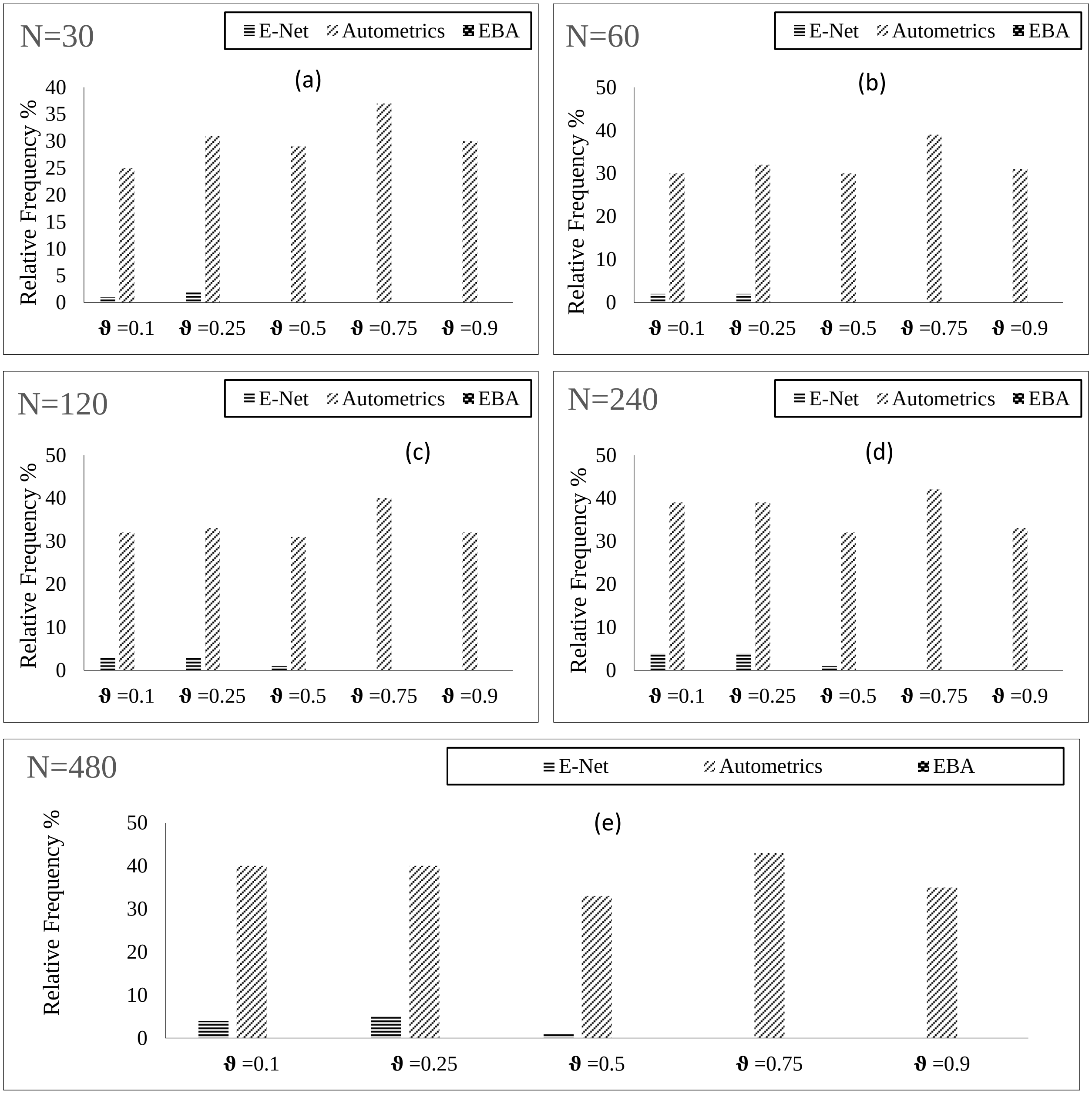


Figure 4.6: Relative Frequencies of DGP when regressors are serially correlated (%)

By analyzing the results, it is clear that Autometrics has a higher frequency in the case of retaining the true model for all sample sizes. It is also noticeable that mostly with increasing the sample size the frequencies are increasing in all cases.

4.5 SCENARIO 5:

Performance of Model Selection Procedures in Case Moving Average in Error Terms

There are numerous situations in economics when we fail to estimate the model without measurement errors. The models with measurement errors generate the errors in the model. This may have led to serial correlation in residuals of the models, which is problematic. Autocorrelation in error terms is a widespread problem in most statistical analyses for time series. In our current simulation study, we focus on variable selection. To analyze that what will be the effect of autocorrelation on the variable selection we generate the DGP with various levels of autocorrelated error terms. Then apply the variables selection procedure on this DGP to check their performance to retain the DGP. The complete description of DGP is given in section 3.5. The results of DGP in the case of autocorrelation in error terms are given below. Here, the total number of candidate variables is ten of which five are relevant and five are irrelevant. Here we put an equal number of relevant and irrelevant variables.

Table 4.8: Relative Frequencies of DGP when errors are correlated (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / Ø	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
Ø*=0	9	9	10	10	11	46	47	49	50	50	3	4	6	6	7
Ø =0.25	8	8	8	9	10	42	44	46	47	49	3	3	5	6	6
Ø =0.5	7	8	8	9	9	41	41	42	42	43	2	3	4	5	6
Ø =0.75	6	7	7	8	8	30	32	32	34	35	2	2	3	4	4
Ø =0.90	0	3	5	5	6	1	8	18	19	21	0	1	2	3	3

Where '*' represents the level of autocorrelation in error terms

Ø shows the level of autocorrelation in error terms. The zero value of Ø represents that there is no first-order autocorrelation in the error terms. The Ø=0.25 shows an exceptionally low level of autocorrelation. Moreover, as the value of Ø increases the level of autocorrelation increases According to the results the frequencies of DGP decrease by increasing the level of autocorrelation in error terms of all variables selection procedures. There is an increase in frequencies of all model selection procedures by increasing the sample size. The highest frequency to find DGP is 50% of

Autometrics in the case of 480-sample size. Overall, Autometrics has higher frequencies to retain DGP followed by E-Net and EBA.

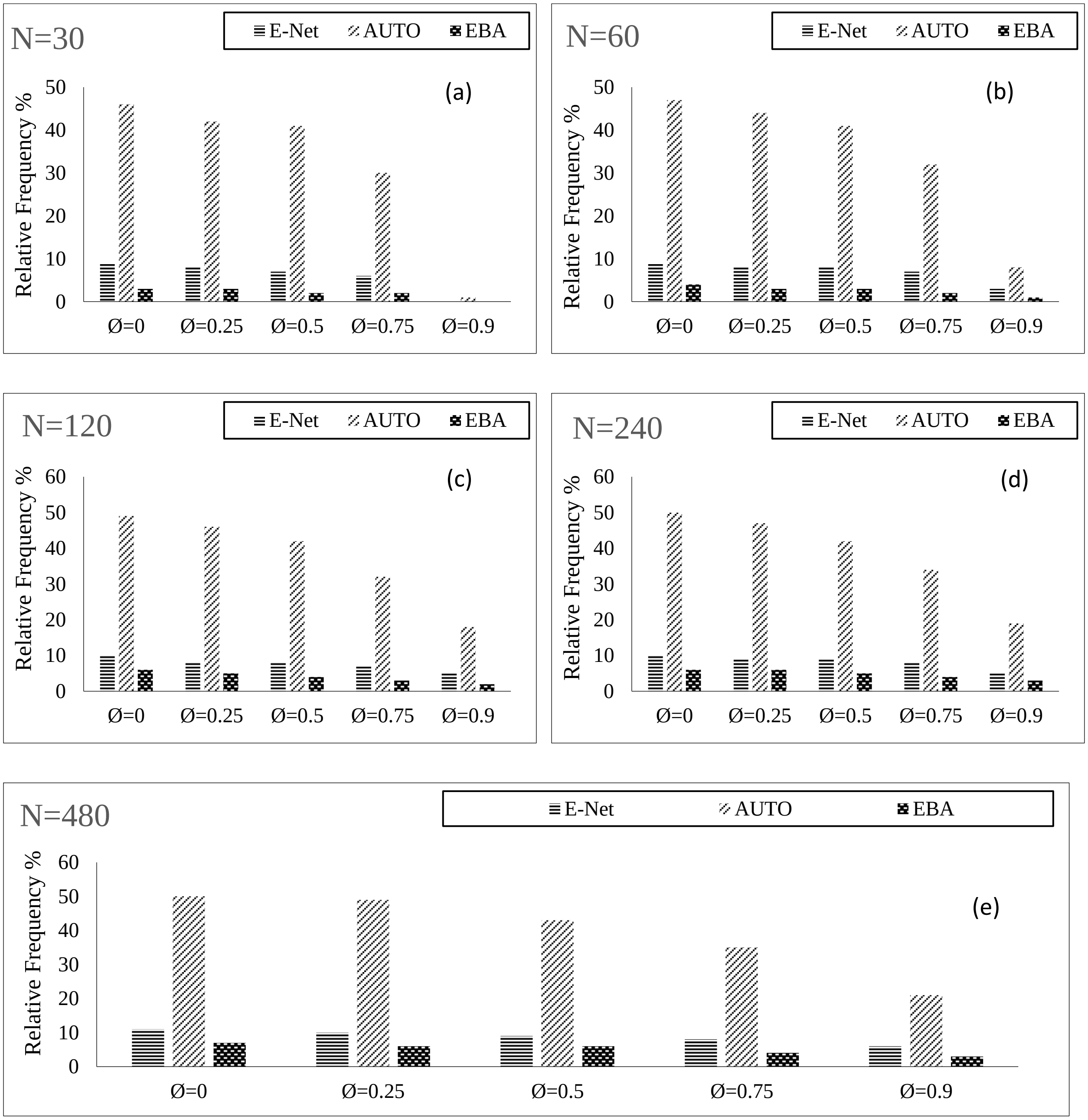


Figure 4.7: Relative Frequencies of DGP when errors are correlated (%)

Figure 4.7 represents the frequencies of DGP for different sample sizes in the case of autocorrelation in error terms. In all sample sizes, Autometrics has a higher frequency for the true model and EBA has a lower frequency for the true model.

4.6 SCENARIO 6:

Heteroscedasticity in Error Variance

Frequencies to find Data Generating Process in case of heteroscedasticity in the variance of errors terms

Heteroscedasticity has profound consequences for the estimation process. Although the OLS estimators remain unbiased, the estimated standard errors (SE) are wrong. Because of this, confidence intervals and hypotheses tests cannot be relied on. In addition, OLS estimators are no longer the best linear unbiased estimators (BLUE). If the form of heteroscedasticity is known, it can be corrected (via appropriate transformation of the data).

Heteroscedasticity tends to produce p-values of estimators that are smaller than they should be. This effect occurs because heteroscedasticity increases the variance of the coefficient estimates. Consequently, OLS calculates the t-values and F-values using an underestimated amount of variance. This problem can lead you to conclude that a model term is statistically significant when it is not significant. DPG is described in detail in section 3.6. The frequencies of DGP are calculated in case of different error variances. Here, the total number of candidate variables is ten, of which five are relevant and five are irrelevant. Here we put an equal number of relevant and irrelevant variables.

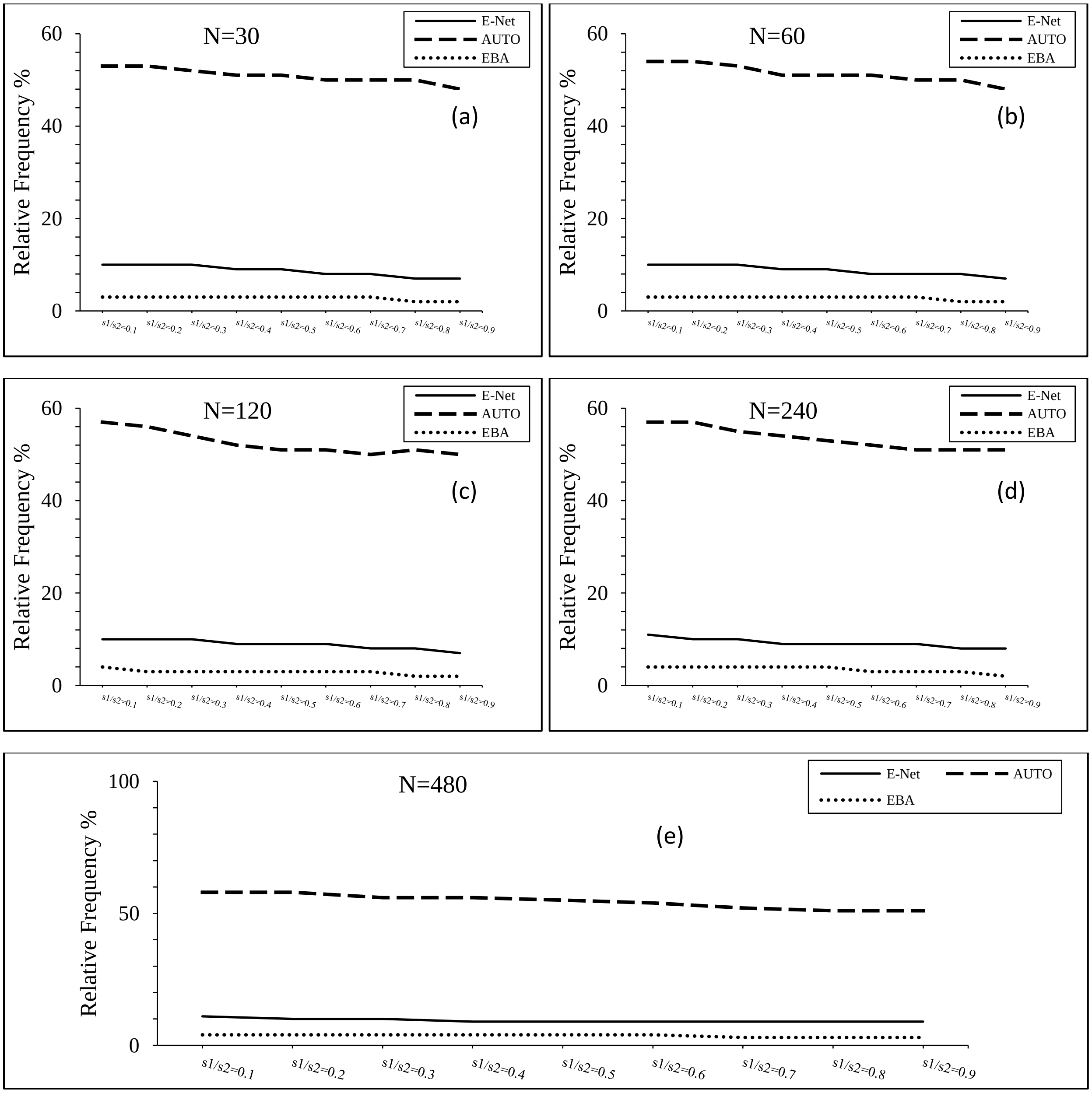
Table 4.9: Relative Frequencies of DGP when errors are heteroscedastic variance (%)

	E-Net					Autometrics					EBA				
Sample size / Hetro level	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
$\sigma1^{*}/\sigma2^{**}$=0.1	10	10	10	11	11	53	54	57	57	58	3	3	4	4	4
σ1/σ2=0.2	10	10	10	10	10	53	54	56	57	58	3	3	3	4	4
σ1/σ2=0.3	10	10	10	10	10	52	53	54	55	56	3	3	3	4	4
σ1/σ2=0.4	9	9	9	9	9	51	51	52	54	56	3	3	3	4	4
σ1/σ2=0.5	9	9	9	9	9	51	51	51	53	55	3	3	3	4	4
σ1/σ2=0.6	8	8	9	9	9	50	51	51	52	54	3	3	3	3	4
σ1/σ2=0.7	8	8	8	9	9	50	50	50	51	52	3	3	3	3	3
σ1/σ2=0.8	7	8	8	8	9	50	50	51	51	51	2	2	2	3	3
σ1/σ2=0.9	7	7	7	8	9	48	48	50	51	51	2	2	2	2	3

Where '*' and '**' represents errors variance of the first part and errors variance of the second part of the sample

Here we check the frequencies of DGP when the variance of error terms is not constant. We divide the variance into two parts, i.e., the variance of the first 50 % sample and the last 50 % sample. We divide the first variance by the second variance and vary this ratio. By increasing the difference in the variance of error terms, the frequency of DGP decreases slightly in the case of all model selection criteria. However, the decrease is slow. So, we can say that heteroscedasticity is a problem, and it may make the statistical techniques less efficient. The above results show that Autometrics has higher frequencies to retain the DGP. EBA has exceptionally low frequencies to retain the DGP and Elastic has slightly higher frequencies than EBA.

The visual description of frequencies is given below.



s1=σ1; s2=σ2

Figure 4.8: Relative Frequencies of DGP when errors are heteroscedastic variance (%)

Figure 4.34 represents the frequencies of DGP for different sample sizes in the case of heteroscedasticity in error variance. Panel (a) shows the frequencies of true model for different model selection procedures in case of 30 sample size, panel (b) shows the frequencies for 60 sample size, panel (c) is for 120 sample size, (d) is for 240 sample size and (e) is 480 sample size. In all sample sizes, Autometrics has a higher frequency for the true model and EBA has lower frequencies to retain the true model.

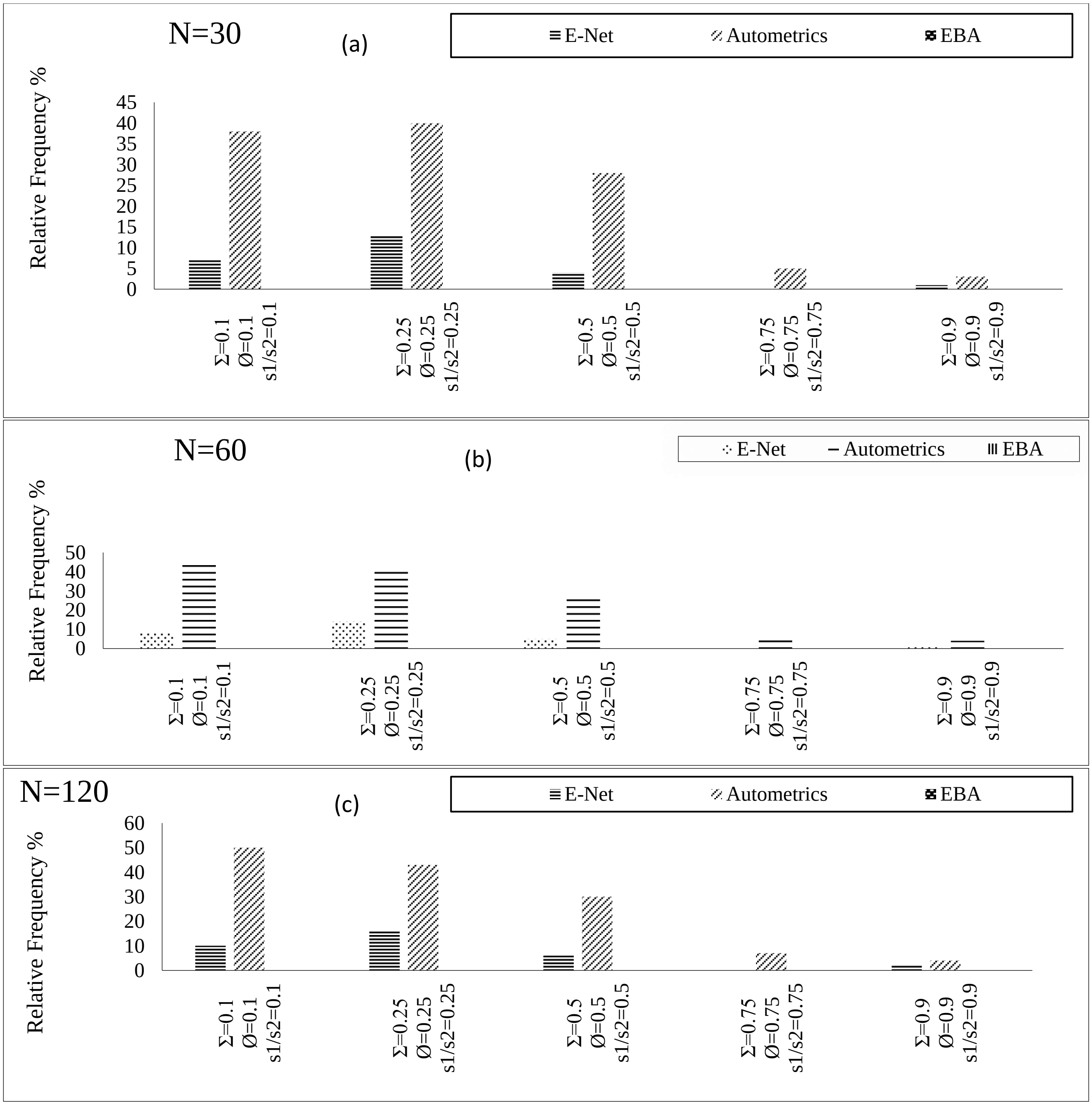
4.7 SCENARIO 7:

PERFORMANCE OF MODEL SELECTION PROCEDURES IN CASE OF MULTIPLE PROBLEMS IN DGP

In this scenario, we introduced multiple problems in DGP at a time because in real-world models there is more than one problem that can occur at a time. In previous scenarios, we introduced only one problem in DGP at a time but now we introduce more than one problem in DGP simultaneously. We introduce autocorrelation in errors terms, Heteroscedasticity in the variance of error terms, and correlation in regressors.

Table 4.10: Relative Frequencies of DGP When There Are Multiple Problems in Data (%)

	E-Net					Autometrics					EBA				
Sample size / No.of Vars	30	60	120	240	480	30	60	120	240	480	30	60	120	240	480
Σ=0.1 Ø=0.1 σ1/σ2=0.1	7	8	10	12	13	38	45	50	54	48	0	0	0	0	0
Σ=0.25 Ø=0.25 σ1/σ2=0.25	13	14	16	16	17	40	41	43	43	43	0	0	0	0	0
Σ=0.5 Ø =0.5 σ1/σ2=0.5	4	5	6	7	8	28	28	30	31	34	0	0	0	0	0
Σ=0.75 Ø =0.75 σ1/σ2=0.75	0	0	0	0	1	5	5	7	7	8	0	0	0	0	0
Σ=0.9 Ø =0.9 σ1/σ2=0.9	1	1	2	2	3	3	4	4	5	5	0	0	0	0	0


N=30
(a)
E-Net
Autometrics
EBA
Relative Frequency %
45
40
35
30
25
20
15
10
5
0
Σ=0.1
Ø=0.1
s1/s2=0.1
Σ=0.25
Ø=0.25
s1/s2=0.25
Σ=0.5
Ø=0.5
s1/s2=0.5
Σ=0.75
Ø=0.75
s1/s2=0.75
Σ=0.9
Ø=0.9
s1/s2=0.9
N=60
(b)
E-Net
Autometrics
EBA
Relative Frequency %
50
40
30
20
10
0
Σ=0.1
Ø=0.1
s1/s2=0.1
Σ=0.25
Ø=0.25
s1/s2=0.25
Σ=0.5
Ø=0.5
s1/s2=0.5
Σ=0.75
Ø=0.75
s1/s2=0.75
Σ=0.9
Ø=0.9
s1/s2=0.9
N=120
(c)
E-Net
Autometrics
EBA
Relative Frequency %
60
50
40
30
20
10
0
Σ=0.1
Ø=0.1
s1/s2=0.1
Σ=0.25
Ø=0.25
s1/s2=0.25
Σ=0.5
Ø=0.5
s1/s2=0.5
Σ=0.75
Ø=0.75
s1/s2=0.75
Σ=0.9
Ø=0.9
s1/s2=0.9

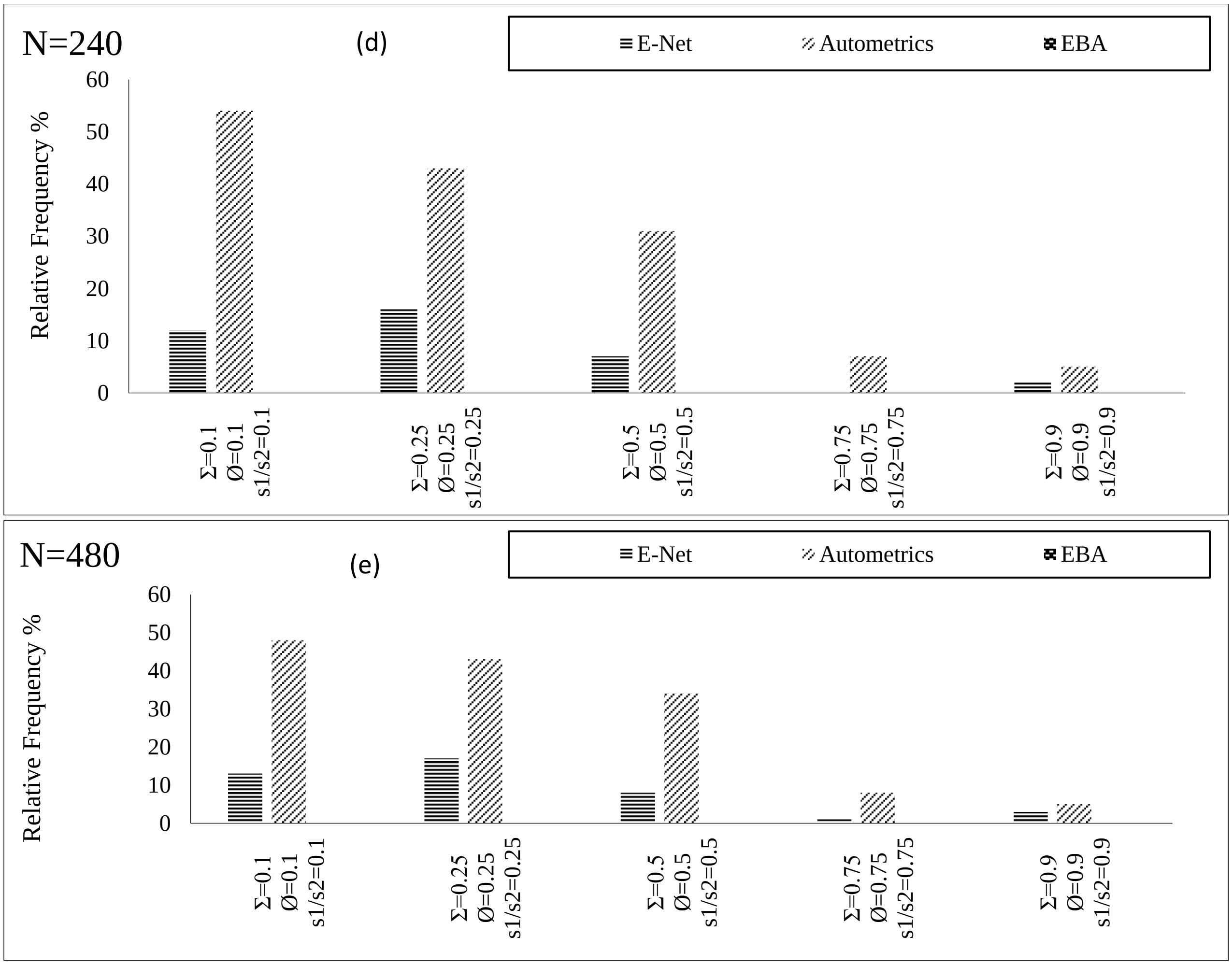


Figure 4.9: Relative Frequencies of DGP When There Are Multiple Problems in Data (%)

In this scenario, we introduced multiple problems in DGP at a time because in real-world models there is more than one problem that occurs at a time. In previous scenarios, we introduced only one problem in DGP at a time but now we introduce more than one problem in DGP simultaneously. We introduce autocorrelation in errors terms, Heteroscedasticity in the variance of error terms, and correlation in regressors. The results show that Autometrics has higher relative frequencies than other model selection procedures when we introduce the multiple problems in DGP. We vary the increase the level of autocorrelation in error, correlation in regressors, and level of heteroscedasticity in error variance from 0.1 to 0.9, the Autometrics again has higher relative frequencies than other variables selection procedures. When the level of autocorrelation in error, correlation in regressors, and level of heteroscedasticity in error variance is 0.9 all model has very exceptionally frequencies to retain the DGP.

CHAPTER 5

EMPIRICAL APPLICATION

In our simulation experiments, we analyzed that some variable selection procedures perform better even in the case of correlation in regressors. In this study, our main concern is variables selection from a large number of candidate variables by different variables selection procedures. We analyzed many Data generating processes to check the robustness of the variable selection procedure. We have taken non-orthogonal variables in simulation experiments, moving average in error, and serial correlation in regressors.

There are many theories of growth with their explanations and justifications. If we take all theories then it will create the problem of variable selection because every theory tells about a different set of variables. In the model of economic growth, the problem of multi-colinearity in DGP (which we discussed in the simulation experiment) is present, like the in economic growth model the regressors are non-orthogonal, they are also serially correlated. So, the model of economic growth is perfect for the real-world application of the above simulation experiments.

Modeling economic growth is a typical case, there are many theoretical models of economic growth that offer possible explanations for this growth and each model has a separate set of regressors. Theories and models of economic growth highlight the different ways in which current economic activity can influence future economic growth and identify the sources that can sustain economic growth. Researchers and economists have confirmed the need for economic growth for the evolution and well-being of generations. Theories of economic growth have evolved from time to time in terms of the evolved period and the dynamics of the economy.

Improvements in mathematical and statistical tools have also had a significant effect on the formation of concepts. Why do we need economic growth? What are the key factors that drive growth? Many researchers, economists, and Nobel laureates have tried to answer these questions. Economic growth can be considered an important factor in the well-being and prosperity of billions of people. Advances in industrialization and technology have left a gap between developed countries and poorer countries. For example, now, in the 21st century, the GDP / per capita of many poor countries is less than the GDP per capita of Europe in the 19th century.

There are many theories of economic growth about which there are some popular theories as follows.

5.1 Some renowned economic growth Models

5.1.1 Harrod Domar model Savings Ratio and Investment

The Harrod-Domar model is a variant of the neoclassical model. It says the growth rate depends on the savings rate. Some growth theories place too much weight on increasing domestic savings. Savings delivers financial support for investments. Therefore, that investment drives further growth. This has been a key growth factor. However, it depends on how effective the investment is. If the savings are too high, that leads to less growth because people are unable to afford the investment.

$$Y = f(K) \tag{5.1}$$

Y is output, K is capital

5.1.2 Neo-Classical model of Solow/Swan

Robert Solow developed a neoclassical theory of economic growth, and Solow won the Nobel Prize in Economics in 1987. He has greatly contributed to our understanding of the factors that determine the economic growth rate of different countries. Growth comes from adding more capital and sources of labor and from ideas and new technologies.

In the aggregate production function, the single good is created by the two elements labor (L) and capital (K).

$$Y(t) = k(t)^{\alpha}(A(t)L(t))^{1-\alpha} \tag{5.2}$$

In equation (5.3) t is for time and $0 < \alpha < 1$ is the elasticity of output concerning capital and Y(t) is total production. A is labor augmented technology AL is for effective labor.

$$L(t) = L(0)e^{nt} \tag{5.3}$$

$$A(t) = A(0)e^{gt} \tag{5.4}$$

Effective units of labor A(t)L(t) and grow at the rate (n+g). According to the neo-classical theory of economic growth, growing capital or labor results in diminishing returns. Therefore, that increment in the capital can stimulate only temporary economic growth. For economic growth in the Solow/Swan model there should be:

- Increment in invested GDP however, is inadequate as more investment results in diminishing returns.
- Technological progress boosts the productivity of capital/labor.

So that the poor countries who are making more investment should see their growth approaching the growth of rich countries.

5.1.3 New Economic Growth Theories (Endogenous growth)

Romer developed endogenous growth theory, emphasizing that technological change is the result of efforts by researchers and entrepreneurs who respond to economic incentives. Anything that affects their efforts, such as tax policy, basic research funding, and education, for example, can potentially influence the long-run prospects of the economy. Romer's fundamental contribution is his clear understanding of the economics of ideas and how the discovery of new ideas lies at the heart of economic growth. His 1990 paper is a watershed (Romer, 1990a), Romer emphasized the centrality of questions such as: "what determines the long-run rate of economic growth in living standards?" This reminder came in the form of his 1983 dissertation (Romer, 1983) and the key growth publication it led to (Romer, 1986). In this way, Romer was a key founder of what came to be known as endogenous growth theory.

$$Y_t = AK_t \tag{5.5}$$

and

$$K_t = S\dot{y}_t - \Delta k_t \tag{5.6}$$

- A is an exogenous and constant productivity parameter and s is an exogenous constant investment rate. In this set-up, K is interpreted as physical capital, but in Romer (1986) K was interpreted as knowledge, and in Lucas (1988) it was replaced by human capital.
- Endogenous growth models of Paul Romer and Robert Lucas, emphasize human capital. With the help of greater knowledge, education, and training workers can increase the advancement of technology.

- This led to the development of a host of different growth models and, together with the widespread availability of data that allowed income comparisons across countries and over time. The emphasis is on the need of government for the encouragement of technological innovation.
- The model puts weight on increment in both capital and labor productivity.
- States labor productivity increment does not have diminishing returns, but, it can increase the returns.
- Argue that capital increasing does not lead to diminishing returns necessarily as by the Solow.
- Spillover benefits have increased the importance of a knowledge-based economy.

Joseph Schumpeter claimed that the intrinsic feature of capitalism was the 'creative destruction' allowing incompetent firms to fail was essential for allowing resources to flow to more well-organized channels.

Apart from the above theories, economic growth can be influenced by

- The domestic rate of saving (Harrod-Domar)
- Investment in the capital (classical model)
- Improvement in technology (Endogenous growth and others)
- Human Capital (Endogenous growth and unified growth)
- Markets openness (Endogenous growth and classical models)

5.2 The general model for Economic Growth can be defined as,

Functional form Economic growth

$$Y = f(FD, FDI, DS, CF, INF, LB, FA, REM, HC, TO) \tag{5.7}$$

Where Y=Economic growth, FD=Financial development, FDI=Foreign direct investment, DS=Domestic investment, CF=Capital formation, INF=Inflation, LB=Labour force, FA=Foreign assets, REM=Remittances, HC=Human Capital, TO=Trade openness

Econometric form of the above model

$$Y = \beta_0 + \beta_1(FD) + \beta_2(FDI) + \beta_3(DS) + \beta_4(CF) + \beta_5(INF) + \beta_6(LB) + \beta_7(FA) + \beta_8(REM) + \beta_9(HC) + \beta_{10}(TO) + \varepsilon \tag{5.8}$$

These are potential determinants of economic growth. We put all determines into one equation called the general unrestricted model (GUM) keeping in mind the properties of GUM.

5.3 The importance of variables of economic growth in literature

5.3.1 Financial Development (FD) and economic growth

A good financial status of a country is one of the main factors for continuous economic development that can be built (Demirguc-Kunt, 2006). One of the issues in the literature of financial economics is the nexus of financial growth. There were two main thoughts. Proponents of the first thought discussed that financial growth was essential to economic growth (Goldsmith, 1969; Levine, 1997; McKinnon, 1973). However, the advocates of the second school of thought, the neoclassical theorists, claimed that finance is not a primary source of growth (Lucas, 1988). Yet, in the recent past, numerous studies have agreed that growth in the financial sector has had a positive effect on growth (Beck, Levine, & Loayza, 2000).

5.3.2 Foreign Direct Investment (FDI) and Economic Growth

Many research papers have analyzed the link between FDI and trade components (exports, import openness, trade restrictions) and growth. Lee and Liu (2005) explored the role of FDI in economic growth for a large sample of both developing and developed countries. The results conclude that FDI directly and positively affects growth. Research by other researchers in the early 2000s suggests that FDI may have a positive correlation between this and economic growth (Lensink and Morrissey, 2006).

5.3.3 Domestic Savings (DS) and Economic Growth

A review of the literature on the relationship between savings and economic growth indicates that there is a positive relationship between household savings and economic growth. This positive relationship can be explained by several assumptions. The first one assumes that increasing savings can boost economic growth through increased investment (Bebczuk 2000). This view was developed by Harrod (1939), Domar (1946), and Solow (1956) growth models. Theories of economic growth require that if there is an increase in investment in human or material capital or scientific research and development (R&D), then the dynamics of the country's economic growth increase. However, if the country has access to international financial markets, it mustn't grow rapidly because of domestic savings, as foreign savings can be used to finance investments (Guterries, Solimano, 2007).

5.3.4 Inflation (INF) and Economic Growth

The economic growth of a country is the result of fiscal, monetary, and other economic policies initiated by its policymakers. There are many factors of economic growth, one is inflation. The relationship between economic growth and inflation is not a simple phenomenon. Numerous studies have examined the complexity of this relationship. Empirical studies on industrialized and developed countries show that there is a negative relationship between economic growth and inflation. In contrast, studies on developing countries show a positive relation.

Nevertheless, research into the relationship between economic growth and inflation (Mamo, 2012) was also considered a central theme in economic research and policy. There is no clarity in the relationship between economic growth and inflation. Various studies (Mamo, 2012) have shown that the relationship between economic growth and inflation can be positive, negative, and neutral. According to Malik and Chowdhury (2001), there is a positive correlation between inflation and economic growth.

5.3.5 Labor force, Human Capital and Economic Growth

One component of employment that directly affects it is the ability to meet the demands of the workers or the needs of the workforce. This requires constant upgrading of skills, especially in areas that experience rapid technological and organizational change, to help their human capital or workforce avoid abandonment.

5.3.6 Foreign Assets (FA) and Economic Growth

Theoretically, the amount of net foreign assets in the process of macroeconomic adjustment can be either exogenous or endogenous. For example, Lane and Milesi-Ferretti (2002) take the form of net foreign assets as an exogenous variable and offer some indication that the size of net foreign assets determines the trade balance and the real exchange rate. According to Lane and Milesi-

Ferretti (2002), a positive stable state net external asset position in a country allows the country to run a permanent trade deficit, and hence the exchange rate is defined.

In contrast, a country with negative net foreign assets has to increase trade to serve its external obligations, thus reducing its exchange rate. Alternatively, Masson and et al. (1994), Cavallo and Ghironi (2002), Ghironi et al. (2008), and Ghironi (2008) consider net foreign assets as the final variable in the process of adjusting the economic system.

5.3.7 Remittances(REM) and Economic Growth

Remittances promote economic growth, promote recipient incomes, reduce debt barriers, accelerate investment, and promote human development through better education and healthcare financing. Reduce poverty Occurs [Calaro (2008) Jongwanich (2007); Stark and Lucas (1988); Taylor (1992); Faini (2002) Gupta et al. (2009)]. However, Chami et al. (2003) found that remittances negatively affect the economic growth of recipients because a significant flow of remittances reduces labor force participation and labor efforts, resulting in reduced productivity. Thus, the effects of remittances on recipient economic growth and development have been controversial.

5.3.8 Trade Openness (TO) and Economic Growth

Over the past three decades, trade (especially in developing economies) has grown as a result of the limited scope of development strategies based on import substitutes and the influence of international financial institutions, such as the International Monetary Fund and the World Bank has expanded rapidly, often demanding their support on trade liberalization. The basic premise of the commitment to the trade reform program is the clear belief that liberalization is a prerequisite for the transition from relatively close to relatively open economies. Economists generally agree that open economies grow faster than their counterparts (Grossman and Helpman, 1991, Edwards, 1993).

We took various countries from Asian, European, American, and African countries. We apply all model selection procedures on data of all countries, calculated the frequencies of different retained variables, and for in-sample forecasting, we have calculated the root mean square error (RMSE) and mean square prediction error (MSPE).

Table 5.1: Selected Countries

Sr. No.	**Abr**	**Country Name**	**Sr. No.**	**Abr**	**Country Name**
Asian Countries					
1	PAK	Pakistan	8	HOK	Hong Kong
2	IND	India	9	JAP	Japan
3	BAN	Bangladesh	10	TUR	Turkey
4	SRI	Sri Lanka	11	KUW	Kuwait
5	CHI	China	12	OMO	Oman
6	RUS	Russia	13	SUA	Saudi Arabia
7	KOR	Korea	14	UAE	United Arab Emirate
European Countries					
15	GER	Germany	19	SPI	Spain
16	UK	United Kingdom	20	SWI	Switzerland,
17	FRA	France	21	AUS	Austria
18	ITL	Italy	22	SWE	Sweden
American Countries					
23	US	United States	26	CAN	Canada
24	BRA	Brazil	27	PER	Peru
25	MEX	Mexico			
African Countries					
28	EGY	Egypt	31	SA	South Africa
29	NIG	Nigeria	32	MOR	Morocco
30	ETH	Ethiopia			

Abr=Abbreviation

5.4 Variables retained for the model of Economic Growth

In the first column of Table 3.1(see in appendix 3), the countries' short names are written in the second column different criteria of variables selection are written in the third column RMSE, and in the fourth column, MPSE is written. All other columns show the potential determinants of Economic Growth.

✓is for the determinants that are retained by different variable selection criteria. By analyzing the whole table 3.1 in the appendix, it is clear that the E-Net mostly retains more variables than other variables selection procedures and this is consistent with earlier results of over-specification in simulations results. Autometrics retains fewer variables and this is consistent with simulation results.

The frequency of different variables that are detected by different variables selection procedures for overall countries and reign specified countries is shown in the following tables.

The frequency for total countries is calculated as

Relative frequency of a variable = number of times a variable is detected from all countries

Let us calculate the relative frequency of a variable X.

The frequency of retention of the variable is calculated in the following way. Let X denote a variable from the general model.

$$Frequency\ (X) = \left(\frac{n(X)}{N}\right) \times 100 \tag{5.9}$$

Here n(x) is the number of times a variable is retained by a variables selection procedure, and N (in our case N=32) is the total number of countries.

All countries

Table 5.2: Frequencies of retained variables in growth model (%)

	Financial Development	FDI	Savings	Capital Formation	Inflation	Labor Force	Foreign Assets	Remittances	Human Capital	Trade Openness
E-Net	66	44	63	50	56	31	50	25	38	59
Autometrics	38	22	41	47	41	19	38	13	41	41
EBA	47	34	34	44	41	19	53	19	38	22

In the first row, the name of variables is written, in the second row, the detection frequency of Elastic Net is given, in the third row, the detection of Autometrics is given, in the fourth row, the detection of Extreme Bound Analysis is given. Here we show the results with percentage, which means variables retained for different countries by variable selection procedure divided by total countries. 0.50 means a variable is retained from half of the countries. The retention of the different determinants of growth with respect to countries is given in table 3.1 in appendix 3. A detailed explanation of retained variables by different variables selection procedure is given below

For all selected countries average of all variable selection procedures, the financial development is the variable that is mostly retained. Financial development is directly related to capital means more financial development in an economy the more capital in that economy. The role of Capital is supported by all economic theories so our results are consistent with economic theories.

The second most retained variables in our selected countries are capital formation and foreign assets. As discussed earlier that capital is the factor for economic growth supported by many economic growth theories, so again our findings are consistent with economic growth models.

The third most retained variable is inflation. In the current era, inflation become a problem of many economies. In earlier eras, inflation remained an important factor for economic growth so again this variable is consistent with earlier studies.

Remittances are a factor for economic growth which is retained minimum times. Remittances are important for developing economies not for developed economies so it is retained in the estimated model a few times.

Elastic net retained the financial development mostly, Autometrics retained capital formation with a high percentage and EBA retained foreign assets with a higher percentage.

Discussion on selection of variables across the regions

Table 5.3: Frequencies of retained variables in growth model (%)

	Financial Development	FDI	Savings	Capital Formation	Inflation	Labor Force	Foreign Assets	Remittances	Human Capital	Trade Openness
Asian Countries										
E-Net	93	64	50	64	71	29	64	29	50	71
Autometrics	43	14	14	50	29	21	50	7	29	36
EBA	71	36	21	43	36	7	71	14	21	14
European Countries										
E-Net	75	38	88	38	63	25	25	13	63	63
Autometrics	63	25	63	38	25	25	38	25	63	13
EBA	50	0	75	25	25	25	13	0	25	38
American Countries										
E-Net	0	60	60	40	100	20	40	0	60	0
Autometrics	0	40	60	40	80	0	20	20	20	40
EBA	0	60	60	40	100	20	40	0	60	0
African Countries										
E-Net	60	60	60	60	40	80	80	60	40	80
Autometrics	0	0	20	60	20	20	40	40	40	40
EBA	20	20	20	60	20	20	60	20	20	80

Discussion on selection of variables across the regions

For Asia, the most retained variable is financial development and the retained frequency is (E-Net=93%, Autometrics=43% EBA=71%). This is consistent with earlier findings and consistent with economic growth theories. For Asian countries, the second most retained variable is foreign assets and the retained frequency is (E-Net=64%, Autometrics=43% EBA=71%). In Asia, most economies are developing so for these countries foreign assets have an important role. The third most retained variable is the capital formation and retained frequency is (E-Net=64%, Autometrics=50% EBA=43%). Capital formation remains an important factor for economic growth in all economies and all growth models so in the case of Asian economies this is also an important factor for economic growth. In the case of Asian countries, the Elastic net retains more variables than Autometrics and EBA, and this is consistent with simulation results. For European economies, the most retained variable is savings and retained frequency is (E-Net=88%, Autometrics=63% EBA=75%). Savings are related to financial development and capital formation and mostly growth theories put importance on capital formation for economic growth.

The second most retained variable for European countries is financial development and retained frequency is (E-Net=75%, Autometrics=63% EBA=50%). This is consistent with earlier findings and consistent with growth models. Here Elastics Net retains more variables than other variables selection procedures. For American countries, the most retained variable is inflation and retained frequency is (E-Net=100%, Autometrics=80% EBA=100%). The second most retained variable in the case of American countries is savings and retained frequency is (E-Net=60%, Autometrics=60% EBA=60%). For African countries, the most retained variable is trade openness and retained frequency is (E-Net=80%, Autometrics=40% EBA=80%). The second most retained variable is foreign assets and retained frequency is (E-Net=80%, Autometrics=40% EBA=60%).

For in-sample forecasting comparison, we used root mean square error (RMSE) and mean square prediction error (MSPE). EBA presents a superior predictive performance in terms of lower RMSE and MPSE that are 1.03 and 0.05 respectively (for more results see table 3.1 in appendix 3). We see in our simulation results that EBA performs better in the case of multi-collinearity, and in the model of economic growth there exists multi-collinearity, so the EBA performs better in the case of in-sample forecasting. If we see the highest value of RMSE that is 8.44 in the case of EBA. Moreover, the highest value of MPSE is 27.09 by Autometrics. Financial development is mostly retained determinant of economic growth by all model selection procedures.

CHAPTER 6

SUMMARY, CONCLUSION, AND RECOMMENDATIONS

In the previous section, we performed many experiments according to our objectives via simulation on several DGPs for variable selection procedures. Our first objective is to find out optimal variables selection procedure that retains the most relevant variables (true DGP) with high frequency. The second objective is to check the robustness of these variables selection procedures in case of multi-collinearity in the regressors, serial correlation in explanatory variables, autocorrelation in the error term, heteroscedasticity in the variance of error terms, and for several sample sizes. For these objectives, we run the simulation on different settings of DGP in which variables are orthogonal they have some specific level of correlation among each other (non-orthogonal), and the variables are serially correlated, auto-correlated error terms and errors term are heteroscedastic variances. The main findings are; when variables are non-orthogonal to each other, then Extreme Bound Analysis performs better than other model selection procedures it has a higher frequency to retain the true DGP. When the variables are orthogonal then Autometrics has a higher relative frequency to retain the true DGP in simulation analysis. In cases of DGP with autocorrelated errors and heteroscedastic variance, the Autometrics performs better than other variables selection procedures to retain the true DGP.

The complete discussion on findings with respect to DGPs settings is following.

6.1 Findings w.r.t DGPs

6.1.1 Performance of variables selection procedures when Variables are orthogonal

One objective was to check the performance of different variable selection procedures in the case of orthogonal (variables are not correlated to each other) variables. For this objective, we generate the DGP with orthogonal variables. The results show that Autometrics has higher relative frequencies to retain the DGP. Here the second thing to notice is that relative frequency increases with increasing the sample size. Elastic net has a higher frequency in the case of all relevant variables and some non-DGP variables are retained in the final model. It means the Elastic Net retains the overspecified model.

6.1.2 Performance of variables selection procedures when relevant variables are non-orthogonal

To meet the objective of the robustness of the variables section in the case of non-orthogonal (high level of multi-collinearity) variables, we have generated the DGP with correlated variables and applied the variables selection procedures on this DGP. Then the frequencies of retained DGP are calculated. We repeated this for several sample sizes. The results show that Extreme bound analysis has a higher relative frequency to retain the GDP (all relevant variables). Relative frequency increases as sample size increases in the case of non-orthogonal variables. As the level of correlation among regressors increases the relative frequency to retain the relevant variables of EBA increases.

6.1.3 Performance of variables selection procedures when moving average in error terms (autocorrelated error terms)

In the estimation process, the researcher keeps in mind the autocorrelation of error terms to incorporate this situation with different levels of autocorrelation in error terms. We have done the experiments for different levels of autocorrelation in error terms. After analysis of the simulation results, it is realized that Autometrics has a higher relative frequency for the true DGP and EBA has a lower relative frequency for the DGP. The relative frequencies of E-Net for DGP remain between the two other procedures.

6.1.4 Performance of variables selection procedures when Heteroscedasticity is present in the variance of error terms

In the estimation process, the researchers keep in mind the heteroscedasticity of error variance so for this situation, we generated the DGP with the error of heteroscedastic variance. In our objectives, one objective is to check the robustness of variable selection procedures in case of heteroscedastic error variance. So we have generated the DGPs with different levels of heteroscedasticity in error variance and applied the variable selection procedures. Autometrics has a higher relative frequency to retain the true DGP and EBA has a lower relative frequency to retain the true DGP. The relative frequencies of E-Net remained between the relative frequencies of the other two variable selection procedures.

6.1.5 Impact of multiple problems in DGP on the relative frequency of variable selection

In real data, there may be multiple problems in the data at a time, for this type of data we have generated the DGPs with more than one problem at a time and applied the different variable selection procedures to retain the true DGP. The simulation results show that the E-Net has higher frequencies in the case of DGP plus one and two non-DGP variables retained. The EBA has higher relative frequencies in case all variables (DGP and non-DGP variables) are retained. In all cases and all sample sizes, Autometrics has lower relative frequencies for non-DGP variables, and EBA has higher relative frequencies for non-DGP variables. The relative frequency of E-Net in cases of non-DGP variables decreases as the number of non-DGP variables increases and for all non-DGP variables the relative frequency is near to zero.

6.2 Summary Table of Simulation Results

Sr.No	DGP	Finding (s)
1	All variables are orthogonal	Automertics performs better
2	There is multi-collinearity in variables	Extreme bound analysis performs better in this case
3	There is Autocorrelation in error terms	Automertics performs better
4	There is heteroscedasticity in error variance	Automertics performs better
5	There is heteroscedasticity, auto correlation in error terms, and multi-collinearity in variables	Elastic net performs better than other variables selection procedures

6.3 Recommendations

After concluding the simulation results, these are some recommendations:

1. As it has been shown in simulation results the Extreme Bound Analysis performs better in retaining, the DGP when the regressors are non-orthogonal, so Extreme Bound Analysis is recommended in the case of non-orthogonal variables.
2. The results of the simulation show that Autometrics has higher relative frequencies in the case of orthogonal variables, it is recommended to apply Autometrics whenever the variables are orthogonal. While there are a large number of irrelevant variables in the candidate variables list, Autometrics is recommended because with an increasing number of irrelevant variables there is no effect on the relative frequencies of relevant variables(true DGP).

6.4 Limitations and Further Research

We have done many simulations using different settings of DGPs, several sample sizes, using different combinations of relevant and irrelevant variables. However, not all the combinations of econometrics models can be handled in a single study. The simulation results become so lengthy and we could not analyze other situations (following situations) because of so extension of the results.

1- We have checked the performance of variables selection criteria for common situations (problems in data) like autocorrelation in error terms, heteroscedasticity in error variance, correlation in regressors, and serial correlation in regressors, researchers can take further scenarios including endogeneity, non-linearity, structural breaks in data, lags selection.
2- We focus on static models further research can be done on dynamic models in which all the above scenarios can be repeated.
3- We compared the Elastic Net to Extreme Bound Analysis and Autometrics, hence future research works can focus on the other latest versions of the regularized type methods of variable selection.
4- In our study, the simulations and real data applications are for low dimensional data, so future research works would consider the high dimensional data sets, in which the number of variable exceeds the observations.

REFERENCES

Abu-Bader, S., & Abu-Qarn, A. S. (2008). Financial development and economic growth: The Egyptian experience. *Journal of Policy Modeling*, *30*(5), 887-898.

Abu-Eideh, O. M. (2014). Factors of economic growth in Palestine: an empirical Analysis during the period of (1994-2013). *International Journal of Business and Economic Development (IJBED)*, *2*(2).

Singh, T. (2008). Financial development and economic growth nexus: a time-series evidence from India. *Applied economics*, *40*(12), 1615-1627.

Acosta, P. (2006). Labor supply, school attendance, and remittances from international migration: the case of El Salvador. *The World Bank.*

Acosta, P., Calderon, C., Fajnzylber, P., & Lopez, H. (2008). What is the impact of international remittances on poverty and inequality in Latin America?. *World Development*, *36*(1), 89-114.

Acs, Z. J., & Varga, A. (2002). Geography, endogenous growth, and innovation. *International Regional Science Review*, *25*(1), 132-148.

Acs, Z. J., Audretsch, D. B., Braunerhjelm, P., & Carlsson, B. (2004). The missing link: The knowledge filter and entrepreneurship in endogenous growth.

Adeel-Farooq, R. M., Bakar, N. A. A., & Raji, J. O. (2017). Trade openness, financial liberalization and economic growth. *South Asian Journal of Business Studies*.

Aggarwal, R., Demirgüç-Kunt, A., & Pería, M. S. M. (2011). Do remittances promote financial development?. *Journal of development economics*, *96*(2), 255-264.

Aghion, P., Comin, D., Howitt, P., & Tecu, I. (2016). When does domestic savings matter for economic growth?. *IMF Economic Review*, *64*(3), 381-407.

Aghion, P., Ljungqvist, L., Howitt, P., Howitt, P. W., Brant-Collett, M., & García-Peñalosa, C. (1998). *Endogenous growth theory*. MIT press.

Ahmad, J., & Kwan, A. C. (1991). Causality between exports and economic growth: empirical evidence from Africa. *Economics letters*, *37*(3), 243-248.

Aho, K., Derryberry, D., & Peterson, T. (2014). Model selection for ecologists: the worldviews of AIC and BIC. *Ecology*, *95*(3), 631-636.

Aigner, D. J., & Judge, G. G. (1977). Application of pre-test and Stein estimators to economic data. *Econometrica: Journal of the Econometric Society*, 1279-1288.

AKAIKE, H. (1973). Information theory and an extension of the maximum likelihood principle. In *2nd International Symposium on Information Theory, 1973* (pp. 267-281).

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716-723.

Akaike, H. (1992). Information Theory and an Extension of the Maximum Likelihood Principle. In *Breakthroughs in Statistics* (pp. 610-624). Springer, New York, NY.

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199-213). Springer, New York, NY.

Alam, K., & Hawkes, J. S. (1978). Estimation of regression coefficients. *Scandinavian Journal of Statistics*, 169-172.

Ang, J. B. (2008). What are the mechanisms linking financial development and economic growth in Malaysia?. *Economic Modelling*, *25*(1), 38-53.

Awan, A. G. (2014). Shifting Global Economic Paradigm. *Asian Business Review*, *4*(3), 113-118.

Awokuse, T. O. (2008). Trade openness and economic growth: is growth export-led or import-led?. *Applied economics*, *40*(2), 161-173.

Ayres, C. E. (1962). The Theory of Economic Progress: A Study of the Fundamentals of. *Economic Development and Cultural Change. 2d ed. New York: Schocken Books*.

Backhouse, R. E., & Boianovsky, M. (2016). Theories of stagnation in historical perspective. *European Journal of Economics and Economic Policies: Intervention*, *13*(2), 147-159.

Baharumshah, A. Z., & Almasaied, S. W. (2009). Foreign direct investment and economic growth in Malaysia: interactions with human capital and financial deepening. *Emerging Markets Finance and Trade*, *45*(1), 90-102.

Baldwin, R. E., & Forslid, R. (2000). Trade liberalisation and endogenous growth: A q-theory approach. *Journal of International Economics*, *50*(2), 497-517.

Bardhan, P. (1995). The contributions of endogenous growth theory to the analysis of development problems: an assessment. *Handbook of development economics*, *3*, 2983-2998.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, *106*(2), 407-443.

Baumol, W. J. (1952). Yet Another Note on the Harrod-Domar Model. *The Economic Journal*, *62*(246), 422-427.

Bebczuk, R. N. (2000). Productivity and saving channels of economic growth as latent variables: An application of confirmatory factor analysis. *Estudios de Economía*, *27*(2), 243-257.

Bebczuk, R. N., & Musalem, A. (2006). Pensions and saving: new international panel data evidence. *Documentos de Trabajo*.

Bebczuk, R. N., Gasparini, L., Garbero, M. N., & Amendolaggine, J. (2015). Understanding the determinants of household saving: micro evidence for Latin America. *Documentos de Trabajo del CEDLAS*.

Bebczuk, R., & Cavallo, E. (2016). Is business saving really none of our business?. *Applied Economics*, *48*(24), 2266-2284.

Beck, T., Demirgüç-Kunt, A., & Levine, R. (2007). Finance, inequality and the poor. *Journal of economic growth*, *12*(1), 27-49.

Bernanke, B. S., & Gürkaynak, R. S. (2001). Is growth exogenous? taking mankiw, romer, and weil seriously. *NBER macroeconomics annual*, *16*, 11-57.

Besomi, D. (1999). *The making of Harrod's dynamics*. London: Macmillan.

Bloch, H., & Tang, S. H. K. (2003). The role of financial development in economic growth. *Progress in Development Studies*, *3*(3), 243-251.

BM, P. (1991). Effects of Model Selection on Inference. *Econometric Theory*, *7*(2), 163-185.

Boianovsky, M. (2018). Beyond capital fundamentalism: Harrod, Domar and the history of development economics. *Cambridge Journal of Economics*, *42*(2), 477-504.

Boldeanu, F. T., & Constantinescu, L. (2015). The main determinants affecting economic growth. *Bulletin of the Transilvania University of Brasov. Economic Sciences. Series V*, *8*(2), 329.

Bonnal, M., & Yaya, M. E. (2015). Political institutions, trade openness, and economic growth: new evidence. *Emerging Markets Finance and Trade*, *51*(6), 1276-1291.

Borensztein, E., De Gregorio, J., & Lee, J. W. (1998). How does foreign direct investment affect economic growth?. *Journal of international Economics*, *45*(1), 115-135.

Bredahl Kock, A., & Teräsvirta, T. (2016). Forecasting macroeconomic variables using neural network models and three automated model selection techniques. *Econometric Reviews*, *35*(8-10), 1753-1779.

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate behavioral research*, *24*(4), 445-455.

Browne, M. W., & Rock, D. A. (1978). The choice of additive constants in ridge regression. *South African Statistical Journal*, *12*(1), 65-74.

Buchanan, B. G., Le, Q. V., & Rishi, M. (2012). Foreign direct investment and institutional quality: Some empirical evidence. *International Review of financial analysis*, *21*, 81-89.

Calderón, C., & Liu, L. (2003). The direction of causality between financial development and economic growth. *Journal of development economics*, *72*(1), 321-334.

Calero, C., Bedi, A. S., & Sparrow, R. (2009). Remittances, liquidity constraints and human capital investments in Ecuador. *World Development*, *37*(6), 1143-1154.

Campos, J., & Ericsson, N. R. (1999). Constructive data mining: Modeling consumers' expenditure in Venezuela. *The econometrics journal*, *2*(2), 226-240.

Campos, J., Ericsson, N. R., & Hendry, D. F. (2005). General-to-specific modeling: an overview and selected bibliography. *FRB International Finance Discussion Paper*, (838).

Campos, J., Hendry, D. F., & Krolzig, H. M. (2003). Consistent model selection by an automatic Gets approach. *Oxford Bulletin of Economics and Statistics*, *65*, 803-819.

Cassel, G. (1967). Capital and income in the money economy. *The Theory of Social Economy*, 51-63.

Castle, J. (2004). Evaluating PcGets and RETINA as Automatic Model Selection Algorithms.

Castle, J. L. (2005). Evaluating PcGets and RETINA as automatic model selection algorithms. *Oxford Bulletin of Economics and Statistics*, *67*, 837-880.

Castle, J. L., & Hendry, D. F. (2005). Extending the boundaries of PcGets: Non-linear models. *Manuscript, Economics Department, University of Oxford*.

Castle, J. L., Clements, M. P., & Hendry, D. F. (2013). Forecasting by factors, by variables, by both or neither?. *Journal of Econometrics*, *177*(2), 305-319.

Castle, J. L., Doornik, J. A., & Hendry, D. F. (2011). Evaluating automatic model selection. *Journal of Time Series Econometrics*, *3*(1).

Castle, J. L., Doornik, J. A., & Hendry, D. F. (2011). Evaluating automatic model selection. *Journal of Time Series Econometrics*, *3*(1).

Castle, J. L., Doornik, J. A., & Hendry, D. F. (2012). Model selection when there are multiple breaks. *Journal of Econometrics*, *169*(2), 239-246.

Castle, J. L., Fawcett, N. W., & Hendry, D. F. (2009). Nowcasting is not just contemporaneous forecasting. *National Institute Economic Review*, *210*(1), 71-89.

Castle, J. L., Qin, X., & Reed, W. R. (2009). *How to pick the best regression equation: A review and comparison of model selection algorithms* (No. 09/13).

Cavallo, M., & Ghironi, F. (2002). Net foreign assets and the exchange rate: Redux revived. *Journal of Monetary Economics*, *49*(5), 1057-1097.

Cetin, M. C., & Aydin, E. R. A. R. (2002). Variable selection with Akaike information criteria: a comparative study. *Hacettepe Journal of Mathematics and Statistics*, *31*, 89-97.

Chandra, R., & Sandilands, R. J. (2005). Does modern endogenous growth theory adequately represent Allyn Young?. *Cambridge Journal of Economics*, *29*(3), 463-473.

Chang, R., Kaltani, L., & Loayza, N. V. (2009). Openness can be good for growth: The role of policy complementarities. *Journal of development economics*, *90*(1), 33-49.

Chatfield, C. (1995). Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *158*(3), 419-444.

Cheshire, P., & Magrini, S. (2000). Endogenous processes in European regional growth: convergence and policy. *Growth and Change*, *31*(4), 455-479.

Chia, W. M., Jinjarak, Y., Rana, P., & Xie, T. (2014). Net foreign assets and macroeconomic volatility. *Journal of Asian Economics*, *34*, 42-53.

Chirwa, T. G., & Odhiambo, N. M. (2018). Exogenous and endogenous growth models: a critical review. *Comparative Economic Research*, *21*(4), 63-84.

Choong, C. K. (2009). Venus khim-sen Liew Ph. D,(2009)"Impact of foreign direct investment volatility on economic growth of asean-5 countries". *Economics Bulletin*, *29*(3), 1829-1841.

Chowdhury, M. (2016). Financial development, remittances and economic growth: Evidence using a dynamic panel estimation. *Margin: The Journal of Applied Economic Research*, *10*(1), 35-54.

Combes, J. L., Kinda, T., Ouedraogo, R., & Plane, P. (2019). Financial flows and economic growth in developing countries. *Economic Modelling*, *83*, 195-209.

Crafts, N. (1996). *Endogenous growth: Lessons for and from economic history* (No. 1333). CEPR Discussion Papers.

Crafts, N. F. (1995). Exogenous or endogenous growth? The industrial revolution reconsidered. *The Journal of Economic History*, *55*(4), 745-772.

Cragg, J. G. (1997). Using higher moments to estimate the simple errors-in-variables model. *Rand Journal of Economics*, S71-S91

Cull, R., Demirgüç-Kunt, A., & Lyman, T. (2012). Financial inclusion and stability: What does research show?.

De Vita, G., & Kyaw, K. S. (2009). Growth effects of FDI and portfolio investment flows to developing countries: a disaggregated analysis by income levels. *Applied economics letters*, *16*(3), 277-283.

Deprez, J., & Milberg, W. S. (1990). Cycle and trend in the dynamics of advanced market economies. *Journal of Economic Issues*, *24*(2), 513-521.

Domar, E. D. (1946). Capital expansion, rate of growth, and employment. *Econometrica, Journal of the Econometric Society*, 137-147.

Domar, E. D. (1949). Capital accumulation and the end of prosperity. *Econometrica: Journal of the Econometric Society*, 307-314.

Dooley, M. P., Folkerts-Landau, D., & Garber, P. M. (2004). *The US current account deficit and economic development: collateral for a total return swap* (No. w10727). National Bureau of Economic Research.

Doornik, J. A. (1999). Object-Oriented Matrix Programming using Ox, 3rd edn, London: Timberlake Consultants Ltd. *See http://www. nuff. ox. ac. uk/Users/Doornik*.

Doornik, J. A. (2007). Econometric model selection with more variables than observations. Working paper.

Doornik, J. A. (2008). Encompassing and automatic model selection. *Oxford Bulletin of Economics and Statistics*, *70*, 915-925.

Doornik, J. A. (2009). Autometrics. *Castle, and Shephard (2009)*, 88-121.

Doornik, J. A., & Hansen, H. (1994). *A practical test for univariate and multivariate normality*. Discussion paper, Nuffield College.

Doppelhofer, G., Hansen, O. P., & Weeks, M. (2016). Determinants of long-term economic growth redux: A measurement error model averaging (MEMA) approach. *NHH Dept. of Economics Discussion Paper*, (19).

Dowrick, S., & Rogers, M. (2002). Classical and technological convergence: Beyond the Solow-Swan growth model. *Oxford Economic Papers*, *54*(3), 369-385.

Edison, H. J., Levine, R., Ricci, L., & Sløk, T. (2002). International financial integration and economic growth. *Journal of international money and finance*, *21*(6), 749-776.

Edwards, S. (1992). Trade orientation, distortions and growth in developing countries. *Journal of development economics*, *39*(1), 31-57.

Edwards, S. (1993). Openness, trade liberalization, and growth in developing countries. *Journal of economic Literature*, *31*(3), 1358-1393.

Ekanayake, E. M., & Ledgerwood, J. R. (2010). How does foreign direct investment affect growth in developing countries? An empirical investigation. *The International Journal of Business and Finance Research*, *4*(3), 43-53.

Engel, C. (2000). *Comments on Obstfeld and Rogoff's" The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?"* (No. w7818). National Bureau of Economic Research.

Ericsson, N. R., & Kamin, S. B. (2009). Constructive data mining: Modelling argentine broad money demand. *Chapter*, *17*, 412-439.

Ericsson, N., Campos, J., & Hendry, D. (2005). *General-to-specific modeling: an overview and selected bibliography* (No. 838). Board of Governors of the Federal Reserve System (US).

Ewubare, D. B., & Ogbuagu, A. R. (2015). Capital Accumulation and Economic Growth in Nigeria “Endogenous Growth Approach”. *IOSR Journal of Economics and Finance*, *6*(6), 49-64.

Faden, V., & Bobko, P. (1982). Validity shrinkage in ridge regression: A simulation study. *Educational and Psychological Measurement*, *42*(1), 73-85.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, *96*(456), 1348-1360.

FDI may have a positive link between it and economic growth (Lensink and Morrissey, 2006)

Firth, L., & Mellor, D. (2000). Learning and the new growth theories: policy dilemma. *Research Policy*, *29*(9), 1157-1163.

Fraley, C., & Raftery, A. E. (1999). MCLUST: Software for model-based cluster analysis. *Journal of classification*, *16*(2), 297-306.

Freckleton, M., Wright, A., & Craigwell, R. (2012). Economic growth, foreign direct investment and corruption in developed and developing countries. *Journal of economic studies*.

Furnival, G. M., & Wilson, R. W. (2000). Regressions by leaps and bounds. *Technometrics*, *42*(1), 69-79.

Ghirmay, T. (2004). Financial development and economic growth in Sub-Saharan African countries: evidence from time series analysis. *African Development Review*, *16*(3), 415-432.

Ghironi, F. (2006). Macroeconomic interdependence under incomplete markets. *Journal of International Economics*, *70*(2), 428-450.

Gong, G., Greiner, A., & Semmler, W. (2004). Endogenous growth: estimating the Romer model for the US and Germany. *Oxford Bulletin of Economics and Statistics*, *66*(2), 147-164.

Granger, C. W., & Hendry, D. F. (2005). A dialogue concerning a new instrument for econometric modeling. *Econometric Theory*, 278-297.

Grossman, G. M., & Helpman, E. (1990). Trade, innovation, and growth. *The American economic review*, *80*(2), 86-91.

Grossman, G. M., & Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European economic review*, *35*(2-3), 517-526.

Gutiérrez, M. A. (2007). *Savings in Latin America after the mid 1990s: determinants, constraints and policies*. ECLAC.

Gutierrez, M., & Solimano, A. (2007). Savings, Investment and Growth in the Global Age: Analytical and Policy Issues. *American University of Paris, Working Paper*, (43).

Hagemann, H. (2009). Solow's 1956 contribution in the context of the Harrod-Domar model. *History of Political Economy*, *41*(Suppl_1), 67-87.

Hall, J., & Ludwig, U. (2010). Neo-liberalism, the changing German labor market, and income distribution: an institutionalist and Post Keynesian analysis. *Journal of Economic Issues*, *44*(2), 345-358.

Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, *41*(2), 190-195.

Hansen, B. E. (2005). Challenges for econometric model selection. *Econometric Theory*, 60-68.

Hansen, P. R., & Timmermann, A. (2012). Choice of sample split in out-of-sample forecast evaluation.

Hanson, G. H. (2007). *Emigration, remittances and labor force participation in Mexico (Working Paper ITD= Documento de Trabajo ITD; 28)*. BID-INTAL.

Harrod, R. F. (1939). An essay in dynamic theory. *The economic journal*, *49*(193), 14-33.

Hassen, S., & Anis, O. (2012). Foreign direct investment (FDI) and economic growth: an approach in terms of cointegration for the case of Tunisia. *Journal of Applied Finance and Banking*, *2*(4), 193.

Haughton, D. M. (1988). On the choice of a model to fit data from an exponential family. *The Annals of Statistics*, *16*(1), 342-355.

Helpman, E. (1992). Endogenous macroeconomic growth theory. *European Economic Review*, *36*(2-3), 237-267.

Hendry, D. F. (2002). Applied econometrics without sinning. *Journal of Economic Surveys*, *16*(4), 591-604.

Hendry, D. F. (2006). A comment on "Specification searches in spatial econometrics: The relevance of Hendry's methodology". *Regional Science and Urban Economics*, *36*(2), 309-312.

Hendry, D. F. (2009). The methodology of empirical econometric modeling: Applied econometrics through the looking-glass. In *Palgrave handbook of econometrics* (pp. 3-67). Palgrave Macmillan, London.

Hendry, D. F. (2009). The methodology of empirical econometric modeling: Applied econometrics through the looking-glass. In *Palgrave handbook of econometrics* (pp. 3-67). Palgrave Macmillan, London.

Hendry, D. F. The Methodology of Empirical Econometric Modeling: Applied Econometrics Through the Looking-Glass.

Hendry, D. F. The Methodology of Empirical Econometric Modeling: Applied Econometrics Through the Looking-Glass. *Palgrave Handbook of Econometrics*, 1.

Hendry, D. F., & Doornik, J. A. (2014). *Empirical model discovery and theory evaluation: automatic selection methods in econometrics*. MIT Press.

Hendry, D. F., & Krolzig, H. (2011). Improving on'Data Mining Reconsidered'by KD Hoover and SJ Perez.

Hendry, D. F., & Krolzig, H. (2011). New Developments in Automatic General-to-Specific Modeling.

Hendry, D. F., & Krolzig, H. M. (1999). Improving on 'Data mining reconsidered'by KD Hoover and SJ Perez. *The econometrics journal*, *2*(2), 202-219.

Hendry, D. F., & Krolzig, H. M. (2001). Automatic Econometric Model Selection Using PcGets London: Timberlake Consultants Ltd.

Hendry, D. F., & Krolzig, H. M. (2001). *Automatic econometric model selection using PcGets 1.0*. Timberlake Consultants Press.

Hendry, D. F., & Krolzig, H. M. (2004). Automatic model selection: a new instrument for social science. *Electoral Studies*, *23*(3), 525-544.

Hendry, D. F., & Krolzig, H. M. (2004). Resolving three 'intractable'problems using a Gets approach. *Unpublished paper, Economics Department, University of Oxford*.

Hendry, D. F., & Krolzig, H. M. (2005). The properties of automatic Gets modelling. *The Economic Journal*, *115*(502), C32-C61.

Hendry, D. F., & Mizon, G. E. (2011). Econometric modelling of time series with outlying observations. *Journal of Time Series Econometrics*, *3*(1).

Hendry, D. F., & Nielsen, B. (2007). *Econometric modeling: a likelihood approach*. Princeton University Press.

Hendry, D. F., & Richard, J. F. (1987). *Recent developments in the theory of encompassing* (No. 1987022). Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Hendry, D. F., Johansen, S., & Santos, C. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, *23*(2), 317-335.

Hendry, D. F., Pagan, A. R., & Sargan, J. D. (1984). Dynamic specification. *Handbook of econometrics*, *2*, 1023-1100.

Hendry, D., Johansen, S., & Santos, C. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, *23*(2), 337-339.

Hendry, D., Santos, C., & Johansen, S. (2008). Automatic selection on indicators in a fully saturated regression.

Hill, R. C., & Ziemer, R. F. (1982). Small sample performance of the Stein-Rule in non-orthogonal designs. *Economics Letters*, *10*(3-4), 285-292.

Hill, R. C., & Ziemer, R. F. (1984). The risk of general Stein-like estimators in the presence of multicollinearity. *Journal of Econometrics*, *25*(1-2), 205-216.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55-67.

Hoerl, R. W. (1985). Ridge analysis 25 years later. *The American Statistician*, *39*(3), 186-192.

Hoover, K. D., & Perez, S. J. (1997). Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search.

Hoover, K. D., & Perez, S. J. (1999). Data mining reconsidered: encompassing and the general-to-specific approach to specification search. *The econometrics journal*, *2*(2), 167-191.

Hoover, K. D., & Perez, S. J. (2000). Truth and Robustness in Cross-country Growth Regressions.

Hoover, K. D., & Perez, S. J. (2004). Truth and robustness in cross-country growth regressions. *Oxford bulletin of Economics and Statistics*, *66*(5), 765-798.

Hoover, K., & Perez, S. J. (1997). *Data Mining Reconsidered: Encompassing and the General-To-Specific Approach to Specification Search* (No. 97-27). Working Paper.

Hoover, K., & Perez, S. J. (2003). *Data Mining Reconsidered: Encompassing And The General-To-Specific Approach To Specification Search* (No. 200).

Huchet, M., Le Mouël, C. L. M., & Vijil, M. (2011, August). The relationship between trade openness and economic growth: some new insights on the openness measurement issue.

Hunt, S. D. (1997). Evolutionary economics, endogenous growth models, and resource-advantage theory. *Eastern Economic Journal*, *23*(4), 425-439.

Hussein, K., & Thirlwall, A. P. (2000). The AK model of "new" growth theory is the Harrod-Domar growth equation: investment and growth revisited. *Journal of Post Keynesian Economics*, *22*(3), 427-435.

Iheanacho, E. (2016). The impact of financial development on economic growth in Nigeria: An ARDL analysis. *Economies*, *4*(4), 26.

Javid, M., Arif, U., & Qayyum, A. (2012). Impact of remittances on economic growth and poverty. *Academic Research International*, *2*(1), 433.`

Jennrich, R. I., & Oman, S. D. (1986). How much does Stein estimation help in multiple linear regression?. *Technometrics*, *28*(2), 113-121.

Johansen, L. (1959). Substitution versus fixed production coefficients in the theory of economic growth: a synthesis. *Econometrica: Journal of the Econometric Society*, 157-176.

Johnson, A. (2006). The effects of FDI inflows on host country economic growth. *The Royal Institute of technology*. *Centre of Excellence for studies in Science and Innovation http://www. infra. kth. se/cesis/research/publications/working papers*.

Johnsson, T. (1992). A procedure for stepwise regression analysis. *Statistical Papers*, *33*(1), 21-29.

Jones, C. I. (2019). Paul Romer: Ideas, nonrivalry, and endogenous growth. *The Scandinavian Journal of Economics*, *121*(3), 859-883.

Jones, H. G. (1975). *An introduction to modern theories of economic growth*. London: Nelson.

Jones, T. A. (1972). Multiple regression with correlated independent variables. *Journal of the International Association for Mathematical Geology*, *4*(3), 203-218.

Jongwanich, J. (2007). *Workers' remittances, economic growth and poverty in developing Asia and the Pacific countries* (No. WP/07/01). United Nations Economic and Social Commission for Asia and the Pacific (ESCAP).

Kaminsky, G. L., Reinhart, C. M., & Végh, C. A. (2004). When it rains, it pours: procyclical capital flows and macroeconomic policies. *NBER macroeconomics annual*, *19*, 11-53.

Karagöz, K. (2009). WORKERS'REMITTANCES AND ECONOMIC GROWTH: EVIDENCE FROMTURKEY. *Journal of Yaşar University*, *4*(13), 1891-1908.

Karlsson, P. S., Behrenz, L., & Shukur, G. (2019). Performances of Model Selection Criteria When Variables are Ill Conditioned. *Computational Economics*, *54*(1), 77-98.

Kennedy, P. E. (2002). Sinning in the basement: What are the rules? The ten commandments of applied econometrics. *Journal of Economic Surveys*, *16*(4), 569-589.

Ketkar, K. W. (1993). Public sector banking, efficiency and economic growth in India. *World Development*, *21*(10), 1685-1697.

Keynes, M. (Ed.). (1975). *Essays on John Maynard Keynes*. Cambridge University Press.

Kim, D. H., & Lin, S. C. (2009). Trade and growth at different stages of economic development. *Journal of Development Studies*, *45*(8), 1211-1224.

King, R. G., & Levine, R. (1993). Finance and growth: Schumpeter might be right. *The quarterly journal of economics*, *108*(3), 717-737.

Khalaf, G., & Shukur, G. (2005). Choosing ridge parameter for regression problems.

Koopman, S. J., Harvey, A. C., Doornik, J. A., & Shephard, N. (2000). STAMP 6.0: Structural time series analyser, modeller and predictor. *London: Timberlake Consultants*.

Kopf, D. A. (2007). Endogenous growth theory applied: Strategies for university R&D. *Journal of Business Research*, *60*(9), 975-978.

Kose, M. A., Prasad, E. S., & Terrones, M. E. (2006). How do trade and financial integration affect the relationship between growth and volatility?. *Journal of international Economics*, *69*(1), 176-202.

Krolzig, H. M. (2003). General-to-specific model selection procedures for structural vector autoregressions. *Oxford Bulletin of Economics and Statistics*, *65*, 769-801.

Krolzig, H. M. (2003). *General-to-Specific Model Selection Procedures for Structural Vector Autoregressions* (No. 2003-W15). Economics Group, Nuffield College, University of Oxford.

Krolzig, H. M., & Hendry, D. F. (2001). Computer automation of general-to-specific model selection procedures. *Journal of Economic Dynamics and Control*, *25*(6-7), 831-866.

Krolzig, H. M., & Hendry, D. F. (2005). The Properties of Automatic GETS Modelling. *The Economic Journal*, *115*, C32-C61.

Lane, P. R., & Milesi-Ferretti, G. M. (2001). The external wealth of nations: measures of foreign assets and liabilities for industrial and developing countries. *Journal of international Economics*, *55*(2), 263-294.

Lane, P. R., & Milesi-Ferretti, G. M. (2002). External wealth, the trade balance, and the real exchange rate. *European Economic Review*, *46*(6), 1049-1071.

Lane, P. R., & Milesi-Ferretti, G. M. (2002). Long-run determinants of the Irish real exchange rate. *Applied Economics*, *34*(5), 549-553.

Lane, P. R., & Milesi-Ferretti, G. M. (2003). International financial integration. *IMF Staff Papers*, *50*(1), 82-113.

Lane, P. R., & Milesi-Ferretti, G. M. (2007). The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities, 1970–2004. *Journal of international Economics*, *73*(2), 223-250.

Lane, P. R., & Milesi-Ferretti, G. M. (2008). International investment patterns. *The Review of Economics and Statistics*, *90*(3), 538-549.

Langlois, R. N. (2001). Knowledge, consumption, and endogenous growth. In *Escaping satiation* (pp. 97-113). Springer, Berlin, Heidelberg.

Lartey, E. K., Mandelman, F. S., & Acosta, P. A. (2012). Remittances, exchange rate regimes and the Dutch disease: A panel data analysis. *Review of International Economics*, *20*(2), 377-395.

Lavezzi, A. (2003). Division of labour and economic growth: Paul Romer's contribution in an historical perspective. *Old and New Growth Theories. An Assessment, Cheltenham: Edward Elgar Publishing*.

Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review*, *73*(1), 31-43.

Leamer, E. E. (1983). Model choice and specification analysis. *Handbook of econometrics*, *1*, 285-330.

Leamer, E. E. (1985). Sensitivity analyses would help. *The American Economic Review*, *75*(3), 308-313.

Leamer, E. E. (1991). Let's Take the Con out of Econometrics. *Modelling Economic Series: Readings in Econometric Methodology*, 29.

Lederman, D., & Maloney, W. (2002). *Open questions about the link between natural resources and economic growth: Sachs and Warner revisited* (Vol. 141). Banco Central de Chile.

Leeb, H., & Poetscher, B. M. (2000). *The Finite-Sample Distribution of Post-Model-Selection Estimators, and Uniform Versus Non-Uniform Approximations*. University Library of Munich, Germany.

Leeb, H., & Pötscher, B. M. (2003). The finite-sample distribution of post-model-selection estimators and uniform versus nonuniform approximations. *Econometric Theory*, 100-142.

Leeb, H., & Pötscher, B. M. (2005). Model selection and inference: Facts and fiction. *Econometric Theory*, 21-59.

Lensink, R., & Morrissey, O. (2006). Foreign direct investment: Flows, volatility, and the impact on growth. *Review of International Economics*, *14*(3), 478-493.

Leroux, B. G. (1992). Consistent estimation of a mixing distribution. *The Annals of Statistics*, 1350-1360.

Li, X., & Liu, X. (2005). Foreign direct investment and economic growth: an increasingly endogenous relationship. *World development*, *33*(3), 393-407.

Lütkepohl, H. (1985). Comparison of criteria for estimating the order of a vector autoregressive process. *Journal of time series analysis*, *6*(1), 35-52.

Malik, I. A., & Amjad, S. (2013). Foreign direct investment and stock market development in Pakistan. *Journal of International Trade Law and Policy*.

Mallows, C. L. (2000). Some comments on Cp. *Technometrics*, *42*(1), 87-94.

Mamo, F. (2012). Economic Growth and Inflation: A panel data analysis.

Martin, R., & Sunley, P. (1998). Slow convergence? The new endogenous growth theory and regional development. *Economic geography*, *74*(3), 201-227.

Mason, R., & Brown, W. G. (1975). Multicollinearity problems and ridge regression in sociological models. *Social science research*, *4*(2), 135-149.

Masson, P. R., Bayoumi, T., & Samiei, H. (1998). International evidence on the determinants of private saving. *The World Bank Economic Review*, *12*(3), 483-501.

Masson, P. R., Kremers, J., & Horne, J. (1994). Net foreign assets and international adjustment: The United States, Japan and Germany. *Journal of International Money and Finance*, *13*(1), 27-40.

Meinshausen, N., & Yu, B. (2009). Lasso-type recovery of sparse representations for high-dimensional data. *The annals of statistics*, *37*(1), 246-270.

Menyah, K., Nazlioglu, S., & Wolde-Rufael, Y. (2014). Financial development, trade openness and economic growth in African countries: New insights from a panel causality approach. *Economic Modelling*, *37*, 386-394.

Meyer, D., & Shera, A. (2017). The impact of remittances on economic growth: An econometric model. *EconomiA*, *18*(2), 147-155.

Misztal, P. (2011). The Relationship between Savings and Economic Growth in Countries with Different Level of Economic Development (Wspólzaleznosci miedzy oszczednosciami i wzrostem gospodarczym w krajach o róznym poziomie rozwoju gospodarczego). *Finansowy Kwartalnik Internetowy e-Finanse*, *7*(2), 17-29.

Mizon, G. E. (1995). Progressive Modeling of Macroeconomic Time Series The LSE Methodology. In *Macroeconometrics* (pp. 107-180). Springer, Dordrecht.

Moreno, E., Girón, F. J., & Casella, G. (2010). Consistency of objective Bayes factors as the model dimension grows. *The Annals of Statistics*, *38*(4), 1937-1952.

Morrissey, O. (2012). FDI in Sub-Saharan Africa: Few linkages, fewer spillovers. *The European Journal of Development Research*, *24*(1), 26-31.

Mulder, P., De Groot, H. L., & Hofkes, M. W. (2001). Economic growth and technological change: a comparison of insights from a neo-classical and an evolutionary perspective. *Technological Forecasting and Social Change*, *68*(2), 151-171.

Mundry, R., & Nunn, C. L. (2009). Stepwise model fitting and statistical inference: turning noise into signal pollution. *The American Naturalist*, *173*(1), 119-123.

Murinde, V., & Eng, F. S. (1994). Financial development and economic growth in Singapore: demand-following or supply-leading?. *Applied Financial Economics*, *4*(6), 391-404.

Nasir, W. A. (2015). Emerging Theoretical Approach for Sustainable External Debt and Economic Growth. *Malaysian Journal of Business and Economics (MJBE)*.

Ndambendia, H., & Njoupouognigni, M. (2010). Foreign aid, foreign direct investment and economic growth in Sub-Saharan Africa: Evidence from pooled mean group estimator (PMG). *International journal of economics and finance*, *2*(3), 39-45.

Neto, P., Brandão, A., & Cerqueira, A. (2008). The impact of FDI, cross-border mergers and acquisitions, and greenfield investments on economic growth. *The IUP Journal of Business Strategy*, 24-44.

Noy, I., & Nualsri, A. (2007). *What do exogenous shocks tell us about growth theories?* (No. 07-16). Working Paper.

Nussipbekova, G., Bakhtgerey, S., & Sabraliyeva, S. (2015, September). Analysis of the economic growth based on the Solow model: Search of the neoclassical growth model with multiple equilibrium states. In *2015 Twelve International Conference on Electronics Computer and Computation (ICECCO)* (pp. 1-3). IEEE.

Obstfeld, M., & Rogoff, K. (2000). The six major puzzles in international macroeconomics: is there a common cause?. *NBER macroeconomics annual*, *15*, 339-390.

Ohtani, K. (1993). A comparison of the Stein-rule and positive-part Stein-rule estimators in a misspecified linear regression model. *Econometric Theory*, 668-679.

Ohtani, K., & Kozumi, H. (1996). The exact general formulae for the moments and the MSE dominance of the Stein-rule and positive-part Stein-rule estimators. *Journal of econometrics*, *74*(2), 273-287.

Oman, S. D. (1991). Random calibration with many measurements: An application of Stein estimation. *Technometrics*, *33*(2), 187-195.

Pagan, A. (1987). Three Econometric Methodologies: A Critical Appraisal 1. *Journal of Economic surveys*, *1*(1-2), 3-23.

Pagan, A. R. (1991). Three Econometric Methodologies: A Critical Appraisal. *Modelling Economic Series: Readings in Econometric Methodology*, 97.

Palley, T. I. (1996). Growth theory in a Keynesian mode: some Keynesian foundations for new endogenous growth theory. *Journal of Post Keynesian Economics*, *19*(1), 113-135.

Paniccià, R., Piacentini, P., & Prezioso, S. (2013). Total factor productivity or technical progress function? Post-Keynesian insights for the empirical analysis of productivity differentials in mature economies. *Review of political Economy*, *25*(3), 476-495.

Perez Amaral, T., Gallo, G. M., & White Jr, H. L. (2005). A Comparison of Complementary Automatic Modeling Methods: Retina and Pcgets. *Econometric Theory*, *21*(1), 262-277.

Phillips, P. (1995). *Automated Forecasts of Asia-Pacific Economic Activity* (No. 1103). Cowles Foundation for Research in Economics, Yale University.

Phillips, P. (1996). Econometric Model Determination. *Econometrica*, *64*(4), 763-812.

Phillips, P. C. (1995). Automated forecasts of Asia-Pacific economic activity. *Asia Pacific Economic Review*, *1*, 92-102.

Phillips, P. C. (1995). *Automated Forecasts of Asia-Pacific Economic Activity* (No. 1103). Cowles Foundation for Research in Economics, Yale University.

Phillips, P. C. (1996). Econometric model determination. *Econometrica: Journal of the Econometric Society*, 763-812.

Phillips, P. C. (1996). Econometric Model Determination. *Econometrica*, *64*(4), 763-812.

Phillips, P. C. (2003). Laws and limits of econometrics. *The Economic Journal*, *113*(486), C26-C52.

Phillips, P. C. (2003). Laws and limits of econometrics. *The Economic Journal*, *113*(486), 26-26.

Pötscher, B. M. (1991). Effects of model selection on inference. *Econometric Theory*, *7*(2), 163-185.

Pradhan, R. P., Arvin, M. B., Hall, J. H., & Nair, M. (2016). Innovation, financial development and economic growth in Eurozone countries. *Applied Economics Letters*, *23*(16), 1141-1144.

Prasetyia, F. (2013). *The role of government and private sector on economic development in ASEAN 5*.

Prescott, E. C. (1988). Robert M. Solow's neoclassical growth model: An influential contribution to economics. *The Scandinavian Journal of Economics*, *90*(1), 7-12.

Qayyum, A., Javid, M., & Arif, U. (2008). Impact of remittances on economic growth and poverty: evidence from Pakistan.

Rajan, R. G., & Zingales, L. (2003). The great reversals: the politics of financial development in the twentieth century. *Journal of financial economics*, *69*(1), 5-50.

Rajan, R., & Zingales, L. (1998). Financial development and growth. *American Economic Review*, *88*(3), 559-586.

Ramona, S., & Ioana, M. (2014). An investigation of cointegration and causality between investments, exports, openness, industrial production and economic growth: A comparative study for the East European countries. *THE ANNALS OF THE UNIVERSITY OF ORADEA*, *367*.

Rapetti, M. (2011). *Macroeconomic policy coordination in a competitive real exchange rate strategy for development* (No. 2011-09). Working Paper.

Razin, A., & Sadka, E. (1997). *Globalization: Public Economics Policy Perspectives*. Sackler Institute for Economic Studies, Eitan Berglas School of Economics, Tel-Aviv University.

Rehman, N. U. (2016). FDI and economic growth: empirical evidence from Pakistan. *Journal of Economic and Administrative Sciences*.

Rivera-Batiz, L. A., & Romer, P. M. (1991). Economic integration and endogenous growth. *The Quarterly Journal of Economics*, *106*(2), 531-555.

Rodrik, D. (2000). Institutions for high-quality growth: what they are and how to acquire them. *Studies in comparative international development*, *35*(3), 3-31.

Rodrik, D. (2006). Capital Account Liberalization and Growth: Making Sense of the Stylized Facts. In *Remarks at the IMF Center Economic Forum: How Does Capital Account Liberalization Affect Economic Growth*.

Romer, D. (2001). [Is Growth Exogenous? Taking Mankiw, Romer, and Weil Seriously]: Comment. *NBER Macroeconomics Annual*, *16*, 62-70.

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, *98*(5, Part 2), S71-S102.

Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic perspectives*, *8*(1), 3-22.

Romer, P. M. (2015). Mathiness in the theory of economic growth. *American Economic Review*, *105*(5), 89-93.

Sachs, J. D., & Warner, A. (1999). Natural resource intensity and economic growth. *Development policies in natural resource economies*, 13-38.

Sahoo, D., Mathiyazhagan, M. K., & Parida, P. (2002). Is foreign direct investment an engine of growth? Evidence from the Chinese economy. *Savings and Development*, 419-440.

Sala-i-Martin, X. (1990). *Lecture notes on economic growth (II): Five prototype models of endogenous growth* (No. w3564). National Bureau of Economic Research.

Sala-i-Martin, X. (1997). I Just Ran Two Million Regressions. *American Economic Review*, *87*(2), 178-183.

Sala-i-Martin, X. (1997). *I Just Ran Four Million Regressions* (No. 6252). National Bureau of Economic Research, Inc.

Sala-i-Martin, X. (2002). 15 years of new growth economics: what have we learnt?.

Sala-i-Martin, X. X. (1997). *I just ran four million regressions* (No. w6252). National Bureau of Economic Research.

Sala-i-Martin, X. X. (1997). *I just ran four million regressions* (No. w6252). National Bureau of Economic Research.

Sala-i-Martin, X., Doppelhofer, G., & Miller, R. I. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *American economic review*, 813-835.

Samuelson, P. A. (1988). The Keynes-Hansen-Samuelson multiplier-accelerator model of secular stagnation. *Japan and the world economy*, *1*(1), 3-19.

Samuelson, P. A., & Solow, R. M. (1956). A complete capital model involving heterogeneous capital goods. *The Quarterly Journal of Economics*, *70*(4), 537-562.

Santos, C., Hendry, D. F., & Johansen, S. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, *23*(2), 317-335.

Sardadvar, S. (2011). Neoclassical Growth Theory and Standard Models. In *Economic Growth in the Regions of Europe* (pp. 9-22). Physica-Verlag HD.

Sato, R. (1964). The Harrod-Domar model vs the Neo-Classical growth model. *The Economic Journal*, *74*(294), 380-387.

Sato, R. (2006). The stability of the Solow-Swan model with biased technical change. *Biased technical change and economic conservation laws*, 25-30.

Scarfe, B. L. (1977). The Harrod Model and the “Knife Edge” Problem. *Cycles, growth, and inflation: A survey of contemporary macrodynamics*, 63-66.

Scarpello, G. M., & Ritelli, D. (2003). The Solow model improved through the logistic manpower growth law. *Annali dell’Universitá di Ferrara*, *49*(1), 73-83.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461-464.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, *52*(3), 333-343.

Shapiro, M. D., Blanchard, O. J., & Lovell, M. C. (1986). Investment, output, and the cost of capital. *Brookings Papers on Economic Activity*, *1986*(1), 111-164.

Shaw, G. K. (1992). Policy implications of endogenous growth theory. *The Economic Journal*, *102*(412), 611-621.

Siddique, A., Selvanathan, E. A., & Selvanathan, S. (2012). Remittances and economic growth: empirical evidence from Bangladesh, India and Sri Lanka. *Journal of development studies*, *48*(8), 1045-1062.

Simut, R., & Mester, I. (2014). An investigation of cointegration and causality between investments, exports, openness, industrial production and economic growth: A comparative study for the East European countries. *Annals of Faculty of Economics*, *1*(1), 369-378.

Singh, T. (2011). International Trade and Economic Growth Nexus in Australia: A Robust Evidence from Time-Series Estimators. *The World Economy*, *34*(8), 1348-1394.

Singh, T. (2015). Trade openness and economic growth in Canada: An evidence from time-series tests. *Global Economy Journal*, *15*(3), 361-407.

Smith, A. F., & Goldstein, M. (1975). Ridge regression: some comments on a paper of Conniffe and Stone. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *24*(1), 61-66.

Smith, K. C., Savin, N. E., & Robertson, J. L. (1984). A Monte Carlo comparison of maximum likelihood and minimum chi square sampling distributions in logit analysis. *Biometrics*, 471-482.

Solow, R. M. (1994). Perspectives on growth theory. *Journal of economic perspectives*, *8*(1), 45-54.

Solow, R. M. (1999). Neoclassical growth theory. *Handbook of macroeconomics*, *1*, 637-667.

Sorger, G. (2003). On the multi-country version of the Solow-Swan model. *The Japanese Economic Review*, *54*(2), 146-164.

Sorina, M. I., & Mihaela, L. (2014). SUSTAINABLE GROWTH: RECENT TRENDS ACROSS CENTRAL AND EASTERN EUROPEAN ECONOMIES. *Annals of the University of Oradea, Economic Science Series*, *23*(1).

Stark, O., Taylor, J. E., & Yitzhaki, S. (1988). Migration, remittances and inequality: A sensitivity analysis using the extended Gini index. *Journal of Development Economics*, *28*(3), 309-322.

Stoneman, P. (1983). The economic analysis of technological change.

Švigir, M., & Miloš, J. (2017). Relationship between inflation and economic growth; comparative experience of Italy and Austria. *FIP-Financije i pravo*, *5*(2), 91-101.

Swan, T. W. (1956). Economic growth and capital accumulation. *Economic record*, *32*(2), 334-361.

Taylor, J. B., Woodford, M., & Uhlig, H. (Eds.). (1999). *Handbook of macroeconomics*. Elsevier.

Tekin, R. B. (2012). Economic growth, exports and foreign direct investment in Least Developed Countries: A panel Granger causality analysis. *Economic modelling*, *29*(3), 868-878.

Terrones, M., Prasad, E., & Kose, A. (2004). How do trade and financial integration affect the relationship between growth and volatility?. In *Federal Reserve Bank of San Francisco Proceedings* (No. Jun).

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.

Tibshirani, R., & Knight, K. (1999). The covariance inflation criterion for adaptive model selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *61*(3), 529-546.

Tourette, J. E. L. (1964). Technological change and equilibrium growth in the harrod-domar model. *Kyklos*, *17*(2), 207-226.

Van den Berg, H. (2013). Growth theory after Keynes, part I: the unfortunate suppression of the Harrod-Domar model. *Journal of Philosophical Economics*, *7*(1), 2-23.

Van den Berg, H. (2013). Growth theory after Keynes, part I: the unfortunate suppression of the Harrod-Domar model. *Journal of Philosophical Economics*, *7*(1), 2-23.

Veall, M. R. (1992). Bootstrapping the process of model selection: An econometric example. *Journal of Applied Econometrics*, *7*(1), 93-99.

Vinod, H. D. (1978). A ridge estimator whose MSE dominates OLS. *International Economic Review*, 727-737.

Vinod, H. D. (1980). Improved Stein-rule estimator for regression problems. *Journal of Econometrics*, *12*(2), 143-150.

Waheed, A. (2007). Foreign Public Capital and Economic Growth of Developing Countries. *India and South Asia: Economic Developments in the Age of Globalization*, 139.

Wang, H., Li, G., & Tsai, C. L. (2007). Regression coefficient and autoregressive order shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *69*(1), 63-78.

White, H. (1989). Learning in artificial neural networks: A statistical perspective. *Neural computation*, *1*(4), 425-464.

White, H. (2006). Approximate nonlinear forecasting methods. *Handbook of economic forecasting*, *1*, 459-512.

Wolde-Rufael, Y. (2009). Re-examining the financial development and economic growth nexus in Kenya. *Economic Modelling*, *26*(6), 1140-1146.

Wood, S. A., & Crerar, D. A. (1985). A numerical method for obtaining multiple linear regression parameters with physically realistic signs and magnitudes: Applications to the determination of equilibrium constants from solubility data. *Geochimica et Cosmochimica Acta*, *49*(1), 165-172.

Yancey, T. A., & Judge, G. G. (1976). A Monte Carlo comparison of traditional and Stein-rule estimators under squared error loss. *Journal of Econometrics*, *4*(3), 285-294.

Yanikkaya, H. (2003). Trade openness and economic growth: a cross-country empirical investigation. *Journal of Development economics*, *72*(1), 57-89.

Zahonogo, P. (2016). Trade and economic growth in developing countries: Evidence from sub-Saharan Africa. *Journal of African Trade*, *3*(1-2), 41-56.

Zaman, A. (2017). Lessons in Econometric Methodology: The Axiom of Correct Specification. *International Econometric Review*, *9*(2).

Zaman, K., Usman, B., Sheikh, S. M., Khan, A., Kosnin, A. B. M., Rosman, A. S. B., ... & Hishan, S. S. (2019). Managing crime through quality education: A model of justice. *Science & Justice*, *59*(6), 597-605.

Zhang, W. B. (1993). Rent and Residential Distribution in a Two-group Neo-classical Growth Model. *Economic Systems Research*, *5*(4), 395-408.

Zhang, Y., Li, R., & Tsai, C. L. (2010). Regularization parameter selections via generalized information criterion. *Journal of the American Statistical Association*, *105*(489), 312-323.

Zhao, P., & Yu, B. (2006). On model selection consistency of Lasso. *Journal of Machine learning research*, *7*(Nov), 2541-2563.

Zheng, X., & Loh, W. Y. (1997). A consistent variable selection criterion for linear models with high-dimensional covariates. *Statistica Sinica*, 311-325.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, *101*(476), 1418-1429.

Zou, H., Hastie, T., & Tibshirani, R. (2007). On the “degrees of freedom” of the lasso. *The Annals of Statistics*, *35*(5), 2173-2192.

Appendix 1

Packages details

The following software and required packages we have used in simulation as well as real data.

We have used the R version 4.1 in our study.

1.1 Package for Elastics Net

Pakage name: glmnet

Author:

Jerome Friedman [aut], Trevor Hastie [aut, cre], Rob Tibshirani [aut], Balasubramanian Narasimhan [aut], Kenneth Tay [aut], Noah Simon [aut], Junyang Qian [ctb]

Maintainer: Trevor Hastie <hastie at stanford.edu>

URL: https://glmnet.stanford.edu, https://dx.doi.org/10.18637/jss.v033.i01, https://dx.doi.org/10.18637/jss.v039.i05

1.2 Package for Autometrics

Package name: Gets

Title: General-to-Specific (GETS) Modelling and Indicator Saturation

Author Genaro Sucarrat [aut, cre], Felix Pretis [aut], James Reade [aut], Jonas Kurle [ctb], Moritz Schwarz [ctb]

Maintainer Genaro Sucarrat <genaro.sucarrat@bi.no>

Description: Automated General-to-Specific (GETS) modelling of the mean and variance of a regression, and indicator saturation methods for detecting and testing for structural breaks in the mean, see Pretis, Reade and Sucarrat (2018) <doi:10.18637/jss.v086.i03>.

getsFun

Auxiliary function (i.e. not intended for the average user) that enables fast and efficient GETSmodelling with user-specified estimators and models, and user-specified diagnostics and goodnessof-fit criteria. The function is called by and relied upon by getsm, getsv, isat and blocksFun.

1.3 Package for Extreme Bound Analysis

Package name: ExtremeBounds

Title: Extreme Bounds Analysis (EBA)

Version 0.1.6

Date 2018-01-03

Author: Marek Hlavac <mhlavac@alumni.princeton.edu>

Maintainer: Marek Hlavac <mhlavac@alumni.princeton.edu>

Description: An implementation of Extreme Bounds Analysis (EBA), a global sensitivity analysis that examines the robustness of determinants in regression models. The package supports both Leamer's and Sala-i-Martin's versions of EBA, and allows users to customize all aspects of the analysis.

Appendix 2

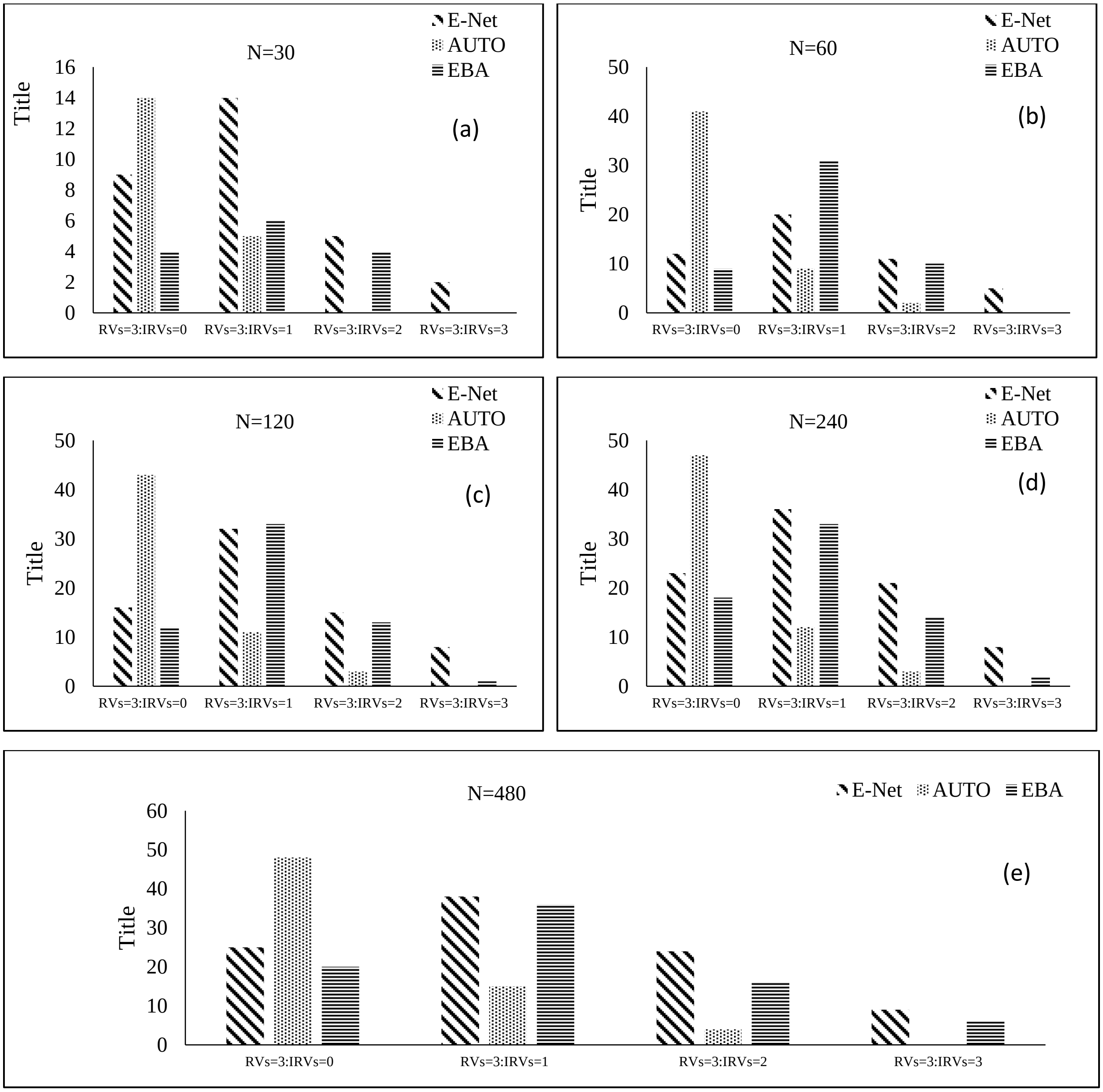
Tables and figures for chapter 4

RELEVANT VARIABLES=3; IRRELEVANT VARIABLES=3

Table 2.1: Relative Frequencies of DGP Orthogonal Variables (%)

	E-Net					Autometrics					EBA				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
DGP	9	12	16	23	25	14	41	43	47	48	4	9	12	18	20
DGP+1_IRV	14	20	32	36	38	5	9	11	12	15	6	31	33	33	36
DGP+2_IRVs	5	11	15	21	24	0	2	3	3	4	4	10	13	14	16
DGP+3_IRVs	2	5	8	8	9	0	0	0	0	0	0	0	1	2	6
Total frequency	**30**	**48**	**71**	**88**	**96**	**19**	**52**	**57**	**62**	**67**	**14**	**50**	**59**	**67**	**78**

IRV=Irrelevant variables



RVs=3 is equal to DGP

Figure 2.1: Relative Frequencies of DGP Orthogonal Variables (%)

The following table shows the relative frequencies of irrelevant variable in case of orthogonal variables DGP. And total number of irrelevant variables in DGP are 3.

Table 2.2: Relative Frequencies of Irrelevant Orthogonal variables (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
IRVs=1	24	66	70	71	75	10	12	13	15	16	21	23	50	60	70
IRVs=2	8	22	25	29	33	0	1	2	3	3	10	23	24	25	27
IRVs=3	2	4	8	8	9	0	0	0	0	0	0	0	1	4	6

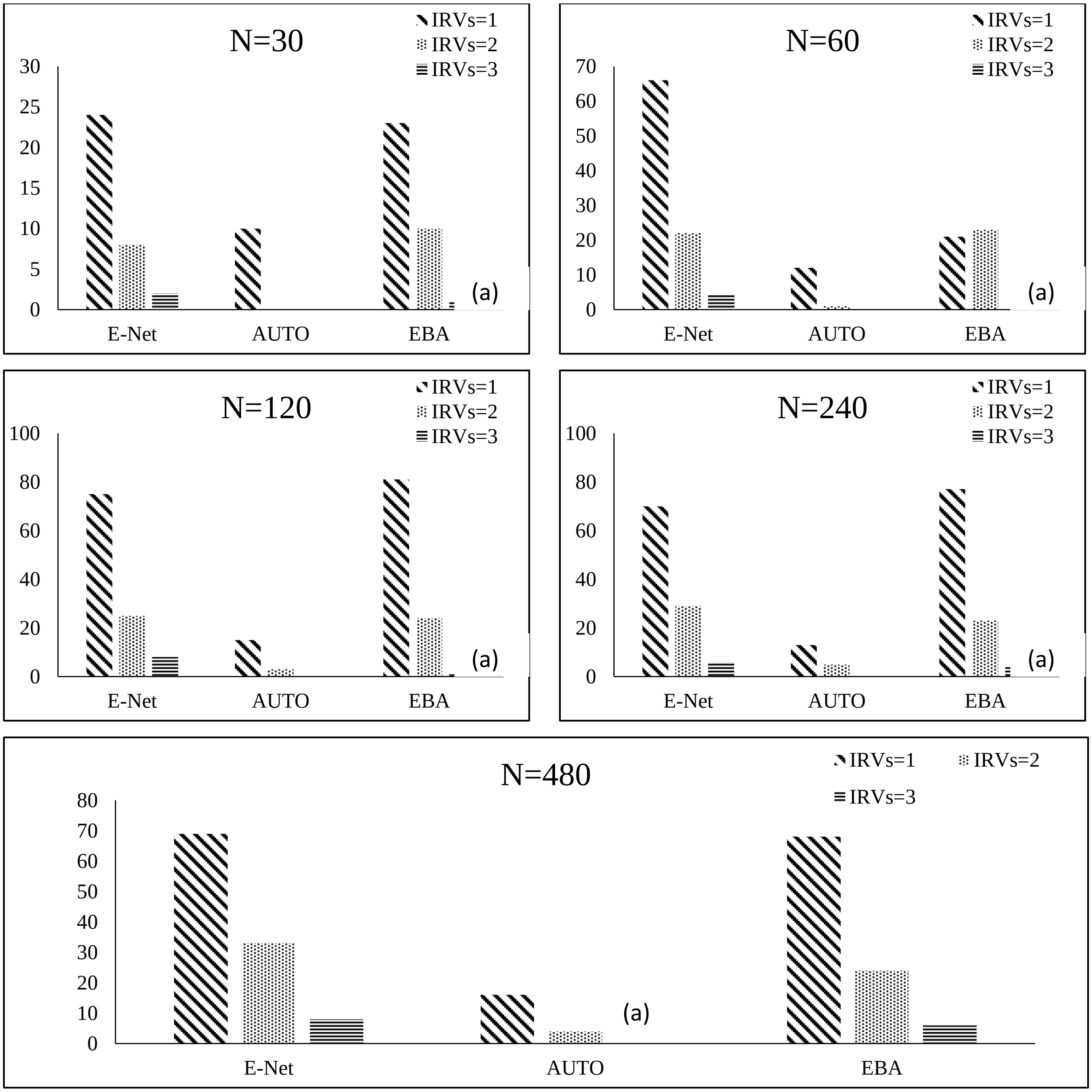


Figure 2.2: Relative Frequencies of Irrelevant Orthogonal variables (%)

The following table shows the relative frequencies of irrelevant variables in case of orthogonal variables DGP. And the total number of irrelevant variables in DGP is 10.

Table 2.3: Relative Frequencies of Irrelevant Orthogonal variables (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
IRVs=1	96	98	98	99	100	35	40	48	51	55	100	100	100	100	100
IRVs=2	88	89	90	98	99	13	16	18	29	37	100	100	100	100	100
IRVs=3	62	68	70	75	77	3	5	7	9	12	98	99	100	100	100
IRVs=4	43	46	48	50	60	1	1	2	2	3	94	97	99	100	100
IRVs=5	23	25	25	29	30	0	0	0	0	0	74	89	91	99	100
IRVs=6	11	12	13	16	19	0	0	0	0	0	45	66	76	85	89
IRVs=7	4	5	5	7	5	0	0	0	0	0	12	43	47	48	50
IRVs=8	1	2	4	4	4	0	0	0	0	0	1	10	15	21	22
IRVs=9	0	0	0	0	0	0	0	0	0	0	0	0	4	6	7
IRVs=10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

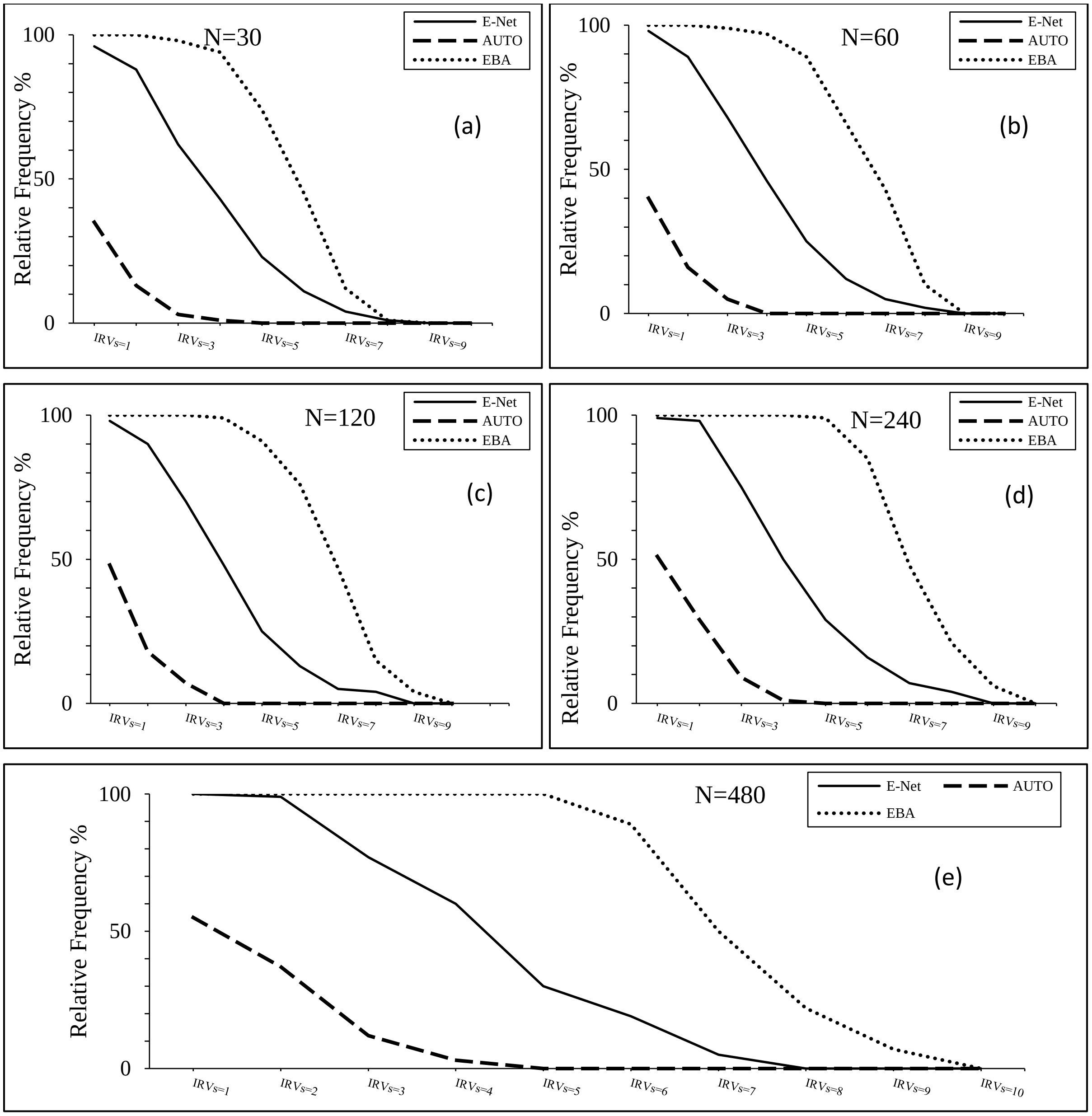


Figure 2.3: Relative Frequencies of Irrelevant Orthogonal variables (%)

4.2 SCENARIO 2: Relevant Variables are Mutually Correlated

The following table shows the relative frequency of DGP when the regressors are mutually correlated

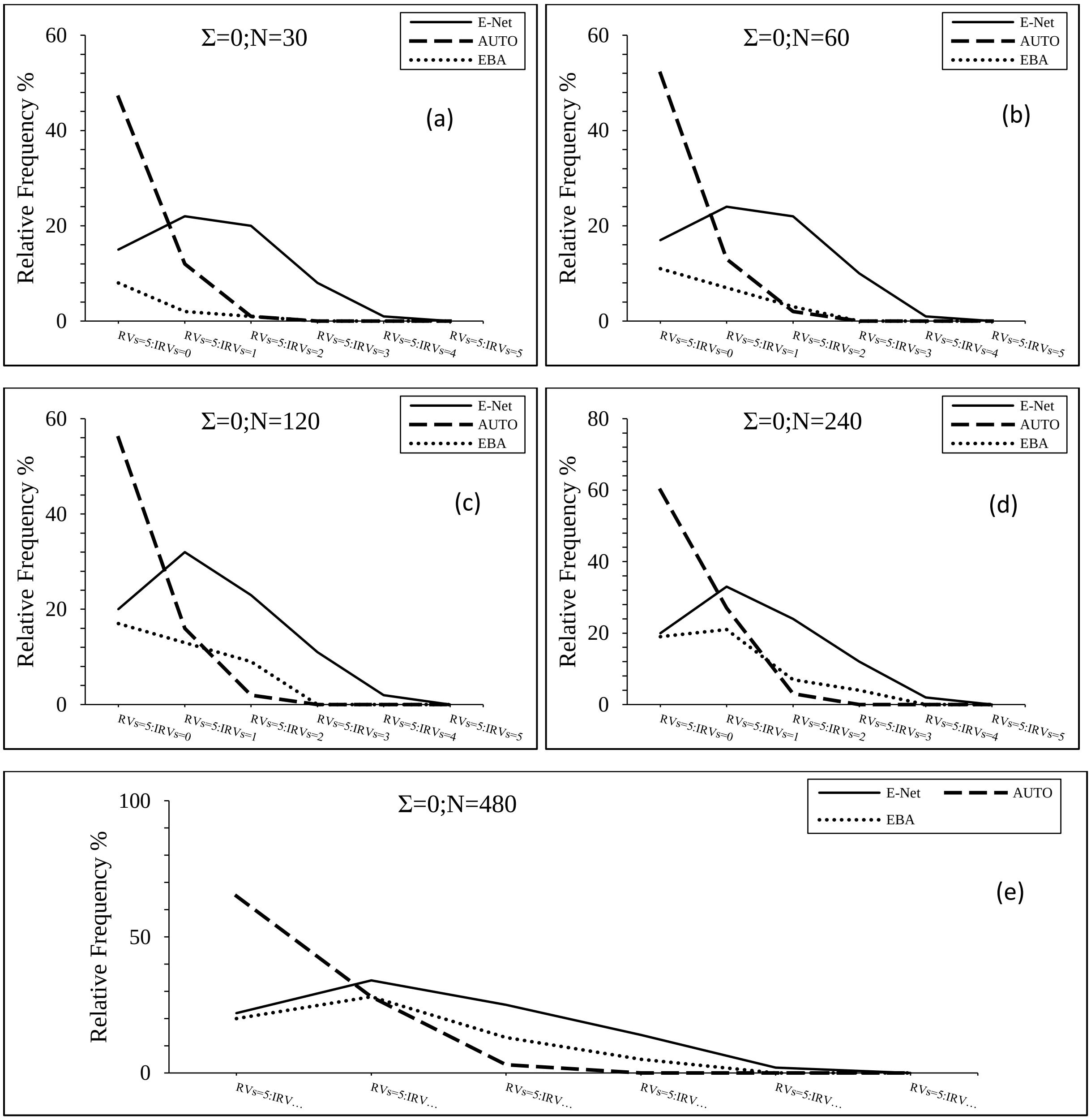
Table 2.4: Relative Frequencies of DGP Non-Orthogonal Variables (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
Σ =0															
DGP	15	17	20	20	22	47	52	56	60	65	8	11	17	19	20
DGP+1_IRV	22	24	32	33	34	12	13	16	27	28	2	7	13	21	28
DGP+2_IRVs	20	22	23	24	25	1	2	2	3	3	1	3	9	7	13
DGP+3_IRVs	8	10	11	12	14	0	0	0	0	0	0	0	0	4	5
DGP+4_IRVs	1	1	2	2	2	0	0	0	0	0	0	0	0	0	0
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total Frequency	**66**	**74**	**88**	**91**	**97**	**60**	**67**	**74**	**90**	**96**	**11**	**21**	**39**	**51**	**66**
Σ =0.1															
DGP	14	16	25	27	28	49	50	54	55	55	16	33	42	38	39
DGP+1_IRV	18	23	29	31	36	12	19	27	29	32	18	21	23	26	27
DGP+2_IRVs	14	15	21	22	24	1	2	2	9	12	2	6	14	15	22
DGP+3_IRVs	2	3	3	4	4	0	0	0	0	0	0	1	1	4	5
DGP+4_IRVs	0	1	1	1	2	0	0	0	0	0	0	0	0	0	0
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total Frequency	**48**	**58**	**79**	**85**	**94**	**62**	**71**	**83**	**93**	**99**	**36**	**61**	**80**	**83**	**93**
Σ =0.25															
DGP	23	24	25	27	29	39	41	42	45	48	45	47	51	52	53
DGP+1_IRV	32	34	35	35	36	14	17	20	23	24	22	24	27	29	30
DGP+2_IRVs	10	14	15	16	18	4	3	4	5	7	2	2	3	9	11
DGP+3_IRVs	4	7	9	12	13	1	1	1	2	3	0	0	0	0	0
DGP+4_IRVs	0	1	1	2	2	0	0	0	0	0	0	0	0	0	0
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total Frequency	**69**	**80**	**85**	**92**	**98**	**58**	**62**	**67**	**75**	**82**	**69**	**73**	**81**	**90**	**94**

Σ =0.5															
DGP	32	34	35	40	48	31	35	39	40	40	58	60	61	62	64
DGP+1_IRV	19	20	21	25	26	8	9	11	18	20	21	22	25	27	30
DGP+2_IRVs	12	14	17	19	20	2	3	3	4	4	2	2	3	3	4
DGP+3_IRVs	1	2	2	2	3	0	0	0	0	0	0	0	0	0	0
DGP+4_IRVs	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
DGP+5_IRVs	0		0	0	0	0	0	0	0	0	0	0	0	0	0
Total Frequency	**64**	**70**	**76**	**87**	**98**	**41**	**47**	**53**	**62**	**64**	**81**	**84**	**89**	**92**	**98**
Σ =0.75															
DGP	33	35	36	37	39	27	29	30	30	31	68	70	72	76	76
DGP+1_IRV	16	21	22	26	28	2	3	5	3	4	9	12	13	15	17
DGP+2_IRVs	10	14	15	16	17	0	3	4	5	7	1	1	2	5	6
DGP+3_IRVs	6	7	8	9	10	0	0	0	0	0	0	0	0	0	0
DGP+4_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total Frequency	**65**	**77**	**81**	**88**	**94**	**29**	**35**	**39**	**38**	**42**	**78**	**83**	**87**	**96**	**99**
Σ =0.9															
DGP	29	30	31	32	34	14	15	16	16	17	66	67	69	70	74
DGP+1_IRV	20	21	22	23	24	2	3	4	5	6	9	15	15	16	18
DGP+2_IRVs	10	11	13	14	15	0	1	1	1	2	0	1	1	2	2
DGP+3_IRVs	5	6	7	9	11	0	0	0	0	0	0	0	0	0	0
DGP+4_IRVs	0	1	1	2	3	0	0	0	0	0	0	0	0	0	0
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total Frequency	**64**	**69**	**74**	**80**	**87**	**16**	**19**	**21**	**23**	**25**	**75**	**83**	**85**	**88**	**94**

Σ=level of correlation

The following figures show the relative frequencies of DGP when the regressors are mutually correlated



(RVs=3) is equal to DGP

Figure 2.4: Relative Frequencies of DGP Non-Orthogonal Variables (%) {Σ =0}

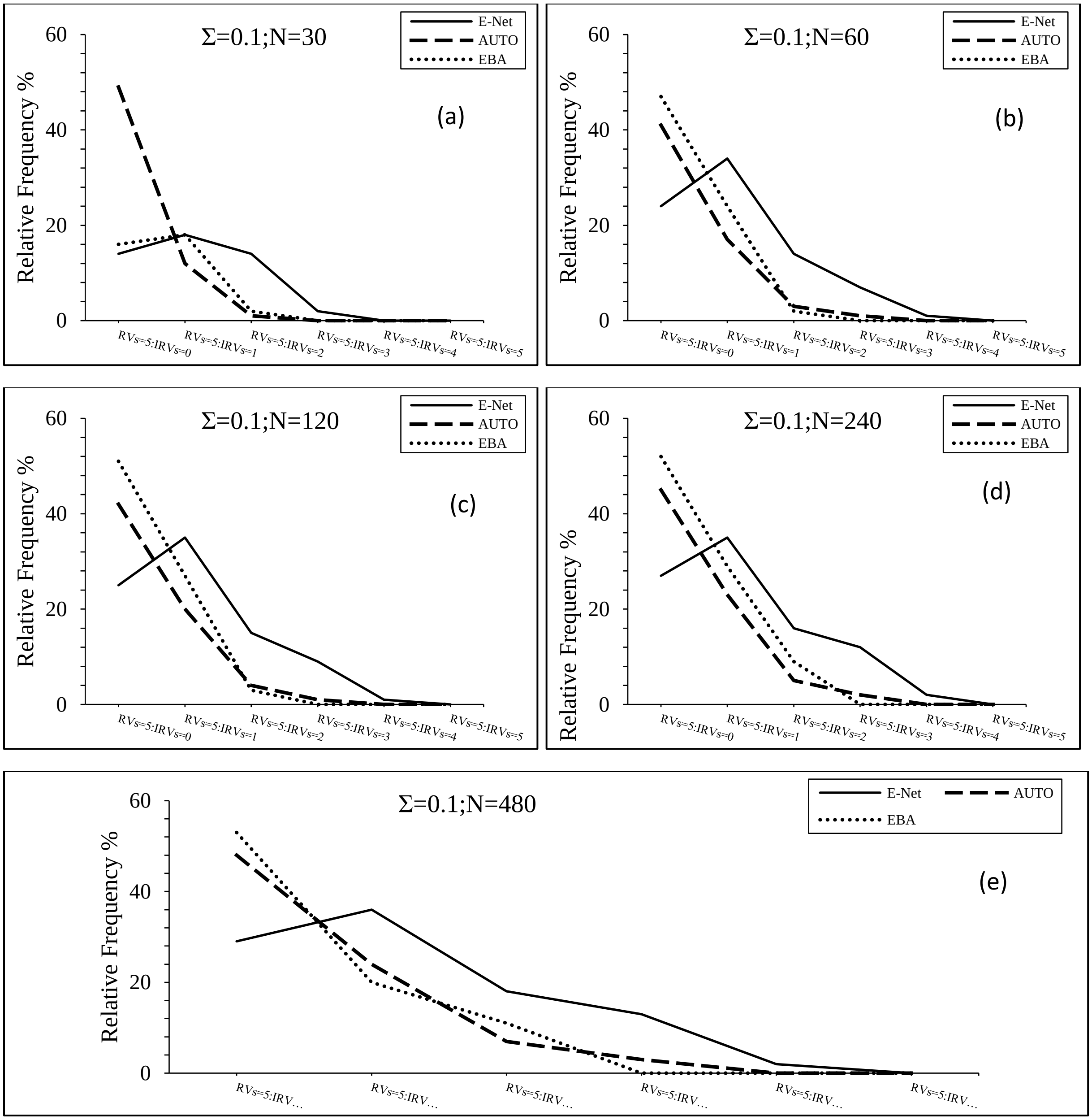


Figure 2.5: Relative Frequencies of DGP Non-Orthogonal Variables (%) {Σ =0.1}

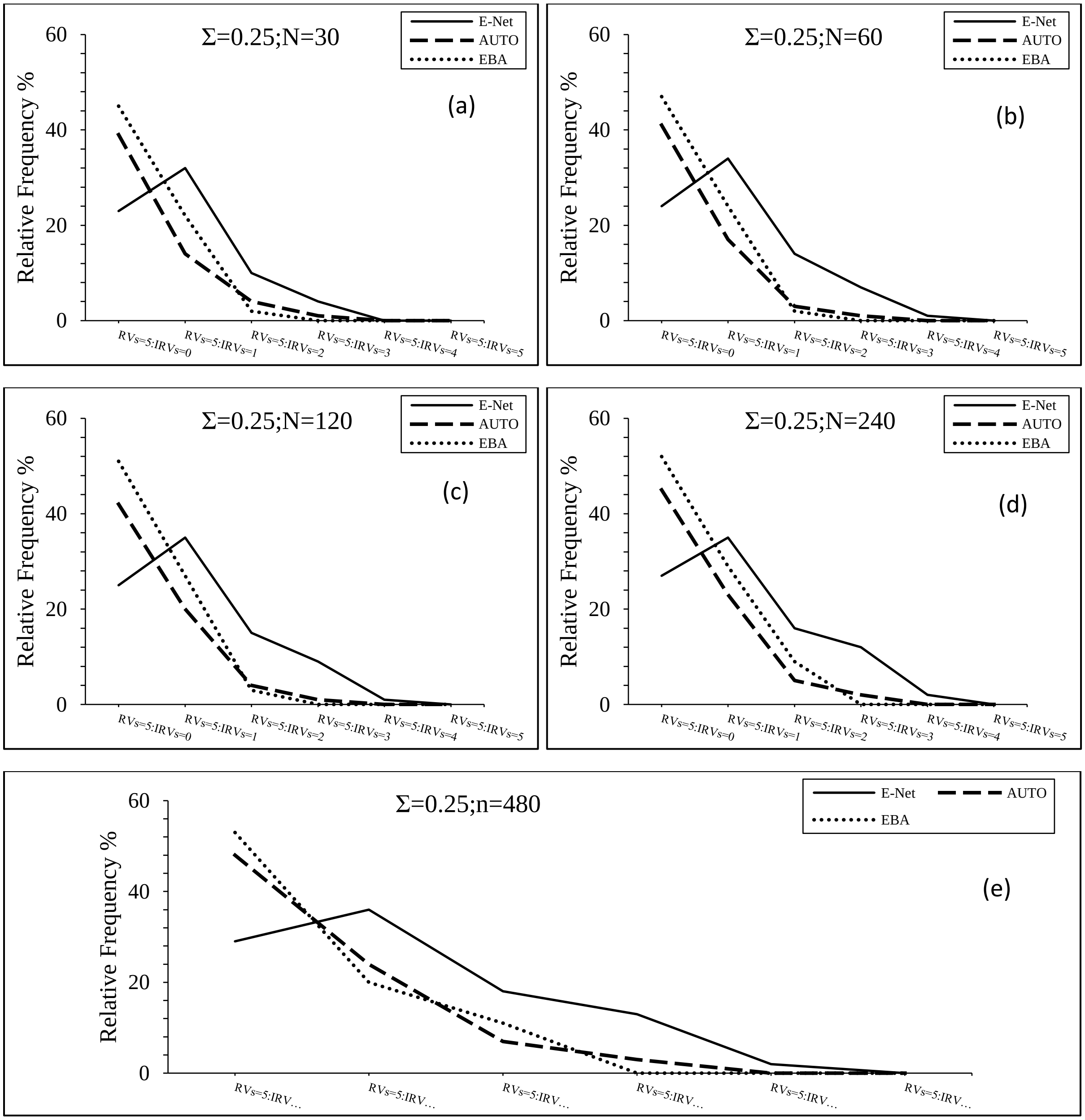


Figure 2.6: Relative Frequencies of DGP Non-Orthogonal Variables (%) {**Σ** =0.25}

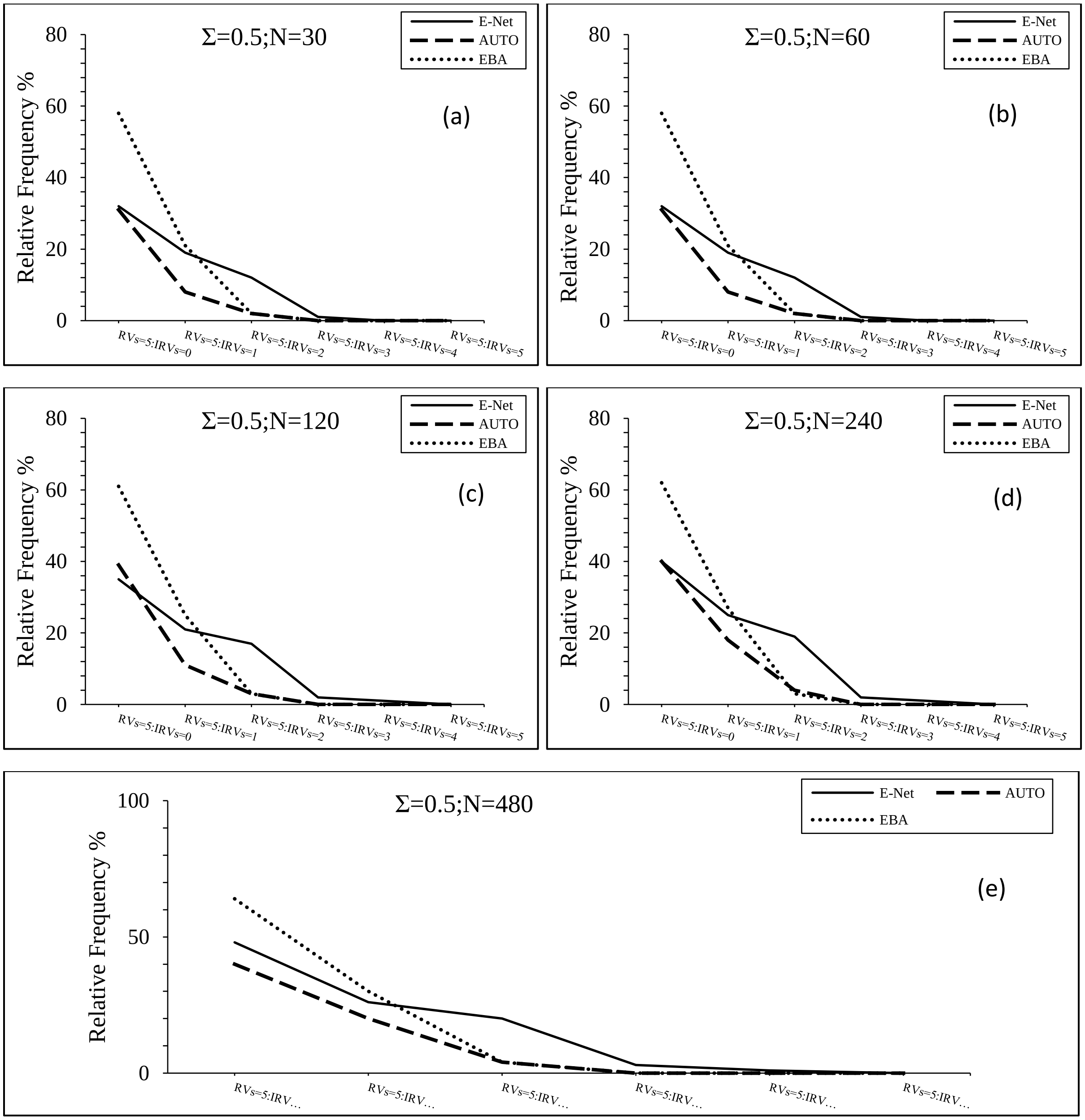


Figure 2.7: Relative Frequencies of DGP Non-Orthogonal Variables (%) {**Σ** =0.5}

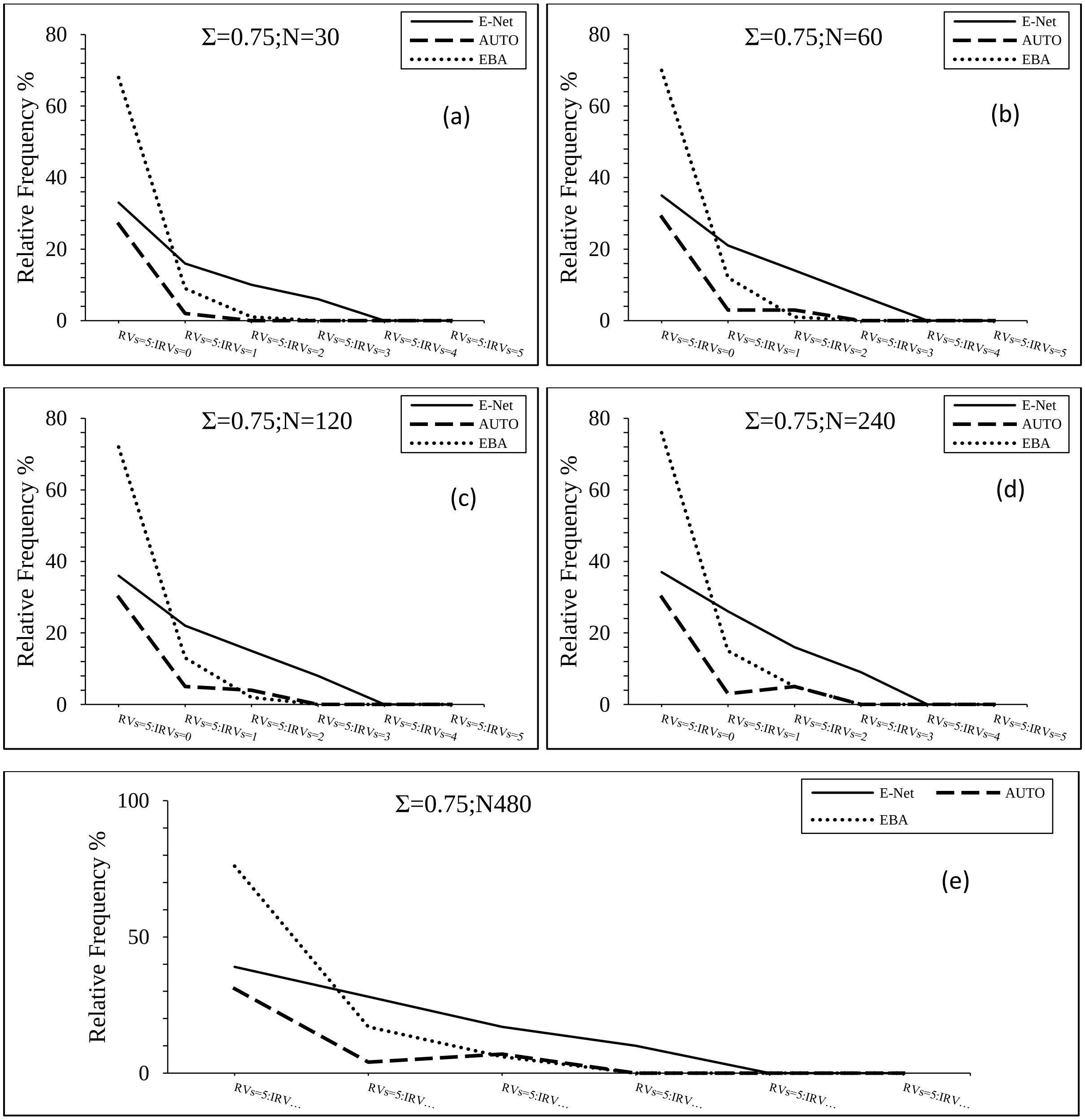


Figure 2.8: Relative Frequencies of DGP Non-Orthogonal Variables (%) {**Σ** =0.75}

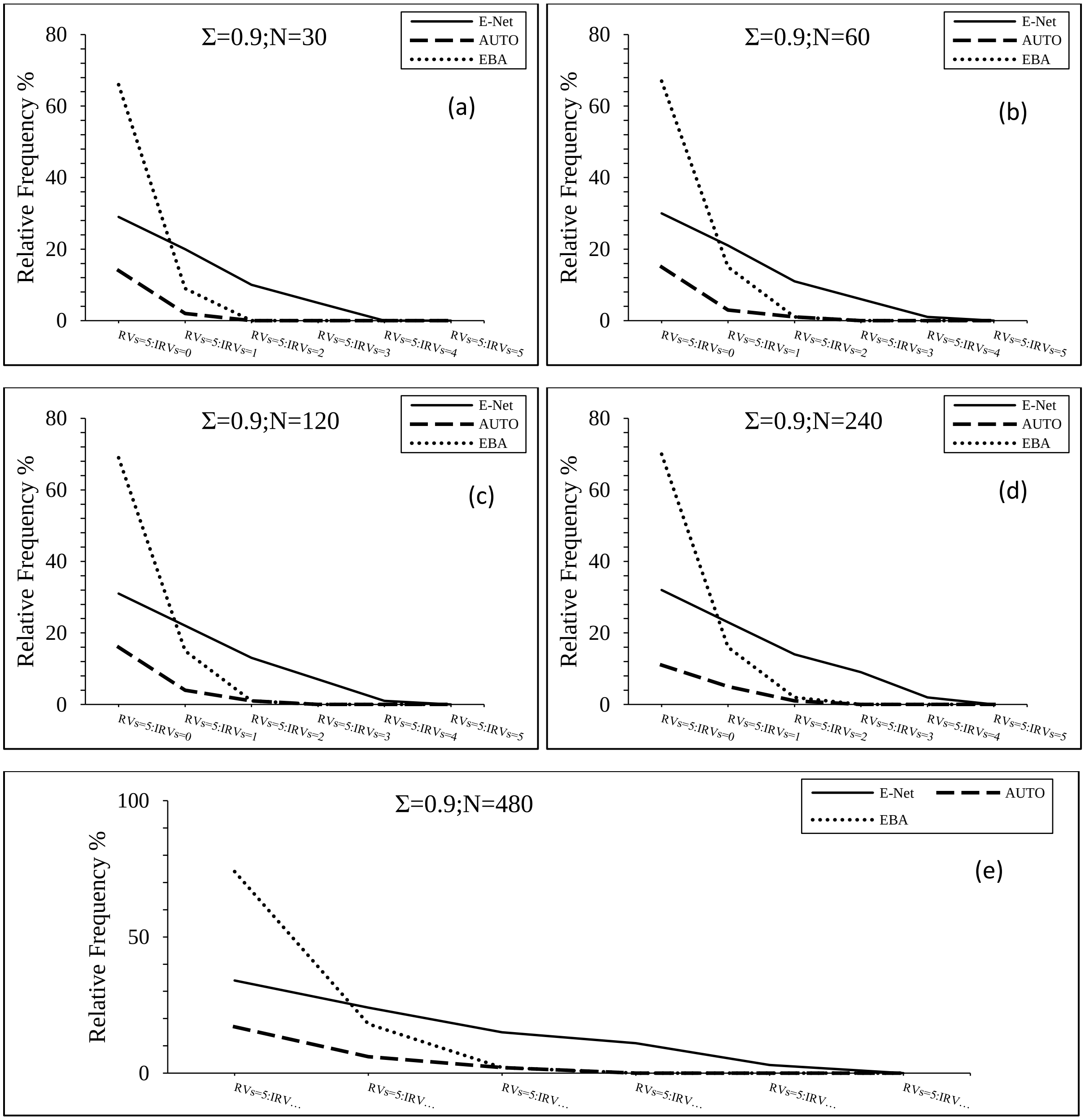


Figure 2.9: Relative Frequencies of DGP Non-Orthogonal Variables (%) {**Σ** =0.9}

4.3 SCENARIO 3: Irrelevant Variables are Correlated to Relevant Variable

The following figure shows the frequencies of the irrelevant variable to the relevant variables

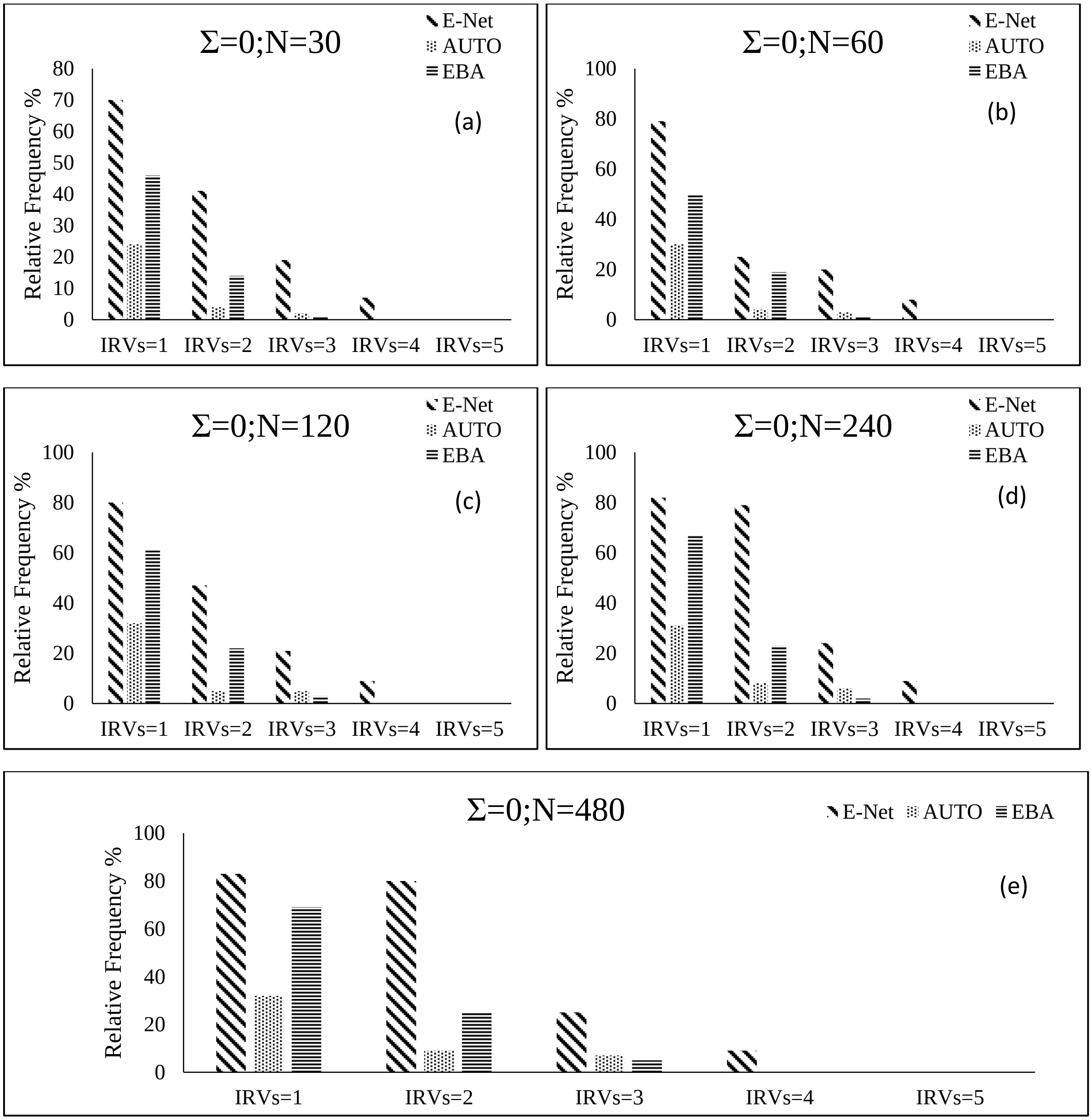


Figure 2.10: Frequencies of Irrelevant variables which are correlated to relevant variables (%) (%) {Σ =0}

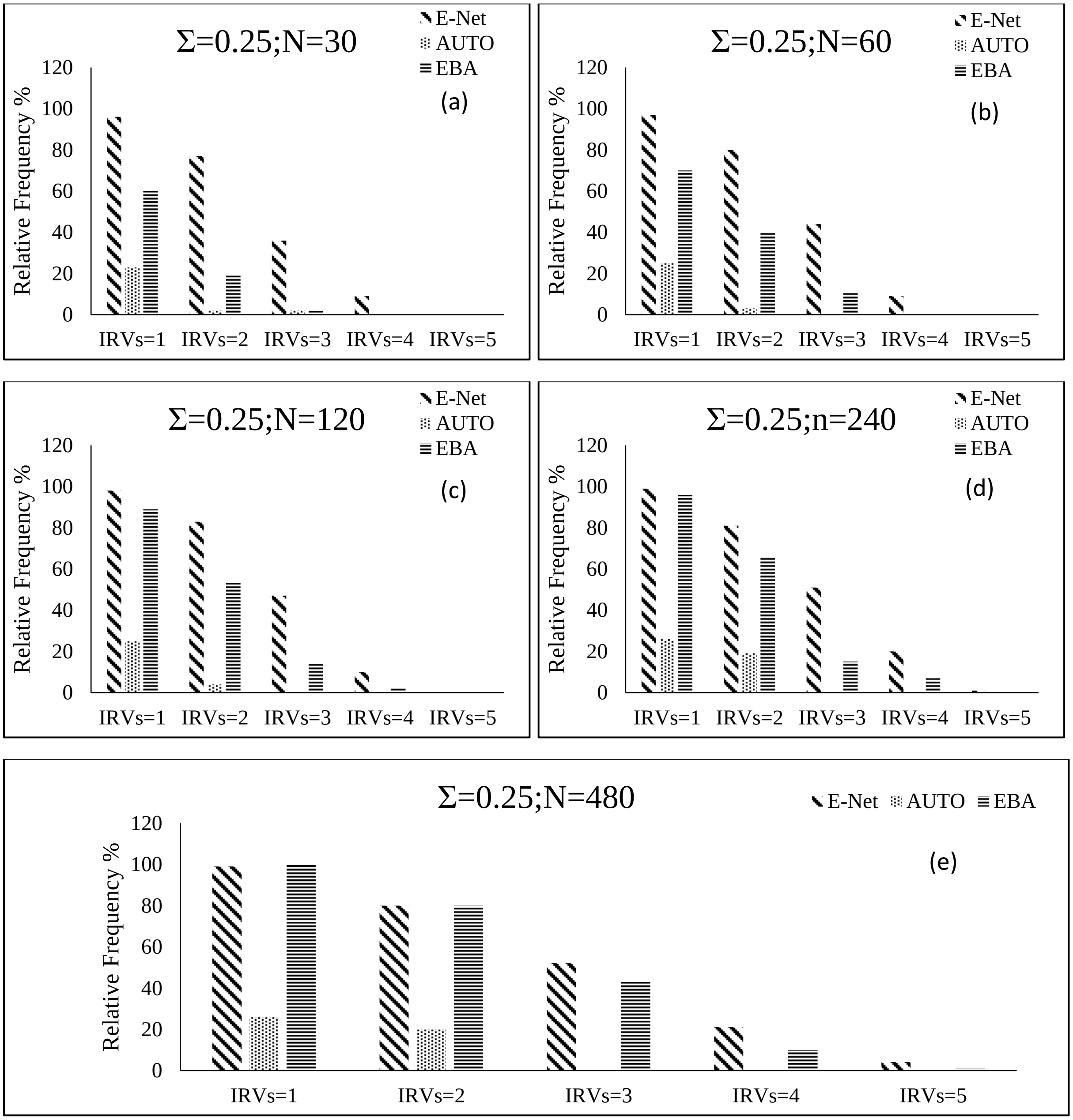


Figure 2.11: Frequencies of Irrelevant variables which are correlated to relevant variables (%) {Σ =0.25}

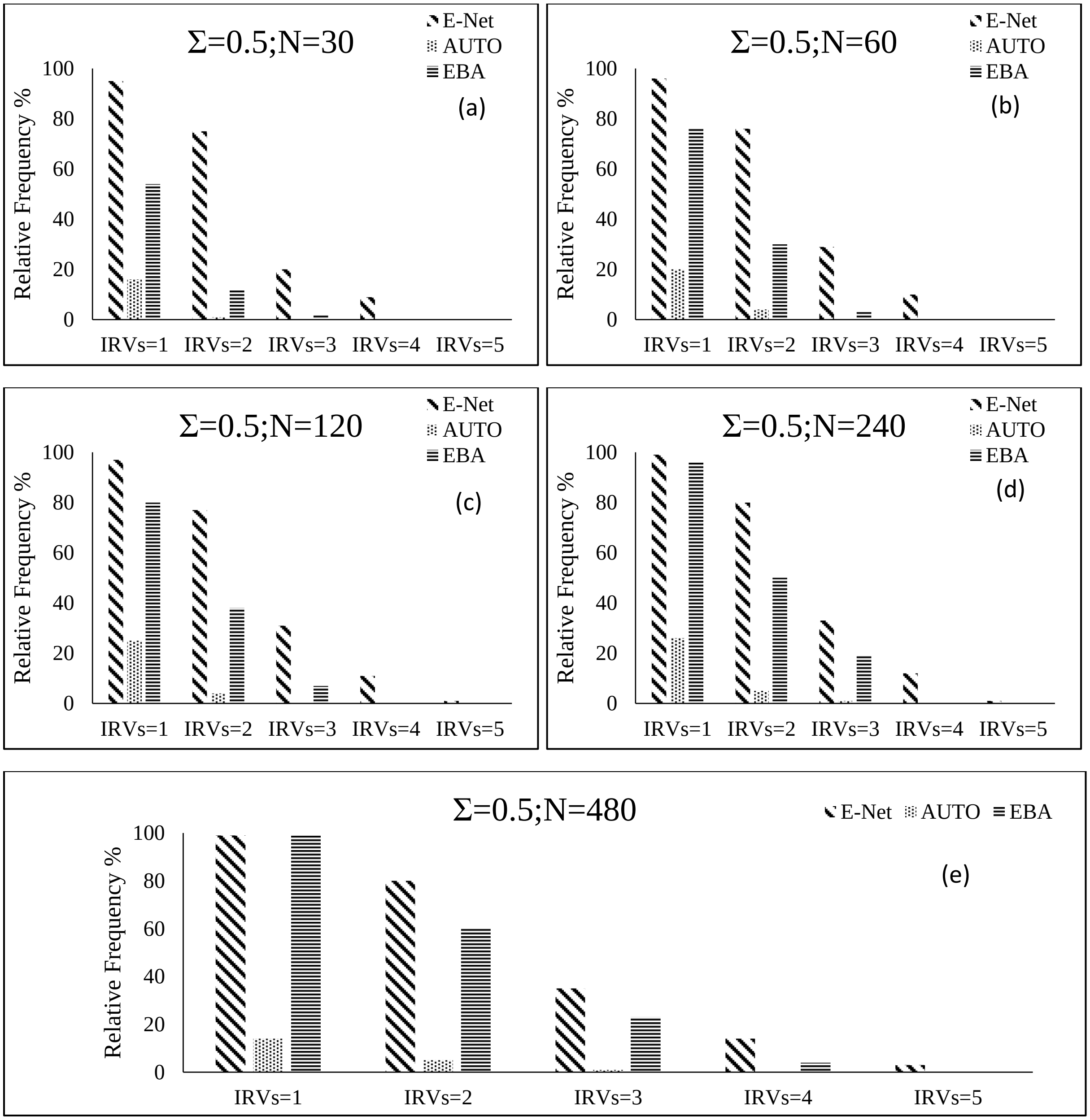


Figure 2.12: Frequencies of Irrelevant variables which are correlated to relevant variables (%) {Σ =0.5}

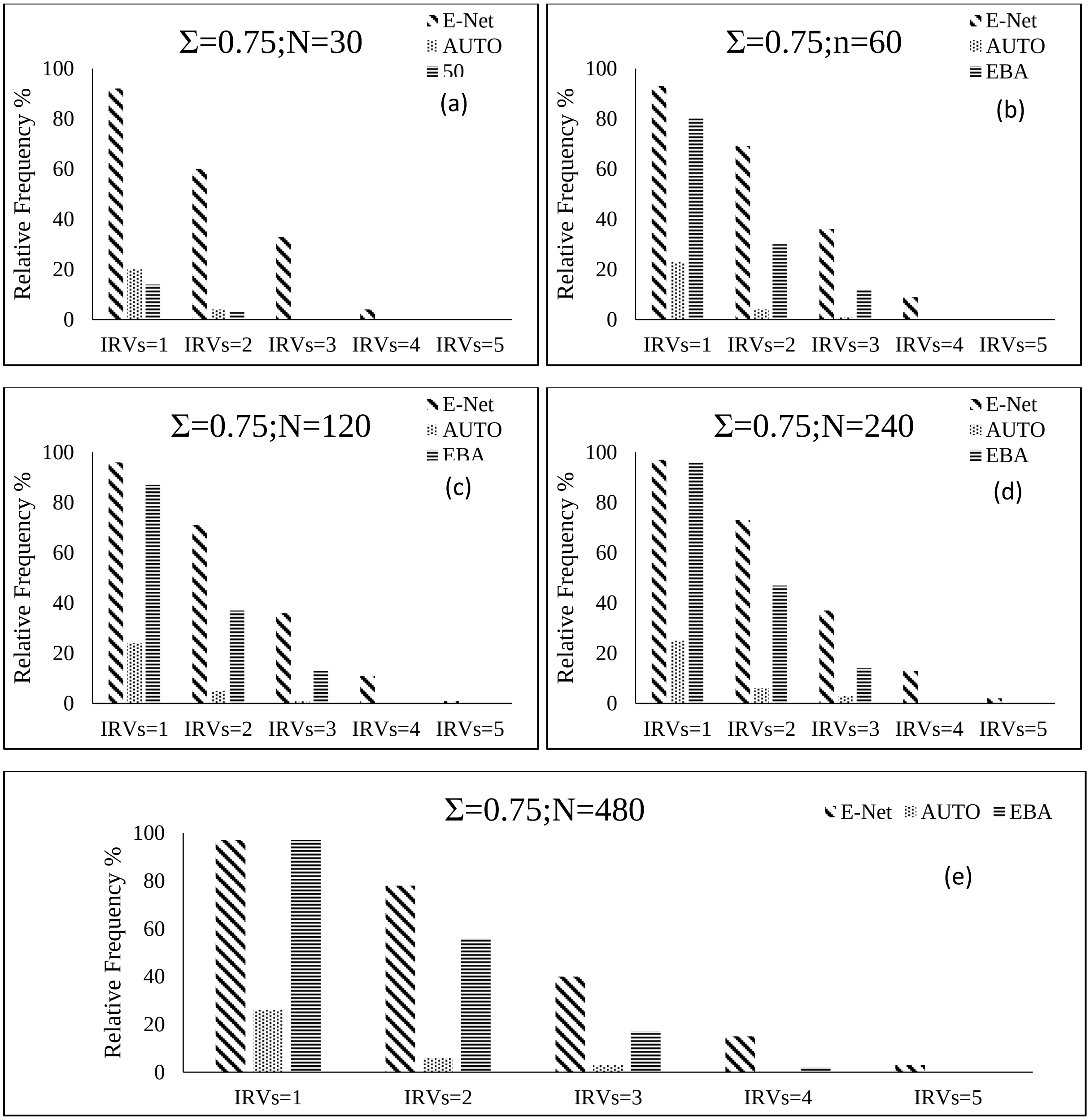


Figure 2.13: Frequencies of Irrelevant variables which are correlated to relevant variables (%) {Σ =0.75}

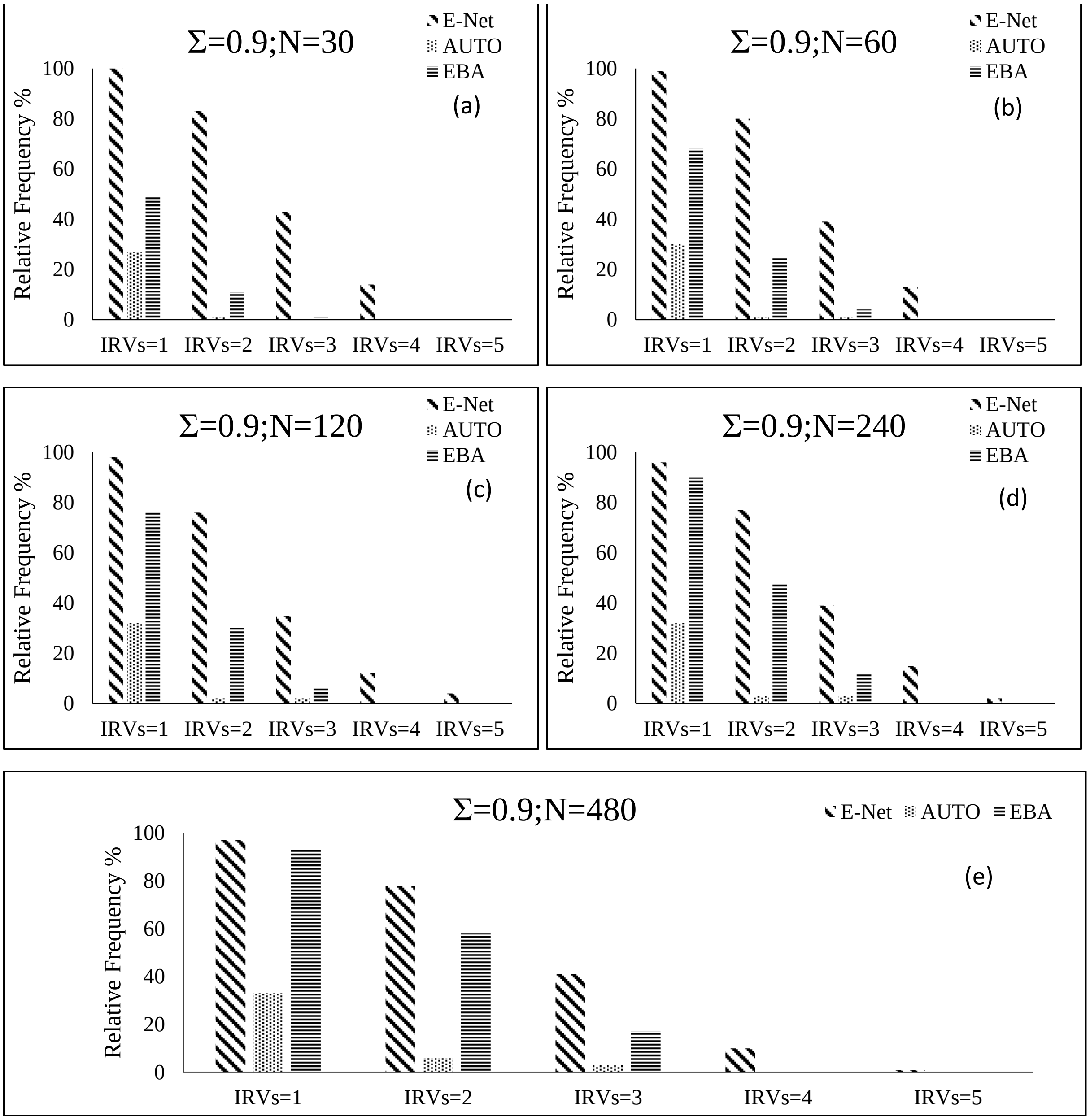


Figure 2.14: Frequencies of Irrelevant variables which are correlated to relevant variables (%) {Σ =0.9}

4.4: SCENARIO 4

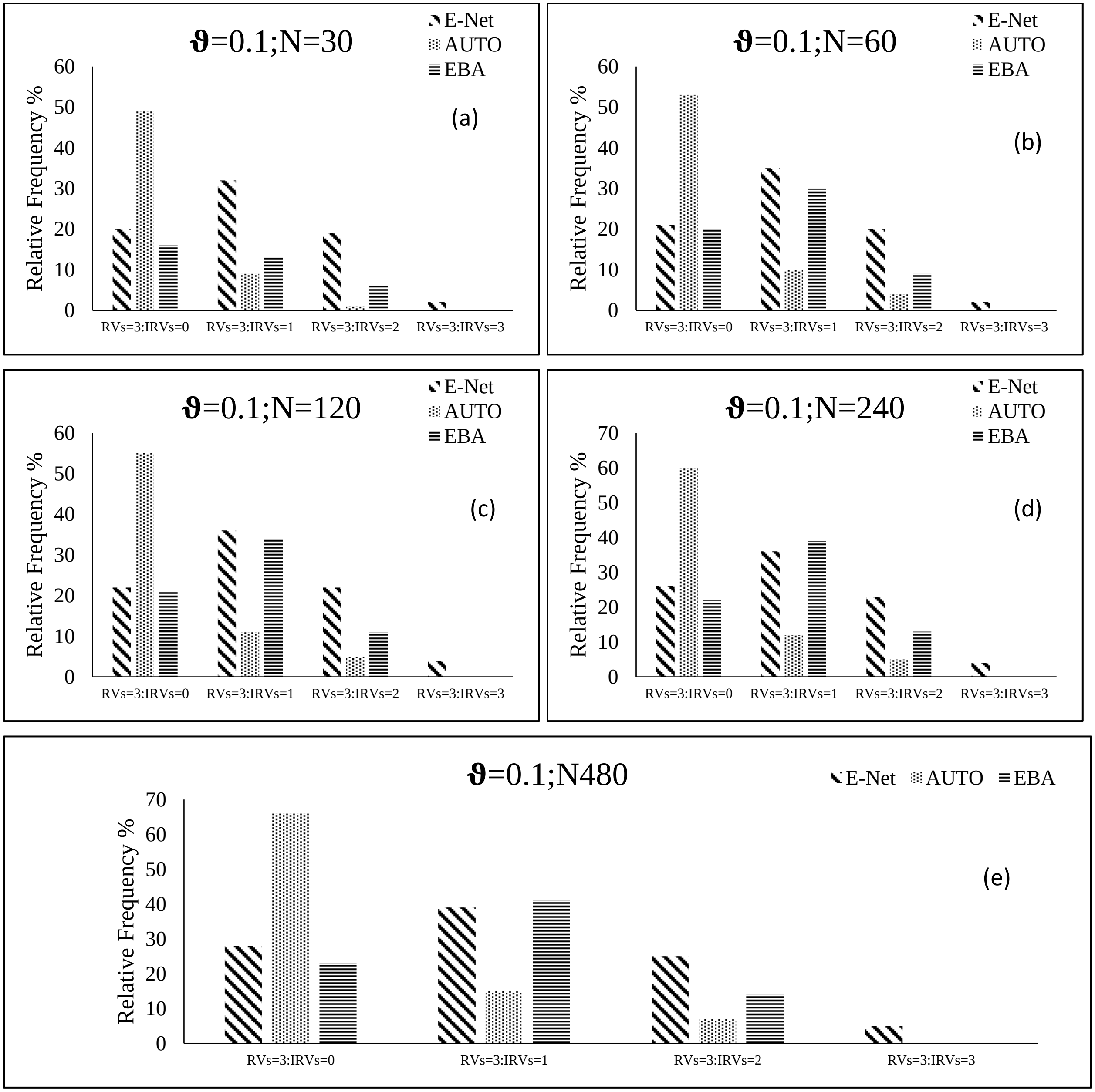
Relevant Variables are Serially Correlated

RELEVANT VARIABLES=3; IRRELEVANT VARIABLES=3

Table 2.5: Relative Frequencies of DGP when regressors are serially correlated (%)

	E-Net					Autometrics					EBA				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
								ϑ *=0.1							
DGP	20	21	22	26	28	49	53	55	60	66	16	20	21	22	23
DGP+1_IRV	32	35	36	36	39	9	10	11	12	15	13	30	34	39	41
DGP+2_IRVs	19	20	22	23	25	1	4	5	5	7	6	9	11	13	14
DGP+3_IRVs	2	2	4	4	5	0	0	0	0	0	0	0	0	0	0
Total frequency	**73**	**78**	**84**	**89**	**97**	**59**	**67**	**71**	**77**	**88**	**35**	**59**	**68**	**74**	**78**
								ϑ =0.25							
DGP	25	28	30	31	31	49	51	52	55	61	16	21	22	23	25
DGP+1_IRV	32	36	38	39	40	10	13	14	14	16	13	29	33	37	39
DGP+2_IRVs	16	20	21	22	23	1	2	3	4	6	5	9	10	12	15
DGP+3_IRVs	2	2	2	3	3	0	0	0	0	0	0	0	1	1	4
Total frequency	**75**	**86**	**91**	**95**	**97**	**60**	**66**	**69**	**73**	**83**	**34**	**59**	**66**	**73**	**83**
								ϑ =0.5							
DGP	20	21	24	28	30	41	49	55	59	62	17	19	20	23	24
DGP+1_IRV	16	30	34	37	43	13	14	16	17	18	18	26	27	28	39
DGP+2_IRVs	11	12	14	15	17	4	4	5	6	6	8	12	16	21	22
DGP+3_IRVs	2	2	4	6	7	0	0	0	0	0	0	0	0	0	0
Total frequency	**49**	**65**	**76**	**86**	**97**	**58**	**67**	**76**	**82**	**86**	**43**	**57**	**63**	**72**	**85**
								ϑ =0.75							
DGP	23	28	32	32	33	66	66	67	67	68	15	17	18	20	23
DGP+1_IRV	32	32	35	36	38	11	15	18	20	21	23	30	33	36	40
DGP+2_IRVs	13	13	14	16	20	2	2	3	4	4	3	15	19	29	30

DGP+3_IRVs	0	0	4	4	5	0	0	0	0	0	1	1	1	2	2
Total frequency	**68**	**73**	**85**	**88**	**96**	**79**	**83**	**88**	**91**	**93**	**42**	**63**	**71**	**87**	**95**
ϑ =0.9															
DGP	17	19	21	26	29	39	60	63	64	66	7	10	11	15	17
DGP+1_IRV	24	25	30	32	33	14	15	16	20	25	22	27	31	37	40
DGP+2_IRVs	9	13	17	21	22	2	2	2	3	3	3	13	16	21	22
DGP+3_IRVs	2	3	6	9	12	0	0	0	0	0	1	2	3	9	10
Total Frequency	**52**	**60**	**74**	**88**	**96**	**55**	**77**	**81**	**87**	**94**	**33**	**52**	**61**	**82**	**89**



(RVs=3) is equal to DGP

Figure 2.15: Relative Frequencies of DGP when regressors are serially correlated (%){ **ϑ** =0.1}

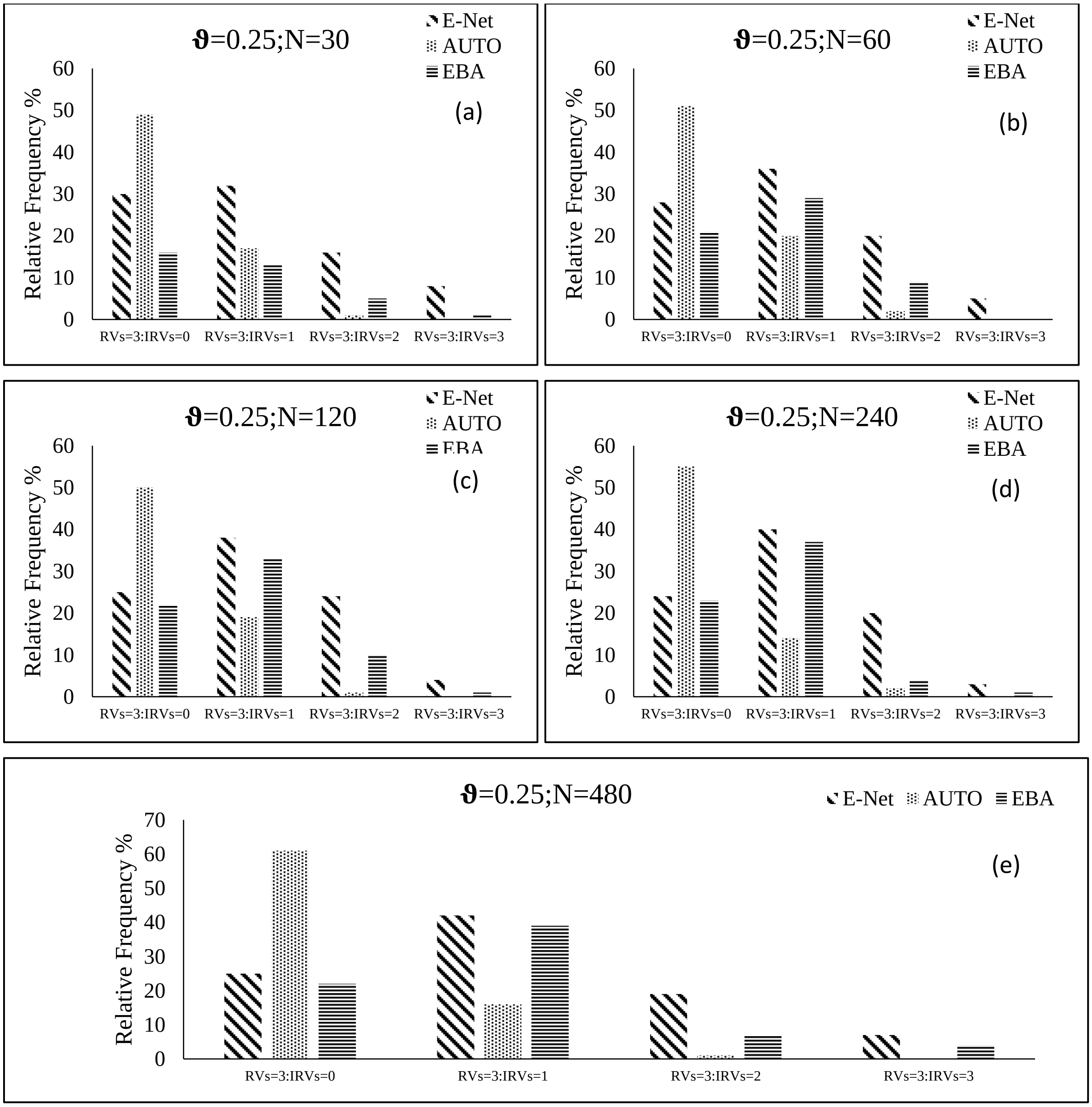


Figure 2.16: Relative Frequencies of DGP when regressors are serially correlated (%){**ϑ** =0.25}

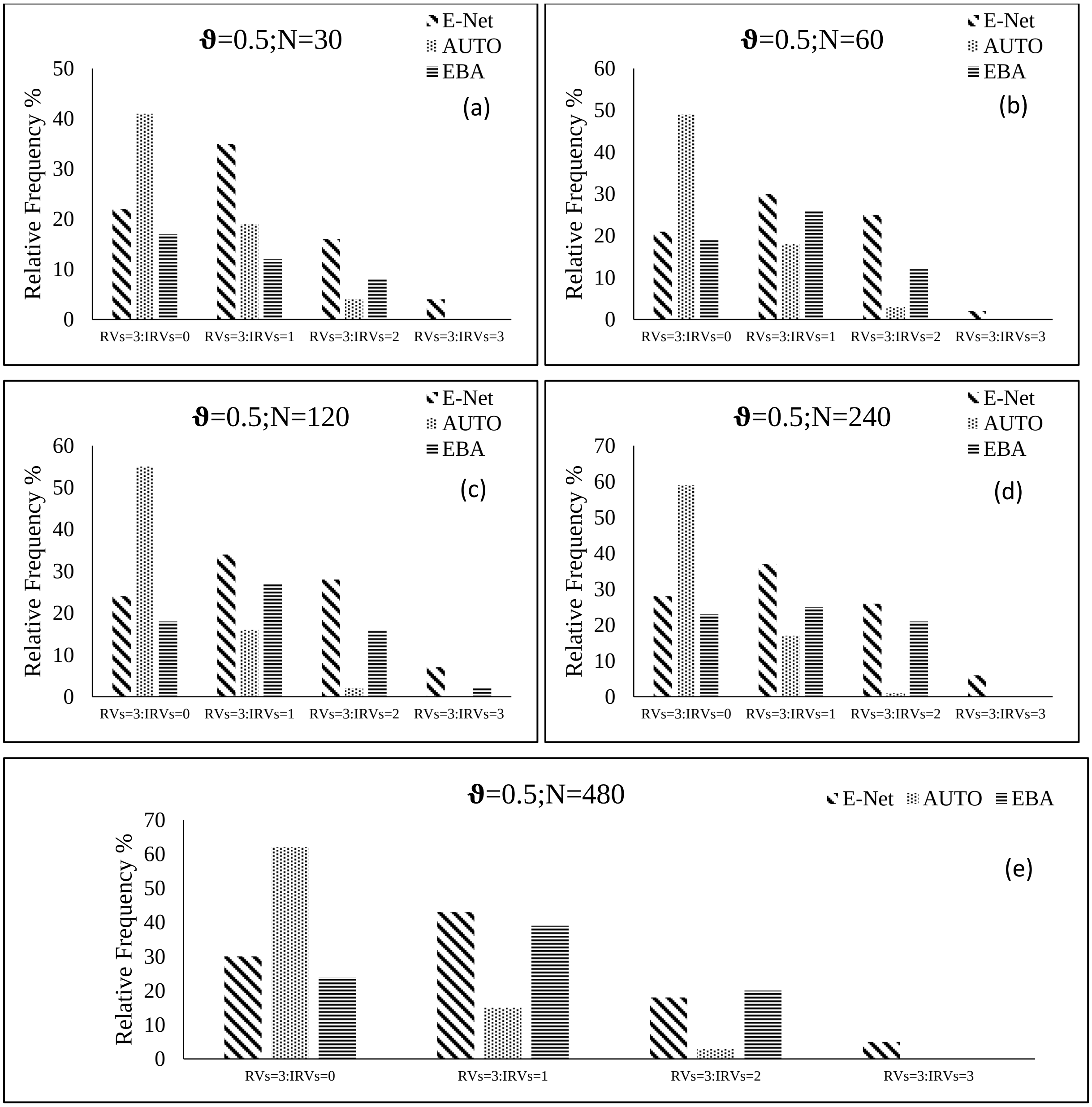


Figure 2.17: Relative Frequencies of DGP when regressors are serially correlated (%){**ϑ** =0.5}

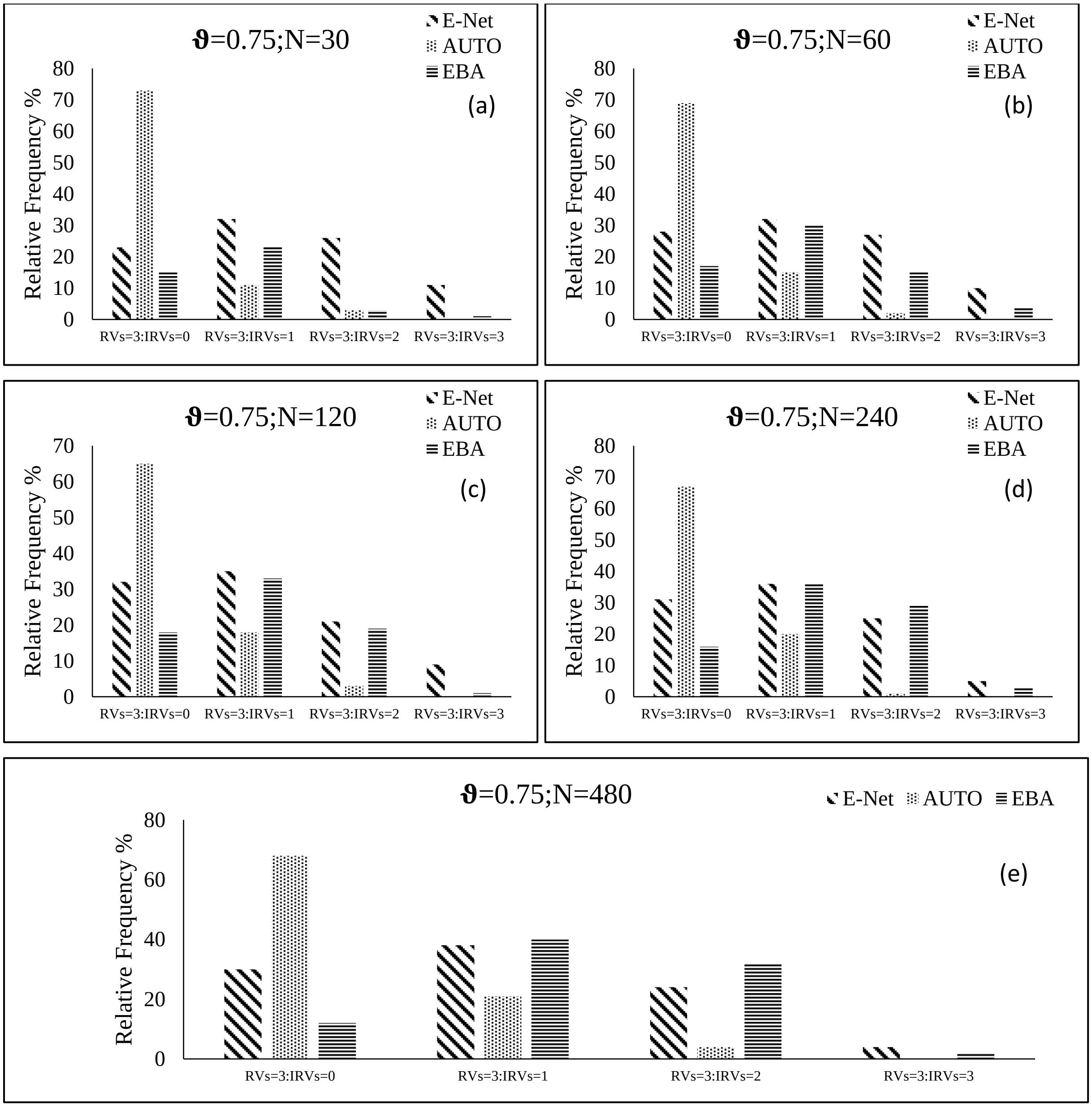


Figure 2.18: Relative Frequencies of DGP when regressors are serially correlated (%){ϑ =0.75}

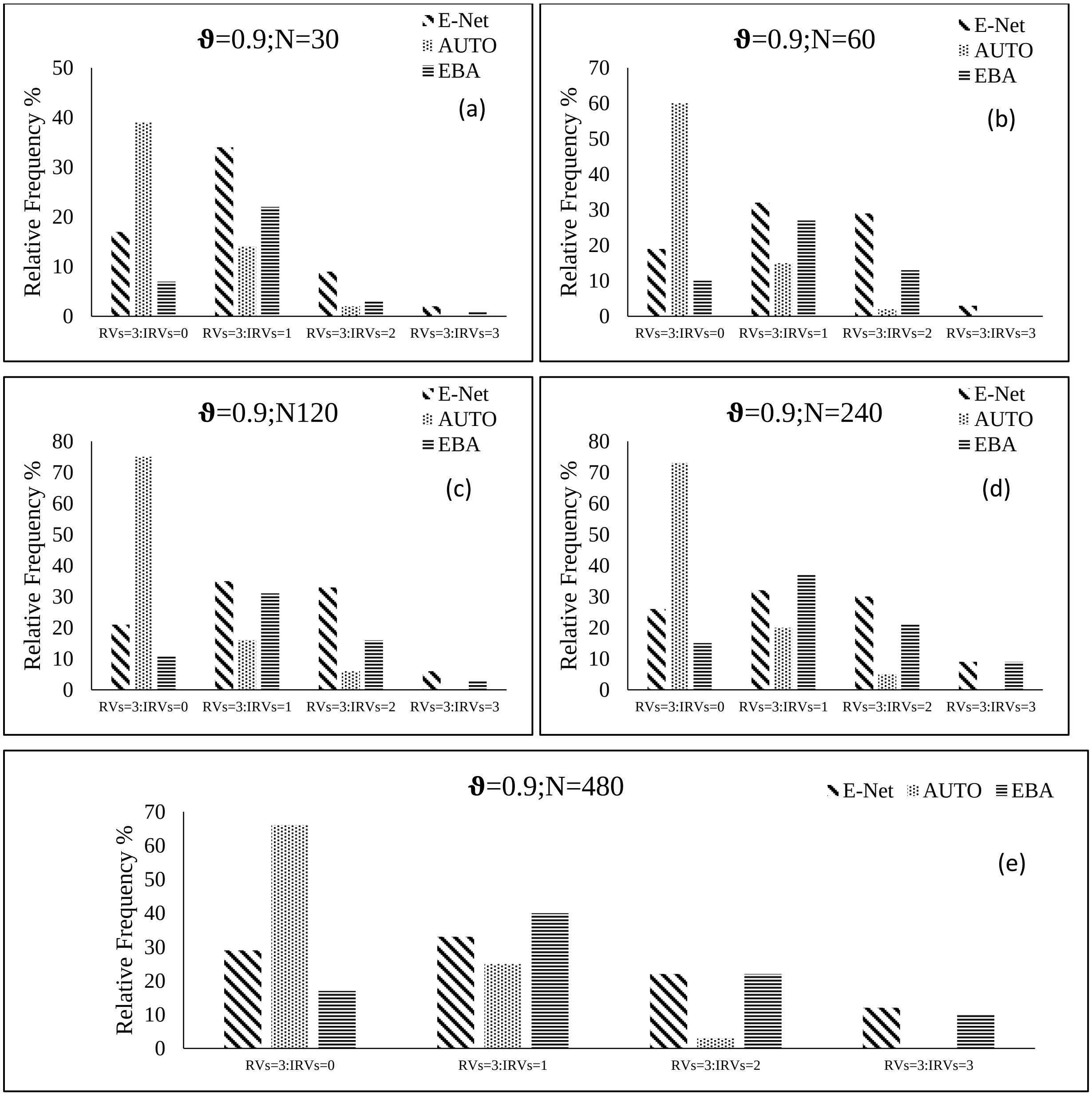


Figure 2.19: Relative Frequencies of DGP when regressors are serially correlated (%) {**ϑ** =0.9}

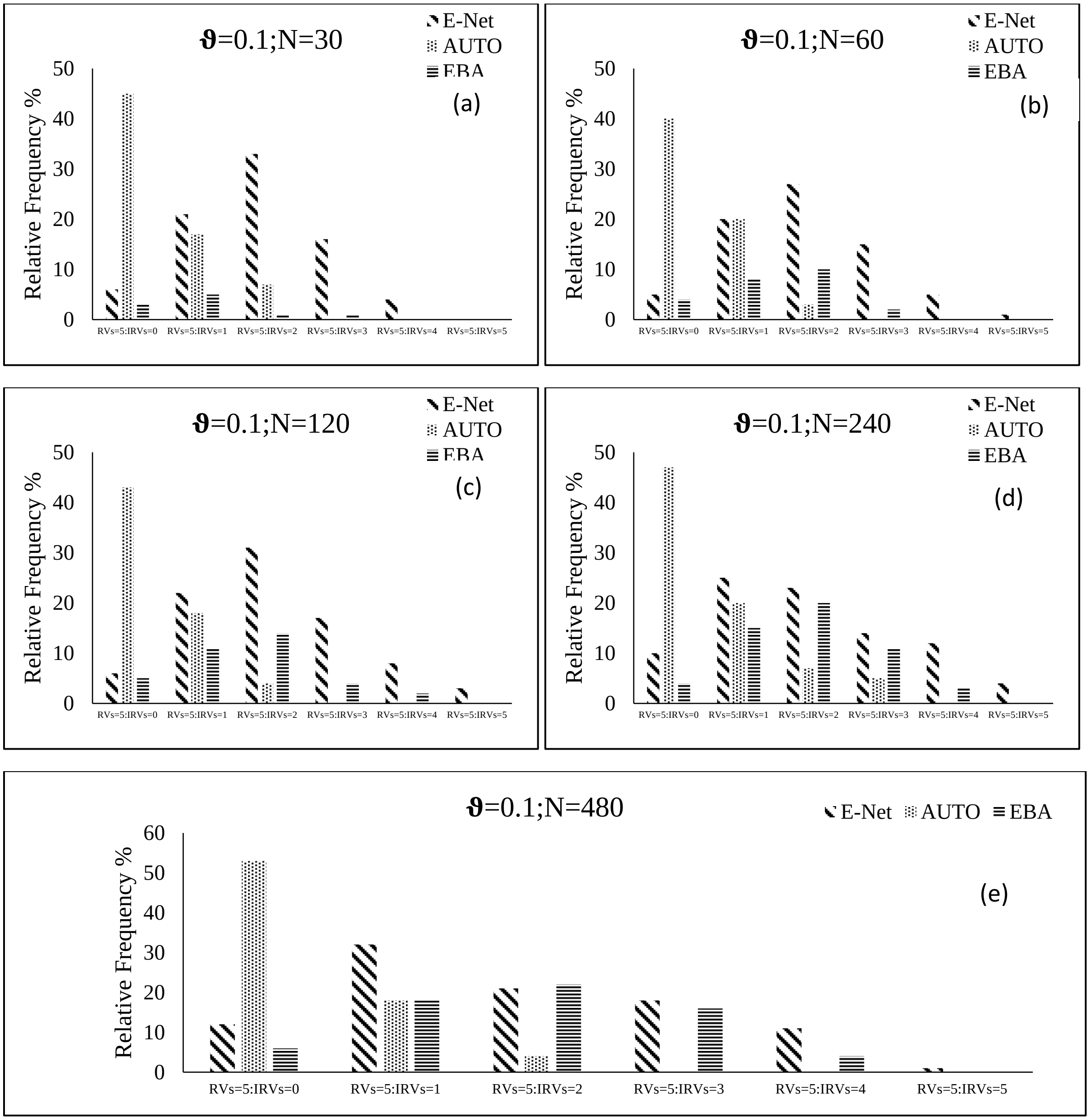
RELEVANT VARIABLES=5; IRRELEVANT VARIABLES=5

Selection of correct and incorrect variables in case of serial correlation in regressors

Table 2.6: Relative Frequencies of DGP when regressors are serially correlated (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
ϑ =0.1															
DGP	5	5	6	10	12	38	40	43	47	53	3	4	5	6	6
DGP+1_IRV	20	21	22	25	27	16	17	18	20	22	5	8	11	15	18
DGP+2_IRVs	26	27	31	33	35	2	3	4	7	9	9	10	14	20	22
DGP+3_IRVs	5	7	10	11	12	0	0	0	0	0	1	2	4	11	16
DGP+4_IRVs	4	5	8	8	9	0	0	0	0	0	0	0	0	3	4
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total frequency	**61**	**64**	**77**	**87**	**95**	**56**	**60**	**65**	**74**	**84**	**18**	**24**	**34**	**55**	**66**
ϑ =0.25															
DGP	7	8	11	13	15	49	51	52	52	56	1	4	5	9	9
DGP+1_IRV	20	21	21	22	22	19	20	22	25	27	4	7	11	20	25
DGP+2_IRVs	31	32	33	35	37	2	3	4	4	7	2	9	10	19	22
DGP+3_IRVs	7	8	10	12	14	0	0	0	0	0	0	2	4	10	11
DGP+4_IRVs	5	5	5	6	6	0	0	0	0	0	0	0	2	2	4
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total frequency	**70**	**74**	**80**	**88**	**94**	**70**	**74**	**78**	**81**	**90**	**7**	**22**	**32**	**60**	**71**
ϑ =0.5															
DGP	4	5	9	10	12	42	45	52	54	56	3	4	5	6	6
DGP+1_IRV	18	20	20	22	23	15	16	16	17	18	3	6	7	10	18
DGP+2_IRVs	20	21	22	27	34	3	3	4	4	5	2	6	8	20	24
DGP+3_IRVs	18	18	21	23	24	0	0	0	0	0	1	3	5	23	14
DGP+4_IRVs	2	2	3	3	4	0	0	0	0	0	0	0	2	3	4
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Total frequency	**62**	**66**	**75**	**85**	**97**	**60**	**64**	**72**	**75**	**79**	**9**	**19**	**27**	**62**	**66**
ϑ =0.75															
DGP	3	4	7	10	11	39	40	42	43	53	0	0	1	3	4
DGP+1_IRV	15	17	19	20	21	22	23	24	30	31	2	3	3	4	5
DGP+2_IRVs	24	27	35	36	39	5	10	13	13	14	4	4	13	20	25
DGP+3_IRVs	5	8	10	13	15	0	0	0	0	0	2	5	9	15	20
DGP+4_IRVs	3	3	4	5	5	0	0	0	0	0	0	2	4	10	14
DGP+5_IRVs	0	0	1	3	4	0	0	0	0	0	0	0	0	0	0
Total frequency	**50**	**59**	**76**	**87**	**95**	**66**	**73**	**79**	**86**	**98**	**8**	**14**	**30**	**52**	**68**
ϑ =0.9															
DGP	4	4	6	10	12	52	55	62	66	66	0	0	0	0	0
DGP+1_IRV	14	15	17	19	22	17	19	20	20	21	2	4	4	9	10
DGP+2_IRVs	25	28	32	33	33	5	6	6	7	8	4	6	11	13	15
DGP+3_IRVs	9	10	13	14	15	0	0	0	0	0	5	9	10	22	25
DGP+4_IRVs	3	4	6	8	10	0	0	0	0	0	1	4	6	14	19
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total frequency	**55**	**61**	**74**	**84**	**92**	**74**	**80**	**88**	**93**	**95**	**12**	**23**	**31**	**58**	**69**



(RVs=5) is equal to DGP

Figure 2.20: Relative Frequencies of DGP when regressors are serially correlated (%) {**ϑ** =0.1}

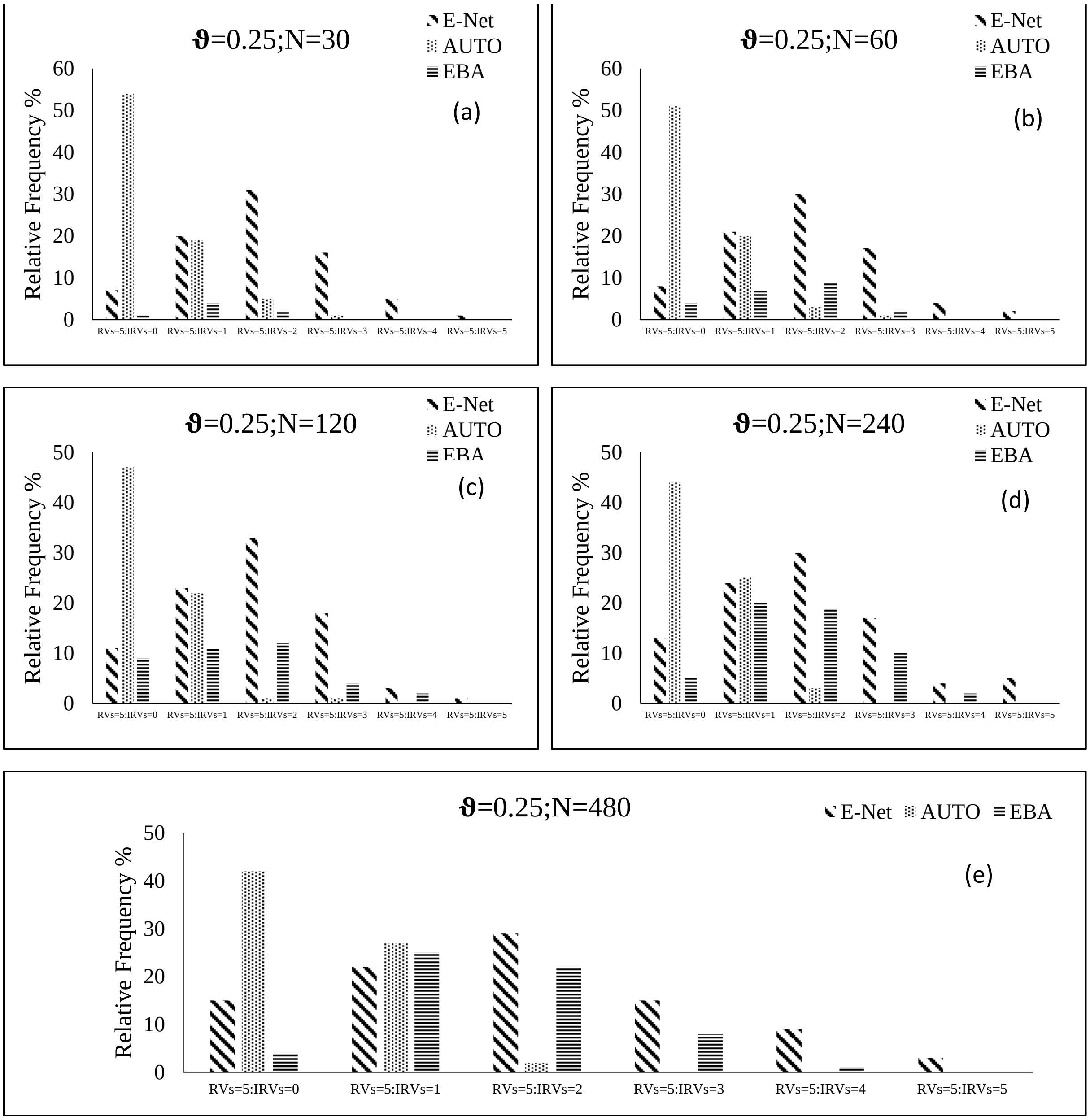


Figure 2.21: Relative Frequencies of DGP when regressors are serially correlated (%){**ϑ** =0.25}

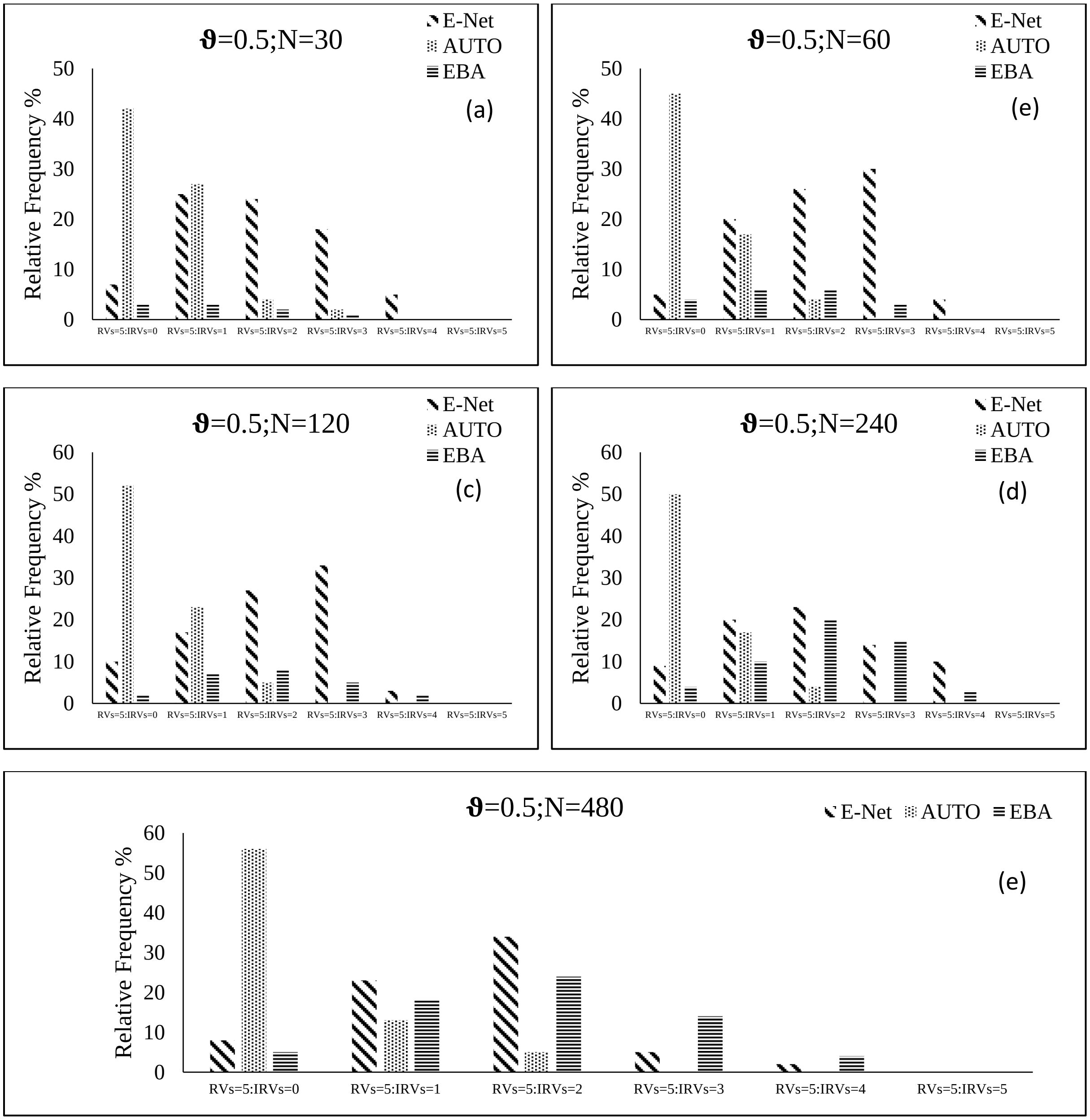


Figure 2.22:Relative Frequencies of DGP when regressors are serially correlated (%) {**ϑ** =0.5}

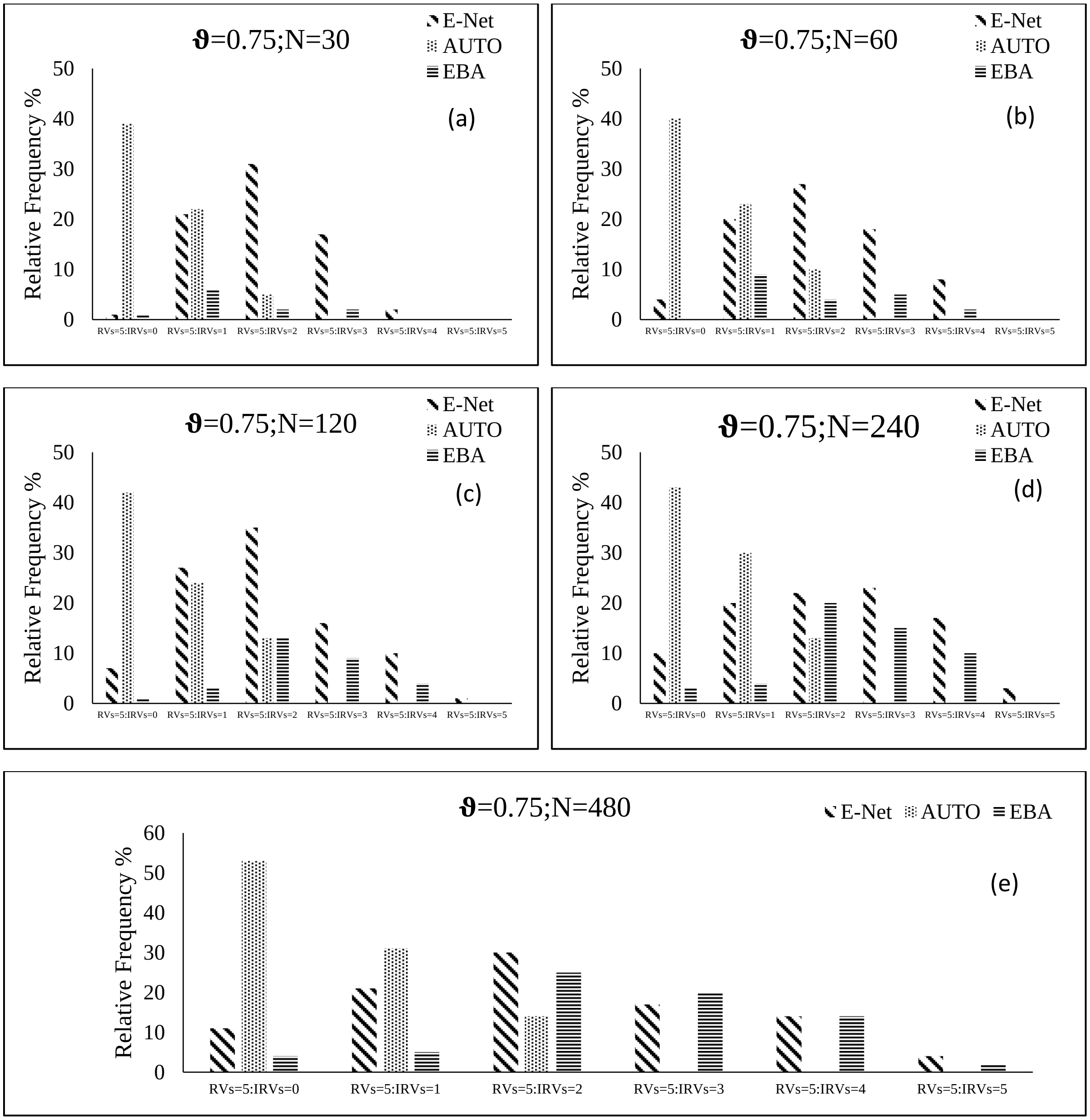


Figure 2.23:Relative Frequencies of DGP when regressors are serially correlated (%) {**ϑ** =0.75}

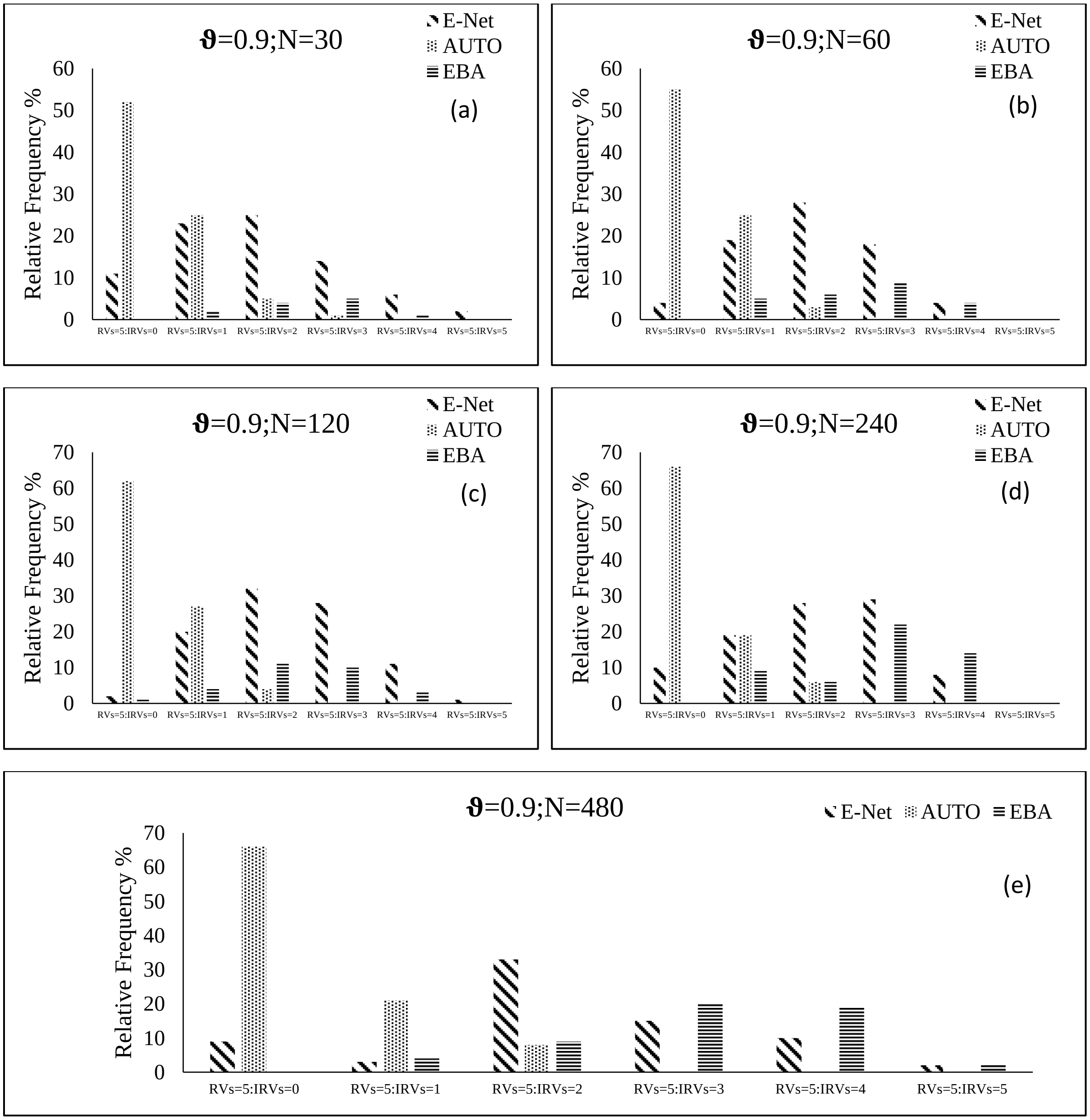


Figure 2.24: Relative Frequencies of DGP when regressors are serially correlated (%) {**ϑ** =0.9}

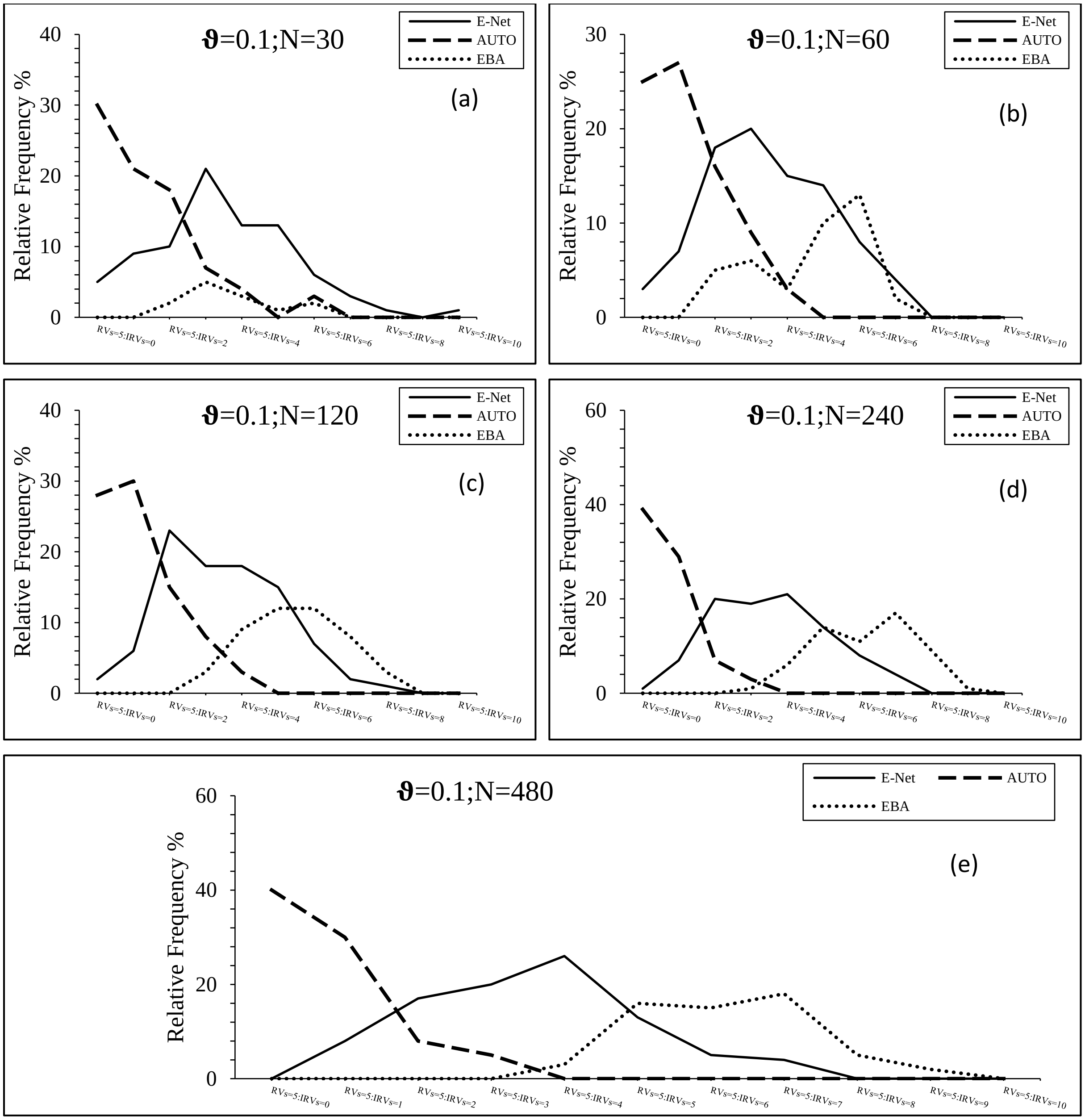
RELEVANT VARIABLES=5; IRRELEVANT VARIABLES=10

Table 2.7: Relative Frequencies of DGP when regressors are serially correlated (%)

	E-Net					**Autometrics**					**EBA**				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
								ϑ =0.1							
DGP	1	2	3	4	4	25	30	32	39	40	0	0	0	0	0
DGP+1_IRV	5	6	6	7	8	21	27	30	30	31	0	0	0	0	0
DGP+2_IRVs	10	13	14	16	20	7	10	15	15	16	0	0	0	0	0
DGP+3_IRVs	16	16	17	18	22	7	8	8	8	8	1	2	3	3	4
DGP+4_IRVs	13	15	16	17	18	2	3	3	4	4	3	3	4	5	5
DGP+5_IRVs	13	13	14	15	16	0	0	0	0	0	5	10	12	14	16
DGP+6_IRVs	6	7	7	8	5	0	0	0	0	0	4	11	12	14	15
DGP+7_IRVs	1	2	2	4	4	0	0	0	0	0	2	2	8	10	13
DGP+8_IRVs	0	0	0	0	0	0	0	0	0	0	3	4	6	9	9
DGP+9_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DGP+10_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total frequency	**65**	**74**	**79**	**89**	**97**	**62**	**78**	**88**	**96**	**99**	**18**	**32**	**45**	**55**	**62**
								ϑ =0.25							
DGP	2	2	3	4	5	31	32	33	39	40	0	0	0	0	0
DGP+1_IRV	4	4	5	6	7	24	26	28	31	32	0	0	0	0	0
DGP+2_IRVs	11	12	13	14	17	10	12	12	13	14	0	0	0	0	0
DGP+3_IRVs	9	10	22	15	19	8	9	10	11	11	1	1	2	2	3
DGP+4_IRVs	8	9	10	12	13	0	0	0	0	0	2	2	2	3	4
DGP+5_IRVs	7	8	10	11	13	0	0	0	0	0	2	2	3	3	5
DGP+6_IRVs	7	8	9	10	11	0	0	0	0	0	3	6	7	9	8
DGP+7_IRVs	6	7	8	9	9	0	0	0	0	0	5	9	10	20	28
DGP+8_IRVs	2	2	3	3	4	0	0	0	0	0	3	5	9	21	24
DGP+9_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	1	3	8

DGP+10_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total frequency	**56**	**62**	**83**	**84**	**98**	**73**	**79**	**83**	**94**	**97**	**16**	**25**	**34**	**61**	**80**
ϑ =0.5															
DGP	0	0	1	1	1	29	30	31	32	33	0	0	0	0	0
DGP+1_IRV	4	6	6	7	8	24	25	28	29	30	0	0	0	0	0
DGP+2_IRVs	11	12	17	18	19	10	12	14	20	21	0	0	0	0	0
DGP+3_IRVs	12	16	18	20	22	5	5	6	9	7	0	0	1	2	2
DGP+4_IRVs	10	12	13	15	17	1	2	2	2	3	1	2	3	3	3
DGP+5_IRVs	9	11	13	15	15	0	0	0	0	0	3	6	10	11	14
DGP+6_IRVs	8	8	9	10	11	0	0	0	0	0	4	9	12	13	20
DGP+7_IRVs	1	2	2	3	3	0	0	0	0	0	2	3	11	15	19
DGP+8_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	2	8	15
DGP+9_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DGP+10_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total frequency	**55**	**67**	**79**	**89**	**96**	**69**	**74**	**81**	**92**	**94**	**10**	**20**	**39**	**52**	**73**
ϑ =0.75															
DGP	0	0	0	0	0	37	39	40	42	43	0	0	0	0	0
DGP+1_IRV	4	4	5	5	6	18	30	32	33	35	0	0	0	0	0
DGP+2_IRVs	12	15	15	16	17	7	9	11	12	13	0	0	0	0	0
DGP+3_IRVs	11	13	13	14	15	2	4	4	5	5	0	0	0	0	0
DGP+4_IRVs	10	11	12	12	14	0	0	0	0	0	0	0	0	0	0
DGP+5_IRVs	5	7	9	11	12	0	0	0	0	0	3	4	4	5	6
DGP+6_IRVs	3	3	4	5	6	0	0	0	0	0	5	10	12	13	18
DGP+7_IRVs	1	1	2	2	3	0	0	0	0	0	8	20	20	21	23
DGP+8_IRVs	0	0	0	0	0	0	0	0	0	0	7	17	19	20	22
DGP+9_IRVs	0	0	0	0	0	0	0	0	0	0	1	1	1	2	5
DGP+10_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total frequency	**46**	**54**	**60**	**65**	**73**	**64**	**82**	**87**	**92**	**96**	**24**	**52**	**56**	**61**	**74**
ϑ =0.9															

DGP	0	0	0	0	0	30	31	32	33	35	0	0	0	0	0
DGP+1_IRV	3	4	4	5	5	14	19	20	20	22	0	0	0	0	0
DGP+2_IRVs	7	8	10	12	15	12	17	18	18	19	0	0	0	0	0
DGP+3_IRVs	12	13	15	17	22	4	5	6	6	7	0	0	0	0	0
DGP+4_IRVs	11	12	15	16	20	0	0	0	0	0	1	1	3	4	4
DGP+5_IRVs	10	10	14	15	17	0	0	0	0	0	2	2	5	9	9
DGP+6_IRVs	5	6	9	10	11	0	0	0	0	0	3	6	7	10	12
DGP+7_IRVs	2	2	8	9	0	0	0	0	0	0	4	7	7	12	13
DGP+8_IRVs	0	0	0	0	0	0	0	0	0	0	6	9	10	23	28
DGP+9_IRVs	0	0	0	0	0	0	0	0	0	0	5	7	9	12	14
DGP+10_IRVs	0	0	0	0	0	0	0	0	0	0	0	0	3	6	7
Total frequency	**50**	**55**	**75**	**84**	**90**	**60**	**72**	**76**	**77**	**83**	**21**	**32**	**44**	**76**	**87**



(RVs=5) is equal to DGP

Figure 2.25: Relative Frequencies of DGP when regressors are serially correlated (%) {**ϑ** =0.1}

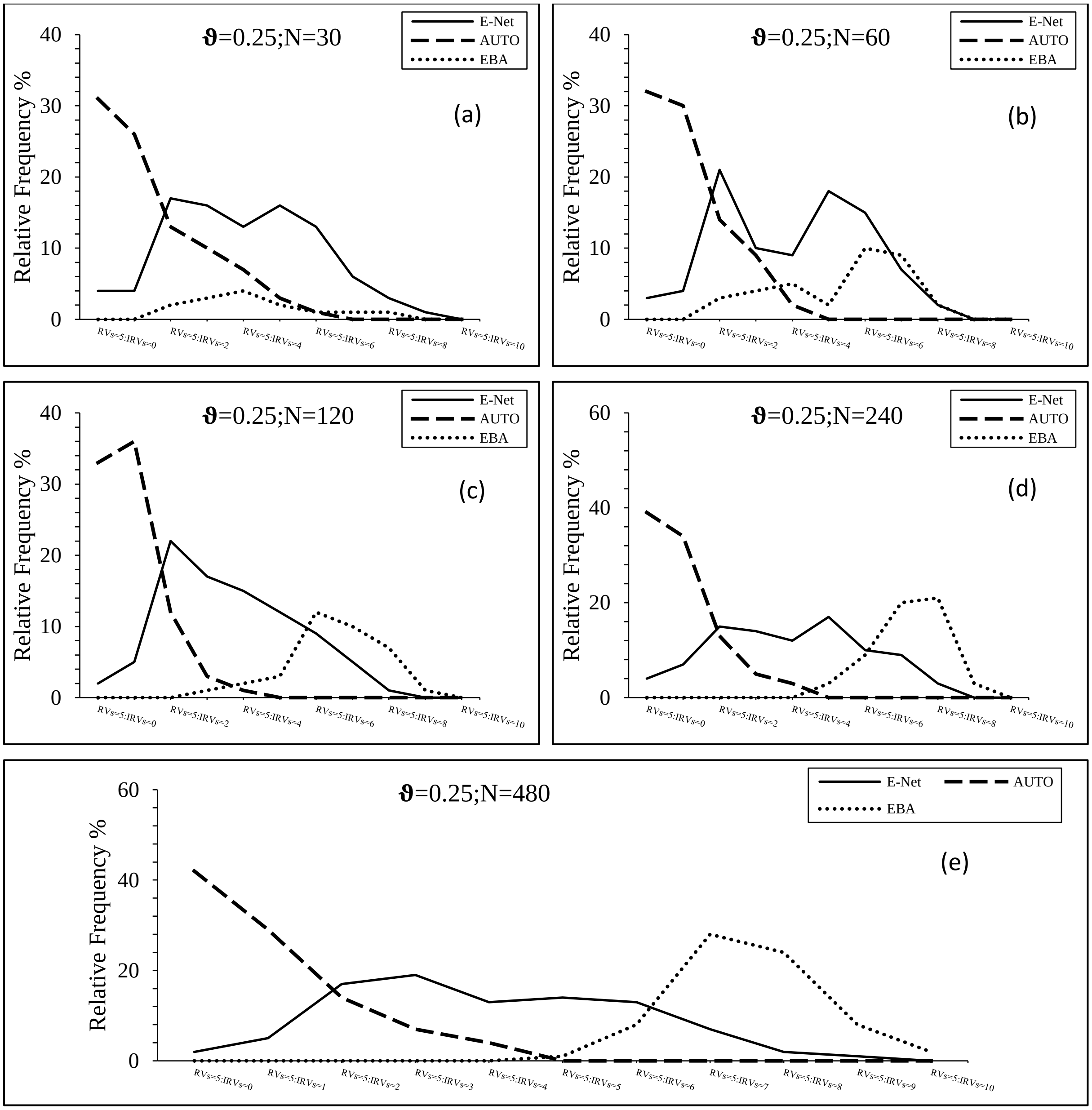


Figure 2.26: Relative Frequencies of DGP when regressors are serially correlated (%)

{**ϑ** =0.25}

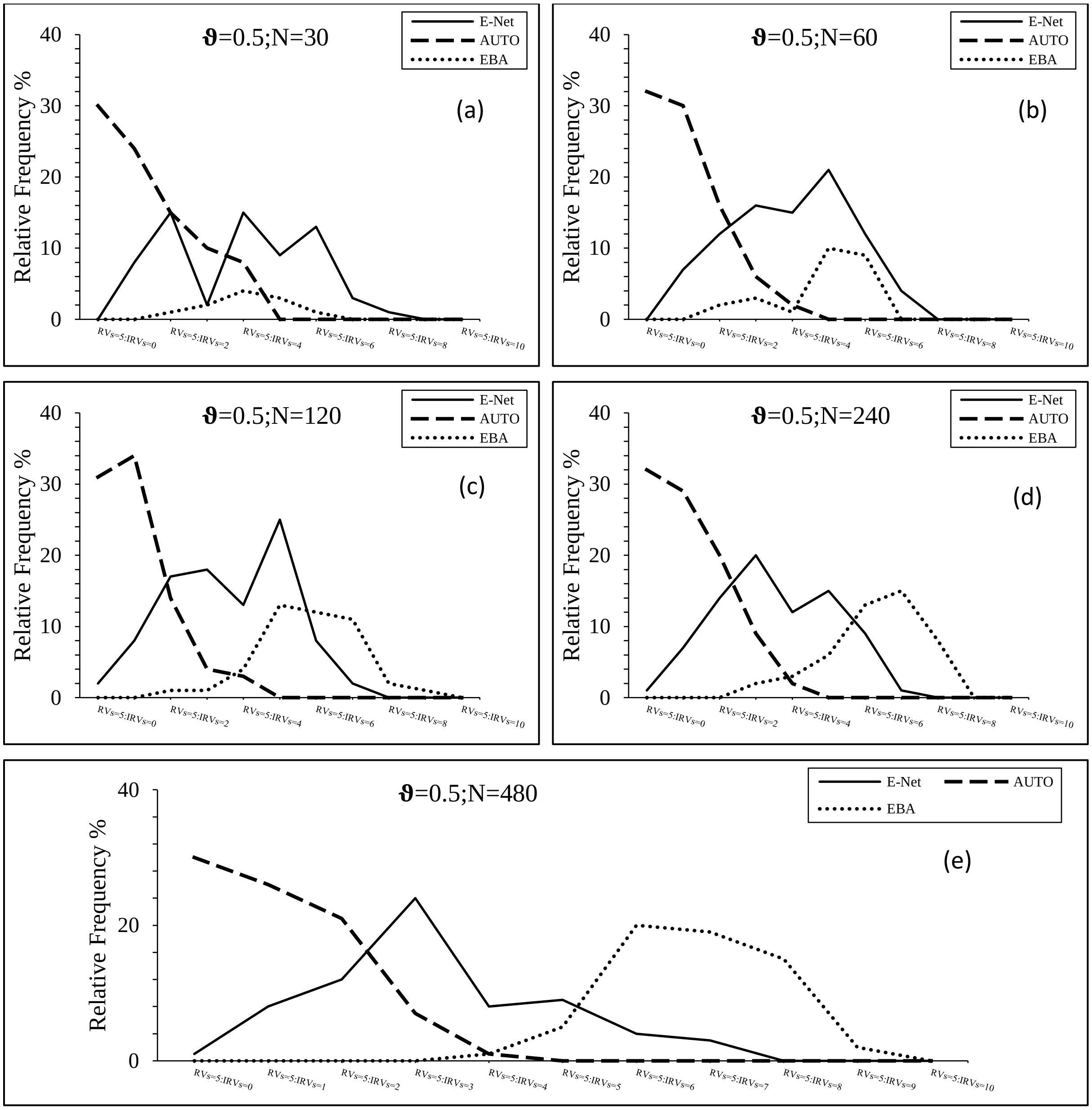


Figure 2.27: Relative Frequencies of DGP when regressors are serially correlated (%)

{**ϑ** =0.5}

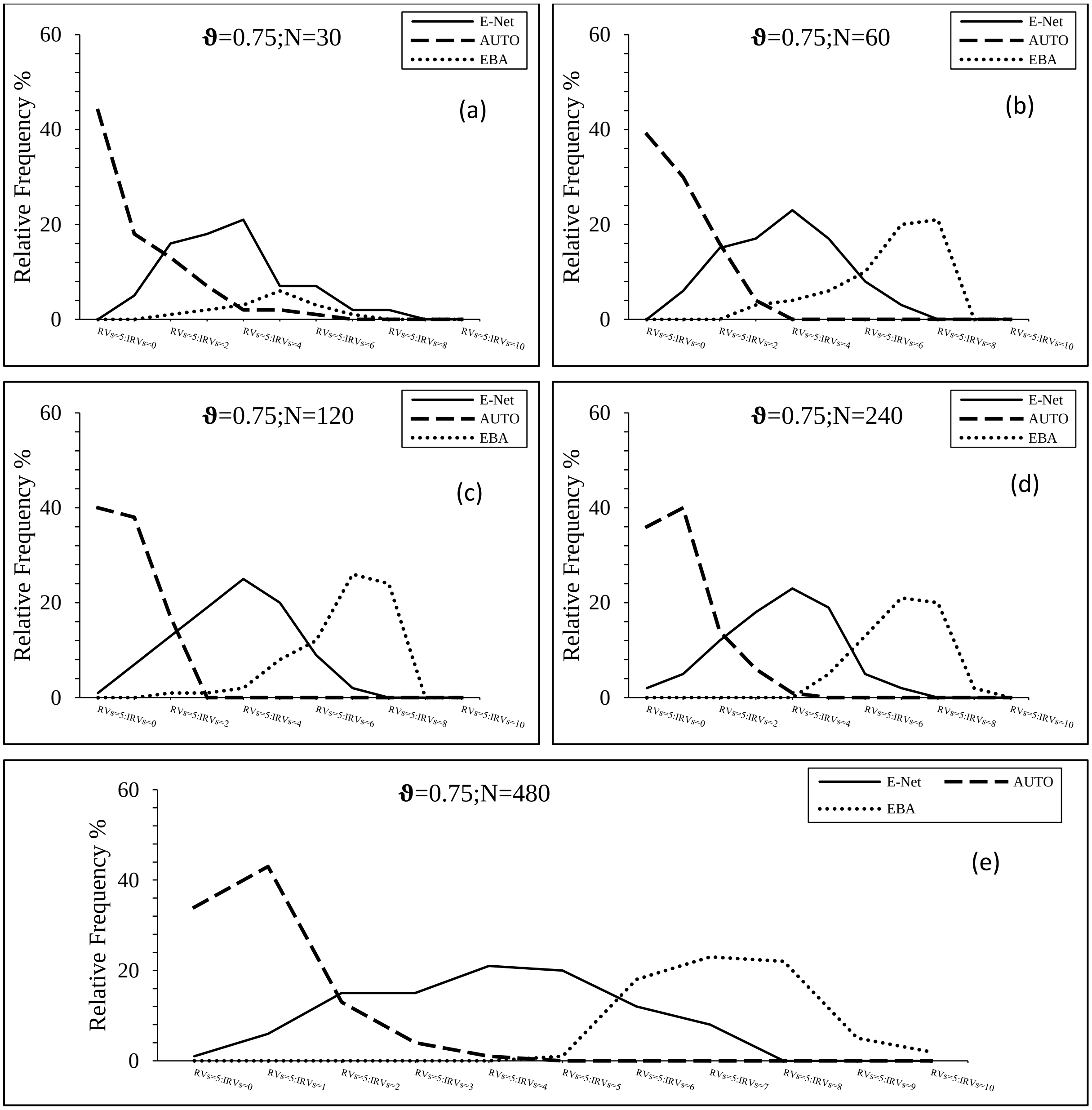


Figure 2.28: Relative Frequencies of DGP when regressors are serially correlated (%)

{**ϑ** =0.75}

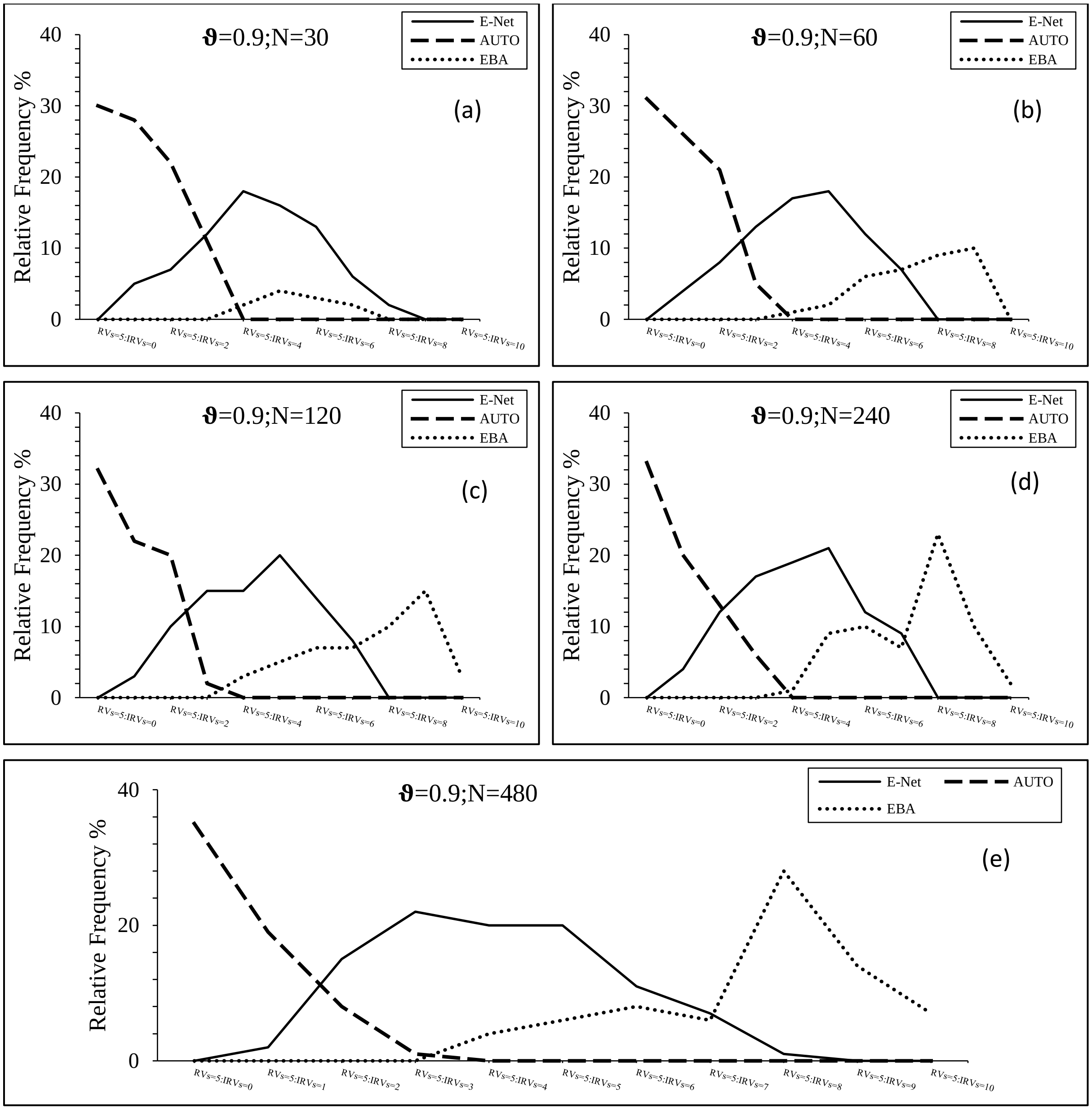


Figure 2.29: Relative Frequencies of DGP when regressors are serially correlated (%)

{**ϑ** =0.9}

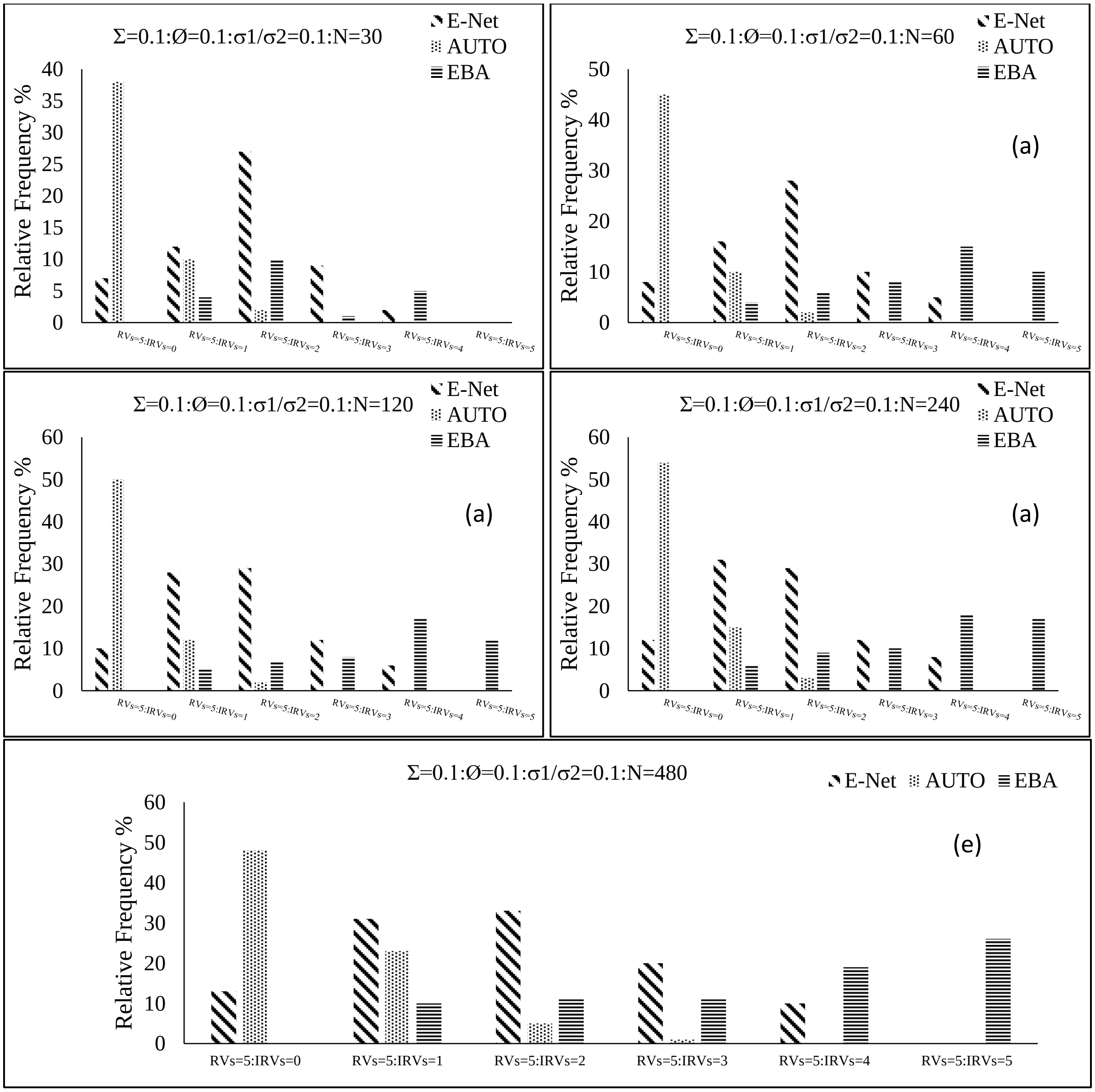
4.7 SCENARIO 7:

Performance of Model Selection Procedures in Case of Multiple Problems in DGP

Table 2.8: Relative Frequencies of DGP When There Are Multiple Problems in Data (%)

	E-Net					Autometrics					EBA				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
Σ=0.1 Ø=0.1 σ1/σ2=0.1															
DGP	7	8	10	12	13	38	45	50	54	48	0	0	0	0	0
DGP+1_IRV	12	16	28	31	31	10	10	12	15	23	4	4	5	6	10
DGP+2_IRVs	27	28	29	29	30	2	2	2	3	5	5	6	7	9	11
DGP+3_IRVs	9	10	12	12	13	0	0	0	0	1	1	8	8	10	11
DGP+4_IRVs	2	5	6	8	10	0	0	0	0	0	5	15	17	18	19
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	0	10	12	17	26
Total frequency	**57**	**67**	**85**	**92**	**97**	**50**	**57**	**64**	**72**	**77**	**15**	**43**	**49**	**60**	**77**
Σ=0.25 Ø=0.25 σ1/σ2=0.25															
DGP	13	14	16	16	17	40	41	43	43	43	0	0	0	0	0
DGP+1_IRV	20	22	23	24	24	16	17	20	22	23	0	0	0	0	0
DGP+2_IRVs	19	20	22	22	30	0	0	2	3	5	0	2	3	4	4
DGP+3_IRVs	10	11	14	15	19	0	0	0	0	0	0	4	5	6	6
DGP+4_IRVs	2	2	2	4	4	0	0	0	0	0	2	15	16	17	18
DGP+5_IRVs	0	0	0	0	0	0	0	0	0	0	66	67	67	67	68
Total frequency	**64**	**69**	**77**	**81**	**94**	**56**	**58**	**65**	**68**	**71**	**68**	**88**	**91**	**94**	**96**
Σ=0.5 Ø =0.5 σ1/σ2=0.5															
DGP	4	5	6	7	8	28	28	30	31	34	0	0	0	0	0
DGP+1_IRV	14	14	19	21	22	7	8	9	9	11	0	0	0	0	0
DGP+2_IRVs	23	24	25	25	26	3	3	4	5	6	0	3	3	4	4
DGP+3_IRVs	14	15	18	18	19	0	1	1	2	2	0	3	4	5	7
DGP+4_IRVs	5	5	9	12	13	0	0	0	0	0	2	4	6	7	7
DGP+5_IRVs	0	2	3	3	5	0	0	0	0	0	59	66	70	75	79

Total frequency	**60**	**65**	**80**	**86**	**93**	**38**	**40**	**44**	**47**	**53**	**61**	**76**	**83**	**91**	**97**
Σ=0.75 Ø =0.75 σ1/σ2=0.75															
DGP	0	0	0	0	1	5	5	7	7	8	0	0	0	0	0
DGP+1_IRV	4	6	10	12	15	4	5	8	10	14	0	0	0	0	0
DGP+2_IRVs	9	12	18	19	21	0	1	2	2	3	0	0	1	1	2
DGP+3_IRVs	16	17	18	20	20	0	0	0	0	0	2	4	5	6	6
DGP+4_IRVs	10	10	11	12	12	0	0	0	0	0	3	5	7	8	8
DGP+5_IRVs	7	7	9	10	10	0	0	0	0	0	70	71	75	80	81
Total frequency	**46**	**52**	**66**	**73**	**79**	**9**	**11**	**17**	**19**	**25**	**75**	**80**	**88**	**95**	**97**
Σ=0.9 Ø =0.9 σ1/σ2=0.9															
DGP	1	1	2	2	3	3	4	4	5	5	0	0	0	0	0
DGP+1_IRV	2	2	4	5	6	0	0	0	0	0	0	0	0	0	0
DGP+2_IRVs	9	10	10	11	12	0	0	0	0	0	2	3	3	4	4
DGP+3_IRVs	9	10	11	11	12	0	0	0	0	0	7	8	8	10	11
DGP+4_IRVs	9	10	12	13	13	0	0	0	0	0	15	15	17	18	19
DGP+5_IRVs	6	6	7	9	10	0	0	0	0	0	50	53	56	57	58
Total frequency	**36**	**39**	**46**	**51**	**56**	**3**	**4**	**4**	**5**	**5**	**74**	**79**	**84**	**89**	**92**



(RVs=5) is equal to DGP

Figure 2.30: Relative Frequencies of DGP When There Are Multiple Problems in Data (%)

{Σ=0.1: Ø=0.1:σ1/σ2=0.1}

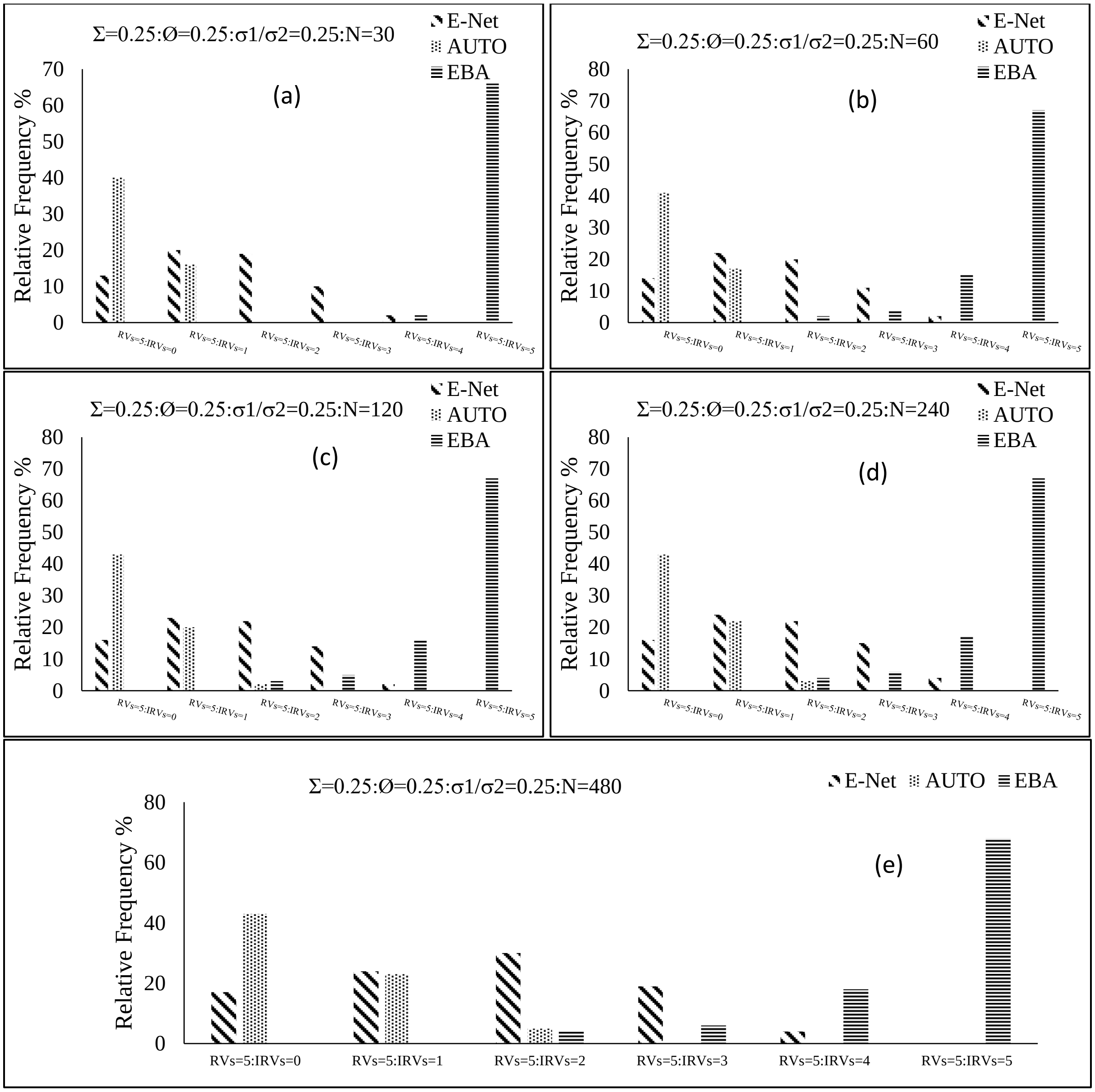


Figure 2.31: Relative Frequencies of DGP When There Are Multiple Problems in Data (%)

{Σ=0.25: Ø=0.25:σ1/σ2=0.25}

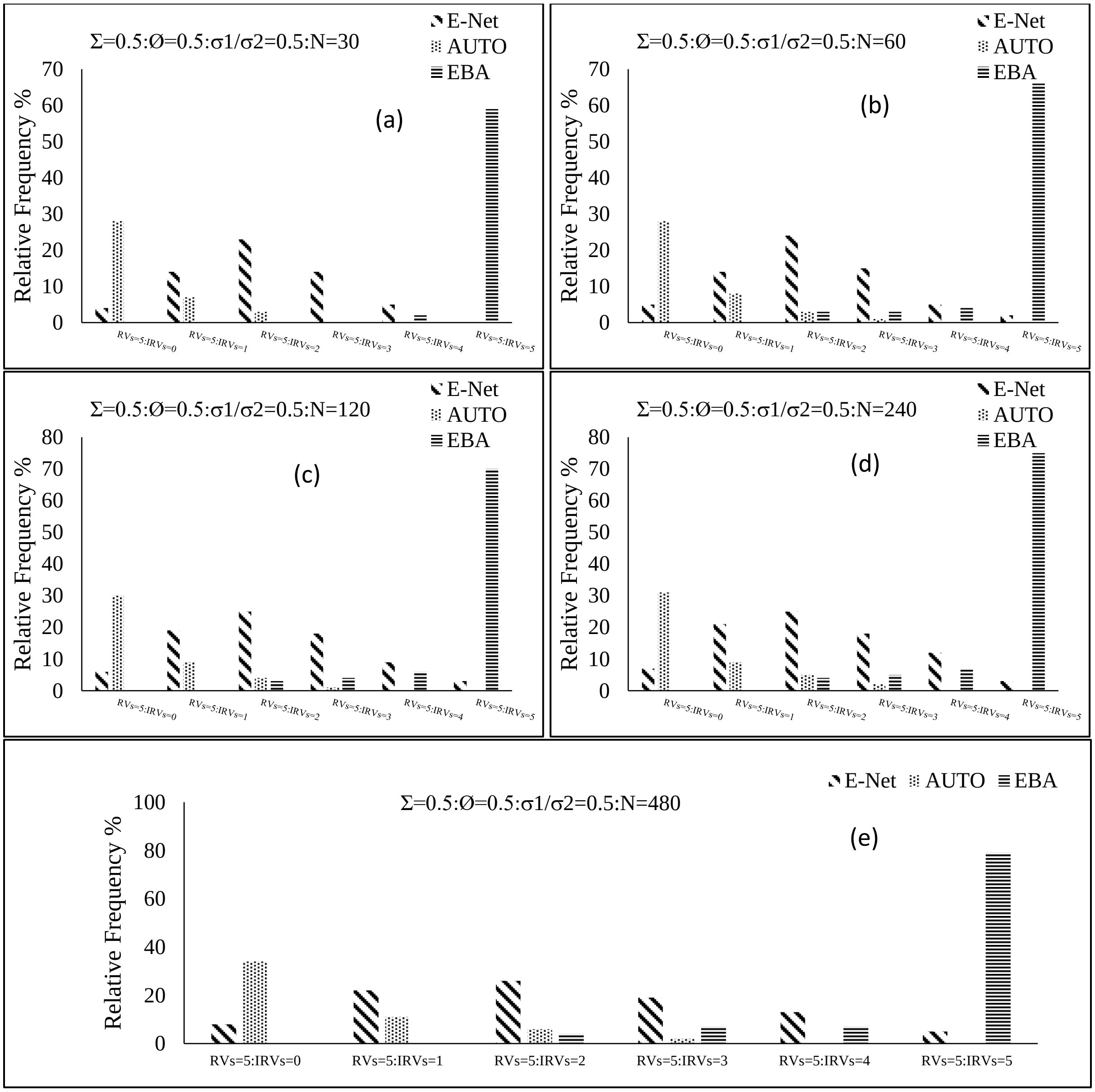


Figure 2.32: Relative Frequencies of DGP When There Are Multiple Problems in Data (%) {Σ=0.5: Ø=0.5:σ1/σ2=0.5}

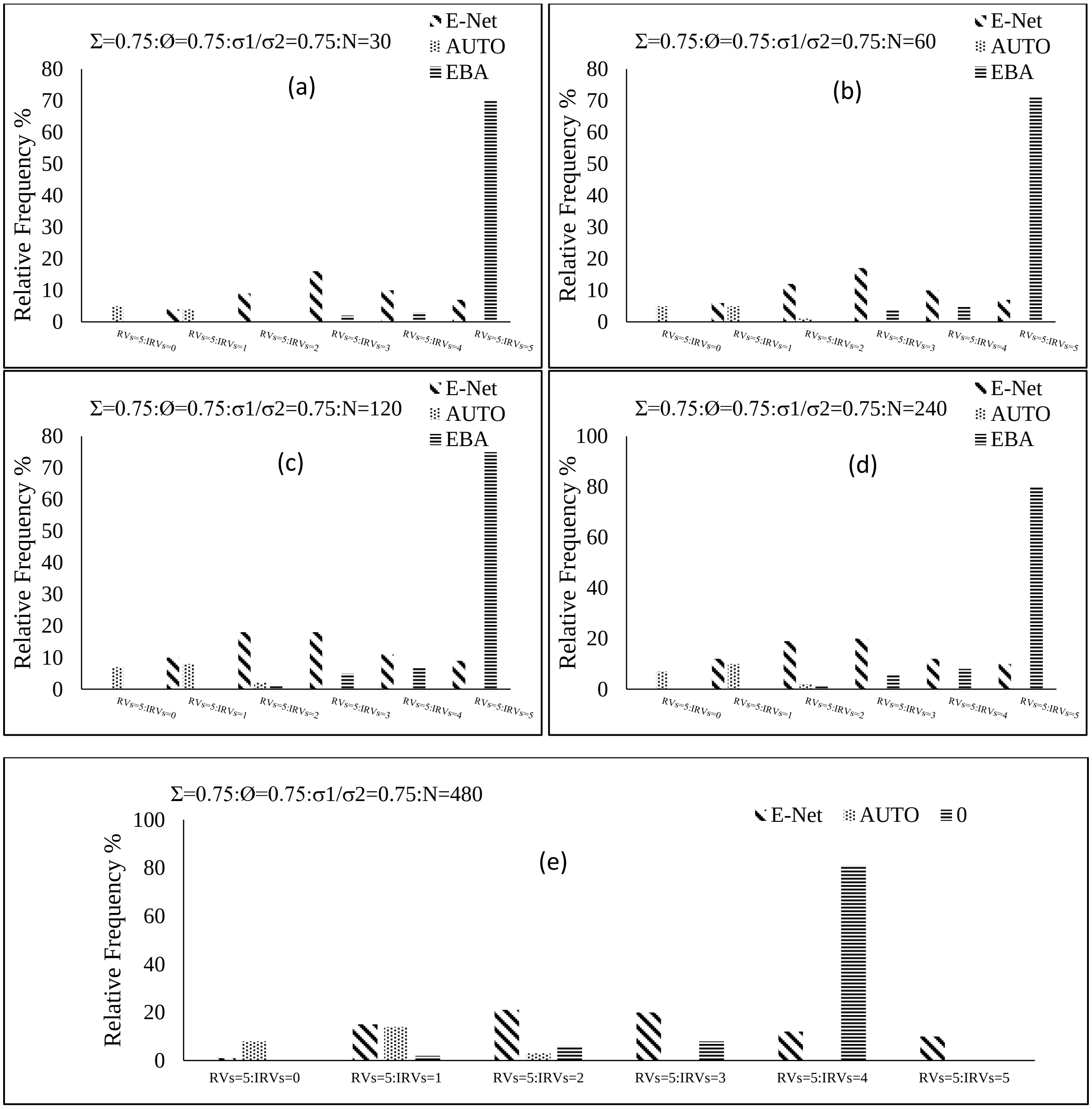


Figure 2.33: Relative Frequencies of DGP When There Are Multiple Problems in Data (%)

{Σ=0.75: Ø=0.75:σ1/σ2=0.75}

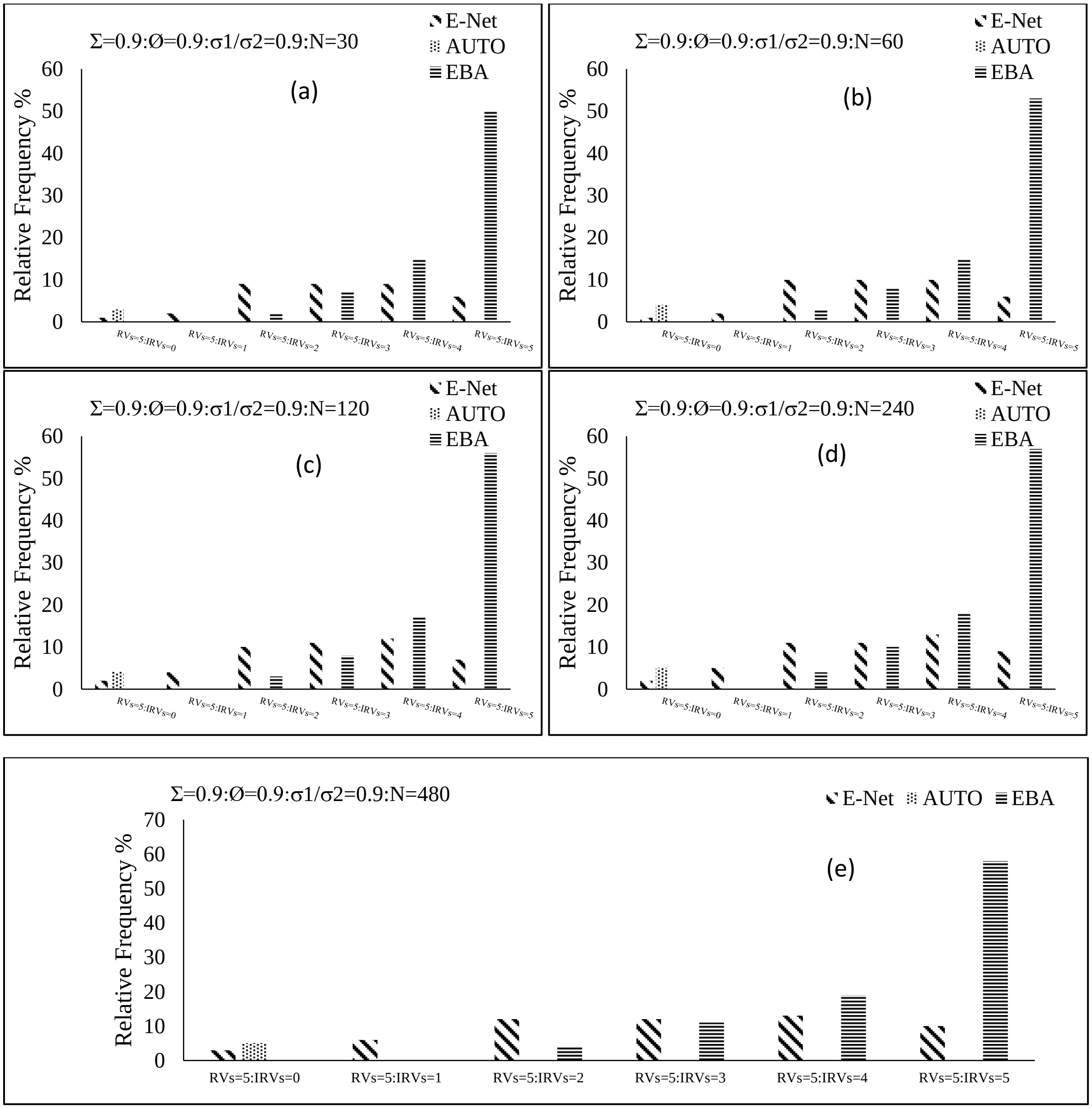


Figure 2.34: Relative Frequencies of DGP When There Are Multiple Problems in Data (%)

{Σ=0.9:Ø=0.9:σ1/σ2=0.9}

The following table shows the frequencies of irrelevant variables in case of multiple problems in DGP

Table 2.9: Relative Frequencies of Irrelevant Variables when there are multiple problems in data (%)

	E-Net					Autometrics					EBA				
Sample size / No.of Vars	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**	**30**	**60**	**120**	**240**	**480**
Σ =0.1 Ø =0.1 σ1/σ2=0.1															
IRVs=1	80	81	85	88	91	15	16	18	18	29	80	90	100	100	100
IRVs=2	40	50	52	53	61	2	2	3	3	6	46	50	60	94	96
IRVs=3	10	12	19	22	26	0	0	0	0	1	6	40	55	80	90
IRVs=4	4	4	5	7	10	0	0	0	0	0	1	10	44	51	68
IRVs=5	1	1	2	2	2	0	0	0	0	0	0	0	13	21	26
Σ =0.25 Ø =0.25 σ1/σ2=0.25															
IRVs=1	84	85	85	86	88	18	19	19	20	20	98	100	100	100	100
IRVs=2	57	60	61	65	66	2	2	3	4	6	90	95	100	100	100
IRVs=3	20	22	24	25	26	0	0	0	0	0	65	70	100	100	100
IRVs=4	5	5	6	9	10	0	0	0	0	0	32	55	90	100	100
IRVs=5	0	0	0	0	1	0	0	0	0	0	7	50	70	95	98
Σ =0.5 Ø =0.5 σ1/σ2=0.5															
IRVs=1	90	91	94	95	95	10	11	12	14	14	100	100	100	100	100
IRVs=2	60	61	69	70	73	1	1	3	4	5	100	100	100	100	100
IRVs=3	25	28	32	40	41	0	0	1	1	4	98	100	100	100	100
IRVs=4	10	10	12	14	17	0	0	0	0	0	87	90	100	100	100
IRVs=5	0	0	3	3	4	0	0	0	0	0	60	70	99	100	100
Σ =0.75 Ø =0.75 σ1/σ2=0.75															
IRVs=1	90	92	93	95	96	5	5	7	8	10	100	100	100	100	100
IRVs=2	60	62	65	66	68	0	0	0	1	1	100	100	100	100	100
IRVs=3	30	31	33	35	36	0	0	0	0	0	100	100	100	100	100
IRVs=4	20	22	23	25	29	0	0	0	0	0	99	100	100	100	100
IRVs=5	1	1	2	3	3	0	0	0	0	0	85	88	90	97	99

$\Sigma = 0.9$ $\varnothing = 0.9$ $\sigma 1/\sigma 2 = 0.9$															
IRVs=1	80	82	90	91	92	1	1	3	5	6	100	100	100	100	100
IRVs=2	67	70	82	83	84	0	0	0	0	0	98	98	99	100	100
IRVs=3	45	50	57	58	60	0	0	0	0	0	93	94	97	97	98
IRVs=4	30	33	34	34	35	0	0	0	0	0	78	79	80	80	81
IRVs=5	15	17	21	22	22	0	0	0	0	0	55	56	65	66	69

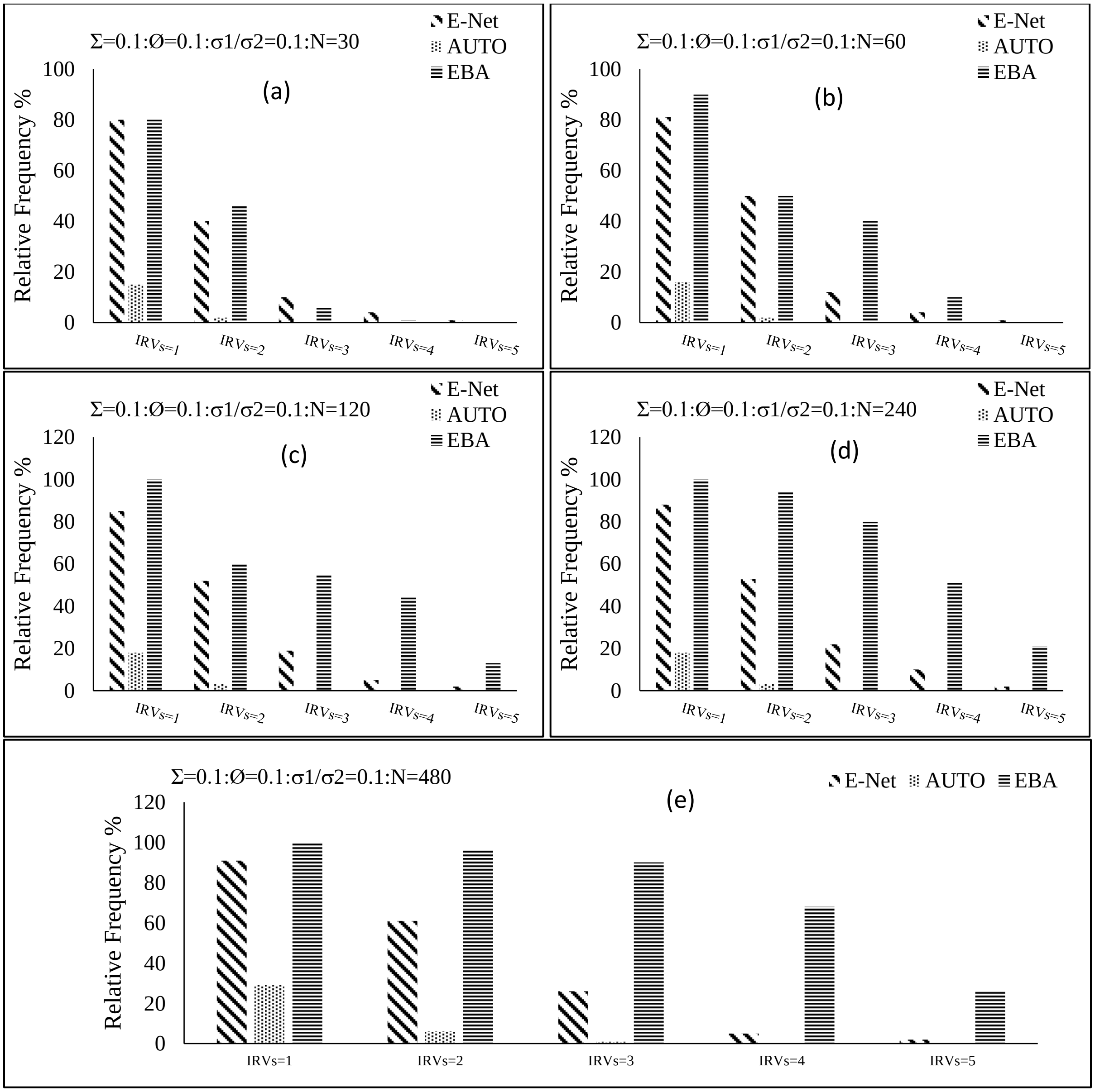


Figure 2.35: Relative Frequencies of Irrelevant Variables When There Are Multiple Problems in Data (%)

{Σ=0.1:Ø=0.1:σ1/σ2=0.1}

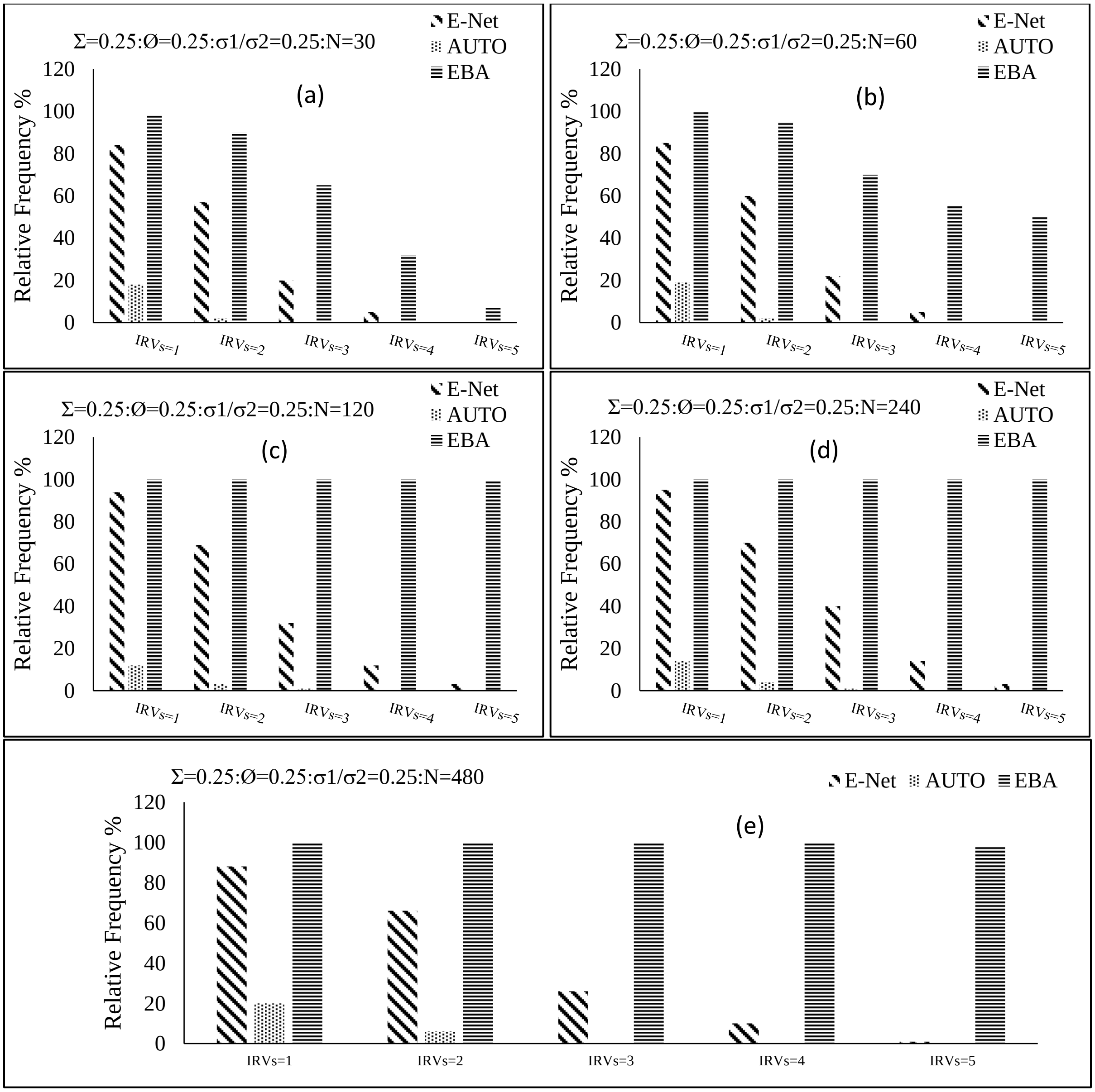


Figure 2.36: Relative Frequencies of Irrelevant Variables When There Are Multiple Problems in Data (%)

{Σ=0.25:Ø=0.25:σ1/σ2=0.25}

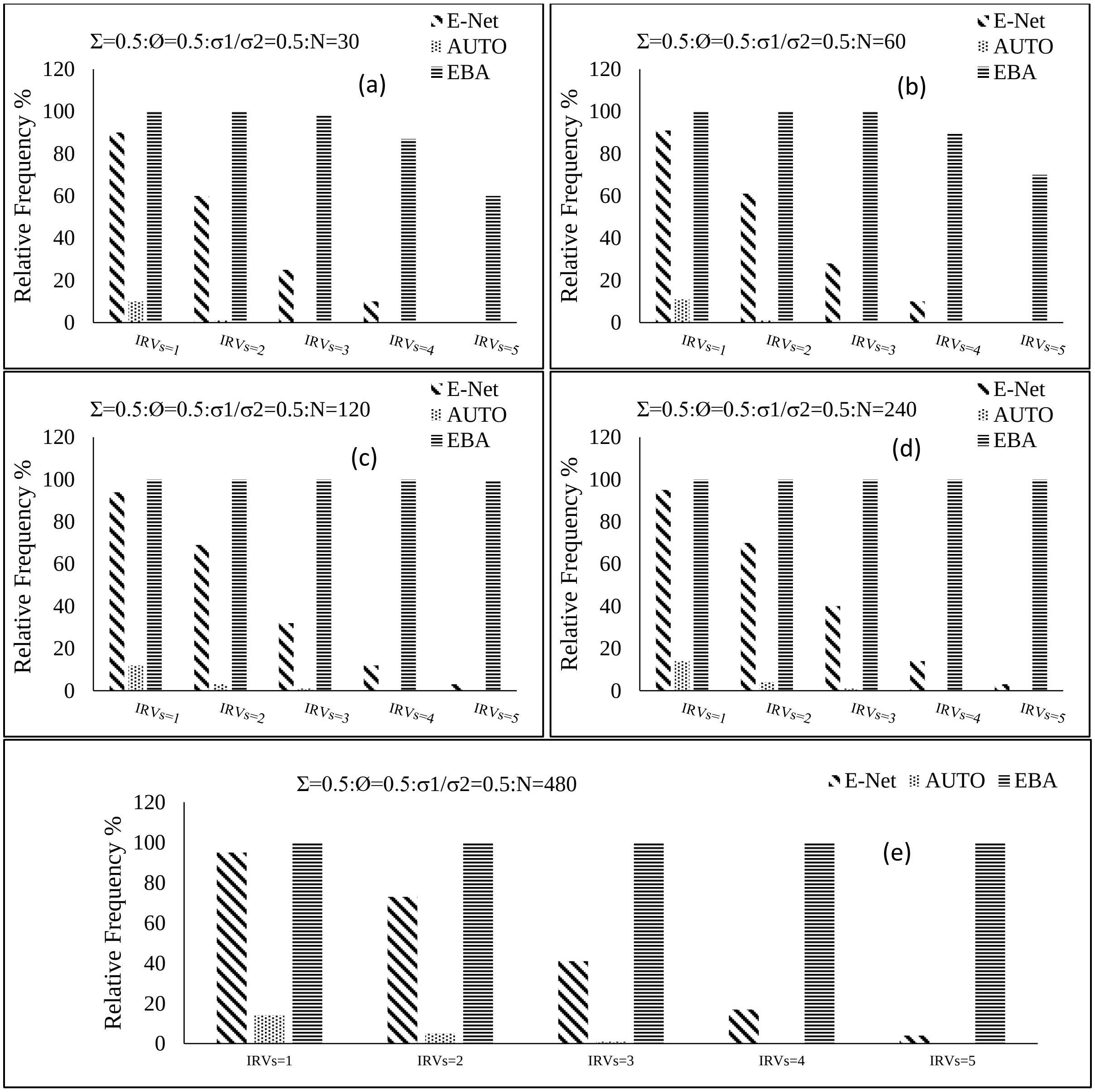


Figure 2.37: Relative Frequencies of Irrelevant Variables When There Are Multiple Problems in Data (%)

{Σ=0.5:Ø=0.5:σ1/σ2=0.5}

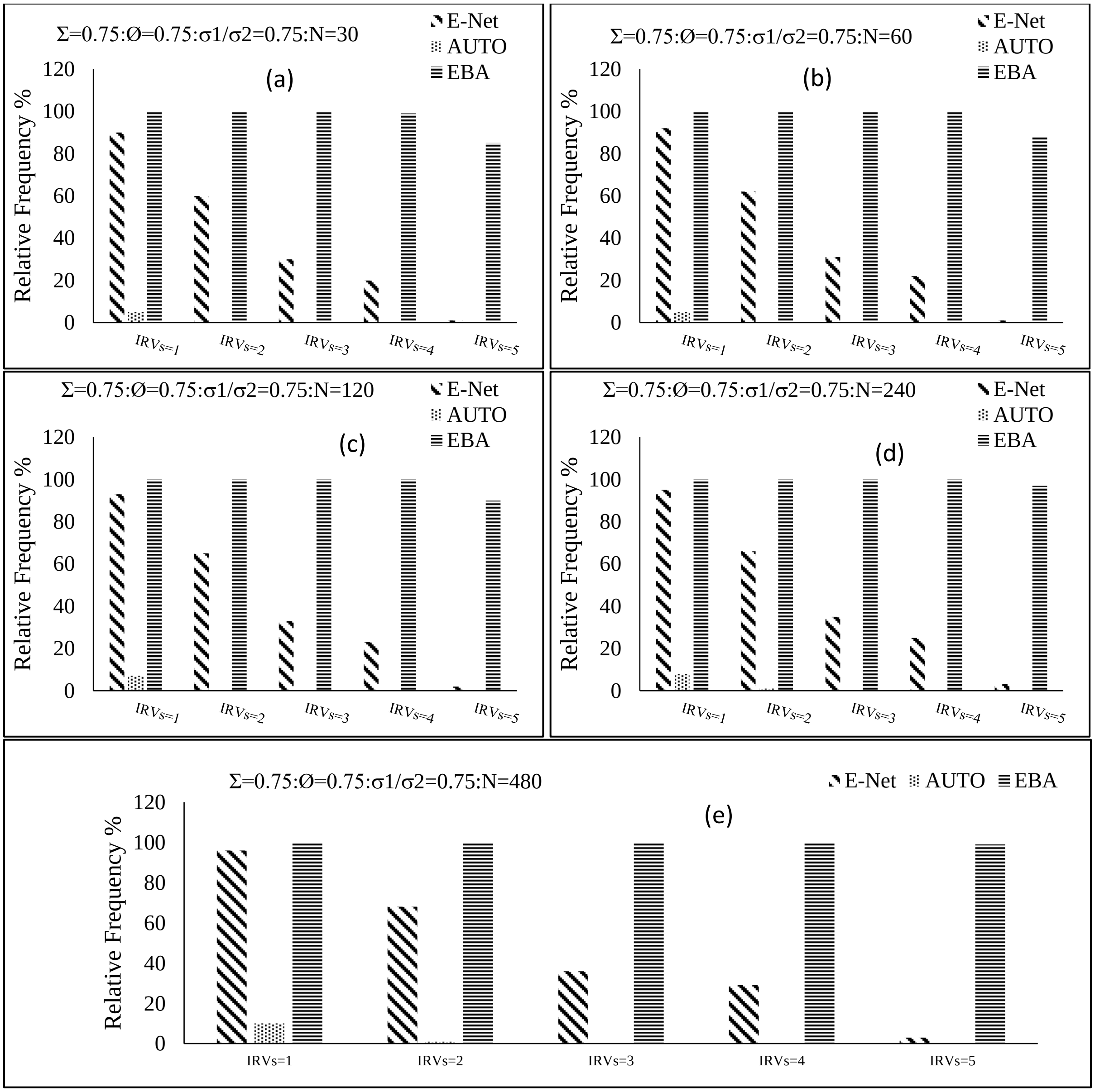


Figure 2.38: Relative Frequencies of Irrelevant Variables When There Are Multiple Problems in Data (%)

{Σ=0.75:Ø=0.75:σ1/σ2=0.75}

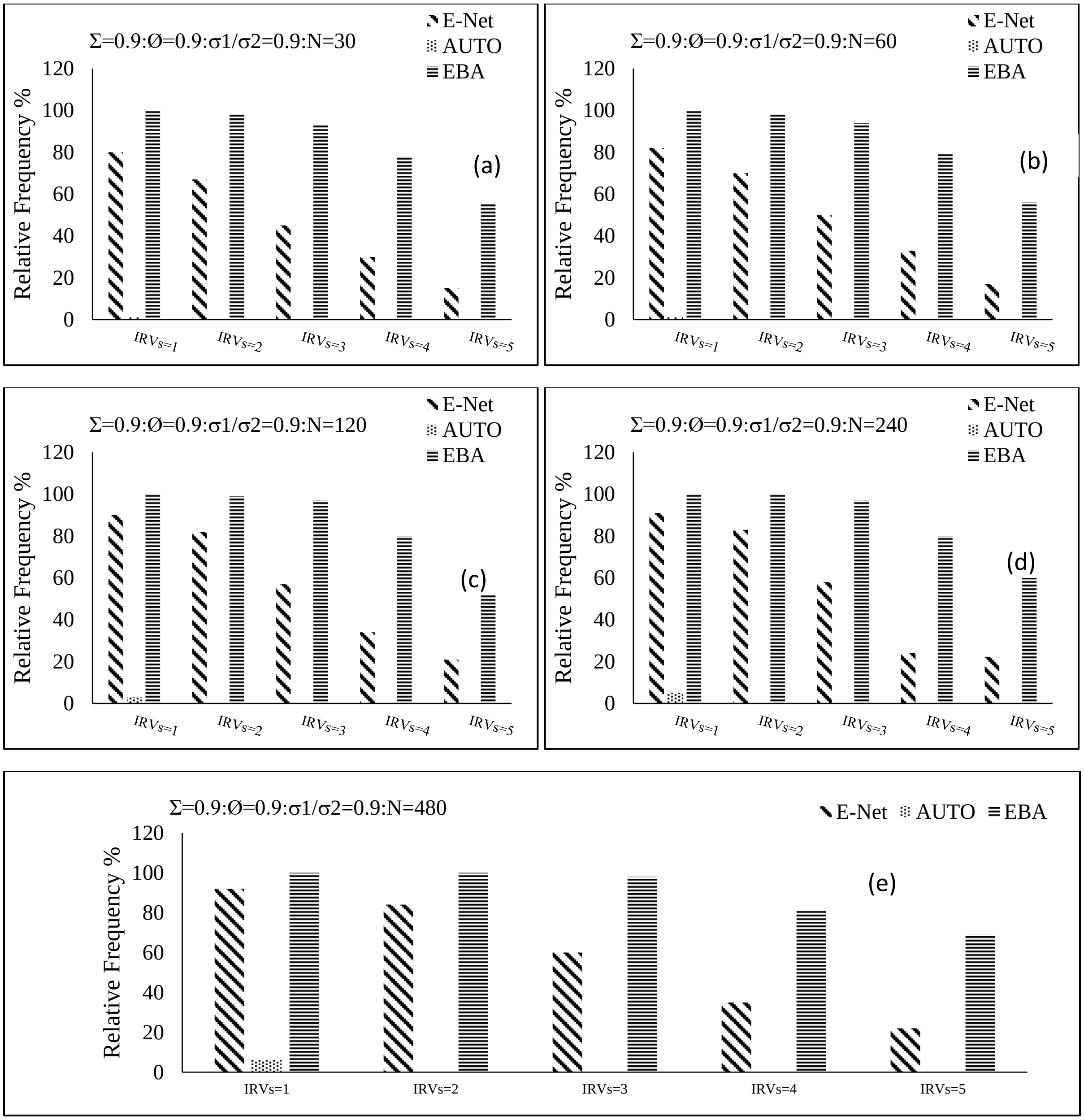


Figure 2.39: Relative Frequencies of Irrelevant Variables When There Are Multiple Problems in Data (%)

{Σ=0.9:Ø=0.9:σ1/σ2=0.9}

Appendix 3

Variables are retained for the model of Economic Growth and ✓ are used for the variable that is retained by the variable selection procedure

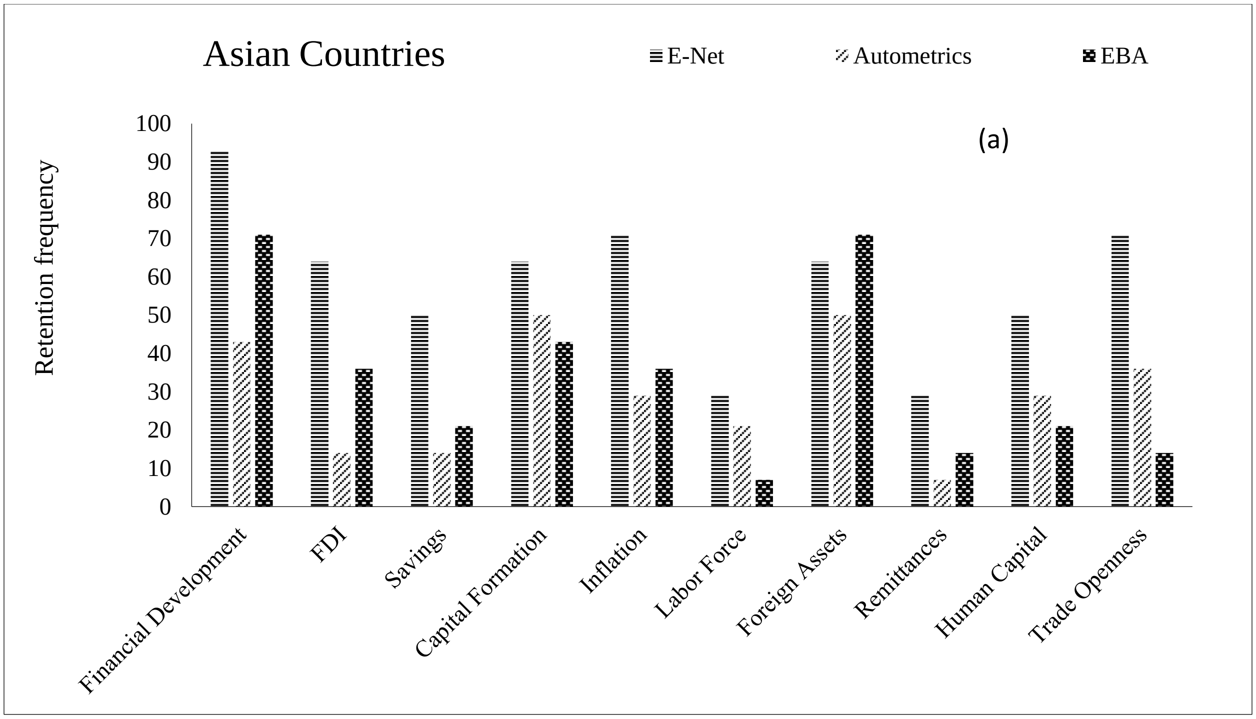
Table 3.1: Variables retained for the model of Economic Growth

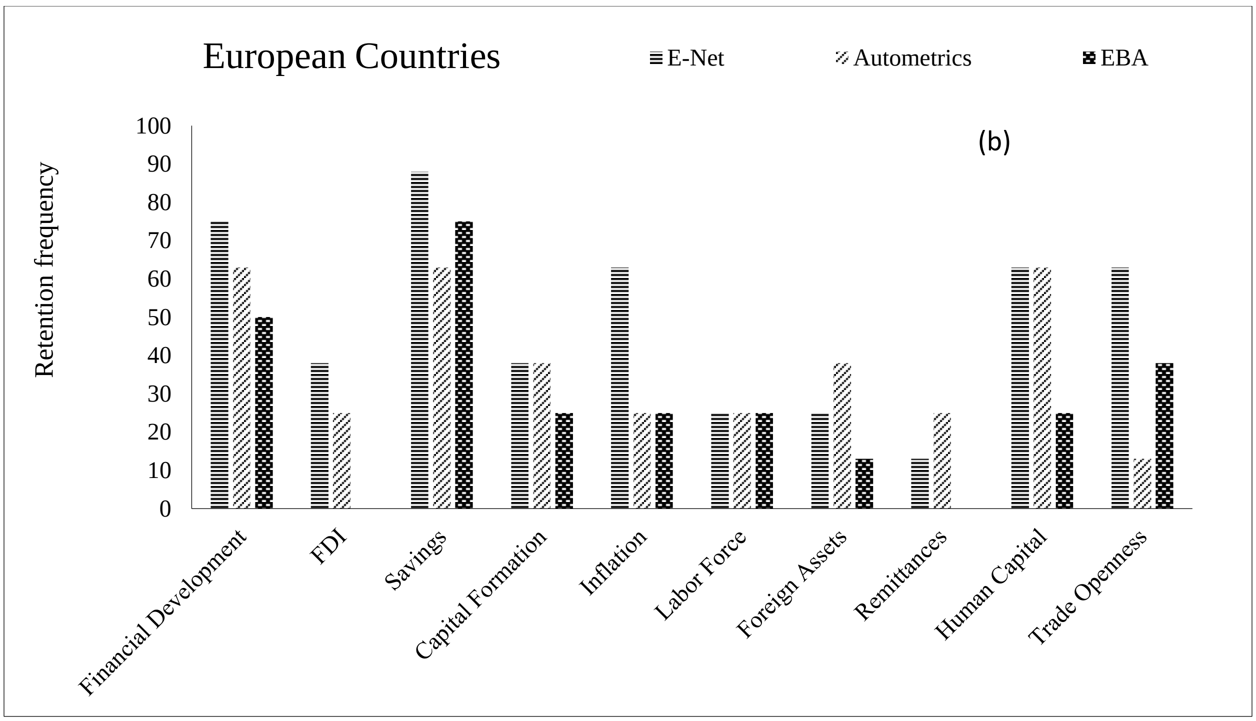
		RMSE	MSPE	Financial Developm	FDI	Savings	Capital Formati	Inflation	Labor Force	Foreign Assets	Remitta nces	Human Capital	Trade Opennes
PAK	E-Net	1.84	14.4	✓	✓	✓	✓	✓		✓	✓	✓	
	Auto	1.80	0.83				✓	✓			✓		
	EBA	1.81	0.78	✓			✓	✓			✓		
IND	E-Net	1.76	3.55	✓			✓	✓		✓			✓
	Auto	2.09	6.44	✓		✓							✓
	EBA	3.78	8.99	✓				✓				✓	
BAN	E-Net	1.08	1.67	✓	✓	✓	✓			✓		✓	✓
	Auto	1.90	1.30						✓				✓
	EBA	1.03	0.05	✓		✓	✓			✓			
SRI	E-Net	1.87	5.20				✓			✓			
	Auto	1.89	27.09				✓			✓			
	EBA	1.80	3.17		✓		✓			✓			
CHI	E-Net	2.41	13.48	✓	✓	✓	✓	✓		✓		✓	✓
	Auto	2.39	2.22		✓				✓	✓			
	EBA	3.20	10.90	✓	✓			✓		✓			
RUS	E-Net	4.55	7.36	✓	✓			✓				✓	
	Auto	3.77	8.99	✓		✓						✓	✓
	EBA	4.97	10.76	✓			✓			✓			
KOR	E-Net	3.62	15.0	✓	✓			✓		✓			✓
	Auto	3.94	5.96				✓	✓		✓			
	EBA	4.05	13.33	✓				✓		✓			✓
HOK	E-Net	3.22	5.60	✓		✓			✓				✓
	Auto	5.37	5.26	✓						✓		✓	
	EBA	7.88	7.46		✓		✓			✓			✓
JAP	E-Net	1.65	1.23	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Auto	1.69	0.34	✓			✓						
	EBA	1.70	0.76	✓	✓	✓				✓			
TUR	E-Net	3.46	0.28	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Auto	3.75	1.67		✓		✓	✓		✓			
	EBA	4.11	1.93						✓	✓		✓	

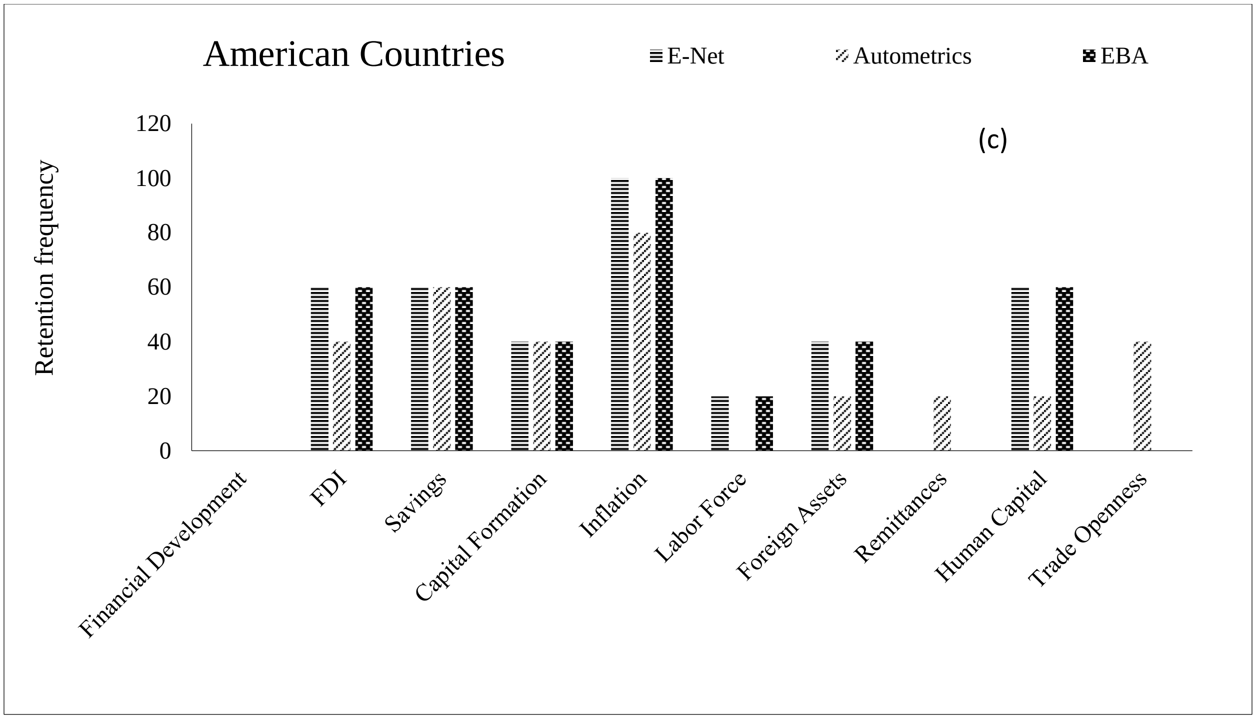
✓=retained

		RMSE	MSPE	Financial Developm	FDI	Savings	Capital Formati	Inflation	Labor force	Foreign Assets	Remitta nces	Human Capital	Trade Opennes
KUW	E-Net	5.64	20.22	✓									✓
	Auto	5.83	2.91	✓								✓	
	EBA	7.09	14.63	✓						✓			
OMN	E-Net	4.38	14.41	✓		✓	✓	✓	✓		✓	✓	✓
	Auto	4.16	5.22	✓					✓				✓
	EBA	4.62	3.80	✓							✓	✓	
SAU	E-Net	6.91	18.39	✓	✓		✓	✓		✓			
	Auto	6.69	14.27				✓			✓		✓	
	EBA	6.86	1.04		✓		✓	✓					
UAE	E-Net	4.33	6.77	✓	✓			✓					✓
	Auto	5.34	8.55				✓	✓		✓			✓
	EBA	8.44	10.44	✓		✓				✓			
GER	E-Net	3.22	6.44			✓	✓	✓					✓
	Auto	4.33	3.55		✓		✓			✓		✓	
	EBA	5.67	6.23	✓				✓					✓
UK	E-Net	1.30	0.32	✓	✓	✓	✓	✓				✓	
	Auto	1.25	0.15	✓			✓	✓					
	EBA	1.44	0.26	✓		✓		✓					
FRA	E-Net	1.05	0.73	✓		✓	✓	✓		✓		✓	✓
	Auto	1.21	0.21			✓	✓			✓		✓	
	EBA	1.18	0.08			✓						✓	
ITL	E-Net	1.48	7.22			✓							
	Auto	1.82	5.67			✓		✓			✓	✓	
	EBA	1.42	0.32			✓						✓	
SPI	E-Net	1.63	0.79	✓	✓	✓			✓		✓	✓	
	Auto	1.74	1.36	✓		✓			✓				
	EBA	1.58	0.62	✓		✓			✓	✓			
SWI	E-Net	3.44	1.33	✓		✓		✓				✓	✓
	Auto	2.47	2.12	✓	✓	✓				✓			
	EBA	6.77	6.38	✓			✓						✓

		RMSE	MSPE	Financial Developm	FDI	Savings	Capital Formati	Inflation	Labor force	Foreign Assets	Remitta nces	Human Capital	Trade Opennes
AST	E-Net	1.36	2.44	✓					✓				✓
	Auto	1.45	0.72	✓					✓		✓	✓	✓
	EBA	1.90	3.45			✓			✓				
SWE	E-Net	1.44	3.03	✓	✓	✓		✓		✓		✓	✓
	Auto	1.54	4.32	✓		✓						✓	
	EBA	2.33	5.26			✓	✓						✓
US	E-Net	1.52	0.38		✓	✓	✓	✓		✓		✓	
	Auto	1.79	0.11			✓		✓				✓	
	EBA	1.63	0.21			✓	✓	✓					
BRA	E-Net	3.19	10.24			✓		✓		✓		✓	
	Auto	3.20	15.23			✓							✓
	EBA	3.11	13.78			✓				✓		✓	
MEX	E-Net	3.11	0.62		✓		✓	✓	✓			✓	
	Auto	2.91	0.58		✓		✓	✓					
	EBA	2.88	0.23				✓	✓				✓	
CAN	E-Net	2.0	0.54			✓		✓					
	Auto	1.75	0.16			✓	✓	✓		✓	✓		
	EBA	1.95	0.44		✓	✓							
PER	E-Net	4.43	3.43		✓			✓					
	Auto	5.24	3.97		✓			✓					✓
	EBA	4.43	0.70		✓			✓					✓
EGY	E-Net	1.87	1.88		✓	✓	✓		✓	✓		✓	✓
	Auto	1.72	0.28				✓						✓
	EBA	1.70	0.21		✓	✓	✓						✓
NIG	E-Net	4.22	7.01			✓	✓	✓	✓	✓		✓	✓
	Auto	5.53	14.75				✓					✓	
	EBA	3.97	8.76				✓					✓	✓
ETH	E-Net	5.91	8.72	✓					✓	✓	✓		
	Auto	6.41	7.86						✓		✓		
	EBA	5.49	7.58						✓	✓	✓		
SA	E-Net	1.48	1.09	✓	✓	✓	✓	✓	✓	✓	✓		✓
	Auto	1.75	1.15					✓		✓	✓	✓	✓
	EBA	1.72	0.04					✓		✓			✓
MOR	E-Net	2.31	4.15	✓	✓						✓		✓
	Auto	3.21	5.10			✓	✓			✓			
	EBA	1.32	3.40	✓			✓			✓			✓


Asian Countries
E-Net
Autometrics
EBA
(a)
Retention frequency
100
90
80
70
60
50
40
30
20
10
0
Financial Development
FDI
Savings
Capital Formation
Inflation
Labor Force
Foreign Assets
Remittances
Human Capital
Trade Openness


European Countries
E-Net
Autometrics
EBA
(b)
Retention frequency
100
90
80
70
60
50
40
30
20
10
0
Financial Development
FDI
Savings
Capital Formation
Inflation
Labor Force
Foreign Assets
Remittances
Human Capital
Trade Openness



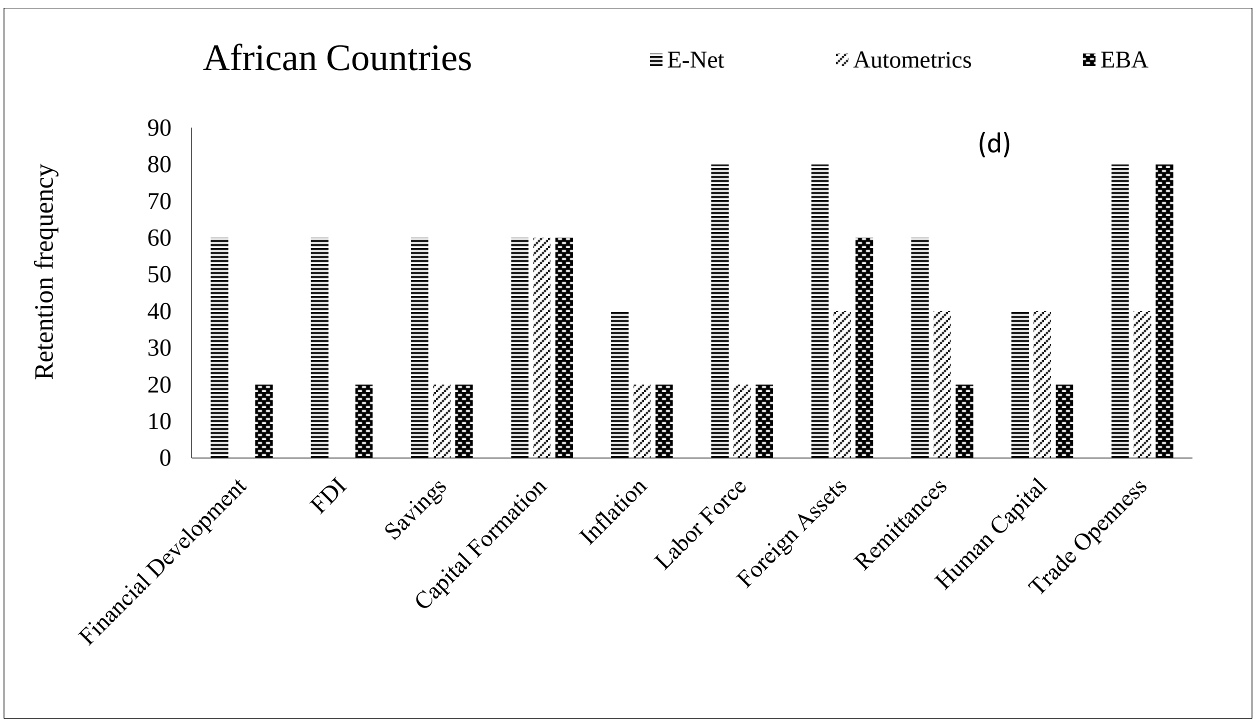


***Figure 3.1*:** Variables retained for the model of Economic Growth