TESTS FOR CAUSALITY IN TIME SERIES: MODIFICATION, COMPARISON AND APPLICATION

Submitted by:

Rizwan Fazal

PIDE2015FPHDETS03

Supervisor:

Dr. Atiq-ur-Rehman

Department of Economics and Econometrics

PAKISTAN INSTITUTE OF DEVELOPMENT ECONOMICS ISLAMABAD

2020

Certificate of Approval

This is to certify that the research work presented in this thesis, entitled: **"Tests for Causality in Time Series: Modification, Comparison and Application"** was conducted by Mr. Rizwan Fazal under the supervision of Dr. Atiq-ur-Rehman. 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. Rizwan Fazal
PIDE2015FPHDETS03

Signature: ________________

Examination Committee:

a) **External Examiner: Dr. Abdul Jabbar**
Assistant Professor
IIIE, International Islamic University, Islamabad

Signature: ________________

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

Signature: ________________

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

Signature: ________________

Dr. Karim Khan
Head, Department of Economics & Econometrics
PIDE, Islamabad

Signature: ________________

Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled **"Tests for Causality in Time Series: Modification, Comparison and Application"** 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: ________________

Name: Mr. Rizwan Fazal

PIDE2015FPHDETS03

Author's Declaration

I Mr. Rizwan Fazal hereby state that my PhD thesis titled **"Tests for Causality in Time Series: Modification, Comparison and Application"** 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. Rizwan Fazal

PIDE2015FPHDETS03

ACKNOWLEDGEMENTS

First of all, I would like to pay my profound thanks to Almighty Allah with those infinite blessings; I have been able to complete this thesis. I offer my courteously respectful and heartfelt thanks to the Holy Prophet Hazrat Muhammad (Peace be upon Him) who urges his followers to “seek knowledge from cradle to grave”.

I am deeply indebted to my supervisor Prof. Dr. Atiq ur Rehman, for his persistent guidance, stimulating criticism, keen interest and personal involvement, and ever smiling cooperation during the entire period of my PhD.

I would like to express my gratitude to my respectable teachers: Dr. Saud Ahmad Khan, Dr. Hafsa Hina, and Dr. Amna Urooj who remain extremely supportive in my research work. I would also like to acknowledge the support of my friends Zaid Ashraf, Tariq Majeed and Asad Shahbaz who have contributed to my work and gave me support in MATLAB Programming. Thanks also to Numan Ahmad, Zeeshan Ahmad and Shakeel Shahzad who have contributed to my work indirectly by making me fresh through jokes and Dawati Tableeg.

I would also like to acknowledge the support of Kevin D. Hoover (Professor of Economics, Duke University) for giving some brief comments and basic tips.

I am thankful to my wife *Hira*, my daughter Zunaira and my son Zakwan, who have suffered more than any other and contributed a lot to completing my PhD. During my coursework and research, every Friday Zunaira called and I remember her words ***“Baba Ap Es Friday ko Aaengy”***.

Last but not the least I am thankful to my parents for their prayers, sacrifices, physical contributions, patience and immortal love that encouraged me at every flash of study period.

RIZWAN FAZAL

ACRONYMS

AR	Autoregressive
ARDL	Autoregressive Distributed Lag
ARMA	Autoregressive Moving Average
CPI	Consumer Price Index
DGP	Data Generating Process
EG	Economic Growth
FDI	Foreign Direct Investment
FER	Foreign Exchange Reserve
GDP	Gross Domestic Product
GMD	Geweke, Meese and Dent
GTA	Graph Theoretic Approach
IFS	International Financial Statistics
MPC	Modified Peter and Clark
MTM	Monetary Transmission Mechanism
NX	Net Export
OECD	Organization for Economic Co-operative and Development
PC	Peter and Clark
POM	Potential Outcome Model
SBP	State Bank of Pakistan
SEM	Structural Equation Models
SLRS	Static Long Run Solution
SVAR	Structural Vector Autoregressive
VAR	Vector Autoregressive

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table No. **Description** **Page No.**

ABSTRACT

Causality is a central problem in all social sciences and primary question facing researcher is to find the casual direction. This central question has no reliable answer. There have been several approaches to test causality i.e. Simon (1953) approach, Wold (1954) approach, Granger causality (1969), Sims causality (1972) and Peter and Clark (PC) algorithm of Graph Theoretic Approach (GTA) developed by Spirtes et al (1993) and Pearl (2000). But there are serious theoretical and empirical weaknesses attached to some of these causality tests. After development of Granger causality, it was thought, initially, that the issue of determining the causal relation would be resolved, but it, too, has major flaws, as Granger causality determines predictability, not the causality; sometime the cause occurs later than the consequences. Among these approaches; theoretically PC algorithm of GTA looks sound and can be held as a preferred approach for testing causality. Because the recent development in graphical models and logic of causality show potential for alleviating difficulties of causal modeling (Pearl, 1998). But how it performs empirically? The literature carries no answer to this question. As it is not known, to what extent the PC algorithm is capable to differentiate between genuine and spurious causal assumption. So current study investigated the size and power properties of PC algorithm of GTA. This study also modifies the PC algorithms with different measure of correlations and evaluated the performance of Modified PC algorithm that how much it is capable of uncovering the true and spurious causal relationship.

This study used Monte Carlo simulations to evaluate performance of PC and Modified PC algorithms of graph theoretic approach. Results of Monte Carlo simulations indicate that PC algorithm (treating VAR residuals as original variables) and Modified PC algorithm (treating Haugh-ARMA residuals as original variables) continuously maintains the size but does not have reasonable power. It is also evident from the simulated results that stationary and non-stationary series with different specifications (drift and trend) and autoregressive coefficients do not affect the size of PC algorithm (using VAR residuals) and Modified PC algorithms (using Haugh-ARMA residuals). The size of Modified PC algorithm (using Modified R recursive residuals) inflates for nonstationary series but it has good power. In case of stationary series when the auto regressive coefficients are near to unity, it also performs well, having high power. But when the auto regressive coefficients in the data generating process tend towards zero, Modified PC algorithm (using Modified R recursive residuals) fails to maintain power. The performance of these procedures is also evaluated when there is confounding variable in the data generating process. The results indicate that performance of Modified PC algorithm (using Modified R recursive residuals) in finding the correct causal path is better than PC algorithm (using VAR residuals) and Modified PC (using Haugh-ARMA residuals). After evaluating the performance of PC and Modified PC algorithms, causal determinants of inflation are estimated using appropriate causality approach having optimal statistical size and power properties.

Key words: Econometrics, Time Series, Graph Theoretic Approach, Causality.

CHAPTER 1

INTRODUCTION

Testing Causality among variables is one of the most important and, yet, one of the most difficult issues in econometrics and economics. The difficulty arises from the non-experimental nature of social sciences. For natural sciences, researcher can perform controlled experiments and identify the causal structures among variables. There is no such privilege for social sciences, and controlled experiments are often impossible, therefore, one has to deal with observational data, for causal analysis (Lin, 2008). It is then necessary to discover causal relations by analyzing statistical properties of purely observational data (Glymour, et al 2019). In observational data, causal inferences are among the most difficult inferences, and several issues arise in it (Le, et al 2019). First and the most important; causality cannot be observed in non-experimental data. Zaman (2009) mentioned that in observational data, only correlation and timing are observable, and causality is inherently unobservable. Freedman et. al. (2007) discuss many cases in which observational studies led to wrong conclusions, sometimes with disastrous consequences. Secondly, one cannot control basic confounding factors in observational data, and a significant type of bias arises from confounding variables[1] (Landeiro and Culotta, 2017). Thirdly, common measures of relationship are symmetric, therefore one cannot find the direction of causality i.e. whether x causes y or y causes x. Finally, like every other econometric test, problems of size and power are associated with each test of causality. Researchers have faced these difficulties in the causal inferences and proposed different causality

[1] Landeiro and Culotta (2017) argued that a confounder z is a variable that is correlated both with the input variables x and the target variable y. When z is not included in the model, the true relationship between x and y can be improperly estimated.

tests for determining it. These include Simon (1953) approach, Wold (1954) approach, Granger causality (1969), Sims causality test (1972), Graph Theoretic Approach developed by Spirtes et al (1993) and many others.

Simon (1953) approach is closely associated with the approach of Cowles commission founded in 1932. Simon says that causality can be checked in structural models not only between exogenous and endogenous variables, but also among the endogenous variables themselves. Wold (1954) gave the concept of controlled experiment for understanding the meaning of causality and suggest that definition of causality is simple in experimental than observational studies.

Granger (1969) used predictability as an approximation to a loose concept of causality called Granger causality. Thus, according to Granger definition of causality "x Granger Cause y" when x_{t-1}can predict y_t. Zellner (1979) and many others criticize Granger (1969) for different reasons, but most important reason is that this approach is atheoretical. Researchers must impose restrictions when Granger approach is implemented practically. Zellner (1979) says, if restriction imposed by researcher are not valid, theoretically, then this approach will discover only accidental regularities.

Another test of causality was originated by Sim (1972) and was further developed by Geweke et al (1983) and came to be known as Sims-GMD test Charemza (1997). According to Sims, one can regress variable Y on lag and lead values of variable X, and if causality runs from X to Y only, future values of X in the regression should have coefficients equal to zero. So, Sims argues that future cannot cause current or past. This concept of causality refers to lagged and lead relationship among economic series.

Among these approaches for determining causality, as discussed above, Pearl (2000) considers the work of Simon (Structural model) as a major contribution in the field of econometrics. Structural equation models are adapted in social sciences and behavioral science, and utilized, but their use in economics is not so common. Wright (1921) a well-known genetics expert and developer of SEM and econometricians, Haavelmo (1943) and Koopmans (1950) are the prime mover of SEM. Like other social variables, economic variables are connected in multiple ways, and the relationship is better described by set of structural equations representing multiple causal paths. Wright (1921) developed SEM for the purpose of investigating quantitative cause and effect by combining theoretical assumption of cause and effect with statistical data. The dominant feature of SEM compared to other approaches is that it checks causality in multiple directions. Other approaches like Granger causality and Sim's Causality check causal inferences in single equation setup with one directional causality. It often happens that apparent causality between two variables *A* and *B* is because of third variable *C* which is causing both *A* and *B*, but *A* and *B* have no direct causal linkage. The Granger causality does not take account of this kind of common cause. Having capability to take care of multiple causal paths, SEM seems better, theoretically.

However, causality in SEM is also considered to be controversial and enigmatic. Many SEM researchers are having difficulty articulating the causal content of SEM, and are seeking foundational answers (Pearl, 1998). The confusion is vividly portrayed in the influential work of Wilkinson (1999) in his study 'Statistical methods in psychology journals: Guidelines and explanations', says that SEM works on basis of correlation coefficients and "Correlation does not show Causation" but in the same study, the author mentions that "The parameters estimated by the use of SEM have interpretation as causal effect". Thus, question arises that if SEM does not prove causation, then under what conditions structural parameters yield causal interpretation? Spirtes et al (1993) offer solutions to the controversial and enigmatic causal content of SEM, which brings development in the area of graphical models or graph theoretic approach (GTA). The recent development in graphical models and logic of causality show potential for alleviating such difficulties and thus revitalizing SEM as the primary language of causal modeling *(for detail discussion see Pearl, 1998)*

Graph-Theoretic Approach is progressively applied in natural sciences and most of research on Graph theoretic approach were generally not designed for time series data, and assumes the causal ordering in cross sectional data *(for detail discussion see Demiralp, 2003)*. Though the methodology is a better match for causality analysis in economics. Swanson and Granger's (1997) for the first time applied it in economics. They realized that vector autoregressive model (VAR) residuals carry causal ordering, and adapt graph theoretic causal search to the problem of finding the correct causal order of structural vector autoregressive model (SVAR). They assumed that information about causal ordering of contemporaneous variables of SVAR may be present in the covariance matrix of VAR error terms. They have used method by estimating VAR and use the residuals as original variables to find causal order, which were used for SVAR.

As mentioned above Swanson and Granger estimate VAR model and use the residuals to find causal order. However, using VAR residuals is likely to remove the non-stationarity problem from the data and correlation can be used to determine causality with very low chance of being spurious. However in VAR model with variable X_t and Y_t, X_t is assumed the function of Y_{t-1} and others i.e. $X_t = f(Y_{t-1} \ldots.)$

and conversely Y_t is assumed the function of X_{t-1} and others i.e. $Y_t = f(X_{t-1} \ldots.)$. So, in equation of Y_t only effect of X_t could be there, effect of X_{t-1} and past values (X_{t-i} where i > 1) are removed. Thus, VAR residuals carry only contemporaneous information about cross variable effect.

On the other hand, the univariate model assumes $Y_t = f(Y_{t-1})$ and residual should carry effect of X_t as well as the Past value i.e. X_{t-i}. Therefore, it should be more powerful to used univariate model residuals in Peter and Clark (PC) algorithm of graph theoretic approach. Therefore, the first contribution of this study is to use univariate residuals in the PC algorithm.

There is also lack of studies on the size and power properties of these causal algorithm. It is not known, to what extent the PC algorithm and Modified PC of graph theoretic approach is capable of uncovering the true causal relationship and how good it is to differentiate between genuine and spurious causal assumption. The second contribution of the study is to evaluate the size and power properties of PC algorithms and Modified PC algorithm by replacing VAR residuals with procedure presented by Haugh[2] (1976) ARMA-residuals and Rehman and Malik[3] (2012) modified R recursive residuals using Monte Carlo simulation.

The final contribution of the study is that after evaluating the performance of PC and Modified PC, causal determinants of inflation will be estimated using appropriate causality approach having optimal statistical size and power properties.

---

[2] Haugh 1976 for the first time introduced a test to check the independence of two stationary autoregressive moving average (ARMA) series based on the residual cross correlation. Haugh test is based on two steps procedure. First step is to fit ARMA models to each of the series and find residuals series of all variables. Second step is to find the cross correlation between two resulting residual series.

[3] introduced modified R which measure the correlation between recursive forecast error of AR model fitted to both series. If we analyze the association between recursive forecast residuals it is expected to give more valid measure of correlation between time series. This new statistic is given the name Modified R

1.1 Research Objectives

The objectives of the study are as follows:

1) To evaluate the performance of PC algorithms of Graph Theoretic approach for causal structure of a VAR using Monte Carlo simulation.
2) To evaluate the performance of Modified PC algorithms of Graph Theoretic approach by replacing VAR residuals with procedure presented by Haugh (1976) and Rehman and Malik (2014) or Modified R test residuals using Monte Carlo simulation.
3) To Compare PC and Modified PC algorithm of Graph Theoretic approach.
4) To find causal determinants of inflation using appropriate causality approach.

1.2 Research Outline

Chapter two provides a brief discussion on different proposed tests for causality, causal search algorithms and development in graph theoretic approach. It also contains a brief discussion on empirical tools used for testing causality.

Chapter three contains discussion on methodologies which are used for empirical analysis. The methodology is based on two components; Data generating Process and Monte Carlo Simulations. It also explains notion, terminologies, and different methods for Graph theoretic approach to determine the causal pattern.

Chapter four provides a brief discussion on simulations results (in term of size distortion under different specifications) obtained by using PC and Modified PC algorithms of graph theoretic methods.

Chapter five and six contain empirical results which are obtained by using PC and Modified PC algorithms. The comparison is made between PC and Modified PC algorithms in term of Power. The former anlayse the case when we have confounding variable, and the later analyses when we do not have it, in the data generating process.

In chapter seven we have summarized our findings from the simulation results in the previous chapters. Eight and final chapter is about the real data analysis using the appropriate algorithm having minimum size distortion and high power.

CHAPTER 2

LITERATURE REVIEW ON CAUSALITY

2.1 Pre- Historic Era

The concept of causality is as old as human history itself. Usually Democritus (460-370 BC) is said to have recognized the importance of causal reasoning and causal explanations. However, it is evident that causal reasoning and causal explanations came very early in human development. The Holy Book Quran while explaining the story of Adam and Eva Says

[2:36] 'O Adam, dwell you and your wife in the Janah, and eat there from abundantly wherever you will, but approach not this tree, lest you be of the wrongdoers.'

[2:37] But Satan caused them both to slip by means of it and drove them out of the state in which they were. And we said: 'Go forth; some of you are enemies of others, and for you there is an abode in the earth and a provision for a time.

In the above verse [2:36] Allah gives warning to Adam while in verse [2:37] cause and effect are realized. That after eating from the tree Allah forbid Adam (AS) from and sent to earth as a consequence. Adam became expert in causal explanations because when Allah asks "Did you eat from the tree"? Adam answered "The female (Eva) that Allah bestowed to be with me, gave me the fruit and I ate". Now the point here is that Allah did not called explanations, but this was Adam who felt the need to explain. So, this is a clear picture that causal explanation was a product of human intellect.

2.1.1 Aristotle

Aristotle's name is among those great philosophers who also contributed to philosophy of causality. Aristotle for the first-time presented theory of causality, and gave four different modes of causation called material cause, formal cause, final cause and efficient cause. The first two causes (material and formal) are among the concerns of economic ontology while causal modeling in economics deals in the way what Aristotle called efficient cause. He argues that there may be multiple causes, but there is final one cause (Nodelman et al, 2003).

2.1.2 Hume (1752)

Theoretical discussions on causality can be found in the works of David Hume, who made distinction between analytical claims and empirical claims. He classifies that the causal claim to be on empirical side and all empirical claims are originated from experiences. It means that causes too originate with experiences. Hume was skeptical of any causal inference and believed that causal events were ontologically reducible to non-causal events, and causal relation where not directly observable but could be known by means of experiences (Asghar, 2007), Demiralp and Hoover (2003).

2.2 Modern Causality: A Review

2.2.1 Theoretical Background

Improvement in statistical methods like regression, multiple correlation and development of structural models, is closely connected to causal inferences. It is fairly understood that, unlike correlation, regression has a natural direction: therefore, people often treat regression as causation while in fact regression does not contain any causal information. As discussed in Chapter 1, Cowles Commission members, with the inception of structural models, tried to overcome the difficulties in causal order. But first the problem of simultaneity bias, and later on serious problem of identification were noted in their approach, which led to insensible estimation "if the equation was not identified". Thus, with the econometrics work of Cowles Commission, two main approaches on the issue of causality emerged.

First modern approach to causality which is given the name ''*process analysis''* by Herman Wold (1954) also known as Wold approach. This approach is based on temporal precedence that cause must occur before effect (or cause and effect will follow pattern or sequence). Wold approach analysis belong to the time series tradition that ultimately produced Granger causality and vector autoregression (see section 2.2.3 and 2.4 below). The second approach which is called "*Simon approach*" attributed to Simon (1953). This approach is closely associated with Cowles commission approach founded in 1932, which shows that causality could be found for structural models.

2.2.2 Simon's (1953)

Simon pointed out that

> "The notion of causality is a deductive logical concept relation to model's characteristics not to empirical features of the world that require statement of inductive logic"

Simon (1953) approach is closely associated with the approach of Cowles commission founded in 1932. Consider the bivariate system:

$$Y_t = \theta X_t + \varepsilon_{1t} \ldots (1)$$

$$X_t = \varepsilon_{2t} \ldots (2)$$

Simon argued that X_t cause Y_t, because X_t is recursively ordered ahead of Y_t. One knows all about X_t without knowing about Y_t, but one must know the value of X_t

to determine the value of $Y_{t.}$ Both equations (1) and (2) also appear to show that any shock in equation (2) would transfer to equation (1); while any shock in (1), say a change in θ would not transfer to (2). Apparently X_t could be used to control $Y_{t.}$[4]
He showed that causality could be found through structural models. Simon says that causality can be checked among endogenous variables as well, in structural models.

2.2.3 Granger Causality (1969)

The most popular approach in economics used for analysis of causal relation in time series data was presented by Granger (1969) called Granger Causality. The Granger definition of causality follow sequence or pattern and based on the idea of predictability. Like Wold, Granger assumes that cause occurs before effect and if past values of x can predict current value of y, then x causes y. Granger (1969) used predictability as an approximation to a loose concept of causality called Granger causality which is based on two important assumptions. The first that cause occur before consequences. While the second that causality will make sense only for stochastic stationary variables.

Granger gave the concept of predictability through which analysis of causality are drawn. Thus, according to his definition of causality "x Granger cause y" when lag values of x can predict current y. Hoover et al (2006) argues that Granger causality is a process approach as well as inferential approach. This is process approach because it is data based, without reference to background economic theory and inferential approach because it was developed for the purpose to apply to time series models.

[4] Kevin D. Hoover (2006)

2.2.4 Criticism on Granger Causality

Granger causality received two major criticism. For instance, first, it determines order or predictability, not the causality. Secondly, sometime the cause occurs later than the consequences. Many authors have given elegant examples that ordering does not necessarily imply causality i.e. Hicks (1979) and many others deny to accept Granger presented causality and gave different examples in which effect occur before cause.

Zellner (1979) and many others criticizes Granger causality (1969) for many different reasons. Zellner (1979) criticizes Granger causality for two reasons. The first and most important reason is that this approach is atheoretical; that researcher must impose restrictions when Granger approach is implemented practically- limit the information set to manageable number of variables, consider only a few moments of probability distribution (in our exposition, just mean), and so forth. Zellner (1979) says that if the restriction imposed by researcher are not valid theoretically then this approach will discover only accidental correlation (regularities). Secondly, Granger-causality does not take into account the structural ordering, and often investigated for bivariate processes. However, different conclusions may be reached when more than two variables are considered. If more than two variables are present, non-causality conditions become more complicated (Song and Taamouti, 2019).

Hoover (2008) mentioned in his article that the concept of Granger causality fails to capture structural causality. "Suppose one finds that a variable *A* Granger-cause another variable *B*. This does not necessarily imply that economic mechanism exists by which *A* can be manipulated to affect *B*. This existence of such a mechanism in turn does not necessarily imply Granger causality either (*for more detail see Hoover 2001, pp. 150-155)".*

2.3 Sims (1972) and Sims-GMD (1983)

Sims (1972) causality is based on the fundamental axiom that "the past and present may cause the future. But the future cannot cause the past" (Kuersteiner, 2010). Geweke, Meese and Dent (1983) further developed Sims (1972) test and therefore called it Sims-GMD test (Charemza, 1997). To test whether x is Granger cause of y, the following equation (3) is consider instead of equation (4).

$$X_t = \sum_{j=1}^{k} \gamma_j X_{t-j} + \sum_{j=-m}^{k} \delta_j Y_{t-j} + v_t \ldots (3)$$

$$Y_t = \sum_{j=1}^{k} \alpha_j Y_{t-j} + \sum_{j=1}^{k} \beta_j X_{t-j} + \varepsilon_t \ldots (4)$$

The above equation (3) contain lead values of ys which can be confirmed from the negative lower summation limit for the variable y_{t-j}. This means that in equation (3) future values (i.e. y_{t+1}, y_{t+2}.....y_{t+m}) as well as past values (i.e. y_{t-1}, y_{t-2}.....y_{t-m}) appear. If the coefficients of future values δ_{-1} = δ_{-2} = ...δ_{-n} ≠ 0? However, future does not cause present, so necessary condition for x to not cause y is that all future values of y_s in equation (3) must be equal to zero. Thus, logical conclusion of finding nonzero values of the coefficients δ_{-1} = δ_{-2} = ...δ_{-n} show that x is Granger cause of y. This concept of causality refers to lagged and lead relationship among economic.

2.4 VAR Approach to Causality

Sims (1980) developed VAR model as a reaction to the method of Cowles Commission. VAR is closely related to Granger analysis for causal prospective. VAR is easily applicable, but difficulties arises when we turn to policy analysis. Starting with the SVAR model of the form:

$$\Gamma Y_t = B(L)Y_{t-1} + e_t \ldots (5)$$

Where Y_{t-1} is $n * 1$ vector of contemporaneous variables, Γ and $B(L)$ represent $n * n$ matrix and polynomial in lag operator respectively. e_t is $n * 1$ vector of uncorrelated disturbance as the covariance matrix Σ is diagonal: that it contains zero elements. "The matrix Γ defines the causal interrelationship among the contemporaneous variables. The system is identified provided there are $n(n-1)/2$ zero restrictions on Γ.

Multiplying Γ^{-1} on both sides of equation (1) yield reduce-form or VAR:

$$Y_t = \Gamma^{-1}B(L)Y_{t-1} + \Gamma^{-1}e_t$$
$$Y_t = B^*(L)Y_{t-1} + U_t \ldots (6)$$

Where $B^*(L) = \Gamma^{-1}B(L)$ and $U_t = \Gamma^{-1}e_t$

A typical problem with VAR (equation (6) is that covariance matrix Σ is not diagonal: that it does not contain any zero elements; which mean that error terms are correlated with each other's, so shock in one is a shock to both. Sims initially advocated solution "Choleski decomposition" to orthogonalize the shocks by choice of recursive order.

Leamer (1985), Cooly and LeRoy (1985) criticized that which recursive order will be chosen, on which the substantive results (Impulse response function) depend. They convinced Sims (1986) that meaningful economic interpretation required identification of Γ. Sims (1982, 1986) accept and consider the point of Leamer et al and introduced SVAR which can be identified through the contemporaneous causal order only[5]. Normally the results of SVAR is interpreted by using impulse response

[5] Hoover (2006) "Causality in Economics and Econometrics"

function. If impulse response function of one variable say “x” to another variable “y” is significant this implies that x cause y.

Identification Problem

If we knew Γ, identification problem is reduced and SVAR in equation (5) can be easily recovered from VAR in equation (6) but the covariance matrix is no more diagonal. If we don’t know Γ matrix, then we have to impose restrictions on it to get identification. To achieve identification first we will make the covariance matrix $\Omega = \mathrm{E}(\mathrm{P_i^{-1}U}(\mathrm{P_i^{-1}U})')$ of (equation (6)) diagonal through orthogonalizing transformation. Let P={Pi} are set of orthogonalizing transformations.

The main identification issue in SVAR is that when we don’t have any information about matrix Γ, then selecting one P from set of n that correspond to true data generating process: (Pi = Γ) is not an easy task. “To mitigate the problem of identification we have to impose $n(n-1)/2$ restrictions on Pi. These restrictions which can be imposed in different ways. Many of researchers confine themselves to “just identify” the lower triangular matrices and is given the name choleski decomposition, and there are many such choleski ordering which correspond to Wold causal ordering of the variables. Single choleski ordering can be used for the identification process of Pi”[6]. If someone restrict himself to “just identified” SVAR, and select one choleski order out of n! ordering. Other method given by Blanchard for just identification by keeping economic theory which will tell us that what the causal order should be. Hoover (2005) argues that formal economic theory is rarely decisive about causal order so mostly researcher’s select the order arbitrarily to get just identified SVAR. There are also chance of over identified causal orderings for which

[6] Asghar. Z, Tayyaba. R (2011), “Energy GDP Causal relationship for Pakistan: A Graph Theoretic Approach

identification of Pi, more than $n(n-1)/2$ zero restrictions should be imposed on Pi which is difficult or impossible to do through "Choleski decomposition and through economic theory". Pearl (2000), Spirtes, Glymour and Scheines (1993, 2000) developed new method for selecting Pi (see section 2.6 below).

2.5 Structural Equation Models Approach to Causality

To model cause and effect, two different but mathematically equivalent languages, Path Analysis or SEM and Potential outcome model (POM) have been proposed by Wright (1921), Haavelmo (1943) and Neyman (1923), Rubin (1974) respectively. However, the two approaches are rarely used in standard economic literature.

SEM is conceived for the purpose to investigate quantitative cause and effect by combining theoretical assumption of cause and effect with statistical data. SEM is a set of equations which connect the cause with consequences through all possible causal paths. Pearl (2000) consider the work of Cowles Commission and mostly, the work of Simon (Structural model) as a major contribution in the field of econometrics and argues that Structural model provide valid inferences and better answer to the question of causality. On the contrary, Granger causality is based on single equation, whereas, most often, the relation between economic variables exhibit complex structural path. To make this point clear, let take the example of relation between monetary policy action and its target variables i.e. inflation and output. The text in monetary economics reveals that there are at least six structural paths through which the monetary policy might be affecting the inflation. This can be explained with the help of the following Figure 2.1.

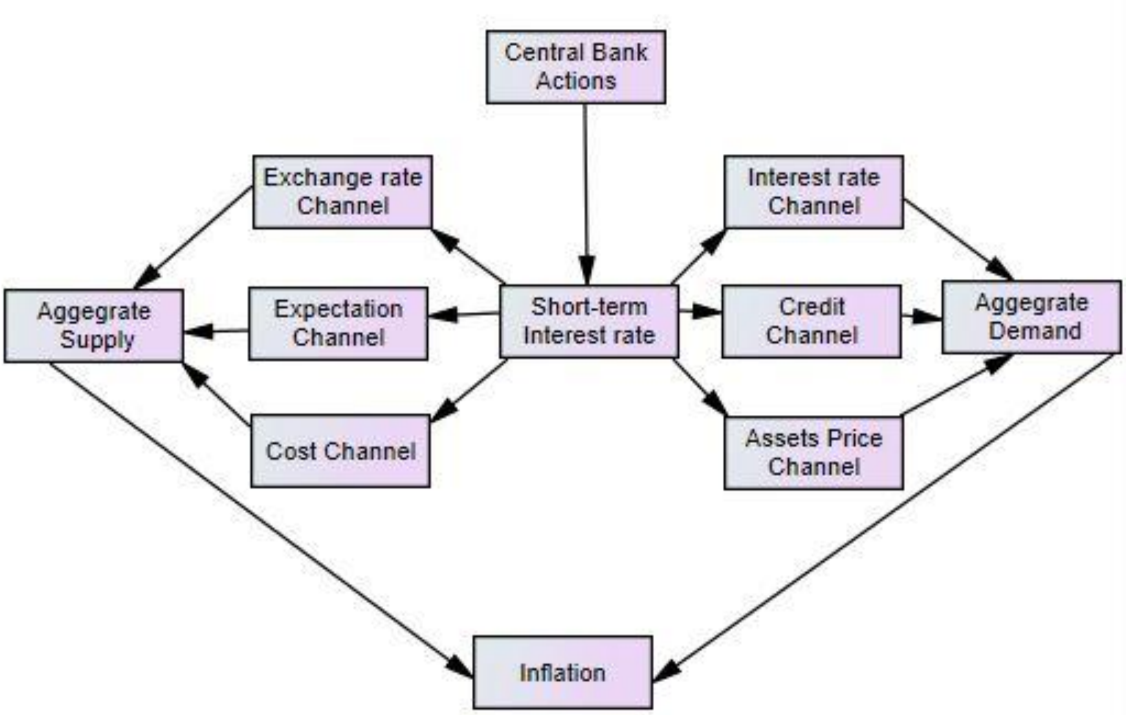


Figure 2.1: Monetary Transmission Mechanism Channels

Usually econometric equation directly regress inflation on central bank policy while ignore other channels. If a researcher, only considers exchange rate channel and take all variables involved in the channel, still his equation would be misspecified because there are multiple paths through which central bank policy affect inflation. This means that many determinants of inflation are missed so the results are expected to be biased. On the other hand, if researchers are taking into account the structural equation modelling, all paths are to be included. This means that all plausible determinants of inflation are included in estimation so the results are expected to be unbiased. Therefore, SEM procedure is expected to provide better results than single equation model.

The dominant feature of SEM, as compared to other approaches is that it checks causality in multiple direction and is capable to model the effect of third variable involved in the relationship. Other approaches like Granger causality and Sim's Causality check causal inferences between two variables. If there is some common cause, Granger causality is unable to take it into the model but it could be captured through SEM. Thus, to the often asked question, "Under what conditions can we give causal interpretation to structural coefficient?" the founding fathers of SEM would have answered, "Al-ways".

Pearl (1998) mentioned that

> "According to Wright and Haavelmo the conditions that make the equation $y = \beta x + \varepsilon$ structural are precisely those that make the causal connection between X and Y have no other value but β and nothing about the statistical relationship between x and ε can ever change this interpretation of β. This basic understanding of SEM has all but disappeared from the literature, leaving modern econometricians and social scientists in a quandary over β".

But most of SEM researchers support that extra conditions are required for SEM to be carrier of causal claim. James et al. (1982) gave condition called self-containment, that correlation between x and e should be zero, $Cov(x, e) = 0$. According to James, whenever $Cov(x, e) \neq 0$, neither the equation nor the functional relation represents causal relation. Econometricians also faced difficulty with the causal reading of structural parameters. Hendry (1995) argues that the status of β may be unclear until the conditions needed to estimate the postulated model are specified. For example, in the model:

$$y_t = z_t\beta + u_t \ldots (7)$$

Until the relationship between z_t and u_t is specified the meaning of β is uncertain since $E(z_t u_t)$ could be either zero or non-zero on the information provided. Freedman (1987) in the critical paper on SEM challenged the causal interpretation of SEM as self-contradictory. These controversies bring alarming tendency among social scientists, economist and econometricians to view SEM as an algebraic object that carries functional and statistical assumptions but void of causal content. Some of economist attribute the decline in the understanding of SEM due to Lucas critique (1976). As economic theory included in the SEM approach do not explicitly account economic agent's expectation. "Agents, according to Lucas, are able to anticipate policy intervention and act contrary to the predication derived from the SEM, since the model usually ignore such anticipation[7]".

Spirtes et al (1993) and (1998) presented that SEM require mathematical and graphical notations to show causal relation so decline in causal content is due to graphical language required in making causal assumption which are basically ignored. Spirtes et al (1993) provide solutions to the problems of SEM which bring development in the area of graphical models. The developments in the areas of graphical models (graph theoretic approach) and logic of causality show potential for mitigating such difficulties and thus revitalizing SEM as the primary language of causal modelling.

2.6 Graph-Theoretic Approach

Graphs have been applied for more than a century to determine the causal pattern. Further improvement in graphical theory provides an effective mathematical language to the researcher to investigate the causal dimension and manipulate them in

[7] Alessio Moneta, Nadine Chlab, Doris Entner, Patrik Hoyer (2011). Causal search in SVAR models. JMLR: Workshop and Conference proceeding 12(2011) 95-118

relation to the associated probability distributions. According to the Cowles Commission, econometric model is the combination of two parts. a) Probability distribution of variable b) Causal structure. Pearl (2000) and Spirtes (2000) showed that there is isomorphism between graphs and probability distribution of variables. This isomorphism allows conclusions about probability distributions to be derived from theorems proven using the mathematical techniques of graph theory.

Graph-Theoretic Approach causal search algorithms are progressively applied in a variety of social sciences other than economics, but are unfamiliar to most economists (Demiralp, 2003). In Graph Theoretic approach, structural model is converted into graph which overcome many problems and bring causality back into the front of researcher and philosophers and have great importance than other Granger causality tests. First, causality tests developed by Granger are applied only to small set of pre-specified and reduced form equations which are only valid with small set of true structural relationships. Second, as argued by Perez et al (2006) through graph theoretic approach one can choose the right regressors. “As shown by Pearl (2000) that incorrect choice of independent variables may results in improper causal inferences. In contrast, the graph-theoretic approach is used to determine the correct set of independent variables.

Graph theoretic approach were generally not conceived with time series data. Swanson and Granger’s (1997) for the first time used Graph-theoretic approaches to causality into the analysis of contemporaneous causal order of SVAR. They assume that information about causal ordering of contemporaneous variables of SVAR is actually contained in the covariance matrix of VAR error terms. Demiralp et al. (2003), Hoover (2005) showed that after estimating VAR model the error terms of that model would be stored and then treated as the original time series variables. But VAR residuals carry only contemporaneous information about cross variable effect which can be explain as under. Consider a VAR model:

$$y_t = \alpha_1 + \beta_1 x_{t-1} + \beta_2 y_{t-1} + \varepsilon_{1t} \ldots (8)$$

$$x_t = \alpha_2 + \beta_3 y_{t-1} + \beta_4 x_{t-1} + \varepsilon_{2t} \ldots (9)$$

After estimating the VAR model, extract residuals series of both equations (8) and (9). However, residual series extracted from equation (8) only effect of x_t could be there, while effect of past values ($x_{t-i}\ where\ i > 1$) are removed. Thus, VAR residuals only contain contemporaneous information about the causal feedback from x to y and vice versa.

On the other hand, there are number of univariate methods, which can eliminate the non-stationarity without purging out the effect of past values. It should have more power if the residuals extract from the univariate methods are used to determine the causal ordering by using PC algorithms of GTA. Therefore, this study intends to modify the original PC algorithm by replacing VAR residuals with univariate models' residuals and give it name Modified PC algorithm. Using the residuals of univariate models (Haugh (1976) and Rehman and Malik (2014)) to find the correct causal ordering is also contribution of the study.

2.7 Empirical Review

Economists are not more familiar with Graph theoretic approach, so empirical literature related to this approach is very rare. We have mentioned very few recently published studies that have used PC algorithms to determine causal paths. The studies which applied Graph Theoretic causal search (PC algorithms) for causality are summarized/arranged topic wise as under.

2.8 Studies Related to Inflation Variable

Yang et al. (2005) investigated international transmission of inflation for G-7 economies using VAR for the analysis of the study. They also discussed PC causal search algorithm and implement it in Tetrad III for the empirical analysis.

Perez et al (2006) applied PC causal search algorithms for the purpose to investigate the causation between agriculture, money, interest rate, prices and real GDP for more than ten economies.

Hoover et al (2008) presented empirical identification of the vector autoregression. The cause and effect of US. M2*. Taking account of cointegration, the methodology combines recent developments in graph-theoretical causal search algorithms with a general-to-specific search algorithm to identify a fully specified structural vector autoregression (SVAR). The SVAR is used to examine the causes and effects of M2 in a variety of ways. The study confirms that M2 is a trivial linkage in the transmission mechanism from monetary policy to real output and inflation[8].

Zahid et al (2010) examined the causal order between money, income and prices by employing graph theoretic causal search algorithms called PC search. Authors concluded that using PC causal search the results support monetarists view that money supply and prices cause income. The authors also mentioned that graph theoretic causal search overcome problem of over identification in vector autoregressive model.

2.9 Studies Related to Economic Growth Variable

Asghar et al (2011) investigated the causal relationship of energy consumption, energy prices, economic growth and gross capital formation for Pakistan applying PC causal search algorithms. The empirical findings of the study show that unidirectional

---

[8] Hoover, K.D; Demiralp, S, and Perez, S.J. (2008),"Empirical Identification of the Vector Autoregression: The Causes and Effects of U.S. M2"presents at the Conference in Honour of David Hendry at Oxford University, 23-25August 2007

causality is running from energy consumption to economic growth which implies that energy conservation policies are in general desirable for the region as a whole.

Nazmus (2016) presented direction of causality between debt and economic growth using a graph theoretic approach for six major OECD countries. The findings of the study suggest that economic growth causes government debt in most economies. The author also noticed that comparison between the full sample and a reduced sample indicates causal direction from growth to debt is a more recent phenomenon.

Li et al (2013) investigated causality among foreign direct investment (FDI) and economic growth (EG). The study has employed the recently developed approach "Directed Acyclic Graph Approach" to examine the causal structure between FDI and EG. The study explores and suggested various findings. First, for developing economies, economic growth causes FDI inflows for whereas FDI induces economic growth. Second, trade is an important intermediary to facilitate the interaction between FDI and other factors. Third, the stock market is found to be an intermediary that amplifies the influence on FDI from many causal variables of FDI for developed countries[9].

2.10 Research Gap

GTA is worth exploring in this context as there is a very rare research that has taken this approach previously in economics. Cooper (1999) and Pearl (2000) provided some evidence of the effectiveness of PC algorithm for graph theoretic approach. Demiralp and Hoover (2003) also present some simulation evidence of the effectiveness in the context of ordering the contemporaneous variables in an SVAR. However, there is lack of studies on the size and power properties of PC algorithm of

---

[9]Li, Yarui ; Woodard, Joshua D. ; Leatham, David J. ; Marchant, Mary A. ; Bosch, Darrell J. (2013). "Causality among Foreign Direct Investment and Economic Growth: A Directed Acyclic Graph Approach"

graph theoretic approach, to what extent the approach is capable of detecting the correct causal relationship and how capable it is to differentiate between genuine and spurious causal assumption. This study also aims to modify PC algorithm of graph theoretic approach (GTA) by using Rehman and Malik (2014) modified R recursive residuals and Haugh (1976) ARMA residuals, that new approach could become valid for finding the true causal ordering in time series. Moreover, this study also aims to investigate size and power properties of Modified PC algorithm to find out how good it is to differentiate between spurious and genuine causal relationship.

CHAPTER 3

METHODOLOGY

This chapter consist of three parts. In the first part, short overview and notions and terminologies used in Graph theoretic method are discussed. In part second, steps involved in PC and Modified PC algorithms to determine the causal structure are discussed. In third part we have presented the simulation design and data generating process.

3.1 Graph Theoretic Approach

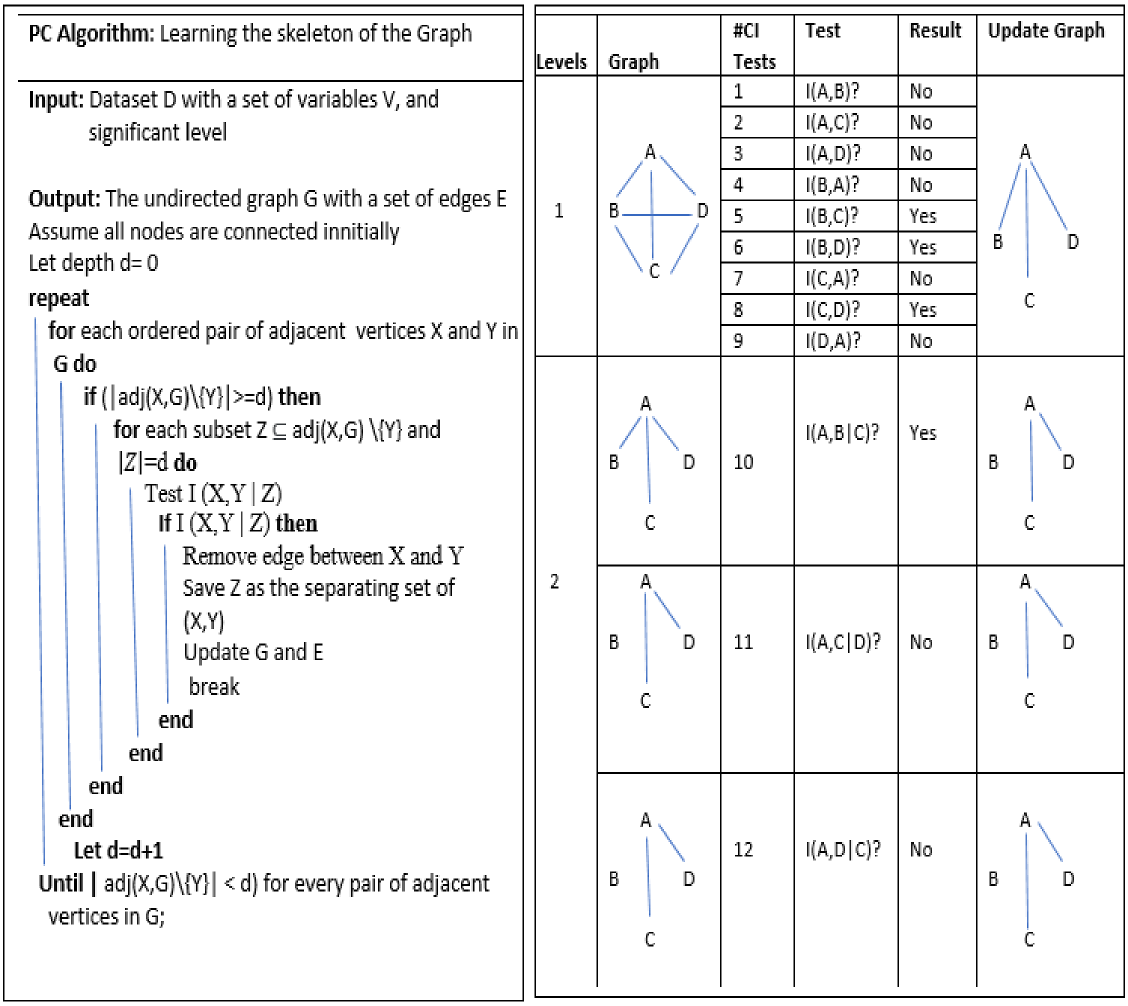
The graph-theoretic approach (GTA) is a novel method that provides an effective mathematical tool for experts and researchers to find the causal direction. Sprites et al (1993) developed graph theoretic approach (GTA) for cross-sectional data. Though the methodology is a better match for causality analysis in economics, economists such as Swanson and Granger (1997), Hoover (2005, 2006) and others have used it for analysis of economic data. Swanson and Granger (1997) for the first time used it for time series data and assumed that information about causal ordering may be present in the covariance matrix of VAR error terms, therefore they treated VAR residuals as original variables in PC algorithm to find causal order.

The use of VAR residuals is likely to remove the non-stationarity problem from the time series data and correlation can be used to determine causality with a very low chance of being spurious. However, in the VAR model with variable X_t and Y_t, X_t is assumed to be a function of Y_{t-1} and others i.e. $X_t = f(Y_{t-1} \dots.)$ and conversely, Y_t has assumed the function of X_{t-1} and others i.e. $Y_t = f(X_{t-1} \dots.)$. So, in the equation of Y_t only effect of X_t could be there, the effect of X_{t-1} and past values

(X_{t-i} where i > 1) are removed. Thus, VAR residuals carries only contemporaneous information about the cross variable effect. This goes against the spirit of various definitions of causality such as Granger Causality. On the other hand, there are several univariate methods, which can eliminate the non-stationarity without removing the effect of cross variable feedback. In this study, we have developed modified PC algorithm of GTA by replacing VAR residuals with modified R recursive residuals.

PC and modified PC algorithms procedure can be divided into two main steps. In the first step, it learns from the data and constructs a skeleton graph based on correlation measure, which contains only undirected edges. In the second step, it orients the undirected edges having arrows head to form the final causal graph. We have displayed the first step; a skeleton graph which explains how the PC algorithm initially constructs the skeleton graph using data information. This can be shown in the following figure 3.

Figure 3.1: PC Algorithms Learning the Skeleton



The right panel of the above figure 3 shows a hypothetical example explaining how the PC algorithm works. We have constructed this scenario in case of the dataset with four variables, A, B, C, and D. Initially the PC algorithm begins with the completely linked graph shown in the first panel. At the first level, the unconditional correlation between all edges is tested. After the implementation of level 1 tests, three links are left. At the level second, the conditional correlation will be checked off each remaining link. For instance, with the edge between A and B, we have at most two tests which are conditioning on C and conditioning on D. If the test returns independence (e.g. I(A, B|C)), we remove the edge from the graph and move to test the other edge. The procedure will stop when there is no test to perform. The final step is the orientation of arrowheads which is based on the screen off the relationship and unshielded collider. Before discussing the modified PC algorithms of GTA in detail, it is helpful to explain some notions and terminologies used in GTA.

The main features of Graph theoretic approach are simple. Any structural equation model can be converted in graph in which arrows represent the causal pattern. The structural equation system provides basis for Graph theoretic approach (*discussed in chapter 2).* In the beginning of Graph theoretic causal search algorithm, all variables are connected through straight lines having no arrowheads as shown in Figure 3.2.

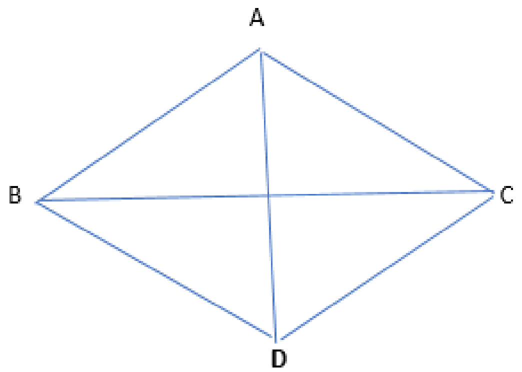


Figure 3.2: Undirected Graph

Graph theoretic causal search algorithms then starts eliminating the links which show insignificant relationship between pair of variables. A hypothetical example, we say that there is insignificant relationship between pair (A and B) and (C and D) so in this case Graph theoretic causal search algorithms will remove the link between them. This can be shown in figure 3.3.

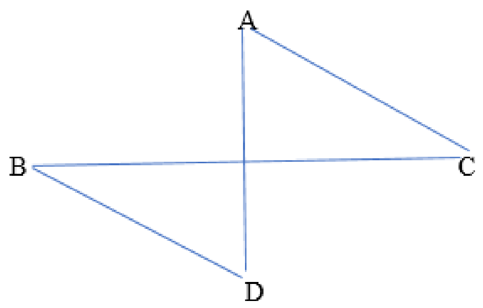


Figure 3.3: Undirected Graph

Similarly, if there is any other pair with insignificant relationship, the related line will be eliminated. Once the elimination is complete, the next step is called orientation in which the algorithm looks for the causal direction and imposes the arrowheads to the lines indicating the direction of causality.

This causal search algorithm has a specific set of terminology, notations and procedures which are mentioned as under.

3.2 Graph Theoretic Approach: Notation and Terminology

In Graph Theoretic approach, structural model is converted into graph. So it is important to show some notation and terms used in graph theoretic approach.

Nodes and edge

A graph consists of a set N of nodes and a set of E edges. The nodes in the graph represent variables while edges show us relationship between pair of nodes or variables. These edges may have arrowheads showing the direction of causation.

Directed and undirected edge

Connection between two nodes through straight line is called *undirected edge* or *adjacent* (A─B). While connection between two nodes through straight line having arrowhead is called *directed edge* (A→B). Edges having arrowheads indicating the direction of causation.

Skeleton Graph

The graph showing only the nodes and strip away all arrowheads from the edges is called skeleton as depicted in Figure 3.1

Path

A path in a graph is chain of connection between two nodes. For example, in Figure 3.4 CAD is a path starting from node C to node D, however, it is not a directed path. In directed path, in a graph, every edge in the path has an arrow which follows the direction of causation. For example, BDE is a directed path from B to E.

Parent and Child

If a node A is linked to another node B by an arrow originating from A to B (A→B) then node A is considering parent of node B, and B is said to be the child of A. This can be explained in the given Figure 3.4 and Figure 3.5.

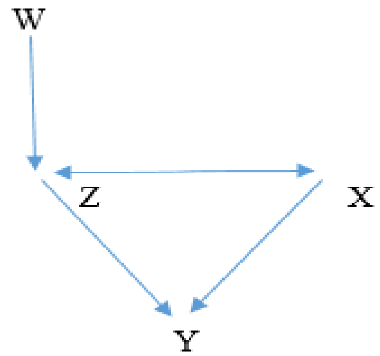


Figure 3.4: Directed Graph

In the above Figure 3.4, Y has two parents' Z and X, three ancestors (X, Z, W) and no children. In figure 3.5. there is a directed path between A and B which means A is ancestor of B and B is descendent of A.

Acyclic graph

The graph is acyclic when there is no feedback causal relationship. If there is arrow head on both ends of an edge, the relationship is called cyclical.

Causally sufficient graph

A graph is causally sufficient when all variables are observable i.e. that there are no latent variables

Condition of faithfulness

A graph and probability distribution is said to be faithful if and only if there is one to one correspondence between conditional independence relationship implied by causal Markov condition.

'Screen-off

This notion can be easy understand as we have two causal graphs *(A→C→B)* and *(A⟵C⟵B)*. In both cases *A* and *B* are not independent because *A* and *B* both depend on *C*, but are independent conditional on *C*. So here *C* is said to screen-off *A* from *B*.

Unshielded collider and shielded collider

The concept of unshielded collider and shielded collider is also important in graph theory. In Figure 4, node *A* and node *B* are unconditionally uncorrelated, however they are correlated conditional on *D*. The classic example given by Demiralp et al (2003) in his paper

> "Suppose *A* = the car battery being charged, *B* = The care starter switch being on, *D* = The car starting. So, *A* and *B* are uncorrelated. Yet, if we know that the car does not start, then knowing that the switch is on increases the probability that the battery is dead".

In the figure 3.5, node *D* is called an unshielded collider on the path *ADB*. As *A* and *B* are unconditionally uncorrelated, but both are correlated conditional on *D*. It is collider because arrow heads are toward *D* i.e. *(→D⟵)* and unshielded because there is no direct link between *A* and *B*. Node *E* is called shielded collider on the path *DEB*. The connection $B \rightarrow D$ acts as shield because *B* and *D* are correlated even without conditional on the common effect [a]. Node E is a shielded collider on the path DEB. The link B to D acts as a shield in that B and D are correlated even without conditioning on the common effect.

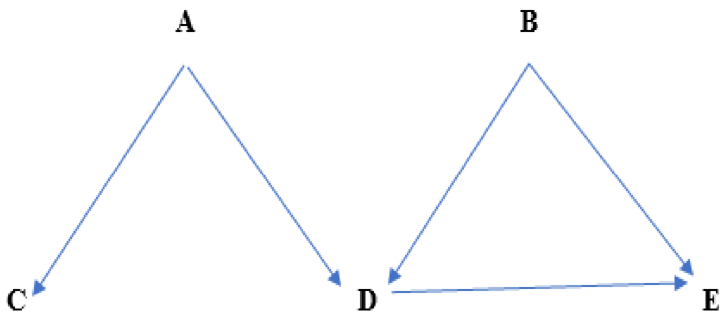


Figure 3.5: Directed Graph

3.3 Steps involved in PC and Modified PC Algorithms of Graph Theoretic Approach

PC and Modified PC causal algorithms have five common steps for calculation of causal ordering. The major difference among them is the use of residuals series. In original PC causal algorithms residuals series of VAR are to be treated as variables, while in Modified PC one step is changed to new version that residuals extract from the univariate methods i.e. Haugh (1976) and Rehman and Malik (2014) instead of VAR model are used to determine the causal ordering. The steps involved in both approaches PC and Modified PC are given below:

Step 1.

Construct the general structure of graph in which all variables are connected through undirected links.

Step 2.

Test unconditional correlation between any two variables. If they are not unconditionally correlated then eliminate that connections i.e. if there are three variables X,Y and Z. we have checked unconditional correlation between $(X, Y)(X, Z)$ and (Y, Z).

Step 3.

Tests correlation between each two variables conditional on a third variable. If each pair of variables are conditionally uncorrelated then again eliminate their connections. If there are three series X, Y and Z the conditional correlation between $(X, Y)(X, Z)$ and (Y, Z) can be shown as:

Conditional correlation between $(X, Y) = Cor(X_{,}Y/Z)$

Conditional correlation between $(X, Z) = Cor(X, Z/Y)$

Conditional correlation between $(Y, Z) = Cor(Y, Z/X)$

If we have more than three variables, suppose four variables X, Y, Z and W then conditional correlation between each pair of variables will be determined in the following way. The conditional correlation between (X, Y) can be determined as: "Conditional correlation between $(X, Y) = Cor(X_{,}Y/Z, W)$ and so on for others. If correlation between X and Y conditional on pair of variables Z, W vanish then eliminate that connections again. Continuing in the same way test correlation conditional on set of three variables and remove links whenever there is no conditional correlation.

The calculated values of conditional correlations will be compared with fisher's z test value. Fisher z test can be applied to test that whether the conditional correlations are significantly different from zero.

$$z(\rho(i,j|k)) = \left[\frac{1}{2}\sqrt{n-|k|-3}\right] ln\left\{\frac{1+\rho(i,j|k|)}{1-\rho(i,j|k|)}\right\}$$

Where n is number of observations, (ρ(i,j|k) is the population correlation between i and j conditional on k. |k| is number of variables that we condition on.

Step 4.

This step is called orientation stage. In the previous step 3 if the pair is correlated conditional on the third variable, the third variable is said to be unshielded collider on that path, and arrows from the pair of variables are oriented toward the third variable.

Step 5.

In this step, arrows are oriented on the basis of screening relationship. If two variables X and Y are not directly linked but are linked through a third variable Z as $X\rightarrow Z$--- Y so that one link points to the third variable say $X\rightarrow Z$ and the other link is undirected Z--- Y. So, orient the second link as $X\rightarrow Z\rightarrow Y$ because orienting the arrow toward Z shows that Z is unshielded collider and if it is true then this should be revealed in step 4. Thus, the intervening variable is a screen and not an unshielded collider, so the arrow cannot point toward it.

3.4 Simulation Design

Objective of the study is to evaluate the performance of PC and Modified PC casual search algorithms to causality for time series data by investigating size and power properties. To achieve these objectives, the study mainly focuses on Monte

Carlo Simulations. We shall analyze the performance of PC and Modified PC causal search algorithms on the basis of Monte Carlo simulation and optimal procedure will be selected. Finally, through that optimal procedure causal determinants of inflation will be explored by taking real data.

The main steps involve in methodology are:

I Data generating process (DGP)

II Methods for Testing Causality

III Size and Power

V Testing and Simulation

The steps involved in methodology are explained and summarized in the following flow chart.

3.5 Procedure for comparing PC and Modified PC Causality approaches

Data Generating Process (DGP)

↓

Three Time series (A_t, B_t, C_t)

↓

Residuals (X, Y, Z)

↓

Graph-Theoretic Causal Search

↓

PC Algorithms

↓

(Size and Power Properties) and (Probability of spurious causation in case of Confounder)

Figure 3.6: Methodology of PC Causality approaches

The above flow chart explains three different methodologies which differ only in residuals. The methodologies are VAR residuals, Modified R recursive residuals and Haugh ARMA residuals.

First, the data is generated from using the data generating process (DGP) and VAR residuals, Modified R residuals and Haugh residuals are extracted by applying VAR model, Recursive AR model and ARMA model respectively. These extracted residuals are then referred to PC and Modified PC algorithms and performance of these procedures are evaluated using size and Power Properties. The performance of these procedures is also evaluated when there is confounding variable in the data generating process (DGP).

3.6 Data Generating Process

The objective of simulation experiment is to find out size and power properties of methodologies for testing causality. Therefore, we need data series with embedded causality (to evaluate power properties) and the data series with no causality (to evaluate size properties). Selection of DGP for Monte Carlo simulation study is very important mostly in comparative analysis. The tests or approaches can be compared in same framework to recommend the superiority of one test or weakness of another test. The data for testing properties of causality tests can be generated from a unified framework which is given as below:

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_1 & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_2 & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_3 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \dots (10)$$

$$\text{Where, } \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho 1 & \rho 2 \\ \rho 1 & 1 & \rho \\ \rho 2 & \rho & 1 \end{bmatrix} \right)$$

The above matrix form equation can be written in the following form:

$$X_t = AX_{t-1} + BDt + \varepsilon_t \qquad \varepsilon_t \sim N(0, \Omega)$$

Where $A = \begin{bmatrix} \theta_1 & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_2 & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_3 \end{bmatrix}$, $B = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}$, $Dt = \begin{bmatrix} 1 \\ t \end{bmatrix}$, $\varepsilon_t = \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$, $\Omega = \begin{bmatrix} 1 & \rho 1 & \rho 2 \\ \rho 1 & 1 & \rho \\ \rho 2 & \rho & 1 \end{bmatrix}$

This equation is general DGP which can take various forms by specifying the parameters A, B and Ω. A can be used to specify conventional Granger type Causality and Omega can be used to specify contemporaneous causality.

As per definition of Granger causality, y is caused by x if lag value of x can be used for predicting y. In DGP (10) suppose $A_{1i} = (\alpha, 0,0)$ and α € (0, 1) then y_{t-1} and z_{t-1} does not appear in the equation of x_t. Therefore y_t and z_t does not Granger cause x_t. On the other hand, if second and third column of the first row are non-zero $\theta_{12} \neq 0$ and/or $\theta_{13} \neq 0$, this means that y_t and z_t Granger cause x_t. Similarly If $A_{2i} = (0, \alpha, 0)$ then x_{t-1} and z_{t-1} does not appear in the equation of y_t. Therefore x_t and z_t does not Granger cause y_t. On the other hand, if second and third column of the second row are non-zero $\theta_{21} = \theta_{23} \neq 0$, this means that x and z Granger cause y. The same causal direction can be examined if we have a case that $A_{1i} = (0,0, \alpha)$.

In addition, if A is null matrix, the three series will be white noise with no auto correlation. θ_1 shows us the autoregressive coefficient of the first series, θ_2 indicate autoregressive coefficient of the second series and θ_3 indicate autoregressive coefficient of the third series. If θ_1= 0, it means that the first series generated is white noise. If $0 < \theta_1 < 1$ then the series generated is stationary and auto-correlated, while if $\theta_1 \geq 1$ then the generated series will become non-stationary.

The data generating process will generate three independent autoregressive series having drift and trend if A= $\begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}$, B= $\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}$ and ρ = 0. Second, the data generating process will generate three independent autoregressive series having drift only, if A= $\begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}$, B= $\begin{bmatrix} a_1 & 0 \\ b_1 & 0 \\ c_1 & 0 \end{bmatrix}$ and ρ = 0. Third, the data generating process will generate three independent autoregressive series without drift and trend, if A= $\begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}$, B= $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ and ρ = 0.

In this study we assume the window size 0.2 and our alternative space become (0.2, 0.4, 0.6, 0.8). All of these comparisons will be done by using different values of matrix A, B and Ω. Next we will generate different but correlated series x_t, y_t, z_t without drift and Trend, with drift only and with drift and trend to check the causal ordering through power of test. These cases can be generated by imposing different restrictions on data generating process. First, the data generating process will generate three dependent series having drift and trend if no restriction is imposed on DGP. Second, the data generating process will generate three dependent series having drift if B= $\begin{bmatrix} a_1 & 0 \\ b_1 & 0 \\ c_1 & 0 \end{bmatrix}$. Finally, the DGP will generate three dependent series having no drift and trend if B= $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

The parameter B is called “nuisance”. The causality does not depend on the matrix of parameter B, however the test statistics for coefficient present in “A” which determine causality is heavily dependent on B and incorrect specification of B may create bias. So, to avoid the biasness we have to include this nuisance term.

CHAPTER 4

SIZE EVALUATION OF PC ALGORITHM

The main focus of the study is to evaluate the performance of causal search PC and Modified PC algorithms of GTA where the following are used as measure of correlation: (i) VAR residuals, (ii) Modified R recursive residuals (iii) Haugh residuals. Size analysis of PC and Modified PC algorithm with non-stationary and stationary series with the use of different specifications are carried out using the data generating process given in equation (5). The following specification will be used: no drift and no trend, drift only and Drift plus trend. These specifications are already been used in (Dickey, 1979) and (Atiq, 2009). This present study will also cover these specifications discussed in section 4.3 and 4.4.

4.1 Size Distortion as Measure of Performance

It is well known that powers of econometric tests/procedures are comparable if the size remain same, and so is the case with the three approaches mentioned above. But size cannot be controlled in the PC algorithm of Graph Theoretic Approach because the causality testing through PC algorithms involve multiple decisions in a chain and standardization of size is not possible when many decisions are involved. Usually, when tests are to be compared, the process starts by finding out the critical values with fixed size, say 5%. These critical values are then used to calculate power curves. However, PC algorithms involves multiple testing, therefore 5% critical values for the entire procedures cannot be calculated. Alternatively, we can measure size distortion where the size of entire procedure can be calculated fixing the size each single step at 5%. The test with minimum size distortion would be the optimal test.

The best performance would be considered as of the procedure having minimum size distortion and highest power.

Let alpha be the size of a test/procedure then

$$\alpha = P(Reject\ H_0/H_0\ is\ True)$$

In our case, the null hypothesis H_0: there is no causality between "x and y" and for calculation of size, the data should be generated such that H_o is true. The alternative hypothesis in our case is "H_1: x causes y". There are multiple decisions involved in causality testing and size of each step involved is assumed 5%. We also assume that the size of entire process will be 5%. At the end, the difference between observed size and the nominal size (5%) can be referred as size distortion.

In the analysis of causality through PC causal search for time series data, calculation of size depends on two nuisance parameters; deterministic part which includes drift and trend and the autoregressive coefficient of the underlying time series. The causality is the feedback from one variable to another and usually there is no interest of researchers in the autoregressive coefficient of the underlying variables. However, the size and power of the causality tests actually depend heavily on the value of autoregressive coefficients of underlying series, therefore these nuisance parameters should be carried throughout the analysis. The size of PC algorithms is calculated for wide range of specifications of these nuisance parameters so that the effect of nuisance parameter on the size can also be analyzed.

Size for PC and Modified PC algorithms are calculated from different stationary and non-stationary series with wide range of specification of deterministic part and autoregressive coefficients. The time series length was kept constant at 100 observation for the entire analysis.

4.2 Data Generating Process

For this analysis, the independent autoregressive stationary and nonstationary time series are being generated with different specifications as discussed in Chapter 3, we have used following data generating process:

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \ldots (10)$$

$$X_t = AX_{t-1} + BD_t + \varepsilon_t$$

$$\text{where, } \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho 1 & \rho 2 \\ \rho 1 & 1 & \rho \\ \rho 2 & \rho & 1 \end{bmatrix} \right)$$

In Chapter 3, it was discussed that the DGP (10) is capable of generating series with various specifications of drift, trend and the autoregressive coefficients. Causality and no causality can also be specified in this DGP. For details see section 3.6. All the estimated results in below tables came after 10000 simulations. The specific details of size in different specifications is given in the section 4.3 and 4.4.

4.3 Size Analysis with Non-Stationary Series

4.3.1 Size Comparison of PC and Modified PC Algorithms using Non-Stationary Series without Drift and Trend:

At first, three independent autoregressive non-stationary series x, y and z are generated through data generating process (DGP) given in equation (10). In DGP (10), we choose $(\theta_{11}, \theta_{22}, \theta_{33})$ = (1,1,1), $(\theta_{12,}\ \theta_{13,}\ \theta_{21,}\ \theta_{23,}\ \theta_{31,}\ \theta_{32})$ = 0 in matrix A, and matrix B = 0 in equation (5). Setting B = 0 means series generated are nonstationary without drift and trend. In DGP, we take A as diagonal matrix which means, we have not put any causal relationship among the three generated series. Using VAR residuals, Modified R recursive residuals and Haugh ARMA residuals, we got various magnitude of probability of spurious causal relationship for all possible directions $x \to y,\ y \to z$, and $x \to z$.

Table 4.1: Probability of Rejection of the hypothesis of no causality using PC and Modified PC algorithms with non-stationary having No Drift and Trend

	PC	Modified PC	Modified PC
	VAR residuals	**MR residuals**	**Haugh residuals**
$x \to y/no\ causality$	0.063	0.481	0.138
$y \to z/no\ causality$	0.079	0.446	0.120
$x \to z/no\ causality$	0.074	0.446	0.120

Table 4.1 summarizes probability of rejection of (true) null hypothesis of no causality using PC and Modified PC algorithms where the measures of correlation are based on three different kind of residuals mentioned in section 4.1 above. The results obtained from PC causal search using VAR residuals indicate about 7% on average significant results against 5% nominal size as shown in column 1 of Table 4.1. This implies there is on the average 2% size distortion which can be regarded as spurious causality because the true DGP doesn't have causality. Column 2 indicates the results of Modified PC causal search algorithm, treating Modified R recursive residuals as original variables showing on average 48% significant result, which means a size distortion of 43%. Using Modified R recursive residuals, probability of incorrect decisions for all possible directions is about 40% which is much higher than the probability of spurious causality obtained from VAR residuals. When Haugh residuals are referred to causal search algorithm, we get significant results for about 12% for all possible directions against 5% nominal size. Thus, the probability of spurious causal relationship using Haugh residuals is on average 7% as shown in column 3.

Hence, Modified R recursive residuals generate a significantly high size distortion than the VAR and Haugh residuals. There is nominal size distortion using

VAR and moderate level of size distortion with Haugh-ARMA residuals. Recursive residuals obtained from Modified R show us spurious causality with high probability when the generated series are non-stationary having no drift and trend as evident from the above Table 4.1.

4.3.2 Size Comparison of PC and Modified PC Algorithms using Non-stationary Series with Drift only:

Using the same data generating process (DGP), three autoregressive non-stationary series were generated. The series were generated with respect to change in autoregressive parameters of DGP (10), while keeping the drift component of deterministic part only. We choose $\theta_{ij} = \begin{cases} 1 & i = j \\ 0 & otherwise \end{cases}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$. Setting second column of matrix B = 0 indicate that no trend, and the non-zero column of matrix B indicates that the three series contain drift which means that series generated are nonstationary with drift and without trend. We also take A matrix as diagonal which means that none of variable enter in the equation of other variables and all variables are independent of each other's. The series x, y and z are independent so, if the results are significant would be indicative of spurious causality. This can be explained in more detail from simulated results given in Table 4.2,

Table 4.2: Probability of Rejection of the hypothesis of no causality using PC and Modified PC algorithms with non-stationary series having only Drift

	PC	**Modified PC**	**Modified PC**
	VAR residuals	**MR residuals**	**Haugh residuals**
$x \to y/no\ causality$	0.071	0.64	0.120
$y \to z/no\ causality$	0.073	0.64	0.104
$x \to z/no\ causality$	0.075	0.64	0.104

We have stated that x, y and z are independent of each other. However, the causality testing PC algorithm using VAR residuals resulted in 7.1%, 7.3% and 7.5%

significant results against 5% nominal size. This implies there is on the average 2.3% probability of spurious causality for all three possible direction i.e. $x \rightarrow y,\ y \rightarrow z$, and $x \rightarrow z$ are found as shown in column 1 of Table 4.2. Column 2 show the results of Modified PC algorithm when modified R recursive residuals are used. we get significant results for about 64% for all possible directions. For Modified PC causal search, the probability of incorrect decisions using modified R recursive residuals is 59% on average for three different causal directions, which is very high than the probability of spurious causality obtained from PC algorithm using VAR residuals. Whereas, the causal search algorithms (Modified PC) using Haugh's ARMA residuals yield on average 6% spurious regression for all three cases at 5% nominal size as displayed in column 3 of Table 4.2.

This shows that Modified PC algorithm using modified R recursive residuals generate a significantly high size distortion as compared to other two methods. There is nominal size distortion in PC causal search algorithm treating VAR residuals while moderate size distortion in Modified PC causal search algorithm using Haugh ARMA residuals. Modified PC using modified R residuals show us spurious relationship with high probability between $x \rightarrow y,\ y \rightarrow z$, and $x \rightarrow z$, but originally all series x, y and z are independent of each other.

4.3.3 Size Comparison of PC Algorithms using Non-stationary Series with Drift and Trend:

Three independent non-stationary series having drift and trend are generated. The coefficients in matrix A, We choose $\theta_{ij} = \begin{cases} 1 & i = j \\ 0 & otherwise \end{cases}$ and $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$, it means series generated are nonstationary having drift and with trend. We take A as diagonal matrix which means that from the data generating process all generated series

x, y and z are independent of each other. so, if the results are significant would be sign of spurious causality.

Table 4.3: Probability of Rejection of the hypothesis of no causality in PC algorithms with non-stationary series having both Trend and Drift

	PC	Modified PC	Modified PC
	VAR residuals	**MR residuals**	**Haugh residuals**
$x \rightarrow y/no\ causality$	0.072	0.632	0.518
$y \rightarrow z/no\ causality$	0.071	0.649	0.52
$x \rightarrow z/no\ causality$	0.064	0.649	0.52

The simulated results of PC and Modified PC causal search algorithms using different measure of correlation is given in Table 4.3. Like previous two cases, we do not impute causality in DGP. Table 4.3 displayed the same results as previously discussed Table 4.1 and 4.2.

From the above simulation results it is clear that when the data series are non-stationary that Modified PC algorithms (using modified R recursive residuals and Haugh's ARMA residuals) generate a significantly high size distortion. There is minimal size distortion using VAR residuals in PC algorithm.

4.4 Size Analysis with Stationary Series

4.4.1 Size Comparison of PC Algorithms using Stationary Series without Drift and Trend:

The three independent stationary series have been generated by using data generating process given in equation (5). In case of stationary series, we choose $\theta_{ij} = \begin{cases} \rho & i = j \\ 0 & otherwise \end{cases}$ $where\ \rho < 1$, we set matrix B = 0, which means stationary series are generated without drift and trend. The generated series x, y and z are independent of each other so, if the results are significant, this would be the indicative of spurious causality.

Table 4.4: Probability of Rejection of the hypothesis of no causality using PC and Modified PC algorithms in case of stationary series with No Drift and Trend

	ρ	PC	Modified PC	Modified PC
		VAR residuals	MR residuals	Haugh residuals
$x \to y/no\ causality$	0.8	0.063	0.358	0.122
$x \to y/no\ causality$	0.6	0.065	0.202	0.11
$x \to y/no\ causality$	0.4	0.05	0.123	0.09
$x \to y/no\ causality$	0.2	0.055	0.09	0.07
$y \to z/no\ causality$	0.8	0.065	0.333	0.136
$y \to z/no\ causality$	0.6	0.063	0.194	0.05
$y \to z/no\ causality$	0.4	0.084	0.116	0.08
$y \to z/no\ causality$	0.2	0.064	0.078	0.08
$x \to z/no\ causality$	0.8	0.074	0.333	0.136
$x \to z/no\ causality$	0.6	0.073	0.194	0.05
$x \to z/no\ causality$	0.4	0.043	0.116	0.08
$x \to z/no\ causality$	0.2	0.078	0.078	0.08

Using different residuals from stationary series in causality testing algorithm (PC algorithm) of graph theoretic approach developed by Pearl and Spirtes the results for all three possible directions are given in Table 4.4. 1st panel of Table 4.4 indicates the results $x \to y$ of PC and Modified PC algorithms, when the value of autoregressive parameters ($\theta_{11}, = \theta_{22} = \ \theta_{33}$) are 0.8, 0.6, 0.4, and 0.2. The panel 2 and 3 summarize the simulation results with the different null hypothesis. The null hypothesis for panel 1 is $x \to y$ whereas for panel 2 and 3 it is $y \to z$ causes and $x \to z$ respectively. The results show that PC algorithm using VAR residuals, the probability of significant results fluctuates around 7% for all three possible causal directions (taken as null hypothesis) at nominal size 5%. It means that there is on average 2% probability of spurious causality in the three possible directions. In the third column, the results of Modified PC algorithms with Haugh's ARMA residuals are summarized, which indicate that the probability of significant results on average is about 8% for all three possible directions. This shows that the average probability of spurious causality is about 3% for all three possible directions. Finally, when we referred Modified R recursive residuals to Modified PC algorithms. The results indicate that when the series is stationary with root close to unity, probability of significant results remains high on average is about 27% for all three possible directions at nominal size 5%. But when the autoregressive parameters are close to zero, the probability of significant results decrease to about 7.8% for all three possible directions. This show on the average 2.8% probability of spurious regression for all three possible direction i.e. $x \rightarrow y,\ y \rightarrow z$, and $x \rightarrow z$. The results also indicate that in case of without drift and trend VAR residuals are independent of autoregressive parameter while Haugh and Modified R procedures the results depend on autoregressive coefficient as the autoregressive coefficient are close to zero.

4.4.2 Size Comparison of Stationary Series with Drift and without Trend:

Autoregressive stationary series with drift and without trend are generated from the data generating process by restricting parameters of matrix A. we choose $\theta_{ij} = \begin{cases} \rho & i = j \\ 0 & otherwise \end{cases}$ $where\ \rho < 1$, i.e. 0.8, 0.6, 0.4, 0.2 and all cross terms are set zero. we set the trend part of matrix B to zero and series are generated. The generated stationary series x, y and z are independent of each other. Using residuals series obtained from different procedures, we got various magnitude of probability of spurious causal relationship given in Table 4.5 as under.

Table 4.5: Probability of Rejection of the hypothesis of no causality using PC and Modified PCs Algorithms in case of stationary series with Drift only

	$\boldsymbol{\rho}$	**PC**	**Modified PC**	**Modified PC**
		VAR residuals	**MR residuals**	**Haugh residuals**
$x \rightarrow y/no\ causality$	0.8	0.063	0.335	0.06
$x \rightarrow y/no\ causality$	0.6	0.077	0.185	0.075
$x \rightarrow y/no\ causality$	0.4	0.072	0.108	0.053
$x \rightarrow y/no\ causality$	0.2	0.066	0.092	0.04
$y \rightarrow z/no\ causality$	0.8	0.069	0.335	0.069
$y \rightarrow z/no\ causality$	0.6	0.07	0.185	0.068
$y \rightarrow z/no\ causality$	0.4	0.062	0.108	0.07
$y \rightarrow z/no\ causality$	0.2	0.056	0.092	0.06
$x \rightarrow z/no\ causality$	0.8	0.066	0.364	0.069
$x \rightarrow z/no\ causality$	0.6	0.054	0.182	0.06
$x \rightarrow z/no\ causality$	0.4	0.054	0.109	0.07
$x \rightarrow z/no\ causality$	0.2	0.077	0.095	0.06

The 1st panel of Table 4.5 indicating the probability of $x \rightarrow y$ for PC using VAR residuals and Modified PC using modified R recursive residuals and Haugh ARMA residuals, when the value of autoregressive parameters ($\theta_{11} = \theta_{22} = \theta_{33}$) are 0.8, 0.6, 0.4, and 0.2. Referring VAR residuals in PC algorithms, the results for all three possible causal directions are given. In panel 2nd and panel 3rd simulated results of $y\ causing\ z$ ($y \rightarrow z$), and $x\ causing\ z$ ($x \rightarrow z$) are given respectively**.** The results of PC algorithm using VAR residuals show that the probability of significant results remains on average about 7.5% for all three possible directions. It means that there is on average 2.5% probability of spurious causality for all three possible direction at nominal size of 5%. Modified PC algorithms using modified R recursive residuals results are given in column 2 of Table 4.5. The results reveal that when the series is stationary with root close to unity i.e. 0.8 and 0.6, the probability of significant results on average is about 26% for all three possible causal directions. But when the autoregressive parameters are close to zero i.e. 0.2, then the probability of significant results on average is about 8.8% for all three possible directions. Using modified R recursive residuals (Modified R), show on the average 3.8% probability of spurious causality for all three possible direction i.e. $x \rightarrow y,\ y \rightarrow z,$ and $x \rightarrow z$.

Finally, the results of Modified PC Algorithms treating Haugh ARMA residuals as original variables are shown in column 3 of the above table. The probability of significant results on average is about 7% for all three possible directions. This show on the average 2% probability of spurious causality for all three possible directions.

4.4.3 Size Comparison of trend Stationary Series with drift and trend

The three independent auto regressive trend stationary series are generated using data generating process given in equation (5). We choose $\theta_{ij} = \begin{cases} \rho & i = j \\ 0 & otherwise \end{cases} where\ \rho < 1,$ We set matrix $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$, which means stationary series are generated with drift and trend.

In the DGP we have not put any causal relationship among the series. The generated stationary series $x, y\ and\ z$ are independent of each other's so any significant result will be the indicative of spurious causal relationship as discussed before. The simulated results of probability of spurious causal relationship among $x, y\ and\ z$ are given in Table 4.6 as under.

Table 4.6: Probability of Rejection the hypothesis of no causality using PC and Modified PCs Algorithms in case of trend stationary series having both Drift and Trend.

	ρ	PC	Modified PC	Modified PC
		VAR residuals	**MR residuals**	**Haugh residuals**
$x \to y/no\ causality$	0.8	0.081	0.42	0.181
$x \to y/no\ causality$	0.6	0.069	0.217	0.138
$x \to y/no\ causality$	0.4	0.082	0.114	0.097
$x \to y/no\ causality$	0.2	0.062	0.079	0.08
$y \to z/no\ causality$	0.8	0.056	0.409	0.201
$y \to z/no\ causality$	0.6	0.078	0.208	0.137
$y \to z/no\ causality$	0.4	0.069	0.137	0.098
$y \to z/no\ causality$	0.2	0.05	0.078	0.066
$x \to z/no\ causality$	0.8	0.075	0.409	0.201
$x \to z/no\ causality$	0.6	0.067	0.208	0.137
$x \to z/no\ causality$	0.4	0.081	0.137	0.098
$x \to z/no\ causality$	0.2	0.07	0.07	0.066

Table 4.6 reveal the simulated results of PC and Modified PC algorithms using residuals obtained from three different procedures. Panel 1st, Panel 2nd and Panel 3rd of Table 4.6 indicating the results $x\ is\ causing\ y\ (x \to y),\ y\ is\ causing\ z\ (y \to z)$, and $x\ is\ causing\ z\ (x \to z)$ respectively. The above results show us that when we referred VAR residuals to PC algorithms, the probability of significant results remains on average about 7.5% for all three possible causal directions at nominal size 5%. It means that there is on average 2.5% probability of spurious causal relationship for all three possible direction i.e. $x \to y,\ y \to z$, and $x \to z$. In column 3rd the results of Modified PC algorithm using Haugh ARMA residuals are given. The simulated results indicate that the probability of significant results is on average 12% for all three possible directions at the same nominal size. This show on the average 7% probability of spurious regression for all three possible directions. Finally, when modified R

recursive residuals are treated as original variables in Modified PC algorithm, the results tell us that when the series is stationary with root close to unity i.e. 0.8 and 0.6 the probability of significant results on average is about 30% for all three possible directions. But when the autoregressive parameters are close to zero i.e. 0.2 then the probability of significant results on average is about 7% for all three possible directions. This show on the average 2% probability of spurious regression for all three possible directions. The results also show that PC algorithm using VAR residuals is independent of autoregressive parameter while Modified PC algorithm using Haugh ARMA residuals show pattern in all three panels. The Modified PC algorithm using Modified R residuals display the same results as displayed in case of no deterministic part.

4.5 Conclusion

In this chapter for evaluation of size properties we have discussed large number of the possible cases with different specifications. For size evaluation various independent non-stationary and stationary series with different complications has been generated from the data generating process (DGP) given in equation 10.

The results indicate that VAR residuals continuously maintains the size. The stationary series, non-stationarity series, autoregressive coefficient, the specification of drift and trend don't affect the size of PC algorithm using VAR residuals.

Keeping in view the results of Modified PC algorithm using Haugh residuals has higher size distortion than using VAR residuals but lesser size distortion than Modified R recursive residuals.

The results of Modified PC algorithm using modified R recursive residuals indicates that the size depends on value of auto regressive coefficient and its distortion reduces when auto regressive coefficients approaches zero. The results indicate that when the series is highly stationary (low memory), then causal algorithm using modified R recursive residuals comparatively perform better than VAR and Haugh residuals in size distortion problem.

CHAPTER 5

PERFORMANCE OF PC ALGORITHMS IN THE PRESENCE OF CONFOUNDING VARIABLE

The definition of confounder has long-standing considered as a result of any third variable that is correlated with the exposure of interest and the outcome of interest. The confounding variable does not exist in the causal pathway between the exposure and outcome, but it is the common cause of both variables which make the investigated causal path complex and bias. David et al (2019) comprehend it in more detail we have directed acyclic graphs in Figure 5.1 and 5.2 showing - confounding bias and collider bias respectively. In Figure 5.1 the red arrows denote in an open back-door path: ***exercise* ← *smoking* → *lung cancer***. “Smoking” is a confounder that naturally leaves the back-door path open. Controlling for “smoking” will close the back-door path, eliminating confounding through this path. In part Figure 5.2 the green arrows represent a closed back-door path: ***shift work* → *sleepiness* ← *obstructive sleep apnea***. “Sleepiness” is a collider that naturally leaves the back-door path closed. Control of “sleepiness” would open the back-door path, introducing confounding through this path[10].

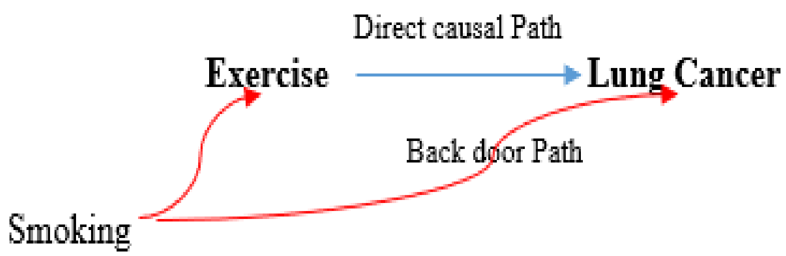


Figure 5.1: Confounding Bias

[10] David et al (2019) Control of confounding and report of results in causal inferences studies. Annals ATS Vol. 16 Number 1|January 2019.

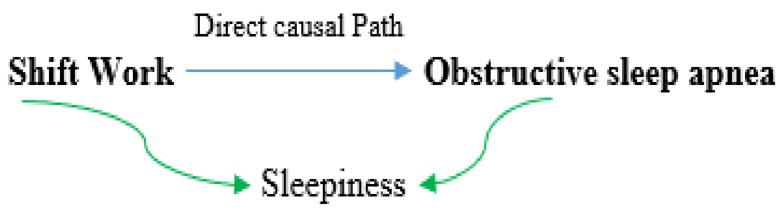


Figure 5.2: Collider Bias

Keeping the above discussion regarding confounding and collider bias, this chapter intends to evaluate the performance of PC and Modified PC algorithms of graph theoretic when we have confounder in the data generating process. From the DGP we have generated three series x, y and z. The two series y and z both are separately caused by the third variable x. To find the probability of spurious causation of these causal search algorithms due to confounding variable Monte Carlo Simulation is conceived as under:

5.1 Monte Carlo Simulation Design

To find the probability of spurious causation of PC and Modified PC causal search due to confounding variable x, three different series x, y and z generated from the given DGP in equation (10) in such a way that variable y and z are independent of each other's but dependent on x which means that variable x occurs in both equations of y and z. This indicates that in the DGP actual path goes from x to y and from x to z. The series y and z are independent so, if the results are significant would be indicative of spurious causation. The DGP is the combination of two nuisance parameters "deterministic" and "autoregressive" part. The deterministic part tells us about the drift and trend while autoregressive part generate correlation between variables when the cross term is non-zero. In this section, we have calculated probability of spurious causation treating VAR residuals, Modified R recursive residuals and Haugh test residuals as original variables in PC and Modified PC algorithms. All possible DGPs for simulation experiment are designed (changing coefficient values of matrix A and B of DGP (10)). The estimated results have been summarized after 10000 times simulations from the following data generating process.

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_1 & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_2 & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_3 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \dots (10)$$

$$X_t = AX_{t-1} + BDt + \varepsilon_t \qquad \varepsilon_t \sim N(0, \Omega)$$

$$\text{where, } \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho 1 & \rho 2 \\ \rho 1 & 1 & \rho \\ \rho 2 & \rho & 1 \end{bmatrix} \right)$$

5.2 Probability of Spurious Causation

With the help of the given data generating process dependency in series can established. If the cross-term coefficients, θ_{21} and θ_{31} in matrix A = $\begin{bmatrix} \theta_{11} & 0 & 0 \\ \theta_{21} & \theta_{22} & 0 \\ \theta_{31} & 0 & \theta_{33} \end{bmatrix}$ is non zero, this means that the lag of variable x is present in equation of y and z, which implies that there is correlation between (x and y) and (x and z). All models are selected in such a way that there is a causality from x to y and from x to z but there is no causality between z and y. After generating the series x, y and z from DGP we used these variables in VAR, AR and ARMA model and residuals extracted from these three models are stored. These residuals are then used in PC causal search algorithms to find the probability of spurious causation due to confounding variable x.

We have conceived three different cases i.e. (drift and trend, drift only and without drift and trend) in DGP and performance of both PC and Modified PC algorithms in the Presence of Confounding Variable are evaluated. These can be shown in the as under:

Stationary Series without Deterministic Part

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_{11} & 0 & 0 \\ \theta_{21} & \theta_{22} & 0 \\ \theta_{31} & 0 & \theta_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} \ldots (11)$$

Stationary Series with Drift only

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_{11} & 0 & 0 \\ \theta_{21} & \theta_{22} & 0 \\ \theta_{31} & 0 & \theta_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} \ldots (12)$$

Trend Stationary Series with Deterministic Part

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_{11} & 0 & 0 \\ \theta_{21} & \theta_{22} & 0 \\ \theta_{31} & 0 & \theta_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} \ldots (13)$$

From the above DGPs (11), (12) and (13) stationary series are generated with respect to change in autoregressive parameters with different specifications (drift and trend). To make the series stationary we put the diagonal entries θ_{11} , θ_{22} and θ_{33} in equation to be smaller than 1. To create cross dependences, we choose some of the non-diagonal entries to be non-zero. We choose $\theta_{21} > 0$ and $\theta_{31} > 0$. The cross terms create association between variables and its value also changes from 0.9, 0.8, 0.6, 0.4 and 0.2 in matrix A=$\begin{bmatrix} \theta_{11} & 0 & 0 \\ \theta_{21} & \theta_{22} & 0 \\ \theta_{31} & 0 & \theta_{33} \end{bmatrix} \neq 0$. It means that in all three equations variable y and z are independent of each other's but dependent on x which means that variable x occurs in both equations of y and z. This setup indicates that in the DGP x is causing y and z i.e. ($x \rightarrow y$) and ($x \rightarrow z$). The series y and z are independent so, if the results are significant would be indicative of spurious causation. On the other hand, if the results for (*x cause y)* and (*x cause z)* are significant this would reveal true causation. This can be explained in more detail from simulated results given in Table 5.1, 5.2 and 5.3.

Table 5.1: Size and Power of PC and Modified PC algorithms having Confounding variable in DGP using Stationary series without Drift and Trend

		PC		Modified PC		Modified PC	
		VAR residuals		MR residuals		Haugh residuals	
	$\boldsymbol{\rho}$	**Correct**	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
				$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 0.8}$			
$x \to y/causality$	0.9	0.06	0.94	0.47	0.52	0.06	0.94
$x \to y/causality$	0.8	0.08	0.92	0.51	0.48	0.07	0.93
$x \to y/causality$	0.4	0.06	0.94	0.54	0.45	0.07	0.93
$x \to y/causality$	0.2	0.07	0.93	0.60	0.39	0.07	0.93
$x \to z/causality$	0.9	0.06	0.94	0.46	0.53	0.09	0.91
$x \to z/causality$	0.8	0.07	0.93	0.50	0.49	0.07	0.93
$x \to z/causality$	0.4	0.07	0.93	0.55	0.44	0.09	0.91
$x \to z/causality$	0.2	0.07	0.93	0.59	0.40	0.09	0.91
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00
				$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 0.6}$			
$x \to y/causality$	0.9	0.07	0.92	0.49	0.50	0.09	0.91
$x \to y/causality$	0.8	0.06	0.93	0.47	0.52	0.09	0.91
$x \to y/causality$	0.4	0.06	0.93	0.47	0.52	0.08	0.92
$x \to y/causality$	0.2	0.06	0.93	0.35	0.64	0.08	0.92
$x \to z/causality$	0.9	0.07	0.92	0.49	0.51	0.07	0.93
$x \to z/causality$	0.8	0.05	0.94	0.52	0.48	0.06	0.94
$x \to z/causality$	0.4	0.06	0.93	0.47	0.52	0.09	0.91
$x \to z/causality$	0.2	0.06	0.93	0.31	0.68	0.09	0.91
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00
				$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 0.4}$			
$x \to y/causality$	0.9	0.05	0.94	0.30	0.69	0.05	0.95
$x \to y/causality$	0.8	0.0	0.94	0.30	0.69	0.06	0.94
$x \to y/causality$	0.4	0.01	0.92	0.25	0.74	0.08	0.92
$x \to y/causality$	0.2	0.06	0.93	0.18	0.81	0.07	0.93
$x \to z/causality$	0.9	0.05	0.94	0.31	0.68	0.08	0.92
$x \to z/causality$	0.8	0.06	0.93	0.34	0.65	0.08	0.92
$x \to z/causality$	0.4	0.05	0.94	0.25	0.74	0.05	0.95
$x \to z/causality$	0.2	0.07	0.92	0.17	0.82	0.06	0.94
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00

Table 5.2: Size and Power of PC and Modified PC algorithms having Confounding variable in DGP using Stationary series with Drift only

		PC		**Modified PC**		**Modified PC**	
		VAR residuals		MR residuals		Haugh residuals	
	$\boldsymbol{\rho}$	**Correct**	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
		$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 0.8}$					
$x \to y/causality$	0.9	0.06	0.93	0.41	0.58	0.06	0.94
$x \to y/causality$	0.8	0.08	0.91	0.42	0.57	0.05	0.95
$x \to y/causality$	0.4	0.06	0.93	0.45	0.5	0.07	0.93
$x \to y/causality$	0.2	0.06	0.93	0.50	0.5	0.08	0.92
$x \to z/causality$	0.9	0.07	0.92	0.42	0.57	0.08	0.92
$x \to z/causality$	0.8	0.06	0.93	0.45	0.55	0.09	0.91
$x \to z/causality$	0.4	0.08	0.91	0.45	0.54	0.05	0.95
$x \to z/causality$	0.2	0.07	0.92	0.52	0.47	0.09	0.91
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00
		$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 0.6}$					
$x \to y/causality$	0.9	0.05	0.94	0.41	0.58	0.09	0.91
$x \to y/causality$	0.8	0.07	0.92	0.43	0.56	0.05	0.95
$x \to y/causality$	0.4	0.08	0.91	0.46	0.54	0.09	0.91
$x \to y/causality$	0.2	0.06	0.93	0.36	0.63	0.08	0.92
$x \to z/causality$	0.9	0.05	0.94	0.39	0.60	0.08	0.92
$x \to z/causality$	0.8	0.06	0.93	0.46	0.53	0.07	0.93
$x \to z/causality$	0.4	0.06	0.93	0.49	0.50	0.09	0.91
$x \to z/causality$	0.2	0.05	0.94	0.32	0.67	0.07	0.93
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00
		$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 0.4}$					
$x \to y/causality$	0.9	0.06	0.93	0.32	0.67	0.07	0.93
$x \to y/causality$	0.8	0.05	0.94	0.31	0.68	0.07	0.93
$x \to y/causality$	0.4	0.07	0.92	0.26	0.73	0.07	0.93
$x \to y/causality$	0.2	0.06	0.93	0.13	0.86	0.06	0.94
$x \to z/causality$	0.9	0.07	0.92	0.29	0.70	0.09	0.91
$x \to z/causality$	0.8	0.06	0.94	0.32	0.67	0.05	0.95
$x \to z/causality$	0.4	0.07	0.92	0.26	0.74	0.07	0.93
$x \to z/causality$	0.2	0.06	0.93	0.18	0.81	0.08	0.92
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00

Table 5.3: Size and Power of PC and Modified PC algorithms having Confounding variable in DGP using trend Stationary series with Drift and Trend

		PC		Modified PC		Modified PC	
		VAR residuals		MR residuals		Haugh residuals	
	ρ	Correct	Omitted	Correct	Omitted	Correct	Omitted
			$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$				
$x \to y/causality$	0.9	0.07	0.92	0.56	0.43	0.07	0.93
$x \to y/causality$	0.8	0.08	0.91	0.56	0.43	0.09	0.91
$x \to y/causality$	0.4	0.05	0.94	0.56	0.43	0.06	0.94
$x \to y/causality$	0.2	0.07	0.92	0.58	0.41	0.08	0.92
$x \to z/causality$	0.9	0.06	0.93	0.56	0.43	0.09	0.91
$x \to z/causality$	0.8	0.07	0.93	0.56	0.43	0.05	0.95
$x \to z/causality$	0.4	0.06	0.93	0.60	0.39	0.08	0.92
$x \to z/causality$	0.2	0.06	0.94	0.57	0.42	0.05	0.95
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00
			$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$				
$x \to y/causality$	0.9	0.06	0.94	0.36	0.63	0.05	0.95
$x \to y/causality$	0.8	0.05	0.94	0.41	0.58	0.05	0.95
$x \to y/causality$	0.4	0.06	0.93	0.49	0.50	0.08	0.92
$x \to y/causality$	0.2	0.05	0.94	0.35	0.64	0.06	0.94
$x \to z/causality$	0.9	0.06	0.93	0.36	0.63	0.06	0.94
$x \to z/causality$	0.8	0.06	0.93	0.38	0.62	0.07	0.93
$x \to z/causality$	0.4	0.07	0.93	0.47	0.52	0.06	0.94
$x \to z/causality$	0.2	0.06	0.93	0.37	0.62	0.09	0.91
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00
			$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$				
$x \to y/causality$	0.9	0.05	0.94	0.32	0.67	0.07	0.93
$x \to y/causality$	0.8	0.05	0.94	0.33	0.66	0.07	0.93
$x \to y/causality$	0.4	0.06	0.93	0.26	0.73	0.05	0.95
$x \to y/causality$	0.2	0.06	0.93	0.19	0.80	0.05	0.95
$x \to z/causality$	0.9	0.08	0.91	0.32	0.67	0.08	0.92
$x \to z/causality$	0.8	0.08	0.91	0.33	0.66	0.06	0.94
$x \to z/causality$	0.4	0.08	0.93	0.26	0.73	0.07	0.93
$x \to z/causality$	0.2	0.08	0.93	0.20	0.79	0.05	0.95
$z \to y/no\ causality$	0.9	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.8	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.4	0.00	0.00	0.00	0.00	0.00	0.00
$z \to y/no\ causality$	0.2	0.00	0.00	0.00	0.00	0.00	0.00

In panel first of Table 5.1, stationary series having autoregressive coefficient value close to one i.e. 0.9 are generated with cross dependence terms θ_{21} and θ_{31}, keeping deterministic part (drift and trend) absent.

The coefficients θ_{21} and θ_{31} vary from 0.9, 0.8, 0.4 and 0.2 in matrix A of DGP given in equation (10). First row of each panel in Table 5.1 corresponds to series where θ_{21} = θ_{31} = 0.9 at different autoregressive coefficients (**$\boldsymbol{\theta_{11}}$ = $\boldsymbol{\theta_{22}}$ = $\boldsymbol{\theta_{33}}$**) **=** (0.8, 0.6, 0.4). In panel first and second of Table 5.1 the results indicate that treating VAR residuals, modified R recursive residuals and Haugh ARMA residuals in PC and Modified PC causal search algorithms, we find zero probability of significant correlation between variables y and z. The conventional results in such scenario indicate that measure of association between y and z would have very high probability of significant results. But using the PC and Modified PC algorithms in case of confounding variable both procedures perform well. It is also important to note that actual causal paths go from x to y and x to z in DGP and the probability of having these paths as significant and have different probabilities. when we use VAR residuals, Modified R recursive residuals and Haugh residuals in causal search algorithms it gives correct results on average about 6%, 46% and 7% respectively and this does not vary significantly when cross term changes from 0.9 to 0.2. But when the auto regressive coefficients (**$\boldsymbol{\theta_{11}}$ = $\boldsymbol{\theta_{22}}$ = $\boldsymbol{\theta_{33}}$**) approaches toward zero i.e. 0.4, the probabilities of significant results go down in case of Modified R recursive residuals, as evident from the Table 5.1, when you move down from panel 1st to panel 3rd.

In Table 5.2 the same procedure is used but the only difference is that the series generated are stationary with cross dependence terms θ_{21} and θ_{31}, keeping drift part of deterministic portion, while keeping trend absent. The coefficients θ_{21} and θ_{31} vary from 0.9, 0.8, 0.4 and 0.2 in matrix A of equation (5). Sub Panel 3rd of each Panel in Table 5.2 indicate that in case of confounder in the DGP, we again find zero probability of significant association between y and z variables. The simulated results of causal algorithms using VAR residuals, Modified R recursive residuals and Haugh ARMA residuals show on average 7%, 45% and 6% correct causal direction and this does not change significantly when cross term value (ρ) varies from 0.9 to 0.2. But when the auto regressive coefficient approaches to zero, the power of Modified PC algorithm using modified R residuals goes down, as evident from the Table 5.2, when you move down from panel 1st to panel 3rd.

In Table 5.3 stationary series are generated with cross dependence terms θ_{21} and θ_{31}, keeping deterministic part present. We found about the same results as displayed in Table 5.2 and 5.2.

5.3 Conclusion

Discussing various cases in this chapter for evaluation of spurious causation when there is confounding variable in DGP. To find the spurious causation, stationary series with different complications has been generated from the data generating process (DGP). In DGP x is causing y and x is causing z, but there is no direct causality between y and z.

Using stationary series with various specifications (drift and trend), simulated results of PC and Modified PC algorithms show us zero probability of significant correlation between variables y and z. It is concluded that in case of confounding variable both PC and Modified PC algorithms perform same. It is also important to note that the performance of Modified PC using modified R recursive residuals in finding the correct causal path is high than using VAR residuals and Haugh ARMA residuals in PC causal algorithms. Because the original causal paths orient from x to y and from x to z only and the probability of having these paths are significant with different probabilities.

CHAPTER 6

POWER COMPARISON

To achieve the proposed objectives, comparison of PC and Modified PC algorithm are made through size and power properties. It is investigated in the previous chapter 4 (Size Evaluation) that PC algorithm using VAR residuals and Modified PC algorithm using Haugh test residuals showed about nominal size distortion. But Modified PC algorithm using Modified R recursive residuals showed high size distortion when the generated series are non-stationary. On the other hand, when we use low memory series (stationary) with weak autoregression, there is no size distortion in PC and Modified PCs algorithms.

In this chapter, the power of the PC and Modified PCs causal search algorithm is analyzed. We know that we have used three kind of residuals of the original time series in the casual search algorithm. The power of any test is defined as the probability of rejecting null hypothesis when it is false i.e.

$$Power = P(Rejecting\ H_0/H_1\ is\ True)$$

We analyze the power of PC for a variety of situations. We know that the power also depends on several nuisance parameters related to the “deterministic part” as well as “stochastic part”. Among the deterministic part are component of drift and trend while among stochastic part, we have the autoregressive coefficient of the three series which also determine the stationarity of the series. This study used sample size of 100 for data generating process under alternative hypothesis to calculate power.

As for calculation of size, to calculate the power, we use the data generating process used in Equation (5) which is as follows

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \dots (10)$$

where, $\begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho 1 & \rho 2 \\ \rho 1 & 1 & \rho \\ \rho 2 & \rho & 1 \end{bmatrix} \right)$

$$X_t = AX_{t-1} + BDt + \varepsilon_t \qquad \varepsilon_t \sim N(0, \Omega)$$

This data generating process will generate causally dependent series when A is non-diagonal matrix. We take $\begin{bmatrix} \theta_{11} & 0 & 0 \\ \theta_{21} & \theta_{22} & 0 \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix} \neq 0$ and ρ1= ρ2 = 0 which implies that there is no contemporaneous correlation among generated series x, y and z. Therefore, x, y and z are serially dependent but have no contemporaneous relationship.

Once the series from DGP in equation (10) are generated, we calculate the residuals of the series through one of the three procedures i.e. VAR model, Modified R and ARMA model which are used subsequently in the PC causal search algorithm. Power properties of these approaches will be calculated by finding the probability of each of the two mentioned scenarios i.e. Correct[11] and Omitted[12]. For this analysis, dependent autoregressive stationary and nonstationary time series are being generated with different complications; (with drift and trend), (with drift only) and (with drift and with trend). Monte Carlo simulations have been carried out in this study. All the estimated results have been summarized after 10000 times simulations from the data generating process given in equation (10).

[11] The link is present both in data generating process and final results of PC algorithm.
[12] The link is present in data generating process but absent in PC algorithm results

6.1 Power of PC and Modified PCs Algorithms with non-stationary Series

First, we have generated non-stationary series x, y and z on the basis of change in both stochastic and deterministic part using the data generating process given in equation (10). we have imputed double cross terms in matrix A (θ_{21} and θ_{32}) and power of PC and Modified PC algorithms are calculated. The cross terms (θ_{21} and θ_{32}) establish correlation between x and y and y and z and its value also changes from 0.9, 0.8, 0.6, 0.4 and 0.2 in matrix A which shows that $x \rightarrow y$ and $y \rightarrow z$ respectively.

i.e. matrix A = $\begin{bmatrix} \theta_{11} & 0 & 0 \\ \theta_{21} & \theta_{22} & 0 \\ 0 & \theta_{32} & \theta_{33} \end{bmatrix} \neq 0$.

Table 6.1: Power of PC and Modified PC algorithms using Non-Stationary Series without Drift and Trend

		PC		**Modified PC**		**Modified PC**	
		VAR residuals		MR residuals		Haugh residuals	
	$\boldsymbol{\rho}$	**Correct**	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
			$\theta_{11} = \theta_{22} = \theta_{33} = 1$				
$x \rightarrow y/causality$	0.9	0.06	0.93	0.65	0.34	0.15	0.84
$x \rightarrow y/causality$	0.8	0.06	0.93	0.63	0.36	0.13	0.86
$x \rightarrow y/causality$	0.6	0.05	0.94	0.60	0.39	0.15	0.84
$x \rightarrow y/causality$	0.4	0.05	0.94	0.62	0.37	0.12	0.87
$x \rightarrow y/causality$	0.2	0.05	0.94	0.60	0.39	0.09	0.90
$y \rightarrow z/causality$	0.9	0.05	0.94	0.57	0.42	0.01	0.98
$y \rightarrow z/causality$	0.8	0.06	0.93	0.55	0.45	0.01	0.98
$y \rightarrow z/causality$	0.6	0.07	0.93	0.53	0.46	0.02	0.97
$y \rightarrow z/causality$	0.4	0.07	0.92	0.51	0.48	0.04	0.95
$y \rightarrow z/causality$	0.2	0.05	0.94	0.49	0.50	0.04	0.95

Table 6.2: Power of PC and Modified PC algorithms using Non-Stationary Series with Drift only

		PC		**Modified PC**		**Modified PC**	
		VAR residuals		MR residuals		Haugh residuals	
	$\boldsymbol{\rho}$	**Correct**	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
			$\theta_{11} = \theta_{22} = \theta_{33} = 1$				
$x \rightarrow y/causality$	0.9	0.07	0.92	0.73	0.26	0.13	0.86
$x \rightarrow y/causality$	0.8	0.06	0.93	0.75	0.24	0.15	0.84
$x \rightarrow y/causality$	0.6	0.06	0.93	0.75	0.24	0.13	0.86
$x \rightarrow y/causality$	0.4	0.07	0.92	0.74	0.25	0.11	0.88
$x \rightarrow y/causality$	0.2	0.05	0.94	0.72	0.27	0.08	0.91
$y \rightarrow z/causality$	0.9	0.07	0.92	0.78	0.21	0.20	0.79
$y \rightarrow z/causality$	0.8	0.05	0.94	0.79	0.20	0.22	0.77
$y \rightarrow z/causality$	0.6	0.06	0.93	0.79	0.20	0.22	0.77
$y \rightarrow z/causality$	0.4	0.06	0.93	0.76	0.23	0.20	0.79
$y \rightarrow z/causality$	0.2	0.06	0.93	0.77	0.22	0.16	0.83

Table 6.3: Power of PC and Modified PC algorithms using Non-Stationary Series with Drift and Trend

		PC		Modified PC		Modified PC	
		VAR residuals		MR residuals		Haugh residuals	
	ρ	**Correct**	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
			$\theta_{11} = \theta_{22} = \theta_{33} = 1$				
$x \rightarrow y/causality$	0.9	0.07	0.92	0.73	0.26	0.68	0.31
$x \rightarrow y/causality$	0.8	0.07	0.92	0.75	0.24	0.67	0.32
$x \rightarrow y/causality$	0.6	0.06	0.93	0.74	0.25	0.66	0.33
$x \rightarrow y/causality$	0.4	0.05	0.94	0.73	0.27	0.68	0.31
$x \rightarrow y/causality$	0.2	0.07	0.93	0.76	0.23	0.60	0.40
$y \rightarrow z/causality$	0.9	0.07	0.92	0.76	0.23	0.89	0.10
$y \rightarrow z/causality$	0.8	0.05	0.94	0.78	0.21	0.89	0.10
$y \rightarrow z/causality$	0.6	0.07	0.92	0.75	0.24	0.88	0.11
$y \rightarrow z/causality$	0.4	0.06	0.93	0.77	0.23	0.88	0.11
$y \rightarrow z/causality$	0.2	0.06	0.93	0.78	0.21	0.86	0.14

In the above Table 6.1 the simulated results of PC and Modified PC algorithms using three different residuals series i.e. VAR residuals, Modified R (MR) residuals and Haugh ARMA residuals are reported. Each outcome is expressed as a proportion of the number of times it might have occurred. First, non-stationary series, $\theta_{11} = \theta_{22} = \theta_{33} = 1$ having no drift and trend are generated from the DGP given in equation (10). The generated series are then analyzed in VAR model, Modified R and Haugh ARMA model and residuals series of the said models are extracted which are then used in PC and Modified PCs algorithm by treating these residuals as original variables.

In row 1 and row 6 of Table 6.1, where the coefficients θ_{21} $(x\ cause\ y) = 0.9$ and θ_{32} $(y\ cause\ z) = 0.9$. Using VAR residuals in PC causal search indicate that the probability of rejection of null hypothesis of no causality (which can be regarded as power, since in DGP null is not true) is about on average 7% and 6% respectively. Using Modified R recursive residuals instead of VAR residuals, indicating that the probability of rejection of null of no causality is about 65% for θ_{21} $(x \rightarrow y)$ and 57% for θ_{32} $(y \rightarrow z)$ Finally, when Haugh test residuals are applied in the same causal search algorithms the results reveal that the probability of rejection of null of no causality is about 15% and 2% for θ_{21} and θ_{32} respectively. These values do not change significantly in all three cases when θ_{21} and θ_{32} changes from 0.9 to 0.2.

In Table 6.2, non-stationary series (θ_{11} = θ_{22} = θ_{33} = 1) with drift are generated from the given DGP (10). The cross dependences terms θ_{21} and θ_{32} also changes from 0.9, 0.8, 0.6, 0.4 and 0.2 in matrix A of DGP given in equation (10), which shows that $x \rightarrow y$ and $y \rightarrow z$ respectively. The simulated results in row 1 and row 6 of Table 6.2, indicate that using VAR residuals in PC causal search the probability of rejection of null of no causality is about on average 7% and 7% respectively, and this does not change significantly when cross terms θ_{21} and θ_{32} changes from 0.9 to 0.2. Using recursive residuals from Modified R, indicating that the probability of rejection of null hypothesis of no causality is on average 74% and 78% and this does not vary when θ_{21} and θ_{32} changes from 0.9 to 0.2. Similarly, when Haugh test - ARMA residuals are used, we found that the probability of rejection of null of no causality is about 13% for θ_{21} =0.9 and 20% for θ_{32} = 0.9 and this follow the same pattern when we change the cross-term value from 0.9 to 0.2.

In Table 6.3, the simulated results of PC and Modified PCs algorithms are given. In this case generated series is non-stationary having both drift and trend. Row 1 and row 6 Table 6.3 indicate that the probability of rejection of null of no causality is about 7% for both the cross terms θ_{21} and θ_{32} and displayed the same pattern when the cross terms θ_{21} and θ_{32} changes from 0.9 to 0.2. Using Modified R recursive residuals instead of VAR residuals the results indicating that the probability of rejection of null hypothesis of no causality is about 74% and 76% for both the cross terms θ_{21} and θ_{32} and this does not change significantly when θ_{21} and θ_{32} changes from 0.9 to 0.2.

Similarly, when Haugh-ARMA residuals are applied, we found that the probability of rejection of null of no causality is about 68% and 89% for θ_{21} and θ_{32} respectively.

It is clear from the above simulated results that PC algorithm (using VAR residuals) perform very bad in power properties, when the generated series are nonstationary with different specifications (drift and trend). Modified PC (using Haugh-ARMA residuals) also badly suffer in first two cases (without drift and trend) and (with drift only) while it performs better when series has both drift and trend. However, Modified PC (using Modified R recursive residuals) is performing better in all cases. Discussing the other scenario: Omitted error is high when VAR residuals are referred to PC causal search in all three Tables 6.1, 6.2 and 6.3. The same results are displayed by Modified PC using Haugh ARMA residuals in the first two Tables 6.1 and 6.2 and omitted error significantly decrease when the series are generated with drift and trend as shown in Table 6.3. Modified PC algorithm using modified R recursive residuals perform good with less omission in all cases. Hence, from the above Tables 6.1, 6.2 and 6.3, it is concluded that Modified PC algorithm using modified R recursive residuals works well in case of power than those algorithms using VAR and Haugh ARMA residuals.

We have also checked the performance of these causal search algorithms when DGP contain triples of cross terms (θ_{21} θ_{32} and θ_{31}). Non-stationary series x, y and z are generated from the given DGP (10). we have imputed triple cross terms in matrix A of DGP in equation (10) i.e. A = $\begin{bmatrix} \theta_{11} & 0 & 0 \\ \boldsymbol{\theta_{21}} & \theta_{22} & 0 \\ \boldsymbol{\theta_{31}} & \boldsymbol{\theta_{32}} & \theta_{33} \end{bmatrix} \neq 0$. This implies that x is causing y $(x \rightarrow y)$ y is causing z $(y \rightarrow z)$ and x is causing z $(x \rightarrow z)$.

Table 6.4: Power of PC and Modified PC algorithms using Non-Stationary Series without Drift and Trend

		Original PC		Modified R		Modified R	
		VAR residuals		MR residuals		Haugh residuals	
	$\boldsymbol{\rho}$	**Correct**	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
		$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 1}$					
$x \to y/causality$	0.9	0.06	0.93	0.63	0.36	0.13	0.86
$x \to y/causality$	0.8	0.08	0.91	0.61	0.38	0.14	0.85
$x \to y/causality$	0.6	0.06	0.93	0.62	0.37	0.14	0.85
$x \to y/causality$	0.4	0.06	0.93	0.61	0.38	0.12	0.87
$x \to y/causality$	0.2	0.07	0.93	0.58	0.41	0.10	0.90
$y \to z/causality$	0.9	0.06	0.93	0.55	0.45	0.10	0.90
$y \to z/causality$	0.8	0.07	0.93	0.53	0.46	0.20	0.80
$y \to z/causality$	0.6	0.08	0.91	0.50	0.49	0.30	0.70
$y \to z/causality$	0.4	0.05	0.94	0.53	0.46	0.20	0.80
$y \to z/causality$	0.2	0.07	0.92	0.48	0.51	0.20	0.80
$x \to z/causality$	0.9	0.06	0.93	0.55	0.45	0.10	0.90
$x \to z/causality$	0.8	0.07	0.93	0.53	0.46	0.20	0.80
$x \to z/causality$	0.6	0.08	0.91	0.50	0.49	0.30	0.70
$x \to z/causality$	0.4	0.05	0.94	0.53	0.46	0.20	0.80
$x \to z/causality$	0.2	0.07	0.92	0.48	0.51	0.20	0.80

Table 6.5: Power of PC and Modified PC algorithms using Non-Stationary Series with Drift only

		Original PC		Modified PC		Modified PC	
		VAR residuals		MR residuals		Haugh residuals	
	$\boldsymbol{\rho}$	**Correct**	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
		$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 1}$					
$x \to y/causality$	0.9	0.05	0.94	0.74	0.25	0.15	0.84
$x \to y/causality$	0.8	0.07	0.93	0.74	0.25	0.13	0.87
$x \to y/causality$	0.6	0.06	0.93	0.74	0.25	0.15	0.84
$x \to y/causality$	0.4	0.06	0.93	0.75	0.24	0.13	0.86
$x \to y/causality$	0.2	0.06	0.94	0.74	0.25	0.09	0.90
$y \to z/causality$	0.9	0.08	0.91	0.78	0.21	0.21	0.78
$y \to z/causality$	0.8	0.06	0.93	0.77	0.22	0.21	0.78
$y \to z/causality$	0.6	0.05	0.94	0.75	0.25	0.24	0.75
$y \to z/causality$	0.4	0.07	0.92	0.78	0.21	0.20	0.79
$y \to z/causality$	0.2	0.06	0.93	0.76	0.23	0.16	0.83
$x \to z/causality$	0.9	0.08	0.91	0.78	0.21	0.21	0.78
$x \to z/causality$	0.8	0.06	0.93	0.77	0.22	0.21	0.78
$x \to z/causality$	0.6	0.05	0.94	0.75	0.25	0.24	0.75
$x \to z/causality$	0.4	0.07	0.92	0.78	0.21	0.20	0.79
$x \to z/causality$	0.2	0.06	0.93	0.76	0.23	0.16	0.83

Table 6.6: Power of PC and Modified PC algorithms using Non-Stationary Series with Drift and Trend

		Original PC		**Modified PC**		**Modified PC**	
		VAR residuals		MR residuals		Haugh residuals	
	$\boldsymbol{\rho}$	**Correct**	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
				$\boldsymbol{\theta_{11} = \theta_{22} = \theta_{33} = 1}$			
$x \to y/causality$	0.9	0.07	0.92	0.74	0.25	0.66	0.33
$x \to y/causality$	0.8	0.06	0.93	0.74	0.25	0.70	0.29
$x \to y/causality$	0.6	0.07	0.92	0.75	0.24	0.70	0.29
$x \to y/causality$	0.4	0.05	0.94	0.74	0.26	0.67	0.32
$x \to y/causality$	0.2	0.05	0.94	0.74	0.25	0.63	0.36
$y \to z/causality$	0.9	0.07	0.92	0.77	0.22	0.86	0.13
$y \to z/causality$	0.8	0.05	0.94	0.76	0.23	0.90	0.09
$y \to z/causality$	0.6	0.06	0.93	0.76	0.23	0.90	0.09
$y \to z/causality$	0.4	0.06	0.93	0.76	0.23	0.88	0.11
$y \to z/causality$	0.2	0.07	0.92	0.76	0.23	0.86	0.13
$x \to z/causality$	0.9	0.07	0.92	0.77	0.22	0.86	0.13
$x \to z/causality$	0.8	0.05	0.94	0.76	0.23	0.90	0.09
$x \to z/causality$	0.6	0.06	0.93	0.76	0.23	0.90	0.09
$x \to z/causality$	0.4	0.06	0.93	0.76	0.23	0.88	0.11
$x \to z/causality$	0.2	0.07	0.92	0.76	0.23	0.86	0.13

In Tables 6.4, when the generated series are non-stationary ($\theta_{11} = \theta_{22} = \theta_{33}$= 1) with no drift and linear trend included in DGP (10) i.e. by setting, $(a_1 \ \ b_1 \ \ c_1 \ \ a_2 \ \ b_2 \ \ c_2) = 0$. The DGP also contain the cross terms θ_{21} = 0.9, θ_{32} = 0.9 and θ_{31} = 0.9 which changes from 0.9 to 0.8, 0.6, 0.4 and 0.2 in matrix A of equation (5). This setup implies that in the DGP x is causing y ($x \to y$), y is causing z ($y \to z$) and x is causing z ($x \to z$)

In Table 6.4, row 1, row 6 and row 11 show that the cross terms θ_{21}, θ_{32} and θ_{31} respectively. The results indicate that using VAR residuals in PC algorithm, the probability of rejection of null of no causality is about 6% on average for all three possible directions and this does not change significantly when the cross terms θ_{21}, θ_{32} and θ_{31} varies from 0.9 to 0.2. In case of Modified PC algorithm using modified R recursive residuals, the probability of rejection of null of no causality is about 60% on average for the given three possible causal directions. Finally, the results of Modified PC algorithm with Haugh ARMA residuals, the probability of rejection of null of no causality is on average 15% and does not change significantly when the values of $\theta_{21},\ \theta_{32}\ and\ \theta_{31}$ changes from 0.9 to 0.2. It is also important that omitted errors in PC algorithm are high when VAR residuals and Haugh residuals are used. But when Modified R residuals are referred to PC algorithms, the proportion of omitted errors decrease as shown.

Table 6.5 consists of simulation results of causal search algorithms when the DGP has only drift. This means we set matrix B = $(\ a_2\ \ b_2\ \ c_2) = 0$. While in Table 6.6 the results of PC and Modified PCs came after generating non-stationary series with both drift and trend. Both Tables 6.5 and 6.6 displayed the same results as discussed in previous Table 6.4 but the results in Table 6.6 differ in case of Modified PC using Haugh ARMA residuals because it performs better when series has both drift and trend. The probability of rejection of null of no causality is about 80% on average for the given three possible causal directions which is very high.

PC algorithm using (VAR residuals) performed very bad in power in all three cases; drift and trend. However, Modified PC using modified R residuals is performing better in all three cases as evident from the Tables 6.4, 6.5 and 6.6.

6.2 Power Comparison of stationary Series with Different Complications:

In previous section the performance of PC and Modified PC algorithms in non-stationary series was analyzed. Tables 6.7, 6.8 and 6.9 summarizes performance of PC and Modified PC causal search algorithms, when the underlying series are stationary. To make the series stationary we put the diagonal entries θ_{11}, θ_{22}, θ_{33} in DGP (10) to be smaller than unity. To create cross dependences, we choose some of the non-diagonal entries to be non-zero. We choose $\theta_{21} > 0$, $\theta_{32} > 0$, and $\theta_{31} > 0$ which make x depending on y, y depending on z and x depending on z respectively.

Table 6.7: Power of PC and Modified PC algorithms using Stationary Series without Drift and Trend

		Original PC		**Modified PC**		**Modified PC**	
		VAR residuals		MR residuals		Haugh residuals	
		Correct	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.9$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.39	0.61	0.06	0.93
$x \rightarrow y/causality$	0.8	0.06	0.93	0.41	0.59	0.05	0.94
$x \rightarrow y/causality$	0.6	0.07	0.92	0.40	0.59	0.05	0.94
$x \rightarrow y/causality$	0.4	0.06	0.93	0.42	0.57	0.05	0.95
$x \rightarrow y/causality$	0.2	0.06	0.93	0.38	0.62	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.29	0.70	0.06	0.94
$x \rightarrow y/causality$	0.8	0.07	0.92	0.28	0.72	0.05	0.94
$x \rightarrow y/causality$	0.6	0.07	0.93	0.28	0.71	0.05	0.94
$x \rightarrow y/causality$	0.4	0.08	0.92	0.26	0.73	0.05	0.94
$x \rightarrow y/causality$	0.2	0.06	0.93	0.30	0.69	0.05	0.94
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \rightarrow y/causality$	0.9	0.05	0.94	0.11	0.88	0.06	0.93
$x \rightarrow y/causality$	0.8	0.05	0.95	0.16	0.84	0.07	0.92
$x \rightarrow y/causality$	0.6	0.06	0.93	0.17	0.82	0.06	0.93
$x \rightarrow y/causality$	0.4	0.06	0.93	0.16	0.83	0.06	0.93
$x \rightarrow y/causality$	0.2	0.06	0.94	0.18	0.81	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$					
$x \rightarrow y/causality$	0.9	0.05	0.94	0.11	0.88	0.07	0.92
$x \rightarrow y/causality$	0.8	0.05	0.94	0.11	0.88	0.05	0.94
$x \rightarrow y/causality$	0.6	0.07	0.93	0.11	0.89	0.05	0.94
$x \rightarrow y/causality$	0.4	0.06	0.93	0.10	0.89	0.05	0.94
$x \rightarrow y/causality$	0.2	0.08	0.91	0.10	0.89	0.07	0.92
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.2$					
$x \rightarrow y/causality$	0.9	0.05	0.94	0.06	0.93	0.06	0.93
$x \rightarrow y/causality$	0.8	0.06	0.94	0.10	0.90	0.05	0.94
$x \rightarrow y/causality$	0.6	0.07	0.92	0.07	0.92	0.05	0.94
$x \rightarrow y/causality$	0.4	0.06	0.94	0.07	0.92	0.05	0.94
$x \rightarrow y/causality$	0.2	0.06	0.93	0.08	0.91	0.06	0.93

Table 6.8: Power of PC and Modified PC algorithms using Stationary Series with Drift only

		Original PC		**Modified PC**		**Modified PC**	
		VAR residuals		MR residuals		Haugh residuals	
		Correct.	**Omitted**	**Correct.**	**Omitted**	**Correct.**	**Omitted**
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.9$					
$x \rightarrow y/causality$	0.9	0.06	0.94	0.49	0.50	0.06	0.93
$x \rightarrow y/causality$	0.8	0.06	0.93	0.50	0.50	0.07	0.92
$x \rightarrow y/causality$	0.6	0.06	0.93	0.50	0.49	0.06	0.93
$x \rightarrow y/causality$	0.4	0.06	0.93	0.51	0.48	0.07	0.92
$x \rightarrow y/causality$	0.2	0.07	0.92	0.49	0.50	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.35	0.64	0.06	0.93
$x \rightarrow y/causality$	0.8	0.06	0.93	0.34	0.65	0.07	0.92
$x \rightarrow y/causality$	0.6	0.06	0.94	0.32	0.67	0.06	0.93
$x \rightarrow y/causality$	0.4	0.08	0.91	0.34	0.65	0.06	0.93
$x \rightarrow y/causality$	0.2	0.06	0.93	0.33	0.66	0.07	0.92
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.16	0.84	0.06	0.93
$x \rightarrow y/causality$	0.8	0.07	0.93	0.15	0.84	0.07	0.92
$x \rightarrow y/causality$	0.6	0.05	0.94	0.16	0.84	0.06	0.93
$x \rightarrow y/causality$	0.4	0.05	0.94	0.18	0.81	0.07	0.92
$x \rightarrow y/causality$	0.2	0.06	0.93	0.18	0.81	0.08	0.91
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$					
$x \rightarrow y/causality$	0.9	0.05	0.94	0.09	0.90	0.06	0.93
$x \rightarrow y/causality$	0.8	0.05	0.94	0.10	0.89	0.06	0.93
$x \rightarrow y/causality$	0.6	0.05	0.95	0.11	0.88	0.07	0.92
$x \rightarrow y/causality$	0.4	0.05	0.94	0.11	0.88	0.05	0.94
$x \rightarrow y/causality$	0.2	0.07	0.92	0.10	0.89	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.2$					
$x \rightarrow y/causality$	0.9	0.05	0.94	0.06	0.93	0.06	0.93
$x \rightarrow y/causality$	0.8	0.05	0.95	0.05	0.94	0.05	0.94
$x \rightarrow y/causality$	0.6	0.06	0.93	0.09	0.91	0.06	0.93
$x \rightarrow y/causality$	0.4	0.06	0.93	0.07	0.92	0.04	0.96
$x \rightarrow y/causality$	0.2	0.06	0.93	0.09	0.90	0.05	0.94

Table 6.9: Power of PC and Modified PC algorithms using Trend Stationary Series with Drift and Trend

		Original PC		Modified PC		Modified PC	
		VAR residuals		MR residuals		Haugh residuals	
		Correct.	**Omitted**	**Correct.**	**Omitted**	**Correct.**	**Omitted**
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.9$					
$x \to y/causality$	0.9	0.07	0.92	0.69	0.30	0.18	0.81
$x \to y/causality$	0.8	0.06	0.93	0.66	0.33	0.18	0.81
$x \to y/causality$	0.6	0.05	0.94	0.69	0.30	0.20	0.79
$x \to y/causality$	0.4	0.06	0.93	0.67	0.32	0.19	0.80
$x \to y/causality$	0.2	0.06	0.93	0.59	0.40	0.23	0.77
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \to y/causality$	0.9	0.06	0.94	0.37	0.62	0.17	0.82
$x \to y/causality$	0.8	0.05	0.94	0.38	0.61	0.17	0.82
$x \to y/causality$	0.6	0.06	0.94	0.37	0.62	0.15	0.84
$x \to y/causality$	0.4	0.08	0.92	0.37	0.62	0.16	0.83
$x \to y/causality$	0.2	0.07	0.92	0.38	0.61	0.17	0.83
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \to y/causality$	0.9	0.05	0.94	0.16	0.83	0.13	0.86
$x \to y/causality$	0.8	0.06	0.93	0.16	0.83	0.12	0.87
$x \to y/causality$	0.6	0.06	0.94	0.17	0.83	0.12	0.87
$x \to y/causality$	0.4	0.08	0.91	0.18	0.81	0.13	0.86
$x \to y/causality$	0.2	0.05	0.94	0.20	0.79	0.14	0.85
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$					
$x \to y/causality$	0.9	0.05	0.94	0.08	0.91	0.10	0.89
$x \to y/causality$	0.8	0.07	0.92	0.10	0.89	0.10	0.90
$x \to y/causality$	0.6	0.07	0.92	0.107	0.89	0.09	0.90
$x \to y/causality$	0.4	0.07	0.92	0.11	0.88	0.04	0.96
$x \to y/causality$	0.2	0.06	0.93	0.12	0.87	0.05	0.94
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.2$					
$x \to y/causality$	0.9	0.06	0.94	0.06	0.93	0.06	0.93
$x \to y/causality$	0.8	0.07	0.93	0.07	0.92	0.05	0.94
$x \to y/causality$	0.6	0.06	0.93	0.06	0.93	0.06	0.93
$x \to y/causality$	0.4	0.06	0.93	0.06	0.93	0.09	0.91
$x \to y/causality$	0.2	0.06	0.93	0.07	0.92	0.06	0.93

In panel first of the Table 6.7 stationary series having coefficient value close to unity (θ_{11} = θ_{22} = θ_{33} = 0.9), with no drift and no trend are generated from the given DGP with cross dependence term $\theta_{21.}$ The cross term θ_{21} changes from 0.9, 0.8, 0.6, 0.4 and 0.2 in matrix A. This setup implies that in the DGP x is causing y ($x \rightarrow y$). The above generated data are then referred PC algorithms, where VAR residuals, Modified R residuals and ARMA residuals are used.

First row of each panel corresponds to series where θ_{21} = 0.9 at different autoregressive coefficient (0.9, 0.8, 0.6, 0.4 and 0.2). In row 1st of Table 6.7, the results indicate that using VAR residuals in PC algorithm and Modified R residuals and Haugh ARMA residuals in Modified PC algorithms, the probability of rejection of null of no causality (which can be regarded as power, since in DGP null is not true) is about 6%, 39% and 6% respectively (given in column **correct**), and this does not change significantly when cross term, θ_{21} changes from 0.9 to 0.2. The table reveals that the power is best for Modified PC using modified R residuals while using VAR and Haugh-ARMA residuals the algorithms perform bad in case of power. But when the auto regressive coefficient goes to zero (i.e. 0.2), the power of Modified PC using MR residuals goes down, as evident from the Table 6.7, when you move down from panel 1st to panel 5th. The other scenario: Omitted errors display different results when the generated series are stationary. The rate of omission in both PC using VAR residual and Modified PC using Haugh ARMA residuals is very high in all three specification (without drift & trend) (with drift only) and (with drift & trend). Discussing the performance of Modified PC using modified R residuals, the proportion of omitted errors are significantly increasing when the auto regressive coefficient approaches to strong stationary.

In Table 6.8, Stationary series are generated having drift and with cross dependence term $\theta_{21.}$ The cross term θ_{21} changes from 0.9, 0.8, 0.6, 0.4 and 0.2 in matrix A of DGP, which implies that $x \rightarrow y$. We have extracted the residuals (VAR residuals, Modified R residuals and ARMA residuals) series from the generated data and these residuals series are treated as original variables in PC and Modified PC algorithms. In Table 6.8, row 1st of each panel corresponds to series where θ_{21} = 0.9 at different autoregressive coefficients (θ_{11}, θ_{22} and θ_{33}) values (0.9, 0.8, 0.6, 0.4 and 0.2). Row 1 results indicate that using VAR residuals, Modified R recursive residuals and Haugh ARMA residuals in causal search algorithms, the probability of rejection of null of no causality is about 6%, 39% and 7% respectively, and this does not change significantly when cross term, θ_{21} changes from 0.9 to 0.2. The table reveals that the power is best for Modified R residuals while VAR and ARMA residuals perform poor in power properties. But when the auto regressive coefficient ($\theta_{11,}$ θ_{22} and θ_{33}) goes from 0.9 to 0.2, the power of Modified R goes down, as evident from the table, when you move down from panel 1st to panel 5th.

The Table 6.9 contain simulated results of PC and Modified PC algorithms using VAR, Modified R and ARMA residuals obtained from stationary series based on different autoregressive coefficients i.e. change in stochastic part in presence of deterministic part (inclusion of drift and linear trend in DGP). In DGP we have also imputed causality between x and y by taking the coefficient θ_{21}value nonzero. The cross term θ_{21} value also changes from 0.9, 0.8, 0.6, 0.4 and 0.2 in matrix A of data generating process. The generated data series x, y and z are then used in VAR model, Modified R and ARMA model, and obtained residuals of the said models. These residuals series are then referred PC and Modified PC algorithms. In Table 6.9, row 1st of each panel 1st corresponds to series where θ_{21} = 0.9 at autoregressive coefficient

(0.9) indicate that using VAR residuals in PC and modified R recursive residuals and Haugh ARMA residuals in Modified PC causal search the probability of rejection of null of no causality is about 8%, 70% and 18% respectively, and this does not change significantly when cross term, θ_{21} changes from 0.9 to 0.2. The table reveals that the power is best for Modified PC using modified R residuals while using VAR and ARMA residuals, the causal search algorithms present poor performance in power properties. But when the auto regressive coefficient (θ_{11}, θ_{22} and θ_{33}) goes from 0.9 to 0.2, the power of Modified PC using modified R residuals goes down, as evident from the table, when you move down from panel 1st to panel 5th.

In Tables 6.7, 6.8 and 6.9, we have generated stationary series with various specifications (drift and trend). We have only imputed causality between only two variables x and y i.e. (θ_{21}) in data generating process and power of various procedures are calculated. In the next round, we have imputed causality not only in the single pair (x, y) but causality between (y, z) is also assigned (i.e. we put two cross terms θ_{21}, θ_{32} in DGP) and then we calculated the power of PC and Modified PC algorithms given in Tables 6.10, 6.11 and 6.12. In the final round we have imputed three cross terms $\theta_{21}, \theta_{32,}$ and θ_{31} in data generating process with different auto regressive coefficient having various complications and power of PC and Modified PCs are calculated given Table 6.13, 6.14 and 6.15 of Appendix II.

Table 6.10: Power of PC and Modified PC algorithms using Stationary Series without Drift and Trend

		Original PC		Modified PC		Modified PC	
		VAR residuals		MR residuals		Haugh residuals	
		Correct.	**Omitted**	**Correct.**	**Omitted**	**Correct.**	**Omitted**
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.45	0.54	0.04	0.95
$x \rightarrow y/causality$	0.8	0.06	0.94	0.49	0.50	0.04	0.95
$x \rightarrow y/causality$	0.4	0.07	0.92	0.57	0.42	0.04	0.95
$x \rightarrow y/causality$	0.2	0.06	0.93	0.50	0.49	0.05	0.94
$y \rightarrow z/causality$	0.9	0.06	0.93	0.43	0.56	0.04	0.95
$y \rightarrow z/causality$	0.8	0.07	0.92	0.47	0.52	0.03	0.96
$y \rightarrow z/causality$	0.4	0.06	0.93	0.59	0.40	0.04	0.96
$y \rightarrow z/causality$	0.2	0.05	0.95	0.51	0.48	0.05	0.95
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.44	0.55	0.04	0.95
$x \rightarrow y/causality$	0.8	0.06	0.93	0.50	0.49	0.03	0.96
$x \rightarrow y/causality$	0.4	0.06	0.93	0.50	0.49	0.03	0.96
$x \rightarrow y/causality$	0.2	0.06	0.93	0.32	0.67	0.05	0.94
$y \rightarrow z/causality$	0.9	0.06	0.93	0.47	0.52	0.04	0.95
$y \rightarrow z/causality$	0.8	0.06	0.93	0.49	0.50	0.03	0.96
$y \rightarrow z/causality$	0.4	0.06	0.93	0.45	0.54	0.05	0.94
$y \rightarrow z/causality$	0.2	0.07	0.92	0.31	0.69	0.05	0.94
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$					
$x \rightarrow y/causality$	0.9	0.05	0.94	0.35	0.64	0.04	0.95
$x \rightarrow y/causality$	0.8	0.05	0.94	0.31	0.68	0.04	0.95
$x \rightarrow y/causality$	0.4	0.07	0.92	0.28	0.71	0.03	0.96
$x \rightarrow y/causality$	0.2	0.08	0.91	0.16	0.83	0.05	0.94
$y \rightarrow z/causality$	0.9	0.05	0.94	0.26	0.73	0.04	0.95
$y \rightarrow z/causality$	0.8	0.06	0.93	0.32	0.67	0.03	0.96
$y \rightarrow z/causality$	0.4	0.07	0.92	0.26	0.73	0.05	0.94
$y \rightarrow z/causality$	0.2	0.06	0.94	0.17	0.82	0.05	0.94
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.2$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.11	0.88	0.04	0.95
$x \rightarrow y/causality$	0.8	0.06	0.93	0.14	0.85	0.04	0.95
$x \rightarrow y/causality$	0.4	0.06	0.93	0.12	0.87	0.04	0.95
$x \rightarrow y/causality$	0.2	0.05	0.94	0.09	0.90	0.06	0.93
$y \rightarrow z/causality$	0.9	0.05	0.94	0.12	0.87	0.05	0.94
$y \rightarrow z/causality$	0.8	0.05	0.93	0.15	0.84	0.04	0.96
$y \rightarrow z/causality$	0.4	0.06	0.93	0.12	0.88	0.04	0.95
$y \rightarrow z/causality$	0.2	0.06	0.93	0.08	0.91	0.05	0.94

Table 6.11: Power of PC and Modified PC algorithms using Stationary Series with Drift only

		Original PC		Modified PC		Modified PC	
		VAR residuals		MR residuals		Haugh residuals	
		Correct	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \to y/causality$	0.9	0.07	0.92	0.42	0.57	0.05	0.94
$x \to y/causality$	0.8	0.07	0.92	0.39	0.60	0.05	0.94
$x \to y/causality$	0.4	0.05	0.94	0.44	0.55	0.06	0.93
$x \to y/causality$	0.2	0.06	0.93	0.50	0.49	0.07	0.92
$y \to z/causality$	0.9	0.05	0.95	0.42	0.57	0.05	0.94
$y \to z/causality$	0.8	0.06	0.93	0.42	0.57	0.05	0.94
$y \to z/causality$	0.4	0.06	0.93	0.42	0.57	0.06	0.93
$y \to z/causality$	0.2	0.06	0.93	0.50	0.49	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \to y/causality$	0.9	0.06	0.93	0.44	0.55	0.06	0.93
$x \to y/causality$	0.8	0.08	0.92	0.44	0.56	0.06	0.93
$x \to y/causality$	0.4	0.05	0.94	0.47	0.52	0.04	0.95
$x \to y/causality$	0.2	0.06	0.93	0.33	0.66	0.06	0.93
$y \to z/causality$	0.9	0.06	0.93	0.43	0.56	0.05	0.94
$y \to z/causality$	0.8	0.07	0.93	0.46	0.53	0.04	0.95
$y \to z/causality$	0.4	0.05	0.94	0.50	0.49	0.07	0.92
$y \to z/causality$	0.2	0.06	0.93	0.33	0.66	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$					
$x \to y/causality$	0.9	0.06	0.93	0.29	0.70	0.06	0.93
$x \to y/causality$	0.8	0.05	0.94	0.33	0.66	0.04	0.95
$x \to y/causality$	0.4	0.06	0.93	0.25	0.74	0.04	0.95
$x \to y/causality$	0.2	0.05	0.94	0.18	0.81	0.06	0.93
$y \to z/causality$	0.9	0.06	0.93	0.32	0.67	0.06	0.93
$y \to z/causality$	0.8	0.07	0.92	0.31	0.68	0.05	0.94
$y \to z/causality$	0.4	0.07	0.92	0.26	0.73	0.07	0.92
$y \to z/causality$	0.2	0.05	0.94	0.16	0.84	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.2$					
$x \to y/causality$	0.9	0.05	0.94	0.14	0.85	0.05	0.94
$x \to y/causality$	0.8	0.05	0.94	0.14	0.85	0.04	0.95
$x \to y/causality$	0.4	0.06	0.93	0.11	0.88	0.05	0.94
$x \to y/causality$	0.2	0.07	0.92	0.09	0.90	0.05	0.94
$y \to z/causality$	0.9	0.08	0.91	0.12	0.88	0.07	0.93
$y \to z/causality$	0.8	0.07	0.92	0.13	0.86	0.06	0.93
$y \to z/causality$	0.4	0.06	0.93	0.10	0.89	0.07	0.92
$y \to z/causality$	0.2	0.05	0.94	0.08	0.91	0.06	0.94

Table 6.12: Power of PC and Modified PC algorithms using Trend Stationary Series with Drift and Trend

		Original PC		Modified PC		Modified PC	
		VAR residuals		MR residuals		Haugh residuals	
		Correct	**Omitted**	**Correct**	**Omitted**	**Correct**	**Omitted**
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \rightarrow y/causality$	0.9	0.07	0.92	0.42	0.57	0.05	0.94
$x \rightarrow y/causality$	0.8	0.07	0.92	0.39	0.60	0.05	0.94
$x \rightarrow y/causality$	0.4	0.05	0.94	0.44	0.55	0.06	0.93
$x \rightarrow y/causality$	0.2	0.06	0.93	0.50	0.49	0.07	0.92
$y \rightarrow z/causality$	0.9	0.05	0.95	0.42	0.57	0.05	0.94
$y \rightarrow z/causality$	0.8	0.06	0.93	0.42	0.57	0.05	0.94
$y \rightarrow z/causality$	0.4	0.06	0.93	0.42	0.57	0.06	0.93
$y \rightarrow z/causality$	0.2	0.06	0.93	0.50	0.49	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.44	0.55	0.06	0.93
$x \rightarrow y/causality$	0.8	0.08	0.92	0.44	0.56	0.06	0.93
$x \rightarrow y/causality$	0.4	0.05	0.94	0.47	0.52	0.04	0.95
$x \rightarrow y/causality$	0.2	0.06	0.93	0.33	0.66	0.06	0.93
$y \rightarrow z/causality$	0.9	0.06	0.93	0.43	0.56	0.05	0.94
$y \rightarrow z/causality$	0.8	0.07	0.93	0.46	0.53	0.04	0.95
$y \rightarrow z/causality$	0.4	0.05	0.94	0.50	0.49	0.07	0.92
$y \rightarrow z/causality$	0.2	0.06	0.93	0.33	0.66	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.2$					
$x \rightarrow y/causality$	0.9	0.06	0.93	0.29	0.70	0.06	0.93
$x \rightarrow y/causality$	0.8	0.05	0.94	0.33	0.66	0.04	0.95
$x \rightarrow y/causality$	0.4	0.06	0.93	0.25	0.74	0.04	0.95
$x \rightarrow y/causality$	0.2	0.05	0.94	0.18	0.81	0.06	0.93
$y \rightarrow z/causality$	0.9	0.06	0.93	0.32	0.67	0.06	0.93
$y \rightarrow z/causality$	0.8	0.07	0.92	0.31	0.68	0.05	0.94
$y \rightarrow z/causality$	0.4	0.07	0.92	0.26	0.73	0.07	0.92
$y \rightarrow z/causality$	0.2	0.05	0.94	0.16	0.84	0.06	0.93
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.2$					
$x \rightarrow y/causality$	0.9	0.05	0.94	0.14	0.85	0.05	0.94
$x \rightarrow y/causality$	0.8	0.05	0.94	0.14	0.85	0.04	0.95
$x \rightarrow y/causality$	0.4	0.06	0.93	0.11	0.88	0.05	0.94
$x \rightarrow y/causality$	0.2	0.07	0.92	0.09	0.90	0.05	0.94
$y \rightarrow z/causality$	0.9	0.08	0.91	0.12	0.88	0.07	0.93
$y \rightarrow z/causality$	0.8	0.07	0.92	0.13	0.86	0.06	0.93
$y \rightarrow z/causality$	0.4	0.06	0.93	0.10	0.89	0.07	0.92
$y \rightarrow z/causality$	0.2	0.05	0.94	0.08	0.91	0.06	0.94

Table 6.10, 6.11 and 6.12 corresponds to stationary series with various complications (drift and trend). The off-diagonal entries θ_{21}, θ_{32} changes from 0.9, 0.8, 0.4 and 0.2 in matrix A of DGP which implies that x is causing y and y is causing z. In Table 6.10, row 1st of panel 1st corresponds to series where θ_{21} *and* θ_{32} = 0.9 at autoregressive coefficient (i.e. 0.8). The results show that PC algorithm using VAR residuals and Modified PC algorithms using modified R residuals and ARMA residuals, the probability of rejection of null of no causality (which can be regarded as power, since in DGP null is not true) is about 6%, 45% and 4% respectively (given column **correct**), and this does not change significantly when off-diagonal entries (θ_{21} *and* θ_{32}) changes from 0.9 to 0.2. The table reveals that the power is best for Modified R residuals while VAR and ARMA residuals perform bad in case of power. But when the auto regressive coefficient goes to zero, the power of Modified R goes down, as evident from the table, when you move down from panel 1st to panel 8th. Similarly, Table 6.11 and 6.12 also display the same results as discussed in Table 6.10. So, the interpretations of these cases are approximately alike as table 6.10.

Finally, we have generated stationary series with different complications and imputed three cross terms θ_{21} θ_{32} and θ_{31} in the DGP given in equation (10). The simulated results are given in Table 6.13, 6.14 and 6.15 (given in Appendix II). The Appendix II Tables reveals that the power is best for Modified PC algorithm using modified R residuals while algorithms using PC algorithm using VAR residuals and Modified PC algorithm using Haugh ARMA residuals present poor performance in power properties. But when the auto regressive coefficient (θ_{11} θ_{22} and θ_{33}) goes from 0.8 to 0.2, the power of Modified PC algorithm (using MR residuals goes down, as evident from Tables 6.13, 6.14 and 6.15, when you move down from panel 1st to panel 8th in each table)

6.3 Conclusion:

It is concluded from the above results that Modified PC algorithm using modified R recursive residuals perform well (with minimum power loss) both in nonstationary and stationary time series with different specifications (drift and trend). But this procedure also loss power when the auto regressive coefficient value approach to 0.2 (decrease the diagonal values from 0.8 to 0.2). PC algorithm using VAR residuals and Modified PC using Haugh-ARMA residuals present very poor performance (low power). It is also important to discuss that Modified PC algorithm using Haugh-ARMA residuals comparatively perform better than PC using VAR residuals in case when the generated series have both drift and trend. The power of these procedures in different specification is given as follows:

When the generated series are without drift and trend then power sequences will be

- *Power of Modified PC/MR residuals > Power of PC/VAR residuals > Power of Modified PC/Haugh residuals*

When the generated series have only drift then power sequences will be

- *Power of Modified PC/MR residuals > Power of PC/VAR residuals > Power of Modified PC/Haugh residuals*

When the generated series have both drift and trend then power sequences will be

- *Power of Modified PC/MR residuals > Power of Modified PC/Haugh residuals > Power of PC/VAR residuals*

CHAPTER 7

DISCUSSION OF RESULTS

7.1 Summary

Causality Modelling is a central problem in all social sciences and primary question facing researcher is to find the casual direction whether X is causing Y or vice versa. It is relatively easy job in case of Natural sciences where controlled experiment can be run. But in case of social science, this central question has no reliable answer. There have been several approaches to test causality, but there are serious theoretical and empirical weaknesses attached to each approach. In this study we have analyzed the performance of graph theoretic approach causal search algorithms PC algorithm (developed by Peter Spirtes, Clark). Theoretically PC causal search algorithms looks sound, but how it performs empirically, especially when we have confounding variable in the data generating process. The literature carries no answer to this question. So current study investigated the size and power properties as well as probability of spurious causation of PC algorithms. As it is not known, to what extent the procedure is capable of uncovering the true causal relationship and how good it is to differentiate between genuine and spurious causal assumption. The PC algorithm of graph theoretic approach was conceived for cross sectional data and extended for Time Series by Swanson, Granger (1997). PC algorithm of Graph theoretic approach can be considered as a preferred approach for testing causality in time series, because the recent development in graphical models and logic of causality show potential for alleviating difficulties of causal modeling (Pearl, 1998). But this approach also requires some modifications, so that causality can be determined with high accuracy. This study extends to modifies/re-designed PC algorithms (with different measure of correlation) by using residuals extracts from univariate methods such as Haugh (1976) ARMA residuals and Rehman and Malik (2014) - Modified R recursive residuals, so that new approach (Modified PC algorithms) could become valid for finding the true causal ordering in time series. Moreover, this study also aims to investigate size and power properties of Modified PC algorithms to find out how good it is to differentiate between spurious and genuine causal relationship. The size and power properties for these procedures are calculated using Monte Carlo simulations and optimal causal algorithm is recommended. The application on real data is also part of study, then the optimal causality test algorithms is applied to real data analysis to find the causal determinants of inflation. The current study also find answer which enable researchers to gain great confidence in the causality modeling.

7.2 Conclusion

We are now in a position to draw some conclusions from our Monte Carlo Simulations results. It is evident that powers of econometric procedures are comparable if the size remain same. But such causal search algorithms involve multiple testing due to which size cannot be controlled. While comparing the tests, the process starts by finding out the critical values with fixed size, say 5%. Therefore 5% critical values for the entire procedures cannot be calculated. Instead, we can measure size distortion which is the difference between nominal and actual size of entire testing procedure; and the test with minimum size distortion and highest power would be the optimal test.

Table 7.1: Summary of Size Distortion

	Non-Stationary Series (Size Distortion)			**Stationary Series** (Size Distortion)		
	No Drift Trend	Drift only	Drift Trend	No Drift Trend	Drift only	Drift Trend
PC/VAR	2.5%	2.5%	2%	2%	2.5%	2.5%
M.PC/MR	44%	64%	63%	3%	3%	3%
M.PC/Haugh	4%	5%	46%	4%	5%	5%

In Table 7.1, using non-stationary series without deterministic part we have seen that PC algorithm using VAR residuals (PC/VAR) and Modified PC algorithms using Haugh ARMA residuals (M.PC/Haugh) are closest to nominal size of 5% with size distortion of 2.5% and 4% respectively, while Modified PC algorithm using modified R residuals (M.PC/MR) badly suffers in size distortion, on average 44% as shown in column 1^{st}. In case of non-stationary series with drift only, approximately the same results are displayed as previous scenario. But when we generated series are non-stationary having deterministic part, we have noticed that size of (PC/VAR) is again closest to nominal size of 5% while (M.PC/Haugh) and (M.PC/MR) badly suffers in size problem as evident from column 3^{rd} of Table 7.1. On the other hand, when residuals series (i.e. VAR, MR and Haugh) obtained from stationary series with no deterministic part are treated as original variables in PC and Modified PCs algorithms, we have noticed that (PC/VAR), (M.PC/MR) and (M.PC/Haugh) with minimum size distortion of 2%, 3% and 4% respectively as shown in column 4^{th}. But in case of (M.PC/MR), it is evident that when the series is stationary with root close to unity i.e. (0.8) the probability of size distortion is about 25% (discussed in chapter 4), but when the autoregressive roots are close to zero i.e. 0.2 then the size distortion decreases to 3% as shown in column 4^{th}. The same results are displayed for stationary series (no drift, with trend) and (with drift, with trend). Finally, we say that when the series is stationary then Modified PC algorithm using modified R recursive residuals (M.PC/MR) comparatively perform well as PC algorithm using VAR residuals.

Table 7.2: Probability of Spurious Causation in case of confounding Variable

	Stationary Series		
	No Drift Trend	Drift only	Drift Trend
PC/VAR	0%	0%	0%
MPC/MR	0%	0%	0%
MPC/Haugh	0%	0%	0%

We have also noticed that PC algorithm and Modified PCs algorithms do not show any spurious causation, when we include confounding variable in the data generating process. It implies that all approaches work well when we have a common cause variable in the DGP.

Table 7.3: Summary of Power

	Non-Stationary Series (Power)			**Stationary Series with smaller root** (Power)		
	No Drift Trend	Drift only	Drift Trend	No Drift Trend	Drift only	Drift Trend
PC/VAR	7%	8%	8%	6.5%	6%	7%
M.PC/MR	48%	64%	62%	62%	74%	76%
M.PC/Haugh	6%	9%	57%	12.5%	14%	70%

The main finding of power analysis is that in both nonstationary and stationary time series, Modified PC algorithm using modified R recursive residuals perform very well as it shows minimum power loss as compare to other two procedures. We have noticed very low power of PC algorithms using VAR residuals and Modified PC using Haugh-ARMA residuals. More importantly, Modified PC using modified R residuals also loss power when the auto regressive roots approach to 0.2 (*detailed discussed in chapter 6*). This show that power of test depends upon relationship between x and y.

Modified PC using Haugh-ARMA residuals also comparatively perform better when the series are generated with deterministic part.

In our Monte Carlo simulation study, we had found Modified PC algorithms using Modified R recursive residuals successful in identifying the correct causal structures with reasonable reliability. So, we selected and consider Modified PC algorithm using modified R recursive residuals the best because it performed better than other measure of correlation used in causal search algorithms.

7.3 Recommendation:

From the outcome of this study we can easily deduce that:

- Modified PC algorithm (using modified R recursive residuals) were successful in identifying the correct causal links presence in both, DGP and final simulated results with high reliability when the auto regressive coefficient is near to unity. But when the auto regressive coefficient in the data generating process tends towards zero, this procedure fails to perform its function to identify correct causal directions. The others two algorithms, PC using VAR residuals and Modified PC algorithm using Haugh ARMA residuals present very poorly performance in finding the correct causal structure.
- Furthermore, it may be noted that the causal links present in DGP is not removed in the simulated results of Modified PC algorithm using modified residuals. This means that the omission errors proportion in this procedure are very low as compared to the other two causal search algorithms using VAR residuals and Haugh ARMA residuals.

CHAPTER 8

APPLICATION OF GRAPH THEORETIC APPROACH TO MONETARY POLICY

8.1 Introduction

After a long phase of evolution, the primary objective of contemporary monetary policy is set to control inflation and sometimes mandated to look for unemployment and economic growth. Today, controlling inflation is central mandate of monetary policy and interest rate has emerged as the major tool. Though the early economists didn't believe in the real impacts of inflation, it appeared as the important worldwide macroeconomic problem. However, what are the effective tools to control inflation; the controversy has remained throughout the history. Monetary policy today is synonymous of interest rate policy and the central banks are using interest rate as a tool to control certain variables including inflation. However, how do interest rate effect inflation, the theoretical literature and practice differ in this regard. The monetary policy practices throughout globe assume that inflation can be decreased by increasing interest rate. On the other hand, the theoretical literature and data-based evidences differ remarkably. There are many evidences showing that increased interest rate adds to inflation, owing to the cost side channels of monetary transmission mechanism, hereinafter referred as MTM (Gibson, 1923; Barth, & Ramey, 2001; Rehman, 2015). This means, if the evidences of cost side effects of monetary policy are realistic, the monetary policy could be counterproductive. Given the growing literature on cost side effects of monetary policy, these effects must be at least part of monetary policy modeling.

The history of literature on the costs side effects of monetary policy is as old as the monetary economics itself. Thomas Tooke (1774-1858), who is considered as the father of monetary economics, is the pioneer of the idea of cost channel of MTM and he argues that there should be positive relationship between inflation and interest rate. Gibson (1923) found strong evidence in support of the idea conceived by Tooke and found positive association between interest rate and inflation in the UK for time period of 200 years. This finding was referred by Keynes as ‘One of most established fact in the whole field of Quantitative economics. There are many other popular studies favoring the exists of cost side effects of monetary policy.

Regardless of the strength of evidences for or against the cost channel, the channel deserves serious attention because if the cost channel was functioning, the use of monetary policy could be counterproductive. That’s why Wright Patman says “the use of interest rate to control inflation is like throwing gasoline on fire. However, despite this strict warning, monetary decision rules used by central banks pay no heed to cost channels.

The theoretical literature informs that there are many causal paths through which interest rate can affect inflation. All these channels have some theoretical support, but which one of these is dominant, there is no consensus on this. Broadly speaking, these channels can be categorized in two streams: The demand side and the cost side channels. The cost side channels predict that rise in interest rate add to inflation. Demand side channel generally agree on that increase in interest rate will reduce inflation. Assuming the demand side effects only, the central banks have designed monetary policy principle such as McCallum and Taylor principle. These rules/principles predict that increasing interest rate can reduce inflation, therefore tight monetary policy has become the most often used anti-inflation measure. However, there are many empirical evidences which contradict to this assumption, Khumalo et al (2017), Abayomi et al (2014) and many authors have found that increase in interest rate leads to increase in inflation. Now, there are many papers which support the existence of demand channels many others supporting the cost side effects of monetary policy, and a third stream indicating no significant relation of interest rate with inflation.

As discussed above, there are many causal paths connecting interest rate and inflation, but most of existing studies focus only on single equation ignoring other parallel channels. This means the methodology is ill-posed. If there are multiple paths connecting the monetary policy action to its objectives, a reasonable model should take all these channels into account to get bias free estimate of the relation. Estimating a relationship that has complex causal paths through a single equation induces both simultaneity bias and missing variables bias.

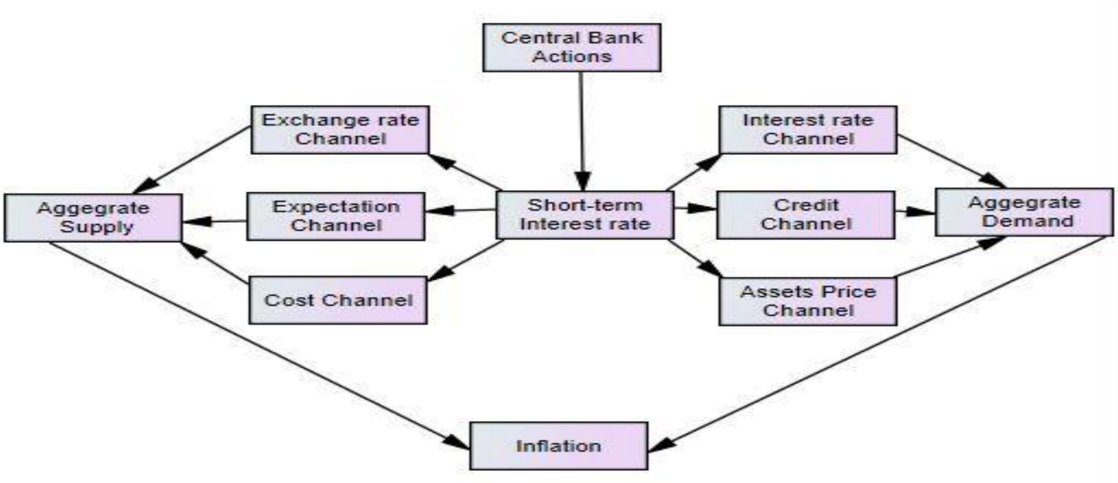
Economic literature usually ignores the presence of multiple causal mechanisms, especially in time series models and offers no satisfactory solution. Spirtes et al (1993) and Pearl (2000) developed Graph theoretic approachassumes that causally ordered data are cross-sectional. Subsequent modification by Swanson and Granger (1997), Hoover (2006), Demiralp and Hoover (2006) and others use it for modeling relationship between time series. In this study we are using the ***Graph-Theoretic approach (PC Algorithms)*** and take all possible theoretical channels that are assumed to link the monetary policy tool and targets. In this approach all complex causal paths are considered simultaneously for investigation of causal relationship. The text in monetary economics reveals that there are at least six structural paths through which the monetary policy might be affecting the inflation.

The aim of this research is to find the causal determinants of inflation using graph theoretic approach

8.2. Monetary Transmission Channels

This section summarizes the causal paths that links monetary policy actions to their targets. As stated earlier, the theoretical literature exhibits that there are many causal paths (monetary transmission channels) which connect the monetary policy actions to monetary policy targets. The popular channels include interest rate channel, credit channel, exchange rate channel, expectation channel, cost channel and assets prices channel These channels can be categorized as demand side channels and supply channels. The demand side channels predict negative relationship between interest rate and inflation, whereas supply side channels predict positive relation between two variables.

In the contemporary monetary decision making, two rules are used most frequently i.e. (i) Tylor rule and (ii) McCullum rule. Both of these are based on assumptions of demand side channel. This is in fact an assumption which has been challenged by all those who find evidences of cost side effects of monetary transmission. But anyhow, the studies on demand side and costs side channels both are questionable because they don't take into account the complex causal mechanism involving all possible channels. A better strategy to study the effects of monetary policy would be to take into accounts all possible channels and causal paths. Therefore, it becomes important to list all popular theoretical channels and discuss their chain of causality. All these causal paths can be tested using Graph theoretic approach to find out which theory and/or channel is supported by data. Therefore, this section lists the monetary transmission channels and the causal paths underlying these channels. These channels and the theory behind their working are mentioned in the following figure.



We will briefly discuss these channels. It is also important to mention that most of channels listed in literature start by change in money supply followed by change in interest rate and ending at inflation, because historically money supply was used as tool of monetary policy. But know a day’s interest rate has become the primary tool of monetary policy. So, in this study to discuss each channel interest rate will be the starting point.

Interest rate Channel:

Interest rate channel is the key MTM in the basic IS-LM model. The views of Keynesian IS-LM MTM state that contractionary monetary policy positively affects real interest rate which leads to decrease investment expenditure and aggregate demand, which finally affect prices in downward direction.

$$i \uparrow \rightarrow i_r \uparrow \rightarrow \mathrm{I} \downarrow \rightarrow Y \downarrow \rightarrow \mathrm{P} \downarrow$$

This increase in real interest rate decrease business fixed investment and consumer durable spending which ultimately affect the aggregate output in downward direction. This will decrease the equilibrium price level Mishkin (1996).

Exchange rate Channel:

The exchange rate is one of the important macroeconomic variables which significantly affect Prices. Taylor (1993) argues that exchange rate plays crucial role in monetary transmission mechanism channels, it explains that how exchange rate channel affects the domestic economy through monetary transmission mechanism. Waseem (2007)[13], Mishkin (1996) argue that rise in nominal interest rate increases real interest rate which attracts foreign investment, thereby increasing foreign exchange reserves (FER). This capital inflow appreciates the local currency, which means that in the foreign market the domestic product will be expensive, thereby decreasing demand for domestic product in the foreign market due to high prices. As a result, the net export (NX) will start decreasing and this will affect aggregate demand and then Prices in negative direction.

$$i \uparrow \rightarrow \ i_r \uparrow \rightarrow \ Capital\ inflow\ \uparrow \rightarrow FER \uparrow \rightarrow ER\ \downarrow \rightarrow \mathrm{NX}\ \downarrow \rightarrow\ Y \downarrow\ \rightarrow \mathrm{P} \downarrow$$

Cost channel:

According to the cost channel of monetary transmission mechanism, when the concerned authority increases money supply as a result short term interest rate also increases, which leads to increase in the cost of working capital. Subsequently, production cost of firm will rise thereby hit inflation/process in upward direction.

[13] Waseem Shahid Malik (2007) PhD dissertation “Three Easy on Monetary Policy in Pakistan”

$$i \uparrow \rightarrow cost\ of\ working\ capital\ \uparrow (\ current\ assets - current\ liabilities)$$

$$\rightarrow Cost\ Of\ Production \uparrow \rightarrow P \uparrow$$

Assets Price Channel:

This channel explore that how monetary policy affect equity prices or stock prices which further affect investment and consumption. There are two channels linking stock prices with investment and consumption (i) Tobin's q Theory of investment and (ii) wealth effect on consumption. According to Tobin's q theory, monetary policy affects the economy through valuation of stock prices (Mishkin 1996). Tobin presented q as market value of business divided by replacement cost of capital. If q is high, new capital is cheap relative to market value of business firm, thereby increase investment expenditure. Question arise that how monetary policy affect stock prices or equity prices? This can be explained with the help of the following channel.

$$i \downarrow \rightarrow SP \uparrow \rightarrow q \uparrow \rightarrow I \uparrow \rightarrow Y \uparrow \mathrm{P} \uparrow$$

This show that increase in money supply reduce the interest rate which make the bonds unattractive, thereby affecting the stock prices in upward direction. It is known that rise in stock prices will lead to rise in Tobin's q which will increase investment spending and finally increase aggregate demand**.** Other than Tobin's q investment channel there exist an alternative channel called wealth effect channel of monetary transmission mechanism. Modigliani (1971) argue that stock is the main component of wealth. Rise in stock value lead to rise in wealth which increase the current consumption and finally increase aggregate demand.

$$i \downarrow \rightarrow SP \uparrow \rightarrow wealth \uparrow \rightarrow consumption \uparrow \rightarrow Y \uparrow \rightarrow \mathrm{P} \uparrow$$

Credit Channel:

Two main channels of MTM that arise as a result of information problems in credit channel (Mishkin 1996).

(i) Bank lending channel:

When central bank adopts expansionary monetary policy, it increases bank reserves and deposits which will, in return, increases the quantity of bank loans available. This will cause to boost investment and spending to rise and affect Prices.

(ii) Balance sheet channel:

Monetary policy can affect firm's balance sheet in several ways. In case of expansionary monetary policy, cost of capital will decrease. Subsequently, equity prices will increase due to rise in net worth of firm's value. It will lead to increase in investment and then aggregate demand and finally affect inflation in upward direction because of decrease in adverse selection and moral hazard problems.

8.3 Literature Review

There has been a long discussion and debates, without any unanimity, that which factors responsible for inflation. Researcher have carried out a lot of empirical research to check the factors affecting inflation and suggested self-selected models.

Mehwish et al (2017) Abayomi et al (2014) Khumalo et al (2017) Asgharpur et al (2007) have written papers that explore interest rate channel in country Pakistan, Nigeria, Switzerland and for 40 Islamic countries respectively. Khumalo et al (2017), Abayomi et al (2014) and Asgharpur et al (2007) have found evidences of positive relationship between interest rates and inflation but Mehwish et al (2017) has found that there is unidirectional causality running from interest rate to inflation with negative sign in case of Pakistan.

Aleem et al (2007), Domaç (1999) and Wulandari (2012) have written papers that explore Credit channel in country Pakistan, Albania and Indonesia respectively. Aleem et al (2007), Domaç (1999) have found evidences that the most important factors of inflation are import prices and private sector credit. The results of the Granger Causality tests show that unidirectional causality run from credit to government to inflation. But Wulandari (2012) has found no evidences.

Dornbusch (1987) for the first time investigated the relationship between exchange rate and inflation. Hunt et al (1998), Rajan et al (2015), Monfared et al (2017) and Choudhri et al (2002) have written papers that explore Exchange rate channel of MTM in small open economy India, Iran and Pakistan respectively. Hunt et al (1998), Rajan et al (2015), Monfared et al (2017) have found evidences that increase in foreign exchange rates add to inflation. But in case of Pakistan Choudhri et al (2002) have found no evidences and conclude that there is no association between rupee devaluations and inflation in Pakistan.

Most of researcher have experienced to answer this debatable issue relating to causal determinants of inflation using Granger causality test and other methods. Granger causality test check causal inferences between two variables. But there may also be a chance that third variable cause both A and B. But Granger causality does not take account of this. So, we extend the previous work on finding the causal determinants of inflation using Graph theoretic approach. The graph-theoretic approaches overcome many problems and good over the more commonly used Granger causality and Sims Causality tests. This effort to investigate the causal determinants of inflation by using conventional techniques in recently developed Graph theoretic method will be a valuable contribution to the literature of inflation.

8.4 Methodology:

This section covers all the required techniques applied in current study. We have referred residuals series (i.e. extracted from univariate model - Modified R recursive residuals) to PC algorithms. Sequentially we will apply the following steps to find the causal determinants of inflation.

- **Modified PC used Modified R recursive residuals**
  - Application of AR model and extract recursive residuals series.
  - Application of Modified PC algorithms in Graph-Theoretic method, treating modified R residuals as original variables in Tetrad software to find the causal determinants of inflation.

Detailed explanation is given in Methodology chapter 3.

Data

Quarterly data from the time period 1985 to 2017 has been used to find the causal determinants of inflation in case of Pakistan. The data has been collected from different sources i.e. State Bank of Pakistan (SBP) and International Financial Statistics (IFS).

8.5 Results and Discussion:

PC Algorithms using Modified R residuals:

Sprites et al (2000) argues that the algorithms do not work directly for time dependent data, so we use recursive residuals extracted from Modified R to remove time dependency as suggested by Swanson and Granger (1997). These residuals are then treated as original variables in TETRAD 4.9.1 which contain PC algorithms. The results of Modified PC algorithm of Graph theoretic approachare given in figure 8.1, 8.2 and 8.3

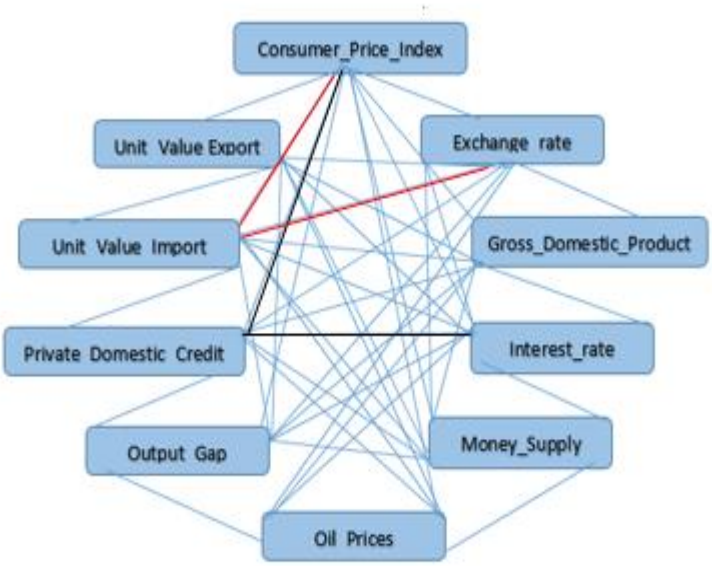


Figure 8.1: Undirected Edge Graph

Figure 8.1 called Skeleton graph, show us the general structure of graph in which all variables are connected through undirected links. The path diagram/Skeleton given in figure 1 covers all theoretical channels of monetary transmission mechanism mentioned in section II, as well as other non-monetary determinants of inflation. For simplicity we have visualized only two of channels i.e. credit channel and exchange rate channel. The red and black lines in figure 1 show us exchange rate channel and credit channel respectively. The algorithm will eliminate some of the links in the skeleton and will orient the remaining links which are significant. If the causal path denoted by red and black lines are significant, this will imply that these channels work to transmit the impact of monetary policy to monetary target.

The algorithms test unconditional correlation between any two variables. If they are not unconditionally correlated, then eliminate that connections for all possible pair. The algorithms also test correlation between each two variables conditional on a third variable. If each pair of variables are conditionally uncorrelated then again eliminate their connections. If the pair is correlated conditional on the third variable, the members of pair are unshielded colliders on that path, and arrows from the pair of variables are oriented toward the third variable. These algorithms also orient arrows based on screening relationship.

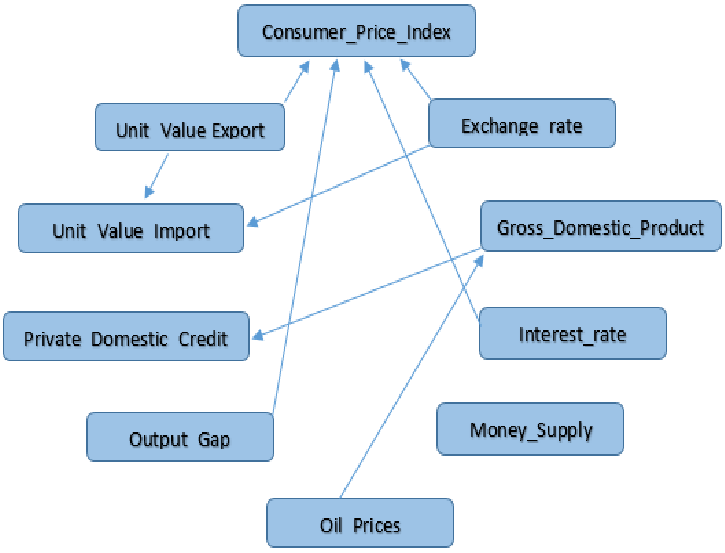


Figure 8.2: Modified PC Algorithm using Modified R recursive Residuals

The results indicated by Modified PC algorithm using modified R recursive residuals in Figure 8.2 show us that output gap cause inflation which support the results of Bolt *et al* (1998) found for European Union. Dominant stream of literature supports this result which include (Watanabe, T. 1997; Claus *et al*, I. 2000; Clark *et al. 1996;* Gerlach, S. *et al* 1999*).* Similarly, the causality running from exchange rate to consumer price index and unit value of import is in line with the study of Chung *et al* (2011) and Parker (2014) mentioned that the pass -through from exchange rate to consumer price index is modest and pass through from exchange rate to import prices is high and quick. This is not surprising because since import prices largely reflect the domestic-currency cost of the good from the foreign supplier, which can be expected to vary in line with the exchange rate. So, the current study supports the results as there exist causal path between exchange rate and imports while the causal path between consumer prices and import is removed. Pakistan is consumer economy having bulk of imports and the prices of import do affect the CPI as well.

The results also indicate that no causality running from money supply to consumer price index. According to Werner Quantity theory of credit, the banks can create money and inflation can have relationship with the money created by banks. The money that is managed by central bank is only a small part of total money and there is no good approximation of total money. Therefore, many monetary models found insignificant coefficient of money supply. Modern Monetary Theory also explain that if money supply is used in the new economic activities, it will not create inflation. This finding is compatible with that of the Asari *et al* (2011) and Saini (1982) for six Asian countries. Saini empirically found that the money stock growth is not the primary source of inflation in these six Asian countries including Pakistan.

The results indicate causality running from interest rate to consumer price index. The earlier findings of Mehwish et al (2017) Abayomi et al (2014) Khumalo et al (2017) Asgharpur et al (2007) that interest rate is significant determinant of inflation is not rejected. This indicate about effectiveness of Monetary policy. If interest rate affect prices in negative direction then monetary policy will be productive, in case of positive direction monetary policy will be counterproductive. We have found the significant determinants of inflation and interest rate is among there. Very important question is unanswered – in which direction interest rate affect inflation? The algorithm is not capable of informing about the dimension of relationship. To solve the issue, we have estimated ARDL model by including all the significant determinants found through graph theoretic approach. The results are given in Table 2

Table 8.1: ARDL Estimated Results

Variables	Coefficient	SE	t-value	t-prob
Consumer price index_1	0.96994	0.01235	78.5	0.0000
Constant	0.0884466	0.03242	2.73	0.0076
Exchange rate	0.00151606	0.00034	4.52	0.0000
Exchange rate_1	-0.0009509	0.00036	-2.67	0.0088
Unit value of export	0.0010847	0.00043	2.5	0.0140
Unit value of export_1	-0.0011979	0.00041	-2.94	0.0041
Interest rate	0.0041354	0.00104	3.99	0.0001
Interest rate_1	-0.0034617	0.00105	-3.3	0.0013
Output Gap	-0.00025197	0.00013	-1.87	0.0643
Output Gap_1	0.00018143	0.00014	1.34	0.1837

Author Calculation

The ARDL model has many advantages, in particular it can be used to find appropriate model when there are multiple theories. After calculating ARDL, we find the static long run solution (SLRS) through the procedure given in Deadman (1997).

Table 8.2: Solved static long-run equation

Variables	Coefficient	SE	t-value	t-prob
Constant	2.94312	0.2327	12.6	0.0000
Exchange rate	0.0188051	0.004124	4.56	0.0000
Unit value of export	-0.00376574	0.004149	-0.908	0.3662
Interest rate	0.0224189	0.01326	1.69	0.0924
Output Gap	-0.00234737	0.006354	-0.369	0.7126

Author Calculation

After estimating static long run coefficients, we found the coefficient of interest rate positive which means that by increasing interest rate the inflation is also going up and causality run from interest rate to inflation. This is matching with the cost side inflation theories and contradicting with the assumption that by increasing the interest rate going inflation down. This also means that current monetary policy counterproductive.

8.6. Conclusion and Recommendations

There has been a long discussion and debates. The classical assumption is that increase in interest rate will decrease inflation but there are many evidences showing that increased interest rate adds to inflation. So, the theoretical literature and data-based evidences differ remarkably. The issue is that the theoretical literature exhibits that there are many causal paths which connect the monetary policy actions to monetary policy targets but previous studies did not take into account all these paths. Therefore, it becomes important to list all popular theoretical channels and discuss their chain of causality. The current study employed Graph theoretic approach and explored significant determinants of inflation. The paper estimates the static long run solution to find the sign of relationship between two variables. We find that interest rate and inflation are positively associated, which implies the monetary policy to be counterproductive. This implies there is need of serious rethinking about current monetary policy regime.

REFERENCES

Abayomi Onanuga and Olaronke Onanuga (2014). The Impact of Interest Rate Channel of Monetary Policy on Output and Prices in Nigeria: An Unrestricted VAR Approach. Munich Personal RePEc Archive

Alessio M. Nadine C. Doris E. Patrik H. (2011). Causal search in SVAR models. JMLR: Workshop and Conference proceeding 12(2011) pp 95-118

Asari, F. F. A. H., Baharuddin, N. S., Jusoh, N., Mohamad, Z., Shamsudin, N., & Jusoff, K. (2011). A vector error correction model (VECM) approach in explaining the relationship between interest rate and inflation towards exchange rate volatility in Malaysia. *World Applied Sciences Journal*, *12*(3), pp 49-56.

Asghar. Z, Tayyaba. R (2011), "Energy GDP Causal relationship for Pakistan: A Graph Theoretic Approach. *Applied Econometrics and international Development*

Asghar. Z. (2007), Sensitivity and Simulation Analysis of Granger Causality: An Empirical Investigation; PhD Dissertations. Pakistan Institute of Development Economics.

Asgharpur, H., Kohnehshahri, L. A., & Karami, A. (2007). The relationships between interest rates and inflation changes: An analysis of long-term interest rate dynamics in developing countries.

Atiq-ur-Rehman, A. U. R., & Zaman, A. (2009). Impact of model specification decisions on unit root tests.

Barth III, M. J., & Ramey, V. A. (2001). The cost channel of monetary transmission. *NBER macroeconomics annual*, *16*, pp 199-240.

Bessler, D.A & S. Lee (2002) Money and price: US data 1869-1914 (A study with directed graphs). Empirical Economics. 27, pp 427-446.

Bolt, W., & Van Els, P. J. (1998). Output Gap and Inflation in the EU (No. 550). Netherlands Central Bank, Research Department.

Charemza, W. W., & Deadman, D. F. (1997). New directions in econometric practice. *Books*.

Chen, B., & Pearl, J. (2014). *Graphical tools for linear structural equation modeling* (Vol. 432). California univ los Angeles dept of computer science.

Choudhri, E. U., & Khan, M. S. (2002). The exchange rate and consumer prices in Pakistan: is rupee devaluation inflationary? *Pakistan Development Review*, *41*(2), pp 107-120.

Chung, E., Kohler, M., & Lewis, C. (2011). The exchange rate and consumer prices. *RBA Bulletin, September*, pp 9-16.

Clark, P., Laxton, D., & Rose, D. (1996). Asymmetry in the US output-inflation nexus. *Staff Papers*, *43*(1), pp 216-251.

Claus, I. (2000). Is the output gap a useful indicator of inflation? *Reserve Bank of New Zealand Discussion Paper No. DP2000/05*.

Cooley, T. F., & LeRoy, S. F. (1985). Atheoretical macroeconometrics: a critique. *Journal of Monetary Economics*, *16*(3), pp 283-308.

Cooper, G. (1999). An overview of the representation and discovery of causal relationships using Bayesian networks. *Computation, causation, and discovery*, pp 4-62.

Cornet, B., & Tulkens, H. (1989). *Contributions to operations research and economics: the twentieth anniversary of CORE*.

David et al (2019) Control of confounding and report of results in causal inferences studies. Annals ATS Vol. 16 Number 1|January 2019.

Demiralp, S., & Hoover, K. D. (2003). Searching for the causal structure of a vector autoregression. *Oxford Bulletin of Economics and statistics*, *65*, pp 745-767.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), pp 427-431.

Domaç, Carlos E (1999). The main Determinants of inflation in Albania. https://doi.org/10.1596/1813-9450-1930.

Dornbusch, R. (1987). Exchange rates and prices. *The American Economic Review*, pp 93-106.

Erdem, O., Ceyhan, E., & Varli, Y. (2014). A new correlation coefficient for bivariate time-series data. *Physica A: Statistical Mechanics and its Applications*, *414*, pp 274-284.

Freedman, D., Pisani, R., & Purves, R. (2007). Statistics (international student edition). *Pisani, R. Purves, 4th edn. WW Norton & Company, New York*.

Gerlach, S., & Smets, F. (1999). Output gaps and monetary policy in the EMU area. *European Economic Review*, *43*(4-6), pp 801-812.

Geweke, J., Meese, R., & Dent, W. (1983). Comparing alternative tests of causality in temporal systems: Analytic results and experimental evidence. *Journal of Econometrics*, *21*(2), pp 161-194.

Gibson, A. H. (1923). The future course of highclass investment values. *Banker's Magazine (London)*, pp *115*, 15-34.

Glymour, C., Zhang, K., & Spirtes, P. (2019). Review of causal discovery methods based on graphical models. *Frontiers in Genetics*, *10*.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424-438.

Granger, C. W., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of econometrics*, *2*(2), pp 111-120.

Guilkey, D. K., & Salemi, M. K. (1982). Small sample properties of three tests for Granger-causal ordering in a bivariate stochastic system. *The Review of Economics and Statistics*, pp 668-680.

Haavelmo, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica, Journal of the Econometric Society*, pp 1-12.

Haugh, L. D. (1976). Checking the independence of two covariance-stationary time series: a univariate residual cross-correlation approach. *Journal of the American Statistical Association*, *71*(354), pp 378-385.

Hendry, D. F. (1995). *Dynamic econometrics*. Oxford University Press on Demand.

Hoover, D. K. (2006). Causality in economics and econometrics.

Hoover, K. D. (2001). *Causality in macroeconomics*. Cambridge University Press.

Hoover, K. D. (2005). Automatic inference of the contemporaneous causal order of a system of equations. *Econometric Theory*, *21*(1), pp 69-77.

Hoover, K.D; Demiralp, S, and Perez, S.J. (2008),"Empirical Identification of the Vector Autoregression: The Causes and Effects of U.S. M2"presents at the Conference in Honour of David Hendry at Oxford University, 23-25August 2007

Hume, D. (1752). Of money. *David Hume: Writings on Economics*.

Hunt, B., Conway, P., Drew, A., & Scott, A. M. (1998). Exchange rate effects and inflation targeting in a small open economy: a stochastic analysis using FPS. *Reserve Bank of New Zealand Discussion Paper*, (G99/4).

James, L., Mulaik, S., & Brett, J. M. (1982). Causal analysis: Assumptions, models, and data.

Khumalo, L. C., Mutambara, E., & Assensoh-Kodua, A. (2017). Relationship between inflation and interest rates in Swaziland revisited.

Koopmans, T. (1950). When Is an Equation System Complete for Statistical Purposes?, Chapter 17 in TC Koopmans (ed.) Statistical Inference in Dynamic Economic Models New York.

Kuersteiner, G. M. (2010). Granger-sims causality. In *Macroeconometrics and time series analysis* (pp. 119-134). Palgrave Macmillan, London.

Landeiro, V., & Culotta, A. (2017, May). Controlling for unobserved confounds in classification using correlational constraints. In Eleventh International AAAI Conference on Web and Social Media.

Le, T., Hoang, T., Li, J., Liu, L., Liu, H., & Hu, S. (2019). A fast PC algorithm for high dimensional causal discovery with multi-core PCs. IEEE/ACM transactions on computational biology and bioinformatics.

Leamer, E. E. (1985, January). Vector autoregressions for causal inference?. In *Carnegie-rochester conference series on Public Policy* (Vol. 22, pp. 255-304). North-Holland.

Li, Yarui ; Woodard, Joshua D. ; Leatham, David J. ; Marchant, Mary A. ; Bosch, Darrell J. (2013). "Causality among Foreign Direct Investment and Economic Growth: A Directed Acyclic Graph Approach" Journal of Agricultural and Applied Economics, 45, 4 pp. 617-637

Lin, J. L. (2008). Notes on testing causality. Institute of Economics, Academia Sinica, Department of Economics. National Chengchi University, pp 3-22.

Lucas Jr, R. E. (1976, January). Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy* (Vol. 1, pp. 19-46). North-Holland.

Mehwish , M. Naeem, Sami U, Salim U. K (2017). “Correlation and Causality between Inflation and Selected Macroeconomic Variables: Empirical Evidence from Pakistan” iBusiness, 2017, 9, pp 149-16

Mishkin, F. S. (1996). *The channels of monetary transmission: lessons for monetary policy* (No. w5464). National Bureau of Economic Research.

Modigliani, F. (1971). Monetary policy and consumption. *Consumer spending and monetary policy: the linkages*, pp 9-84.

Moneta, A. (2003) Graphical Models for Structural Vector Autoregressions. Typescript, S. Anna School of Advance Studies, Italy.

Monfared, S. S., & Akın, F. (2017). The Relationship Between Exchange Rates and Inflation: The Case of Iran. *European Journal of Sustainable Development*, *6*(4), pp 329-340.

Nazmus Sadat Khan, 2016. "In search of causality between debt and growth: a graph theoretic approach," Economics Bulletin, AccessEcon, vol. 36(2), pages pp 677-687.

Neyman, J. S. (1923). On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9.(Tlanslated and edited by DM Dabrowska and TP Speed, Statistical Science (1990), 5, 465-480). *Annals of Agricultural Sciences*, *10*, pp 1-51.

Nodelman, U., Allen, C., & Perry, J. (2003). Stanford encyclopedia of philosophy.

Parker, M. (2014). Exchange rate movements and consumer prices: some perspectives. *Reserve Bank of New Zealand Bulletin*, *77*(1), pp 31-41.

Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods & Research*, *27*(2), pp 226-284.

Pearl, J. (2000). Causality: models, reasoning, and inference. *IIE Transactions*, *34*(6), pp 583-589.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics surveys*, *3*, pp 96-146.

Pearl, J. (2012). *The causal foundations of structural equation modeling* (Vol. 370). California univ los Angeles dept of computer science.

Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London. A*, *185*, pp 71-110.

Perez, S. J., & Siegler, M. V. (2006). Agricultural and monetary shocks before the great depression: A graph-theoretic causal investigation. *Journal of Macroeconomics*, *28*(4), pp 720-736.

Rajan, R. S., & Yanamandra, V. (2015). Impact of Exchange Rate Pass-Through on Inflation in India. In *Managing the Macroeconomy*. Palgrave Macmillan, London. pp. 110-136

Rehman, A. U. (2015). Revival of Legacy of Tooke and Gibson: Implications for Monetary Policy. *Journal of Central Banking Theory and Practice*, *4*(2), pp. 37-58.

Rehman, A. U., & Malik, M. I. (2014). The modified R a robust measure of association for time series. *Electronic Journal of Applied Statistical Analysis*, *7*(1), pp. 1-13.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688.

Saini, K.G., 1982. The monetarist explanation: The experience of six Asian countries". World Development, 10(10): pp. 871-884.

Simon, H. A. (1953). Causal ordering and identifiability. In *Models of Discovery* (pp. 53-80). Springer, Dordrecht [Hood and Koopmans (eds.), Studies in Econometric Methods, New York: John Wiley & Sons, 1953, pp. 49–74]

Simon, H. A. (1954). Spurious correlation: A causal interpretation. *Journal of the American statistical Association*, *49*(267), pp. 467-479.

Sims, C. A. (1972). Money, income, and causality. *The American economic review*, pp. 540-552.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pp. 1-48.

Sims, C. A. (1986). Are forecasting models usable for policy analysis?. *Quarterly Review*, (Win), pp. 2-16.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European economic review*, *36*(5), pp. 975-1000.

Sims, C. A., Goldfeld, S. M., & Sachs, J. D. (1982). Policy analysis with econometric models. *Brookings papers on economic activity*, *1982*(1), pp. 107-164.

Song, X., & Taamouti, A. (2019). A better understanding of granger causality analysis: A big data environment. *Oxford Bulletin of Economics and Statistics*, *81*(4), 911-936.

Spirtes, P., Glymour, C. N., Scheines, R. (1993). Causation, prediction, and search, *New York, N.Y.: Springer-Verlag.*

Spirtes, P., Glymour, C. N., Scheines, R., Heckerman, D., Meek, C., Cooper, G., & Richardson, T. (2000). *Causation, prediction, and search*. MIT press.

Swanson, N. R., & Granger, C. W. (1997). Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, *92*(437), pp. 357-367.

Taylor, J. B. (1993, December). Discretion versus policy rules in practice. In *Carnegie-Rochester conference series on public policy* (Vol. 39, pp. 195-214). North-Holland.

Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of econometrics*, *66*(1-2), pp. 225-250.

Waseem Shahid Malik (2007) PhD dissertation “Three Easy on Monetary Policy in Pakistan”

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. American psychologist, 54(8), 594.

Wold, H. (1954). Causality and econometrics. *Econometrica: Journal of the Econometric Society*, pp. 162-177.

Wright, S. (1921). Correlation and causation. *Journal of agricultural research*, *20*(7), pp. 557-585.

Wulandari, R. (2012). Do credit channel and interest rate channel play important role in monetary transmission mechanism in Indonesia?: A structural vector autoregression model. *Procedia-Social and Behavioral Sciences*, *65,* pp. 557-563.

Yang,J; Guo,H and Zijun,W. (2005), “ International Transmission of Inflation among G-7 Countries: A Data Determined VAR Analysis”, Working Paper 2004-028

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between Time-Series?--a study in sampling and the nature of time-series. *Journal of the royal statistical society*, *89*(1), pp. 1-63.

Zahid. A. Nighat. J. (2010). “Causal order between money, income and prices through graph theoretic approach for Pakistan”. Economics Bulletin, vol. 30 no3 pp. 2515-2523

Zaman, A. (2009). Causal Relations via Econometrics. *International Econometric Review*, *2*(1), pp 36-56.

Zellner, A. (1979, January). Causality and econometrics. In *Carnegie-Rochester conference series on public policy* (Vol. 10, pp. 9-54). North-Holland.

APPENDIX I

Modified R

Correlation coefficient R was developed for handling cross sectional data but currently most people have applied to time series data as well. There is difference between the nature of both time series and cross-sectional data. One of important feature of time series data is serial dependence of observations which results the spurious correlation between two-time series. Using conventional measure of correlation between two-time series exaggerate R which is not valid and reliable measure of association. This is valid only in case of IID series, but mostly time series are subject to non-stationary in which probability of spurious correlation is very high. Later Granger have also proved the probability of spurious correlation in stationary time series. So, R is not a reliable and trustful measure of association for time series.

A new statistic is introduced to measure the correlation between time series which also work well in case when the series are not IID. This new statistic is the correlation between recursive forecast error of AR model fitted to both series. If we analyze the association between recursive forecast residuals it is expected to give more valid measure of correlation between time series. This new statistic is given the name Modified R, denoted by MR. Steps involved in MR is given as under:

For three-time series of length T, let T1 be number smaller than T i.e. T1 < T series

$x = \{x_1, x_2 \ldots x_T\}$, $y = \{y_1, y_2 \ldots\ y_T\}$ and, $z = \{z_1, z_2 \ldots\ z_T\}$

1) For T1 < T, estimate the autoregressive model $x_t = \hat{\alpha}_{T1} + \hat{d}_{T1} x_{t-1} + e_t$ using OLS.
2) Compute $\hat{x}_{T1+1} = \hat{a}_{T1} + \hat{b}_{T1} x_{T1}$

3) Compute $e_{T1+1} = x_{T1} - \hat{x}_{T1+1}$

4) Repeat the process for T1+1, T1+2, T-1 to compute $e_{T+1}, e_{T+2}, e_{T+3} \dots\ e_T$

5) Repeat steps 1-4 for the series $y = \{y_1, y_2 \dots\ y_T\}$ and $z = \{z_1, z_2 \dots\ z_T\}$ to compute forecast error $u_{T+1}, u_{T+2,} u_{T+3} \dots u_T$ and $u^*{}_{T+1},\ u^*{}_{T+2},\ u^*{}_{T+3} \dots\ u^*{}_T$ respectively.

Finally, one can use the computed recursive forecast residuals of three-time series as variables while ignoring the original series x_t, y_t and z_t. Thus, modified R is based on recursive forecast residuals from autoregressive model with generalized form of linear trend thus is capable of producing desired results.

Haugh (1976) test

Haugh 1976 for the first time introduced a test to check the independence of two stationary autoregressive moving average (ARMA) series based on the residual cross correlation. Haugh test is based on two steps procedure.

1. First step is to fit ARMA models to each of the series and find residuals series of all variables.
2. Second step is to find the cross correlation between two resulting residual series.

Let r_{uv} (j) be the residual cross correlation at lag j:

$$r_{uv}(j) = \frac{\sum_{t=j+1}^{n} u_t v_{t-j}}{\left(\sum_{t=1}^{n} u_t^2 v_{t-j} \sum_{t=1}^{n} v_t^2\right)^{1/2}} \qquad \text{For } |j| \le n-1,$$

“Where u_t and v_t represent residual series of variable x and y respectively. n is the length of each time series from t = 1,2,3……. n.

The asymptotic distribution of a fixed number of residual cross-correlation is given by Haugh. He considered a portmanteau statistic which is based on the first M residual cross correlations, where M≤n-1 is a fixed integer more precisely, he introduced the statistics

$$\mathbf{S_M = n \sum_{J=-M}^{M} r^2{}_{uv(j)}}$$

Its asymptotic distribution is chi- square under the null hypothesis of independency, and the hypothesis is rejected for the large value of the test statistics".

APPENDIX II

Table 1

		VAR		Modified R		Haugh	
		Correct.	Omitted	Correct.	Omitted	Correct.	Omitted
Without Drift and Trend							
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \to y/causaliy$	0.9	0.072	0.928	0.328	0.672	0.031	0.969
$x \to y/causaliy$	0.8	0.066	0.934	0.297	0.703	0.026	0.974
$x \to y/causaliy$	0.4	0.068	0.932	0.29	0.71	0.03	0.97
$x \to y/causaliy$	0.2	0.062	0.938	0.367	0.633	0.036	0.964
$y \to z/causaliy$	0.9	0.066	0.934	0.719	0.281	0.029	0.971
$y \to z/causaliy$	0.8	0.066	0.934	0.784	0.216	0.029	0.971
$y \to z/causaliy$	0.4	0.073	0.927	0.729	0.271	0.053	0.947
$y \to z/causaliy$	0.2	0.077	0.923	0.491	0.509	0.048	0.952
$x \to z/causaliy$	0.9	0.066	0.934	0.719	0.281	0.029	0.971
$x \to z/causaliy$	0.8	0.066	0.934	0.784	0.216	0.029	0.971
$x \to z/causaliy$	0.4	0.073	0.927	0.729	0.271	0.053	0.947
$x \to z/causaliy$	0.2	0.077	0.923	0.491	0.509	0.048	0.952
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \to y/causaliy$	0.9	0.064	0.936	0.134	0.866	0.035	0.965
$x \to y/causaliy$	0.8	0.068	0.932	0.126	0.874	0.041	0.959
$x \to y/causaliy$	0.4	0.051	0.949	0.325	0.675	0.043	0.957
$x \to y/causaliy$	0.2	0.058	0.942	0.299	0.701	0.063	0.937
$y \to z/causaliy$	0.9	0.074	0.926	0.816	0.184	0.033	0.967
$y \to z/causaliy$	0.8	0.075	0.925	0.765	0.235	0.029	0.971
$y \to z/causaliy$	0.4	0.065	0.935	0.404	0.596	0.06	0.94
$y \to z/causaliy$	0.2	0.057	0.943	0.309	0.691	0.051	0.949
$x \to z/causaliy$	0.9	0.074	0.926	0.816	0.184	0.033	0.967
$x \to z/causaliy$	0.8	0.075	0.925	0.765	0.235	0.029	0.971
$x \to z/causaliy$	0.4	0.065	0.935	0.404	0.596	0.06	0.94
$x \to z/causaliy$	0.2	0.057	0.943	0.309	0.691	0.051	0.949
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$					
$x \to y/causaliy$	0.9	0.064	0.936	0.112	0.888	0.041	0.959
$x \to y/causaliy$	0.8	0.068	0.932	0.138	0.862	0.048	0.952
$x \to y/causaliy$	0.4	0.061	0.939	0.221	0.779	0.042	0.958
$x \to y/causaliy$	0.2	0.069	0.931	0.175	0.825	0.059	0.941

$y \rightarrow z/causaliy$	0.9	0.063	0.937	0.404	0.596	0.031	0.969
$y \rightarrow z/causaliy$	0.8	0.078	0.922	0.385	0.615	0.028	0.972
$y \rightarrow z/causaliy$	0.4	0.067	0.933	0.228	0.772	0.057	0.943
$y \rightarrow z/causaliy$	0.2	0.074	0.926	0.151	0.849	0.054	0.946
$x \rightarrow z/causaliy$	0.9	0.063	0.937	0.404	0.596	0.031	0.969
$x \rightarrow z/causaliy$	0.8	0.078	0.922	0.385	0.615	0.028	0.972
$x \rightarrow z/causaliy$	0.4	0.067	0.933	0.228	0.772	0.057	0.943
$x \rightarrow z/causaliy$	0.2	0.074	0.926	0.151	0.849	0.054	0.946

Table 2

		VAR		**Modified R**		**Haugh**	
		Correct.	Omitted	Correct.	Omitted	Correct.	Omitted
With Drift only							
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \rightarrow y/causaliy$	0.9	0.072	0.928	0.765	0.235	0.035	0.965
$x \rightarrow y/causaliy$	0.8	0.057	0.943	0.761	0.239	0.042	0.958
$x \rightarrow y/causaliy$	0.4	0.06	0.94	0.584	0.416	0.05	0.95
$x \rightarrow y/causaliy$	0.2	0.062	0.938	0.39	0.61	0.072	0.928
$y \rightarrow z/causaliy$	0.9	0.062	0.938	0.886	0.114	0.041	0.959
$y \rightarrow z/causaliy$	0.8	0.066	0.934	0.886	0.114	0.038	0.962
$y \rightarrow z/causaliy$	0.4	0.063	0.937	0.788	0.212	0.062	0.938
$y \rightarrow z/causaliy$	0.2	0.062	0.938	0.575	0.425	0.065	0.935
$x \rightarrow z/causaliy$	0.9	0.062	0.938	0.886	0.114	0.041	0.959
$x \rightarrow z/causaliy$	0.8	0.066	0.934	0.886	0.114	0.038	0.962
$x \rightarrow z/causaliy$	0.4	0.063	0.937	0.788	0.212	0.062	0.938
$x \rightarrow z/causaliy$	0.2	0.062	0.938	0.575	0.425	0.065	0.935
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \rightarrow y/causaliy$	0.9	0.076	0.924	0.266	0.734	0.05	0.95
$x \rightarrow y/causaliy$	0.8	0.066	0.934	0.229	0.771	0.046	0.954
$x \rightarrow y/causaliy$	0.4	0.065	0.935	0.265	0.735	0.055	0.945
$x \rightarrow y/causaliy$	0.2	0.055	0.945	0.324	0.676	0.057	0.943
$y \rightarrow z/causaliy$	0.9	0.067	0.933	0.907	0.093	0.048	0.952
$y \rightarrow z/causaliy$	0.8	0.053	0.947	0.839	0.161	0.043	0.957
$y \rightarrow z/causaliy$	0.4	0.064	0.936	0.488	0.512	0.071	0.929
$y \rightarrow z/causaliy$	0.2	0.076	0.924	0.293	0.707	0.06	0.94
$x \rightarrow z/causaliy$	0.9	0.067	0.933	0.907	0.093	0.048	0.952
$x \rightarrow z/causaliy$	0.8	0.053	0.947	0.839	0.161	0.043	0.957
$x \rightarrow z/causaliy$	0.4	0.064	0.936	0.488	0.512	0.071	0.929
$x \rightarrow z/causaliy$	0.2	0.076	0.924	0.293	0.707	0.06	0.94

		$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$					
$x \to y/causaliy$	0.9	0.063	0.937	0.103	0.897	0.044	0.956
$x \to y/causaliy$	0.8	0.056	0.944	0.122	0.878	0.059	0.941
$x \to y/causaliy$	0.4	0.064	0.936	0.225	0.775	0.051	0.949
$x \to y/causaliy$	0.2	0.078	0.922	0.185	0.815	0.063	0.937
$y \to z/causaliy$	0.9	0.072	0.928	0.446	0.554	0.051	0.949
$y \to z/causaliy$	0.8	0.074	0.926	0.389	0.611	0.037	0.963
$y \to z/causaliy$	0.4	0.073	0.927	0.226	0.774	0.069	0.931
$y \to z/causaliy$	0.2	0.058	0.942	0.165	0.835	0.061	0.939
$x \to z/causaliy$	0.9	0.072	0.928	0.446	0.554	0.051	0.949
$x \to z/causaliy$	0.8	0.074	0.926	0.389	0.611	0.037	0.963
$x \to z/causaliy$	0.4	0.073	0.927	0.226	0.774	0.069	0.931
$x \to z/causaliy$	0.2	0.058	0.942	0.165	0.835	0.061	0.939

Table 3

		VAR		**Modified R**		**Haugh**	
		Correct.	Omitted	Correct.	Omitted	Correct.	Omitted
		With Drift and Trend					
	$\boldsymbol{\rho}$	$\theta_{11} = \theta_{22} = \theta_{33} = 0.8$					
$x \to y/causaliy$	0.9	0.069	0.931	0.791	0.209	0.035	0.965
$x \to y/causaliy$	0.8	0.079	0.921	0.77	0.23	0.042	0.958
$x \to y/causaliy$	0.4	0.067	0.933	0.679	0.321	0.05	0.95
$x \to y/causaliy$	0.2	0.071	0.929	0.417	0.583	0.072	0.928
$y \to z/causaliy$	0.9	0.083	0.917	0.922	0.078	0.041	0.959
$y \to z/causaliy$	0.8	0.076	0.924	0.899	0.101	0.038	0.962
$y \to z/causaliy$	0.4	0.054	0.946	0.86	0.14	0.062	0.938
$y \to z/causaliy$	0.2	0.061	0.939	0.596	0.404	0.065	0.935
$x \to z/causaliy$	0.9	0.083	0.917	0.922	0.078	0.041	0.959
$x \to z/causaliy$	0.8	0.076	0.924	0.899	0.101	0.038	0.962
$x \to z/causaliy$	0.4	0.054	0.946	0.86	0.14	0.062	0.938
$x \to z/causaliy$	0.2	0.061	0.939	0.596	0.404	0.065	0.935
		$\theta_{11} = \theta_{22} = \theta_{33} = 0.6$					
$x \to y/causaliy$	0.9	0.054	0.946	0.457	0.543	0.05	0.95
$x \to y/causaliy$	0.8	0.051	0.949	0.391	0.609	0.046	0.954
$x \to y/causaliy$	0.4	0.051	0.949	0.187	0.813	0.055	0.945
$x \to y/causaliy$	0.2	0.067	0.933	0.317	0.683	0.057	0.943
$y \to z/causaliy$	0.9	0.078	0.922	0.912	0.088	0.048	0.952
$y \to z/causaliy$	0.8	0.085	0.915	0.897	0.103	0.043	0.957
$y \to z/causaliy$	0.4	0.058	0.942	0.55	0.45	0.071	0.929
$y \to z/causaliy$	0.2	0.068	0.932	0.332	0.668	0.06	0.94
$x \to z/causaliy$	0.9	0.078	0.922	0.912	0.088	0.048	0.952
$x \to z/causaliy$	0.8	0.085	0.915	0.897	0.103	0.043	0.957

$x \to z/causaliy$	0.4	0.058	0.942	0.55	0.45	0.071	0.929
$x \to z/causaliy$	0.2	0.068	0.932	0.332	0.668	0.06	0.94
$\theta_{11} = \theta_{22} = \theta_{33} = 0.4$							
$x \to y/causaliy$	0.9	0.063	0.937	0.11	0.89	0.044	0.956
$x \to y/causaliy$	0.8	0.073	0.927	0.107	0.893	0.059	0.941
$x \to y/causaliy$	0.4	0.061	0.939	0.193	0.807	0.051	0.949
$x \to y/causaliy$	0.2	0.06	0.94	0.265	0.735	0.063	0.937
$y \to z/causaliy$	0.9	0.057	0.943	0.49	0.51	0.051	0.949
$y \to z/causaliy$	0.8	0.063	0.937	0.446	0.554	0.037	0.963
$y \to z/causaliy$	0.4	0.058	0.942	0.231	0.769	0.069	0.931
$y \to z/causaliy$	0.2	0.069	0.931	0.245	0.755	0.061	0.939
$x \to z/causaliy$	0.9	0.057	0.943	0.49	0.51	0.051	0.949
$x \to z/causaliy$	0.8	0.063	0.937	0.446	0.554	0.037	0.963
$x \to z/causaliy$	0.4	0.058	0.942	0.231	0.769	0.069	0.931
$x \to z/causaliy$	0.2	0.069	0.931	0.245	0.755	0.061	0.939

APPENDIX III

Programming

% Data Generating Process (DGP)

```
function[X Y Z]=RIZdgp(A,B,sig)
W=[0,0,0]';
L=chol(sig);
for i=2:100
    m=randn(3,1);
    U=L*m;
W(:,i)= A*W(:,i-1)+B*[5;i]+U;
W2=W';
X=W2(:,1);
Y=W2(:,2);
Z=W2(:,3);
end
```

OLS FIT

```
function yhat=olsfit(y,x)
%X is the variable on which y is conditioned.
[n k]=size(x);
b=ones(n,1);
X=[b x];
betahat=inv(X'*X)*X'*y;
yhat=X*betahat;
```

VAR Residuals

```
function[b1 b2, b3 e1 e2 e3]=var313(x,y,z);
n=length(x)
n=length(y)
n=length(z);
x1=ones(n,1);
y1=ones(n,1);
z1=ones(n,1);
X=[x1 x];
Y=[y1 y];
Z=[z1 z];

X1=[X(1:end-1,1:end) y(1:end-1,1) z(1:end-1,1)]; %var
equation 1
Y1=[Y(1:end-1,1:end) x(1:end-1,1) z(1:end-1,1)]; %var
equation 2
Z1=[Z(1:end-1,1:end) x(1:end-1,1) y(1:end-1,1)]; %var
equation 3
```

```
xt=x(2:end,1);
yt=y(2:end,1);
zt=z(2:end,1);

b1=inv(X1'*X1)*X1'*xt;  %Coeffieients of Var 1 equation
b2=inv(Y1'*Y1)*Y1'*yt;  %Coeffieients of Var 2 equation
b3=inv(Z1'*Z1)*Z1'*zt;  %Coeffieients of Var 3 equation

e1=xt-X1*b1;    %Residual of Var 1 equation
e2=yt-Y1*b2;    %Residual of Var 2 equation
e3=zt-Z1*b3;    %Residual of Var 3 equation
```

Modified R Recursive Residuals

```
function [e]=mr(x)
[a]=size(x);
for i=10:a
    xt=x(2:i,1);
    xtm1=[ones((i-1),1) x(1:(i-1),1)];
    bhat=inv(xtm1'*xtm1)*xtm1'*xt;
    xi_hat=[1, x(i,1)]*bhat;
    e((i-9),1)=x(i,1)- xi_hat;
end
```

Haugh ARMA Residuals

```
model=arima('ARLags',2,'MALags',2);
m=estimate(model,x);
[residuals] = infer(m,x)
```

Conditional correlation

```
function [xygz]=ccorrelation(x,y,z)
[n,k]=size(z);
xgz=olsfit(x,z);
ygz=olsfit(y,z);
cov_xygz=(x-xgz)'*(y-ygz)/(n-1);
var_xgz=(x-xgz)'*(x-xgz)/(n-k);
var_ygz=(y-ygz)'*(y-ygz)/(n-k);
xygz=cov_xygz/(sqrt(var_xgz*var_ygz));
```

Fisher Z test

```
function [Fisher_Z]=fishz(rho,n,k)
% rho is calculated value of conditional correlation
% n is number of observation
% k is number of variables that we condition on.
Fisher_Z=(0.5*sqrt(n-abs(k)-3))*log((abs(1+rho))/(1-rho));
```

```
% CV

for i=1:1000
    % DGP
A=[1,0,0;0,1,0;0,0,1];
B=[0,0;0,0;0,0];
sig=[2,0,0;0,2,0;0,0,2];
[x,y,z]=RIZdgp(A,B,sig);

% Unconditional Correlation
ucxy=corrcoef(x,y);
sim_ucxy(i,:)=ucxy(2,1);
% Conditional Correlation
ccxygz(i,:)=ccorrelation2(x,y,z);

end
cv_up=[prctile(sim_ucxy,2.5),prctile(sim_ucxy,97.5)];
cv_cp=[prctile(ccxygz,2.5), prctile(ccxygz,97.5)];
```

Causal Path

```
function [xnay xnaz znay xcy ycx xcz zcx zcy
ycz]=causalpath(x,y,z)
xnay=0;
xnaz=0;
znay=0;
xcy=0;
ycx=0;
xcz=0;
zcx=0;
zcy=0;
ycz=0;

corxy=corrcoef(x,y);
coryz=corrcoef(z,y);
corzx=corrcoef(x,z);

ccorrxygz=ccorrelation2(x,y,z);
ccorrxzgy=ccorrelation2(x,z,y);
ccorrzygx=ccorrelation2(z,y,x);
c=1;
[n k]=size(x);

rho_xygz=fishz(ccorrxygz,n,c);
rho_xzgy=fishz(ccorrxzgy,n,c);
rho_zygx=fishz(ccorrzygx,n,c);

if abs(ccorrxygz)<0.70 %rho_xygz
        xnay=1;
else
        xcz=1;
```

```
        ycz=1; %unsheilded colider x->z<-y
end

if abs(ccorrxzgy)<0.70 %rho_xzgy
        xnaz=1;
else
        xcy=1;
        zcy=1; %unsheilded colider x->y<-z
end

if abs(ccorrzygx)<0.70 %rho_zygx
        znay=1;
else
        zcx=1;
        ycx=1; %unsheilded colider z->x<-y
end
```

PC and Modified PC Algorithms

```
for i=1:10
    % Data Generation
A=[1,0,0;0,1,0;0,0,1];
B=[0,0;0,0;0,0];
sig=[2,0,0;0,2,0;0,0,2];
[x y z]=RIZdgp(A,B,sig);

dgp_xcy=0
dgp_xcz=0
dgp_ycx=0
dgp_ycz=0
dgp_zcx=0
dgp_zcy=0

% VAR, MR and ARMA Residuals
[b1 b2 b3 e1 e2 e3]=var313(x,y,z);
[e]=mr(x);
[residuals] = infer(m,x)

% Un-correlation between residuals (e1 e2 and e3) from Var
uce1e2=corrcoef(e1,e2);
uce1e3=corrcoef(e1,e3);
uce2e3=corrcoef(e2,e3);

%conditional correlation between residuals (e1 e2 and e3)
from Var
ccorre1e2ge3=ccorrelation2(e1,e2,e3);
ccorre1e3ge2=ccorrelation2(e1,e3,e2);
ccorre3e2ge1=ccorrelation2(e3,e2,e1);


%Power and Size Procedure
```

```
correct=0     %The edge presence and orientation in both,
DGP and final graph will same.
Committed=0  %The edge is removed in the DGP while in the
resultant graph it is present.
Omitted=0     %The edge is present in DGP while removed in
the resultant graph.
Reversed=0   %Edges are present in both DGP and resultant
graph but reversed in directions.

if dgp_xcy=1
    if test_xcy=1
        correct=correct+1
    else
        ommitted=ommitted+1
    end
else
    if test_xcy=1
        comitted=comitted+1
    else
        correct2=correct2+1
    end

end
%.......................

if dgp_xcz=1
    if test_xcz=1
        correct=correct+1
    else
        ommitted=ommitted+1
    end
else
    if test_xcz=1
        comitted=comitted+1
    else
        correct2=correct2+1
    end

end

%.......................

if dgp_ycx=1
    if test_ycx=1
        correct=correct+1
    else
        ommitted=ommitted+1
    end
else
    if test_ycx=1
        comitted=comitted+1
```

```
    else
        correct2=correct2+1
    end

end
%......................

if dgp_ycz=1
    if test_ycz=1
        correct=correct+1
    else
        ommitted=ommitted+1
    end
else
    if test_ycz=1
        comitted=comitted+1
    else
        correct2=correct2+1
    end

end
%......................

if dgp_zcx=1
    if test_zcx=1
        correct=correct+1
    else
        ommitted=ommitted+1
    end
else
    if test_zcx=1
        comitted=comitted+1
    else
        correct2=correct2+1
    end

end
%.....................

if dgp_zcy=1
    if test_zcy=1
        correct=correct+1
    else
        ommitted=ommitted+1
    end
else
    if test_zcy=1
        comitted=comitted+1
    else
        correct2=correct2+1
    end
    end
```