# EVALUATION OF CAUSALITY METHODS AND TESTS FOR PANEL DATA

*By*

**ABIDA NAURIN**

Reg. No. PIDE2015FPHDETS06

***A Dissertation Submitted to the Pakistan Institute of Development Economics, Islamabad, in partial fulfillment of the requirements of the Degree of Doctor of Philosophy in Econometrics***

Supervised By

**Dr. AHSAN UL HAQ**

Assistant Professor

Co. Supervised By

**PROFESSOR Dr. ASAD ZAMAN**

**Ex-Vice Chancellor PIDE**

**PIDE School of Economics**

**Pakistan Institute of Development Economics,**

**Islamabad Pakistan**

**August 2021**

# Certificate of Approval

This is to certify that the research work presented in this thesis, entitled: **"Evaluation of Causality Methods and Tests for Panel Data"** was conducted by **Ms. Abida Nourin** under the supervision of **Dr. Ahsan ul Haq and Co-Supervisor Dr. Asad Zaman**. 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. Abida Naurin
PIDE2015FPHDETS06

Signature: Abida

**Examination Committee:**

a) **External Examiner: Dr. Eatzaz Ahmed**
Professor, SBP Memorial Chair
Qauid-e-Azam University Islamabad.

Signature: ______________

b) **Internal Examiner: Dr. Saud Ahmed Khan**
Assistant Professor
PIDE, Islamabad

Signature: ______________

**Supervisor: Dr. Ahsan ul Haq**
Assistant Professor
PIDE, Islamabad

Signature: ______________

**Co-Supervisor: Dr. Asad Zaman**
Ex-Vice Chancellor
PIDE,
Islamabad

Signature: Asad Zaman

**Dr. Shujaat Farooq**
Head, PIDE School of Economics (PSE)
PIDE, Islamabad

Signature: ______________

# Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled **"Evaluation of Causality Methods and Tests for Panel Data"** 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: Abida

Ms. Abida Naurin
PIDE2015FPHDETS06

# Author's Declaration

I Ms. Abida Naurin hereby state that my PhD thesis titled **"Evaluation of Causality Methods and Tests for Panel Data"** 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.

Abida

Ms. Abida Naurin

PIDE2015FPHDETS06

# ABSTRACT

Causality is the most important concept which is tested frequently in social sciences. Unfortunately, it is not easily detected from observational studies. It is difficult for a researcher to differentiate between cause and effect, which is even more challenging in economics. Researchers have been applied various approaches to test causality, i.e., Regression discontinuity design (1960), Cross correlation-based methodology (1994), Difference in differences approach (2005), Granger causality test (1969), Error correction mechanism (1986), Toda and Yamamoto method (1995) for cross-sectional and time-series data. These usual regression methodologies rely on normality and linearity assumptions not supported by the data used in the analysis and may lead to unreliable results. Therefore, the policies built on poor tests for causality remain unreliable. Some causality techniques are sensitive to distributional assumptions and specification issues. In the literature, no study compares causality methods and tests for all types of data and panel data that can help choose an appropriate testing methodology. The current study evaluates the size and power based on Monte Carlo Simulations of various causality methods for panel data under whole alternative hypotheses for all possible causal combinations. This study also modifies the Sims test (1972) and Final Prediction Error (FPE) test (1981) for the Panel dataset. Comparison of Panel Causality Tests (PCT) has been used with different model specifications (stationary series with drift only, with drift and trend) for different sample sizes (small, medium, and large) under this study. Monte Carlo results reveal that the Granger Non-Causality (GC) test by Dumitrescu and Hurlin (2012) has the least size distortion from nominal compared to size distortion of the Sims test and FPE test for all sample sizes of cross-section units. However, the GC test's power attainment is much better than the other two tests at all alternatives and all sample sizes. Among the Sims test and FPE tests, the former gains lower power at all alternatives than the latter one corresponding to small, medium, and large cross-section units and thus identified as the worst performer. A similar pattern has been observed for almost all tests at different sample sizes; medium sample size (i.e., T=50) and large sample size (i.e., T=200). Based on the comparison of size and power analysis of the PCT, this study concludes that the GC test is a point optimal and performs better at all causal combinations and panel dimensions, whether drift only or both drift and trend have been taken into account. On the other hand, the Sims test with its lowest power gain at all causal combinations and panel dimensions is the worst performer test. However, the FPE test having a power curve between the better and worst performer is graded as the average performer test. We investigate the government and household spending nexus on education in Pakistan for the applied application of our proposed PCT testing procedure. The result is fascinating and useful for policymakers. It indicates that the causality clearly runs from the intensity of government spending on education to the corresponding household intensity, but the effect is only direct.



**Keywords;** Econometrics, Causality, Panel data, Monte Carlo Simulations, Granger Non-Causality test

***Dedicated to My Beloved PARENTS***

***My Trust, My Inspiration, My World***

# *Acknowledgement*

*First and foremost, to my best friend, my **ALLAH**, the most loving and merciful, who created the universe with beauty and perfection and gave me the ability to ponder and refine my thinking. I offer my humblest feelings and thanks to the **Prophet Muhammad** (peace be upon him), who is forever the source of guidance and knowledge for humanity.*

*First of all, I acknowledge with immense pleasure and want to owe my cordial gratitude and sincere thanks and deference to my considerate and esteemed Supervisor, Mentor, and Spiritual father, **Professor Dr. Asad Zaman,** for his sympathetic and encouraging attitude, noteworthy and inspiring suggestions, and zealous interest in supervising this research. There is a debt of gratitude I cannot repay to my adviser, Dr. Ahsan Ul Haq, for his superb guidance and encouragement during the whole course of research. His wise counseling has always stimulated my mind and made me truly understand what this manuscript is all about. I am thankful to Dr. Panayiotis M. Pourides, Professor at Cardiff University, UK, for his remarkable and unprecedented guidance. I am also grateful to my examiners, Professor Dr. Eatzaz Ahmad (Quaid-i-Azam University), Prof. Dr. Terence D. Agbeyegbe (University of New York, USA), and Prof. Dr. Muzafar Shah Habibullah (Putra Business School, Malaysia) for their time and valuable suggestions for further improvement of the manuscript. Special thanks to my dear teacher Dr. Atiq Ur Rahman (AJK University), and my senior fellow Dr. Shahid Akbar (University of Swabi), who helped me complete the thesis and made this easy and accurate.*

*There are no words to say thanks to my parents; without their love, care, and prayers, and I would not be able to accomplish such research work. They have always been there for me. I am also thankful to my brothers and sisters for their unconditional emotional and financial support throughout my life; this degree made my path easier to walk on. I have missed them all during my time at university, away from home. Last but not least, a special thank goes to my Friends; they exchanged their interesting ideas and thoughts and made my experience at PIDE enjoyable. I am grateful to all the people who positively or negatively contributed to accomplishing this task.*

*ABIDA NAURIN*

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ADF | Augmented Dickey-Fuller |
| AR | Autoregressive |
| BNL | Bank Nonperforming Loans |
| CD | Cross-Sectional Dependence |
| DAG | Directed Acyclic Graph |
| DID | Difference in Difference approach |
| DGP | Data Generating Process |
| DH | Dumitrescu and Hurlin |
| DTMC | Discrete-Time Markov Chain |
| ECM | Error Correction Mechanism |
| ECT | Error Correction Term |
| FPE | Final Prediction Error |
| GC | Granger Causality |
| GDP | Gross Domestic Product |
| GEX | Government Expenditure on Education |
| GMM | Generalized Method of Moments |
| HEX | Household Expenditure on Education |
| IFS | International Financial Statistics |
| MCM | Markov Chain Method |
| MCSS | Monte Carlo Sample Size |
| PCT | Panel Causality Test |
| PMG | Pooled Mean Group |
| RDD | Regression Discontinuity Design |
| SBP | State Bank of Pakistan |
| SEM | Structural Equation Models |
| SIM | Sims |
| SUR | Seemingly Unrelated Regression |
| SVAR | Structural Vector Autoregressive |
| VAR | Vector Autoregressive |

# CHAPTER 1

# INTRODUCTION

## 1.1. Background of the Study

Causality is the most important concept which is tested frequently in social sciences. Unfortunately, it is not easily detected from observational studies[1]. In the natural sciences, causality can be determined through controlled experiments, whereas controlled experiments are difficult to be carried out in social sciences. Experimental and observational studies have different statistical tools, which can be explained with various descriptive analyses. Therefore, one has to investigate the causal analysis for observational data. However, causal inferences are among the most challenging in observational data and have several issues. The first and the most critical point is that causality is not directly observable in the non-experimental data. Second, one cannot control fundamental confounding factors in observational data. Third, statistical relationship measures are symmetric and do not directly form causality. Hence, it is difficult for the researcher to differentiate between cause and effect.

It is well-known that different causality methods and techniques are applicable in different scenarios. Therefore, it is necessary to determine which statistical technique/test gives us better statistical properties in a particular data set. Researchers have been applied various approaches to test causality, i.e., Regression discontinuity design (1960), Cross correlation-based methodology (1994), Difference in difference approach (2005), Granger causality test (1969), Error correction mechanism (1986), Toda and Yamamoto method (1995) for cross-sectional and time-series data. However, Econometric analysis based on the regression as mentioned above methodologies rely on normality and linearity assumptions that are not supported by the data used in the research and may lead to unreliable results. Therefore, it becomes impossible to formulate policies based on the findings of the studies that follow conventional methods.

[1] Observational study draws inferences from a sample to a population where the independent variable is not under the control of the researcher because of ethical concerns or logistical constraints.

Researchers working on causality have faced these difficulties and proposed different solutions to determine causality. The most commonly used technique to determine causation is a correlation. However, Correlation is very sensitive to data. Pierce and Haugh (1977) identify that causality cannot be conceptualized as an appropriate correlation measure; rather, it indicates empirical correlation. Later on, economists tried to differentiate causal relation from observed correlation with the evolution of econometrics. Furthermore, the researcher developed several methods to test causality according to the data type.

One another exciting debate is made by Hume (1752) on causality, where he has explored the relation between prices and money. Another fascinating insight has been provided by Adam Smith, in which it has been demonstrated that the concept of causality occupies a central position in economics. Moreover, economists like John Stuart Mill and Ricardo have discussed causality issues in literature.

The primary objective of empirical research is the understanding of causal relations which present among a set of variables. It is generally considered that high correlation among variables does not create any causation; however, variables can be related to each other functionally. They may be correlated, but we cannot say they are causally related. Pierce and Haugh (1977) put the concept in this manner: *"The former impact emerges because correlation is just a measure of linear relationship; the latter comes as a result of each's shared relationship with other components."*

Wold (1954) explained that the necessity of causality as a concept is significant for all science subjects. Wold (1954) defined the concept of causality in terms of controlled experiments. He also described that the definition of causality appears to be very simple in experimental studies compared to observational studies. In contrast, it is difficult to conduct it in economics as economics is based on observational studies. Hence, he concluded that statistical methods might be valid in observational studies if subject matter theory is considered.

Granger (1969) used a predictability test to approximate the concept of causality called Granger Causality (GC). Thus according to Granger's definition of causality, "X Granger cause Y" when $X_{t-1}$ can predict $Y_t$ . Zellner (1979) criticized the GC on many grounds, and the most important criticism was that this approach was atheoretical. The practical implementation of the Granger approach requires the imposition of a lot of assumptions. He also claimed that if the assumptions imposed by the researcher are not valid theoretically, this approach will be discovered only accidental regularities.

Economic theory is mainly concerned with causal relationships among social and other economic variables such as endowments, production of goods, how goods are consumed, how different entities accumulate over time, etc.

The early 20th century reflected the improvement and the link between Causal inference and statistical tools, i.e., correlation and regression, have been observed. It was believed by the researchers that regression has a natural direction, unlike correlation. Likewise, the coefficient estimates from the regressions and the inverse regressions may not be the same. Therefore, it can be said the direction of regression must be according to the principle of causation. (Hoover, 2001)

The data does not reveal the correctness of direction as each direction is similar observationally, even though a regression has a natural approach. The researchers mainly considered issues of simultaneity and problem of identification later on. Cowles Commission identified the solution to these problems[2] , and the issue of causality was set to the side. Later, in the second part of the twentieth century, two major approaches[3] to causality were developed. (Asghar, 2008)

There are four major approaches to addressing the issues of causality in economics. First, Suppes (1970) developed the probabilistic approach to causality. The second approach was developed by Granger (1969), who gave one of the most profound definitions of causality, which depends on statistical grounds in a time-series format. And the third approach here is called Structural Causality, proposed by Hendry and Mizon (1998), which reveals causality cannot be detected from observed data without structural change. Lastly, Pearl (2000) considered the work of Simon and Shemin's (1953) Structural Equation Model (SEM) as a significant contribution to the field of econometrics. However, causality in SEM is also considered to be controversial. SEM is based on correlation, and it shows an association, not causation. Therefore, empirical researchers in this field might be confused about the direction of causation among the variables under consideration. Theorists like Spirtes et al. (2000),

[2] Simon (1953)
[3] One approach by Herman Wold and the other by Herbert Simon

and Pearl (2000)offered solutions for the problems of SEM that enhanced development in this area through the graphical models.

The concept of Causality is fundamental but, unfortunately, cannot be detected easily. Therefore, the exploration of data intensively is essential for detecting causality.

According to the approaches mentioned above, the association between two variables exists due to the other confounding variables in some cases. Therefore, researchers were not able to identify which method should be preferred for testing causality in the presence of confounding variables. Such causality tests could not detect a causal relationship without taking all the confounders under control. However, these can be only controlled in experimental, unlike observational studies. The causation must be supported by theory as compared to statistical techniques.

Freedman et al. (1995) found that to develop causal inference, one needs to engage critically and develop different skills. Natural variations must be recognized in this manner, and data must be gathered. In addition, cofounders must be considered, and other hypotheses must be investigated before concluding. Finally, to find supporting evidence, theory must be consulted to identify true causes rather than focusing only on statistical analysis.

We have various data sets like cross-sectional data covering causality methods, i.e., Regression discontinuity design (RDD), Difference in difference approach (DID), and Cross correlation-based methodology. The second type of data set is the time series which applied methodologies like GC test, Sims test, Final prediction error method, Error correction mechanism, Toda and Yamamoto method, and Structural Causality test to examine this issue. Cross-sectional and time-series both are restricted cases of Panel data. Both techniques can be applied for this, and many methods are designed to handle cross-sectional and time-series simultaneously. However, there is a lack of literature regarding comparative analysis for each kind of data that can help choose the appropriate testing methodology. The specific purpose of this current study is to compare causality tests for each kind of data set using Monte Carlo simulation.

The Mediation models are particularly efficient at demonstrating how one variable (usually, X) influences another variable (Y) through its impact on an intermediate variable (M). Therefore, researchers that seek to investigate the causal links between variables might use such models. However, statistical findings for any causal model are insufficient to establish causation after applying the mediation model. Furthermore, in addition to the statistical findings, Selig and Preacher (2009) conceptual model require theory and evidence.

Some causality techniques are sensitive to distributional assumptions and specification issues. Therefore, this study also aims to test the robustness of causality tests to specification and distributional issues.

Noticeably over the last few decades, massive developments in computational power have allowed feasible and flexible computational approach that was not imagined originally. Therefore, it allowed researchers to apply concepts extensively, even more than Lehmann believed. Interestingly, Zaman (1996) used this methodology to reflect that the famous Durbin-Watson test conducted for autoregressive errors in the regression model was poorly compared to specific alternatives. Similarly, Khan et al. (2016) have utilized the approach to compare normality tests to come up with definitive recommendations. This concept reflected what to focus on while running any tests as contrasted to Lehmann's ideology that tells us that problem is not fully identifiable and solution-oriented; there are always some issues left unresolved. However, it is significant to highlight that Stringency provides the formulation of the problem and tells us what to expect and look for in a good test. Although sometimes, an accurate assessment of stringencies is not possible, many strategies can be used to provide an approximation to this particular number that will help us calculate a single performance measure for all tests. (Rehman et al., 2017)

## 1.2. Problem Statement

As a matter of fact, different causality methods and techniques are applicable in different econometric scenarios. That is why the choice of the causality method is of prime importance to get reliable econometric estimates. Researchers have been applying various approaches to test causality, i.e., Regression discontinuity design (1960), Cross correlation-based methodology (1994), Difference in difference approach (2005), Granger causality test (1969), Error correction mechanism (1986), Toda and Yamamoto method (1995) for cross-sectional and time-series data.

However, Econometric analysis based on the regression as mentioned above methodologies rely on normality and linearity assumptions that are not supported by the data used in the research and may lead to unreliable results.

Therefore, it becomes impossible to formulate policies based on the findings of the studies that follow conventional methods. In the literature, no study provides a detailed comparison of causality methods and tests for panel data that can help choose an appropriate testing methodology. Hence, exploring which conventional causality method is more appropriate to find out the true causal relationship. This necessitates a comparison of available causality methods to identify the robustness of various techniques under varying econometric scenarios.

### 1.3. Objectives of the Study

The key objective of the current study is to analyze the effective functional comparison of all the causality methods regarding the type of data, i.e., time-series data, cross-section data, and panel data. Furthermore, this study will show which methodology detects true causal relation by employing the appropriate method for specific data types.

Based on the motivation laid down in the previous section, the following are the objectives of the present research:

1. To theoretically review all the conventional causality methods and tests. It provides a comprehensive literature survey of all traditional causality methods and tests for all forms of data, including time series, cross-sectional, and panel data. In addition, there is a brief explanation of the notions and terminologies of all causality tests, which can be utilized in testing and simulations.
2. To modify the time series causality tests such as Sims (1972) and Hsiao's Final prediction error method for the panel data set.
3. To evaluate the performance of panel causality methods and tests based on size and power analysis for different model specifications using Monte Carlo Simulations.
4. To test the robustness of causality tests to specification and distributional issues.

5. To examine the appropriate causal methods and tests by using real panel data application. "Analysis of the Causality between Intensity of Government spending and the Intensity of Household spending on Education with the Role of Credit Constraints: Panel Data".

### 1.4. Significance of the Study

For causal inference, it is significant to know that Panel data allows both opportunities and challenges for the researchers. Panel data has the most significant substantial benefit over cross-sectional data in evaluating unobserved time-invariant factors. Whereas, the main issue with panel data is the likely serial correlations that might come in the form of errors for every individual. On the other hand, the present study attempts to provide models and methods for analyzing panel data, focusing on examining how different models and methods handle causality issues. First, by combining cross-section and time-series data, many observations may be represented while maintaining a high degree of freedom and the decreasing possibility of significant linear association between explanatory factors.

Secondly, Panel data can evaluate effects that are difficult to notice in time series data and cross-sections. Thirdly, establishing and examining highly complex behavioural models is preferred (Baltagi, 2008). These reasons open space for the research needed in this domain, allowing us to explore and evaluate comparative studies on different causality methods.

We have found no comprehensive comparison of different causality methods in the literature. As it is known that different causality techniques are applicable in different scenarios. Therefore, it is necessary to determine which statistical technique/test gives us better statistical properties in a particular scenario. Econometric analyses like the error-correction model by Phillips (1986), cointegration, and SIM test, GC tests are applied to examine this issue. However, these existing regression methodologies rely on normality and linearity assumptions not supported by the data used in the analysis and may lead to unreliable results. Therefore, these techniques sometimes may not be able to provide proper/true results. Consequently, it becomes impossible to formulate policies based on the findings of the studies that follow conventional methods. Hence, exploring which traditional causality method is more appropriate to find out the true causal relationship is required.

In the current research, longitudinal data are chosen for evaluating mediation hypotheses for two significant reasons. Firstly, the high quality of the findings obtained after employing a mediation model with cross-sectional data enables evaluating mediation in any field. However, there are several fundamental issues with using conventional mediation models on cross-sectional data. For instance, outline three such cases. (Gollob & Reichardt, 1987)

The central problem is the duration of time that requires understanding the causal relationships while using mediation models. We do receive instantaneous effects through cross-sectional use, but logically, it is a problematic assumption. The second main issue is that when we remove predictors from the conclusion, it can be resulted in promoting errors as variables measured at different instances are not controlled. This further results in either making paths over or underestimated compared to their original values. Thirdly, because effects only emerge with time, the amount of a causal impact should not be expected to be constant across all possible intervals. (Cheong et al., 2003)

Furthermore, the impact size is unaffected by the time interval between the observations. In contrast, Cole and Maxwell (2003) argued that utilizing cross-sectional data for model mediation might be an evident source of systematic bias.

This research proposes that the use of “stringency” can be used as a Gold Standard for evaluating and examining the relative performance of tests. In the case of the one-dimensional parameter, it is easy to assess hypotheses for testing diverse problems. Moreover, it implies that we do not need an alternative hypothesis before conducting a test.  The reason is that we do not have any set standard method of devising comparisons between different hypotheses.

### 1.5. Contribution of the Study

We have found no study in the existing literature that provides a detailed literature review comparing all causality methods for cross-section, time-series data, and panel data. A causal relationship can be checked for different kinds of data. It is considered that a comparison of conventional methods of testing causal relationships used to detect causality will provide whether these methods detect the true causal relationship.

Cross-sectional and time-series both are restricted cases of Panel data. Both techniques can be applied for this, and some methods are designed to handle cross-sectional and time-series simultaneously.

In order to make a comprehensive comparison among panel causality methods and tests, available causality tests are required to make a proper justification for identifying the best and worst performer tests. In view of existing literature for panel data, almost all panel data tests were first developed for a single cross-section with a time series structure and then expended for more than one cross-section. An average of all cross-sections is taken to create a panel data test. Keeping in view the same practice, this study has modified the Sim's time series causality test and the final prediction error method of casualty for time series to panel counterparts to compare with the Granger non-causality test. The comparison between a limited number of tests indicates that researchers explore panel causality literature less after developing the Granger non-causality test.

Some causality techniques are sensitive to distributional assumptions and specification issues. This study also aims to test the robustness of causality tests to specification and distributional issues. It is significant to highlight the stringency of a test is the most defined shortcoming of a test; however, we only found rare applications in the present literature. A considerable gap in the literature tells us that the stringency concept is not researched in-depth, suggesting that most researchers do not understand the use of stringency as a source of finite sample performance. Hence, this particular research study is designed to explain the status of stringency in depth.

### 1.6. Organization of the Study

Chapter 2 is a quick overview of the theoretical and empirical econometrics literature that has been utilized to assess causality. It also provides a theoretical framework for determining the real causal relationship in econometrics literature. Chapter 3 presents a comprehensive literature survey of all traditional causality methods and tests for all forms of data, including time series, cross-sectional, and panel data. Chapter 4 discusses the research methodologies utilized for the empirical analysis of this study. The chapter is divided into the data generation process and

Monte Carlo simulations. It also defines concepts, terminologies, and several tests for causality to establish the causal pattern.

In Chapter 5, the simulation results achieved by employing panel causality tests with heterogeneous panel DGP under various model specifications are briefly discussed. Based on Monte Carlo simulation findings, a size and power comparison is performed between the Granger non-causality test, Modified Sims test, and Modified Hsiao's Final prediction error technique. Chapter 6 is about the real data analysis using the appropriate causality tests and methods: "Analysis of the Causality between Intensity of Government spending and the Intensity of Household spending on Education and the Role of Credit Constraints: Panel Data”. Chapter 7 summarizes the conclusion from the simulation findings in Chapter 5. It also discusses the study's limitations and future research directions.

# CHAPTER 2

# REVIEW OF LITERATURE ON CAUSALITY

This chapter aims to review relevant literature associated with causality in the context of various economic problems. This chapter is divided into two sub-sections. Sub-section 2.1 discusses theoretical literature and sub-section 2.2 presents the empirical literature in the case of International and Pakistan studies. It also provides a theoretical framework for determining the real causal relationship in econometrics literature.

## 2.1. Review of Theoretical Literature

Economists and econometricians have long studied the question of causal laws and causality. The critical issue is establishing a causal relationship between a result and circumstances that may have influenced it. For example, GC test and Sims test (SIM test) (1969; 1972) causality in time series is based on the idea that cause must follow effect and that factor cannot cause another factor if it does not add to the expectation or conditional distribution of the element provided in the past. This is one of the most popular macro-econometric modeling and time-series concepts. It also plays an essential role in the concepts of exogeneity developed by Engle et al. (1983).

For multiple time series, linear dependency and feedback measures have been established by Geweke (1982). The measure of linear dependence is the sum of the portion of linear feedback from the first series to the second, linear feedback from the second to the first and instantaneous linear feedback. The measures are nonnegative and reach zero only when the relevant type of feedback (causality) is absent. The frequency can be used to decompose the measures of linear feedback from one series to another. A simple theory of inference is provided for all of these measures and their decompositions; the calculations required are modest.

The current literature describes causal impacts by comparing possible outcomes on the same unit and evaluating simultaneously but exposed to various treatments. Later, one of these two possible outcomes may be observed, and the causal impact is effectively treated as an inference issue with incomplete data. That is why; causality is a central question but cannot be detected easily.

Hoover (2001) stated that the need of identifying causal relationships among economic variables is the relevant justification for policymaking. The famous theorist Hume (1752) discussed the concept of causality as a philosophical phenomenon. He defines a cause as one item that is followed by another, and all subsequent items identical to the first are followed by objects similar to the second. In other words, if the first thing did not exist, the second would not have been. Moreover, Hume suggested that causal events might be reduced to non-causal occurrences ontologically. Direct causal relationships, on the other hand, are not visible. (Demiralp & Hoover, 2003)

However, they can be traced down through constant conjunctions and the structure of general laws. There is no doubt that inquiries of causality are a central part of understanding economic theory. The problem lies in evaluating the economic measurement of the issues. Haavelmo (1944) developed the structural equation models and other Cowles Commission, econometricians. They focused on how to hypothesize causal relationships. However, the explicit causal interpretations are not considered anymore. These models are often interpreted as reduced representations of joint probability distribution (Pearl, 2000). The empirical malfunction and incapability of such structural models promoted space for multivariate time series methods which do not rely on clear causal interpretation (Heckman, 2000).

In order to identify and evaluate significant counterfactual questions, there is ultimate demand to understand the concept of causal relations so that we can manipulate the effect of one variable on another to prove our stance (Cartwright, 2003).

One of the massive and complex problems faced most of the time by the Econometricians' is their reluctant tendency to exposit clear and robust conclusions only because of the limitation of using observational data. Mill (1884) believed that one could not derive results about causal relationships by using observational data. The problem is heightened because researchers are oriented to focus on the Auxiliary hypotheses most of the time. In contrast, the original problem lies in assessing and questioning the limitations of the theoretical underpinnings, which are blindly followed (Blaug & Mark, 1992). In other words, it can be said that either we verify the existing theories or confirm some already found conclusions of the theorists in the past.

Reichenbach (1956) found that in the case of using random variables, there are explicit causal inferences that can be traced down through statistical data. Moreover, Hausman (1983) elaborated that causal inference in such scenarios becomes feasible. However, there has been another method suggested in the literature about using some algorithms for conducting such inferences. (Glymour and Cooper (1999); Spirtes et al. (2000); and Pearl (2000). It is worth highlighting here that we often see literature related to causality primarily focused on a causal relationship. The manipulation concept is present between two variables, 'A' and 'B', and there is no focus on statistical analysis or prediction (Pearl, 2000); page 85 and Woodward (2003), chapter 2.

Various theorists study the use of causal inference methods. Swanson and Granger (1997), Hoover-Dempsey et al. (2005), and Demiralp and Hoover (2003), proposed that to deduce contemporaneous causal relations among variables during vector autoregressions, we need to use causal inference algorithms. Kirmizi-Alsan et al. (2006) formulated a bootstrap method to evaluate the confidence employed in such results.

In the present scenario, one of the most controversial topics is determining how to characterize the nature of the link between various factors, such as export and economic growth, money and economic growth, and so on. The focus of research is to identify whether these variables or vice versa determine economic performance.

Algorithms of causal inference are designed to be readily adopted in data mining and machine learning setup. First, these algorithms use the values of several variables that can have an association in some aspect and then make inferences about their supposed relationship. This is a sequential procedure and is frequently followed to test the conditional dependence of variables. Since this methodology is a multi-layered process, it is criticized for its high probability of errors that may not be quantified in advance. Evidence for a high unknown probability of error in this approach is observed in several empirical investigations. These experiments have used different data sets and different variables but have observed the same fragility and inverse causal relationships. Monte Carlo simulation studies (Demiralp and

Hoover, 2003) suggest that the probability of having errors is unknown because it varies with types of data sets and choice of parameters. (Bryant et al., 2009)

## 2.2. Review of Empirical Studies

Bayramoğlu and Öztürk (2018) employed the panel causality approach to investigate the twin and triple deficit hypotheses for the economies of 15 developing countries, discovering a relationship between domestic saving, budget deficit/surplus, current account deficit, and fixed investments. The technique Dumitrescu and Hurlin (2012) panel causality approach was employed between 2000 and 2015. The results revealed that, from the public budget balance to the current account balance, there is a one-way causation connection. And, for the national group studied, the notion of twin deficits is correct.

An and Winship (2017) examined causal inference in panel data to assess Race of Interviewer Effects (ROIE) in a general social survey. This study examined advanced parametric methods and non-parametric matching models from 2006 to 2010 of the General Social Survey. In order to estimate the causal impacts in panel data, presented seven methods which could be divided into two sections. The first section is included six parametric models, the fixed-effect model, the first difference model, AR model, MA model, the Random effect model, and the Random Trend and Slope (RTS) model. The other section is contained a nonparametric approach, i.e., Difference-in-Difference estimator for causal inference in panel data. These methods have offered different advantages for estimating causal inference and also suggested both models cross-validate the evidence. The researchers found a statistically insignificant result that ROIE fluctuates by using different interview methods. Further, this study concluded matching method offered a good covariate provision and an additional concentrated causal implication through DID as a remedy to eliminate the impact of unnoticed time-invariant issues.

Attiaoui et al. (2017) researched the causality connections among economic growth (GDP), renewable energy consumption (REC), non-renewable energy consumption (NREC), and CO2 emissions (CE) in Africa. The sample panel data was used between 1990 and 2011 for 22 African countries. The methodology employed the Autoregressive distributed lag model established on the pooled mean group estimation (ARDL-PMG) and GC tests (1969). The findings of this research revealed one-way causality from $CO_2$ emissions to economic growth in the short run, and in the short run, the causal path between CE and REC is not determined. Further, they found one-way short-term causality from renewable energy consumption to economic growth. While evaluating per pair of variables, bidirectional causalities between REC, GDP, and CE were found in the short run and bidirectional long-term causalities, favouring the feedback hypothesis. In this panel, research causality is not strong from GDP to REC.

Chang et al. (2017) conducted a study from 1870 to 2013 on the relationship between quality of living growth and population growth in 21 nations. The methodology employed was the bootstrap causality test for the panel offered by Kónya (2006), which deals with both dependence and heterogeneity crosswise countries. This test determines the causative factors that influence population and living standard growth.

Bayraktar-Sağlam (2016), analyzed economic growth and the phases of human capital. The causal direction matter for poor and rich using panel data was identified from 1970 to 2010 for the 90 countries. For this objective, the panel vector VAR approach under the system GMM method has been employed. For further analysis, the physical capital accumulation has been included to remove the problem of omitted variable bias. This paper revealed that tertiary and secondary education had determined power for economic growth, but economic growth didn't determine human capital accumulation in developing countries. However, in the OECD countries, results showed tertiary education i.e., mean of technological progress, stimulated economic growth.

Bedir and Yilmaz (2016) investigated the causal link between the human development index and $CO_2$ emissions in Organization for Economic Cooperation and Development (OECD) nations from 1992 to 2011. A new technique of panel data established by Kónya (2006), which is founded on a seemingly unrelated regression (SUR) model projected by Zellner (1979) and The Wald test, was used, using country-specific bootstrap critical values. Moreover, three approaches have been employed to find the causal direction for longitudinal data: the GMM (generalized method of moments) approach. Hurlin and Dumitrescu (2008) proposed an additional method that reduced heterogeneity but not cross-sectional dependence. Kónya (2006) developed a method to cope with cross-sectional dependence and panel data heterogeneity (Kar et al., 2011). The findings of this paper implied that protection policies regarding oil consumption, coal gas, and electricity could decrease CO2 emissions and hinder human living standards and economic growth.

The research was carried out to examine the causal link between smoking and happiness for the five developed nations (i.e. Germany, Japan, France, the US, and the UK) from 1961 to 2003. For this analysis, the methodology used was the recently established panel causality bootstrap test suggested by Konya in 2006. This bootstrap panel causality method is more vigorous than other methods because of the country-specific cohort of critical values. The main findings of this study indicated a bidirectional relation for both France and Japan; independence for the other three nations. Further concluded that people were happier while they smoked. (Chang et al., 2016)

Zhang et al. (2011), investigated three industries in Beijing from 1980 to 2008, and the causal link between energy use and economic development was studied. They used the most recently created panel unit root, heterogeneous panel cointegration, and Panel based error correction model to accomplish this. This study concluded that in the short run, two-way Granger causation exists, but in the long run, one-way GC exists, with energy consumption leading to economic growth.

## 2.3 The Gap in the Literature

In the literature, no study provides a detailed comparison of causality methods and tests for panel data that can help choose an appropriate testing methodology. A causal relationship can be checked for different kinds of data. It is considered that a comparison of conventional methods of testing causal relationships used to detect causality will provide whether these methods detect the true causal relationship. Therefore, this is the literature gap to evaluate the performance of the causality testing methods for panel data based on Monte Carlo Simulation Design.

# CHAPTER 3

# LITERATURE SURVEY ON TESTING CAUSALITY METHODS

Researches validate that economist tried to determine the causal relationship among economic variables by measuring its correlation. Initially, Pierce and Haugh (1977) identify that correlation cannot be regarded as an appropriate measure of causality; rather, it indicates empirical correlation. Later on, researchers started trying to differentiate causal relation from empirical correlation. However, it is well-known fact that finding a high correlation among variables does not indicate causality. Moreover, researchers developed several methods to test causality according to the kind of data in recent literature.

## 3.1. Causality Methods in Time Series

One of the major contributions of Granger (1969) was the formulation of GC test which wholly transformed the process of evaluating relationships among variables. Through advanced research using this test over time different dimensions of variables were explored and new techniques were devised. In the following paragraphs, different methodologies have been discussed that were formed out of Granger's Causality test. (Asghar and Abid 2007)

### 3.1.1. Granger Causality test (1969)

The Granger Causality (GC) test is done, in the existence of lagged values of $Y_t$, regress the present value of the time series $Y_t$ against the previous values of the time series $X_t$ (Granger, 1969). Let us now assume a specific autoregressive with lag length k and estimate the following unconstrained equation using ordinary least squares (OLS);

$$Y_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i Y_{t-i} + \sum_{j=1}^{j} \beta_i X_{t-j} + \mu_t \quad (3.1)$$

$$H_{0a}: \beta_1 = \beta_2 = . \ . \ . = \beta_k = 0 \quad (3.2)$$

Conduct an F-test to examine the null hypothesis and determine if the coefficients related with the X's are jointly statistically significant. Run the following equation now:

$$F = \frac{(SSR_{restricted} - SSR_{unrestricted})/k}{SSR_{unrestricted}/(T-2k-1)} \sim F_{k,\ T-2k-1} \quad (3.3)$$

If the F- statistic is larger than the critical value, reject $H_0$ that X does not Granger-cause Y, and instead accept the alternative hypothesis that X does Granger-cause Y. Likewise, after regressing X on its past values and past values of Y, i.e.

$$X_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \beta_j X_{t-j} + \varepsilon_t \quad (3.4)$$

$$H_{0b}: \alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$$

Conduct an F-test to examine the null hypothesis and determine if the coefficients related with the Y's are jointly statistically significant. Run the following restricted equation using OLS;

$$X_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i X_{t-i} + \mu_t \quad (3.5)$$

Now, applying the F-statistic of Eq. (3.3), compared their respective sum of squared of residuals from Eq. (3.4) and Eq. (3.5), and get the following four potential cases: (i) there is a one-way causation from 'Y' to 'X' if $H_{0a}$ is accepted and $H_{0b}$is denied. (ii) There is a one-way causation from 'X' to 'Y' if $H_{0a}$ is rejected and $H_{0b}$ is accepted. (iii) If both $H_{0a}$ *and* $H_{0b}$are denied then there occurs feedback causality among 'X' and 'Y'. (iv) If both $H_{0a}$ *and* $H_{0b}$are accepted then it means there is no causation among 'X' and 'Y'.

This is necessary to highlight here that GC test (1969) assumed that the variables are stationary 'X' and 'Y' and $\mu_t,\ \varepsilon_t$ are uncorrelated. So as a matter of fact, in the above-mentioned equations, assumed that the variables are integrated of order zero at level and $\mu_t,\ \varepsilon_t$ are uncorrelated. The lag length in the above-mentioned equations can be choose while residuals are white noise. However, if the variables are stationary at first difference, and then first to do differencing of the variables. Use the first differences of variables in the multivariate scenario, and if they are stationary at first difference, then apply the following equations to verify causality among three variables.

$$\Delta Y_t = \alpha + lagged\ (\Delta Y_t, \Delta X_t, \Delta Z_t) + \varepsilon_t \quad (3.6)$$

$$\Delta X_t = \alpha + lagged\ (\Delta Y_t, \Delta X_t, \Delta Z_t) + \varepsilon_t \quad (3.7)$$

$$\Delta Z_t = \alpha + lagged\ (\Delta Y_t, \Delta X_t, \Delta Z_t) + \varepsilon_t \quad (3.8)$$

In the above equations, 'Δ' known as the first difference operator expressed as $\Delta Y_t = Y_t - Y_{t-1}$. Now, in the above three equations, we must test the related null hypothesis by assuming that "$\varepsilon_t$" is a white noise procedure. Moreover, it is claimed that real causative processes are removed by using the first difference of variables. So, the GC test (1969) stated that the aforementioned conclusions are not effective in such circumstances. That means that it is recommended to use the Error Correction Mechanism (Phillips 1986) or Toda and Yamamoto (1995) under such situations.

GC test (1969) assumed that all variables are stationary at the level and one variable is regressed on its lags, plus lags of another explanatory variable, then the error term to be white noise. Contrary to this, if the variable is non-stationary at level, then we will have to take appropriate differencing to make it stationary. Thus, the transformation will decrease GC's importance and create the following issues. Firstly, by transforming the variable the functional form of the variable will get change. Secondly, F-statistic is used to test the GC among variables in order to confirm whether the variable is stationary or non-stationary. However, Ender (1995) believed that F-statistics results do not reflect consistency across the stationary and non-stationary processes. Finally, we have witnessed that the results of GC are seen to be sensitive to minor changes in the specifications. For instance, they are sensitive to sample size, base year, lag length variable transformation, and tests used for model selection criteria. (Asghar and Abid 2007) Hence, through discussion, we have witnessed that the GC test (1969) is not appropriate to test the causality of time series. Therefore, we need to find other avenues.

### 3.1.2. Sims Test (1972)

In GC test (1969), explored that often one variable is regressed on its lags and the lags of another explanatory variable. But it is important to note that it does not include lead values of the explanatory variable during the process. The problem is confronted by Sims (1972) in which he argued that if one variable is regressed on its lags then the leading values of the explanatory variable will result in causality run from explanatory to regressed variable and also all the leading values of regressor in the regression will not become statistically significant and different from zero as a group. Henceforth, Sims demonstrated that *"future cannot cause current or past"* (Sims 1972).

Moreover, it assumes that the error term white-noise and variables must be integrated of order zero at level. However, one variable is non-stationary at level but become stationary at first difference, then in that case variable at first difference should be used preferably (Sims 1972).

In such scenarios, we see causality as capable of formulating a relationship between lagged and lead among various kinds of economic variables. But according to the Sims (1972), when we regressed Y variable on the lags and the lead values of X variable. Then If X Causes Y, we argue that the coefficients of X's lead values are thus equal to zero.

Thus, the application of the Sims (1972) will demand an equation for testing regression X to Y;

$$Y_t = \alpha_0 + \sum_{i=1}^{m} \alpha_{1i} X_{t-i} + \sum_{j=1}^{n} b_j X_{t+j} + \nu_{1t} \qquad (3.9)$$

Here tested the following null hypothesis by using the F-test (eq.3.3):

$$H_0: b_{11} = b_{12} = . \ . \ . = b_{1n} = 0$$

Now, If $H_0$is accepted then we will say X cause Y; otherwise, we will say that X does not Granger cause Y. Likewise, to assess causality from Y to X, we have to apply the following equation:

$$X_t = \alpha_1 + \sum_{i=1}^{m} \alpha_{2i} Y_{t-i} + \sum_{j=1}^{n} b_{2j} Y_{t+j} + \nu_{2t} \qquad (3.10)$$

And for the process of conducting a test for the null hypothesis; we will use the following equation:

$$H_0: b_{21} = b_{22} = . \ . \ . = b_{2n} = 0 \qquad (3.11)$$

What is quite significant here is to highlight the limitations of the Sims test (1972). The critical problem is the occurrence of autocorrelation among residuals which, according to Sims, can be handled if we use the filter $(1 - 0.75L)^2$ where 'L' is called a lag operator. But then another serious issue arises: What would be the impact on all series if it became stationary after using this filter?

For economists, one of the serious problems in using the Sims (1972) test and Granger Causality test (1969) is how to deal with non-stationary.

### 3.1.3. Final Prediction Error (FPE) Method (Hsiao 1981)

Hsiao introduced one of the relevant and intelligent compositions to handle the limitations of the Grangers Causality test. Hsiao (1981) merged GC test with Akaike

Final Prediction Error (FPE) criterion. The first step of the process involved the use of the regressed variable by way of a one-dimensional autoregressive procedure. After it regresses a variable only on its lagged then it is capable to calculate the FPE.

In the second step of the process, regress a variable on its lag, plus lags of explanatory variable, and then calculate its FPE. Now, find out that the FPE of the second step is far less than the first step FPE then this will conclude that the causal relationship exists from the explanatory variable to the explained variable under such scenario.

The procedure can repeat the same process to examine the GC (1969) among three variables. It is significant to highlight here that both assumptions and methodology will be similar to that of Sims test (1972). However, the results reflect that FPE will minimize the mean square prediction error, further decreasing the uncertainty at the significance level while using the optimality criterion. Furthermore, Hsiao believed that additional variables are substantially allowed in this method (Hsiao 1981).

Now represent the Hsiao method in the form of the equations. As we can see that in the first step, we have to estimate the following autoregressive equation having this particular form:

$$Y_t = \alpha_0 + \sum_{i=1}^{m} \alpha_{1i} Y_{t-i} + \nu_{1t} \quad (3.12)$$

Now here selected "m" to the greatest extent possible. The FPE was then calculated in the following manner for each regression;

$$FPE_{(m)} = \frac{T+m+1}{T-m-1} Q(m)/T \quad (3.13)$$

In the above equation, we have 'T', the number of observations utilized, m' is the lag order ranging from 1 to m, and Q (m) is the related sum of squared residuals. Assume that the precise value of m, say $m^*$, is the optimal lag length which results in the lowest FPE.

Now in the second stage, treat 'Y' as the regressed variable with the optimal lag order set at $m^*$ and 'X' is regarded here as an regressor variable with the order of lags ranging from 1 to n. After it, have to run the regression of the following:

$$Y_t = \alpha_1 + \sum_{i=1}^{m^*} \alpha_{1i} Y_{t-i} + \sum_{j=1}^{n} b_{1j} X_{t-j} + \nu_{2t} \quad (3.14)$$

As we can see that the corresponding two-dimensional FPE will come out to be:

$$FPE_{(m,n)} = \frac{T + m^* + n + 1}{T - m^* - n - 1} Q(m,n)/T \qquad (3.15)$$

In the above equation, 'n' is the lag order of 'X'. We have witnessed here that once again, the optimal lag order of 'n' say "$n^*$" is picked to reduce FPE (m, n). Hence, this procedure concludes here that X has Granger causality to Y only if *FPE* $(m^*, n^*)$<*FPE* $(m^*)$.

Furthermore, repeat the same process for the following regression lines if you want to run the GC test (1969) between three variables.

| Restricted equation | Unrestricted equation |
|---|---|
| $Y_t = \alpha_1 + lagged\ (Y_t) + \nu_{2t}$ (3.16) | $Y_t = \alpha_1 + lagged\ (Z_{t,} Y_t) + \nu_{2t}$ |
| $Y_t = \alpha_1 + lagged\ (Y_t) + \nu_{2t}$ (3.17) | $Y_t = \alpha_1 + lagged\ (X_{t,} Y_t) + \nu_{2t}$ |
| $Z_t = \alpha_2 + lagged\ (Z_t) + \nu_{3t}$ (3.18) | $Z_t = \alpha_2 + lagged\ (Z_{t,} Y_t) + \nu_{3t}$ |
| $Z_t = \alpha_3 + lagged\ (Z_t) + \nu_{4t}$ (3.19) | $Z_t = \alpha_3 + lagged\ (Z_{t,}, X_t) + \nu_{4t}$ |
| $X_t = \alpha_4 + lagged\ (X_t) + \nu_{5t}$ (3.20) | $X_t = \alpha_4 + lagged\ (Z_{t,}, X_t) + \nu_{5t}$ |
| $X_t = \alpha_5 + lagged\ (X_t) + \nu_{6t}$ (3.21) | $X_t = \alpha_5 + lagged\ (Y_{t,}, X_t) + \nu_{6t}$ |

As we have written above estimated unrestricted and restricted equations, we found out that the associating minimum FPE for precise values of m and n are present, hence we can appeal conclusions. We have witnessed that in all of the above-mentioned regressions, errors are white noise and all variables have been used stationary at their levels.

### 3.1.4. VAR Approach to Causality (1980)

Sims (1980) was able to develop the VAR model out of reactionary attitude against the methodology of the Cowles Commission. There is no doubt in it that VAR is closely related to Granger analysis from the perspective of causal relations. The technicality of using VAR is not much but the problem lies in applying it to the policy analysis scenarios.

Let us begin using the SVAR model for the following form:

$\Gamma Y_t = B(L)Y_{t-1} + e_t$ (3.22)

In the above equation, $Y_{t-1}$ is $n_1^*$ is called the vector of contemporaneous variables, $\Gamma$ and $(L)$ represent $n*n$ matrix and polynomial in the lag operator respectively. Also, note that $e_t$is $n*1$ vector of uncorrelated disturbance as the covariance matrix $\Sigma$ is diagonal which means that it contains zero elements. Furthermore, the matrix $\Gamma$ defines the causal interrelationship among the contemporaneous variables. Therefore, the equation reflects that the system is identified provided that there are $(n-1)/2$ zero restrictions on $\Gamma$.

Now, by Multiplying $\Gamma^{-1}$ on both sides of equation (3.22) yield reduce-form and VAR will come out to be:

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

$$Y_t = \beta^*(L)Y_{t-1} + \mu_t \qquad (3.23)$$

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

From the above equation, we identified that the problem with VAR (equation 3.23) is that covariance matrix $\Sigma$ is not diagonal which means that it does not contain any zero elements; which automatically means that the error terms are correlated with each other. This also establishes the fact that the shock in one will become shocked for both. For this purpose, Sims (1980) came up with a solution that he called Cholesky decomposition. The basic purpose of Cholesky decomposition was to orthogonalize the shocks by providing the choice of recursive order (Bhattacharjee, Schwer et al. 2003).

The exposition of this solution initiated a new debate. Some of the famous theorists including Leamer (1985), Cooley, and LeRoy (1985) condemned by saying that if we have to use recursive then only results will decide their functional usage in any scenario and that they called as Impulsive Response Function. They provided evidence for their stance by saying that any kind of meaningful economic interpretation will first definitely require the identification of Γ*(1986).*

Sims accepted criticism (1982, 1986) and then he introduced SVAR that was identified only through the usage of the contemporaneous causal order. In this method, we can interpret the results of SVAR only after using the impulse response function. It is analyzed as if the impulse response function of one variable say "$x$" to another variable "$y$" comes out to be significant, then this implies that $x$ cause $y$. (Hoover, Demiralp et al. 2006)

### 3.1.4.1. Identification Problem

Now when we knew Γ matrix, then in this case we say that the identification problem is reduced and SVAR in equation (3.22) can be easily recovered from VAR but at the same time we can say that the covariance matrix is no more diagonal. In case, we don't know the matrix, then we can impose restrictions on it so that we get identification.

Now, let see if we have to achieve identification, first, we have to make the covariance matrix$\Omega = E(P_{i-1}U(P_{i-1})')$ of (equation (3.22)) diagonal through orthogonalizing the transformation. For instance, let us say that P = {Pi} is a set of orthogonalizing transformations. (Hoover, Demiralp et al. 2006)

What is significant to highlight here is that the identification issue in SVAR is that if we do not have any information about matrix Γ, then selecting one $P$ from a set of $n$ that correspond to true data generating process: ($Pi$ = Γ) will become quite complex. Therefore, to avoid this problem, we have to impose ($n$−1)/2 restrictions on $Pi$.

There are various ways of imposing restrictions in this particular scenario. One way is to only identify the problem by using Cholesky decomposition. Other ways include the Blanchard method and Single Cholesky Ordering. (Hoover, Demiralp et al. 2006)

Furthermore, Hoover-Dempsey, Walker et al. (2005) justified his argument by saying that in the formal economic theory the identification of the causal order is obvious and that is why most of the researchers select the order arbitrarily just to get identified SVAR. But this approach can result in over-identified causal orderings for which we have identification of $Pi$ and that is more than $n(n-1)/2$, which will further imply that there are zero restrictions that we can impose on $Pi$. This scenario will make it impossible to apply Cholesky decomposition.

Hence, the solution to this problem was demanded which eventually leads to the development of the new method for selecting $Pi$ Pearl (2000), Spirtes, Glymour et al. (2000) and Spirtes, Glymour et al. (1993), and this method is called an Error Correction Mechanism (Phillips 1986).

### 3.1.5. Error Correction Mechanism (1986)

It is critical to realize that both the GC test (1969) and Sims (1972) do not include any error correction term in their methodology because it can lead to misleading results. Originally, Phillips introduced the error correction model (Phillips 1986). The first step of this model identifies if the underlying variables are stationary or not. When we know that the variables are stationary, we can apply the GC test (1969). However, if they have a unit root problem, then to sort this problem out, we can apply Augmented Dickey-Fuller (ADF) test to make our variables stationary. This test helps in determining the order of integration of non-stationary variables and if we came to know that variables are stationary at the same order, only then we continue to carry out this test for cointegration (Phillips 1986).

In order to carry out Co-integration analysis, it is preferred to use the Johansen method and Granger method. The reason is that when we see that the variables are found to be cointegrated with each other, then this means that the variables have a long-run association. And it called error correction mechanism (ECM).

In ECM, the coefficient of the error term's lagged value will define the significance of the long run relationship and the coefficients of the lagged independent variable that is capable of examining the short-run dynamics among the variables.

For instance, if two variables have a long-run equilibrium relationship, we say that ECM is appropriate for them rather than using the simple Vector Autoregressive Model. (Imbens and Wooldridge 2009)

The mechanism is followed in this way. The first step involves the identification of the state of variables. We can proceed only if the variables are stationary. The second step involves the application of the Granger causality test. However, we must know that if the variables come out to be non-stationary, then we have to use Augmented-Dickey Fuller (ADF) test to decide the order of integration. Furthermore, if found out that variables are integrated at the similar order; then there continue to carry out the check of cointegration. Johansen method and the Granger method known as the most famous techniques for analysis of cointegration. (Bhattacharjee, Schwer et al. 2003)

According to the researchers, it is believed that an additional channel of causality is exposed. Hence, we found that the significance of long run relationship is determined by the Coefficients of the lagged value of the error term and it is also found that the coefficients of the lagged independent variables measure only short run dynamics (Imbens and Wooldridge 2009).

In the Granger theorem, if variables X and Y have a long-run relationship, we can say that the relationship among the two variables can be stated as ECM. Let's assume the two variables are exposed to be cointegrated; ECM should be used (Imbens and Wooldridge 2009). Let us consider an ECM model in three variables case:

$$\Delta Y_t = \alpha_0 + b_0\mu_{t-1} + \sum_{i=1}^{p}\delta_{1i}\Delta Y_{t-i} + \sum_{j=1}^{p}\theta_{1j}\Delta X_{t-j} + \sum_{k=1}^{p}\lambda_{1k}\Delta Z_{t-k} + w_{1t} \quad (3.24)$$

$$\Delta X_t = \alpha_1 + b_1\mu_{t-1} + \sum_{i=1}^{p}\delta_{2i}\Delta Y_{t-i} + \sum_{j=1}^{p}\theta_{2j}\Delta X_{t-j} + \sum_{k=1}^{p}\lambda_{2k}\Delta Z_{t-k} + w_{2t} \quad (3.25)$$

$$\Delta Z_t = \alpha_2 + b_2\mu_{t-1} + \sum_{i=1}^{p}\delta_{3i}\Delta Y_{t-i} + \sum_{j=1}^{p}\theta_{3j}\Delta X_{t-j} + \sum_{k=1}^{p}\lambda_{3k}\Delta Z_{t-k} + w_{3t} \quad (3.26)$$

In the above equations, the first difference operator is "Δ", expressed as $\Delta Y_t = Y_t - Y_{t-1}$. And $\mu_t$ is expressed as '$\mu_t = y_t - \alpha_1 x_t - \beta_1 z_t$ and it is presumed that the error-terms $w_{1t}, w_{2t}, and\ w_{3t}$ are white-noise. Now we can witness that two probable sources of causality may be found in the ECM above.

Variations in $X_t$ are partly driven by $\mu_{t-1}$ and also lagged values of the two variables $Y_t$ and $Z_t$. For instance, in the equation (3.24) it leads to 'X' Granger causation 'Y' if $b \neq 0$ or $\theta_{11} = \theta_{12} = \theta_{1p} \neq 0$.

Thus, we can also argue that using an F-test to examine the $H_{0a}$ that 'X' does not Granger causes 'Y' by equation (3.24) is equal to using an F-test to check the hypothesis in the following manner;

$$H_{0a}: \theta_{11} = \theta_{12} = \theta_{1p} = 0$$

After applying a t-test to examine the hypothesis: $H_{0b}: b_0 = 0$

Now, we see that If either of the null hypothesis $H_{0a}$or $H_{0b}$is denied, leads to 'X' Granger causation 'Y'. If both hypotheses are true in another scenario, we may conclude that 'X' does not cause 'Y'. Noticeably, ECT '$\mu_{t-1}$' depicts the long-run effect of one variable on the other. In contrast, the short-run influence of one variable on another variable is quantified by utilizing the variable's lagged values. As a result, when variables have cointegration relationship, researchers show that we must prefer to apply ECM instead of just simple VAR model with first difference in order to evade deceptive results. (Imbens and Wooldridge 2009)

### 3.1.6. Toda and Yamamoto method (1995)

This method was proposed to estimate the VAR model expressed at levels and check some common conditions on the use of matrices' parameter irrespective of fact that if the process is cointegrated or integrated of arbitrary order. (Toda and Yamamoto 1995)

It must be understood that ECM and Granger methods are chosen based on prior information about the cointegration properties and about the integration of a series. Sometimes, there is no prior information about the variables have long-run relationship, (trend) stationary, or integrated, in most applications. Before estimating the VAR model generally required a pre-test for non-stationarity and cointegration in various economic time series, in which statistical inferences are formulated.

What is interesting about Toda and Yamamoto (1995) method is that they employed extended Wald test for the need of constraints for the coefficients of the VAR (k) model wherein "k" is the lag order depicting of this particular method. Toda and Yamamoto (1995) validated that test is followed by an asymptotically chi-square distribution only if a VAR (k+$d_{max}$) model is assessed wherein $d_{max}$ the maximum order of integration is assumed to be part of this method.

In this scenario, the lag length is taken by applying AIC and SBC criterion. Moreover, we have to assume that error terms are white noise. Now, what is so different about this method is that the information of the integration properties is not a requisite here. The specialty of this test is that variables need not be stationary or having long-run equilibrium (Zapata and Rambaldi 1997).

Let us reflect following the VAR (k+$d_{max}$) model for two variables circumstances:

$$Y_t = \alpha_0 + \sum_{i=1}^{k} \delta_{1i} Y_{t-i} + \sum_{j=k+1}^{d_{max}} \alpha_{1j} Y_{t-j} + \sum_{j=1}^{k} \theta_{1j} X_{t-j} + \sum_{j=k+1}^{d_{max}} \beta_{1j} X_{t-j} + w_{1t} \quad (3.27)$$

$$X_t = \alpha_1 + \sum_{i=1}^{k} \delta_{2i} Y_{t-i} + \sum_{j=k+1}^{d_{max}} \alpha_{2j} Y_{t-j} + \sum_{j=1}^{k} \theta_{2j} X_{t-j} + \sum_{j=k+1}^{d_{max}} \beta_{2j} X_{t-j} + w_{2t} \quad (3.28)$$

In above equation, the error terms $w_{1t}$ *and* $w_{2t}$ are uncorrelated between equations and within equation, $d_{max}$ is the maximum order of integration.

By applying AIC and SBC, the lag length in the above two equations can be determined.

In the equation (3.27) 'X' granger causes 'Y', using modified Wald statistic, we can assess the following $H_0$ in the equation (3.27 and 3.28):

$$H_0 : \theta_{11} = \theta_{12} = \cdots = \theta_{1k} = 0 \ \ (X\ does\ not\ Granger\ cause\ Y)$$

$$H_0 : \alpha_{11} = \alpha_{12} = \cdots = \alpha_{1k} = 0 \ \ (Y\ does\ not\ Granger\ cause\ X)$$

### 3.1.7. Structural Causality Test

The Structural Causality Test was introduced because the GC test (1969) could not represent one another aspect of Causality. This concept of causality is borrowed from Hendry and Mizon (1998), Simon and Shemin (1953), and Hoover (2001). The benefit of structural causality as compare to other methods, it includes all possible facts before getting of causal findings and not at all based on the Statistical methods. However, there is one obvious limitation: it cannot compute the magnitude of the impact of a casual variable on the dependent variable. And it cannot be allowed according to our limited understanding of the statistical methods. It further implies that correct causal relations can persist various kinds of structural variation.

Furthermore, this knowledge can be utilized to distinguish among models those are not in the period of structural change and those which are causally correct (Asghar and Abid 2007). This further means that in periods where we have no structural changes and stability, then in such situations models will perform well with incorrect causality (Simon 1953).

Moreover, this method for investigating causality involves much more analysis in the fundamental economic mechanism not merely on the theoretical framework but it traces its roots in the historical perspective. (Asghar and Abid 2007)

Now build on the foundation of the structural causality detection technique. Let (*X*, *Y*) be a sequence of observations with *N* ($\boldsymbol{\mu_t}, \Sigma$) as the *i.i.d.* distribution. There are three approaches to create a series of observations like this.

The first method is elaborated as follows:

Let (V, W) is i.i.d $N\ (0, I_2)$. It is feasible to decompose as $\Sigma = \mathrm{UU}$, since $\Sigma$ is a 2x2 positive definite matrix. The linear transformation may be used to create (X, Y) from (V, W):

$$\binom{X}{Y} = \mu + U\binom{V}{W}$$

Then, according to the required distribution, X and Y will be jointly normal. It is observed that if this is the process of generating data worldwide, then we can say that hidden variables V and W jointly cause X and Y. Moreover, we may also say that neither variable causes the other.

The second method is explained as follows:

Generated X from its marginal distribution $X \sim N(\mu_t, \Sigma_{11})$. Then, generated Y from its conditional $Y/X \sim N(\mu_2 + \Sigma_{21}\Sigma_{22}^{-1}(X - \mu_1), \Sigma_{11} - \Sigma_{21}\Sigma_{22}^{-1}\Sigma_{21})$. It can be written in the regression model as follows:

$$Y = \alpha + bY + \varepsilon \qquad (3.29)$$

In this case scenario, we may claim that X is the cause of Y in (X, Y), and that variations in *X* will induce changes in *Y,* but it is not the case the other way around as proved in the above equation.

The third method can be described as follows:

Let's make Y and X from the marginal and conditional, respectively. In this case scenario, we say that *Y* is the cause of *X*.

Henceforth, we witnessed that all three techniques provide data with the same distribution, observationally identical. This is the apparent reason for the complexity of evaluating causality from observational data. Let us suppose for some reasons like structural change that takes place, which further results in varying the marginal distribution of X's mean and variance. We can see that all three models will now act in distinct ways in such scenarios.

For instance, the mean and variance of X have changed in the first model, the relationship between X and (V, W) must fluctuate. As a result, the equation X=U11V+ U12W affects the parameters U11 and U12. Similarly, the parameters of the X distribution will vary in the second model, but the regression model will remain unchanged, implying that the conditional distribution of Y is given X.

Moreover, we must first assume that the structural change does not affect the marginal distribution of Y in the third model.

$$X = c + dY + \varpi \qquad (3.30)$$

As a result, we can observe that a type relationship cannot remain constant since it must alter to match the change in X's distribution.

Resultantly, the structural change in this association will be evident in the estimations. Even though the structural change has an effect on the distribution of *Y*, it is common for such shifts to cause instability in the conditional distribution of *X* given *Y*.

From the perspective of policymaking, there is a need to search shifts in distributions due to exogenous causation that might also include extra statistical information for its applicability. For instance, we might also need historical information to take interventions in policy-making regimes.

Hoover (2001) intervention has its roots in the historical perspectives and the statistical tests are carried out only to validate the necessity of intervention. Hoover (2001) and Freedman (1991) pointed out that to determine the causal direction, one must acquire detailed and extensive knowledge of the issue to be resolved.

The application of finding the intervention time is significant. Suppose if we know that one variable's intervention time is being estimated, then we can also work on marginal and conditional distributions of the variable. Thus, finding casual direction involves applying some statistical test to identify intervention. Then after the confirmation of the chronological intervention, we can apply regression on two different data sets. Furthermore, we can identify stable conditional distribution in the casual relation provided that intervention existed for only one variable. (Asghar and Abid 2007)

#### 3.1.8. Graph-Theoretic Approach and Fast Causal Inference (FCI) (Pearl (2000) and Spirtes, Glymour et al. (2000)

Another renowned approach is the Graph Theoretic approach that allows converting the structural model into graphs with affinity to overcome many problems. There are two ways of it: Firstly, GC test (1969) applies to only a tiny set of pre-specified and reduced-form equations for small relationships (Granger 1969). Second, as argued by Perez et al (2006) this particular approach can help in identifying true regressors.

Furthermore, Spirtes, Glymour et al. (2000) supported this method by saying that incorrect independent variables can mislead casual inferences. Thirdly, through the graph-theoretic approach, we can evaluate causal orderings that are primarily determined on the basis of data properties and correlation. Finally, the method also involved various kinds of causal search algorithms.

According to the Cowles Commission, econometric model is the combination of two parts: a) Probability distribution of variable, b) Causal structure as we can observe it in different researches as well. Moreover, Pearl (2000) and Spirtes, Glymour et al. (2000) represented that because their isomorphism present between graphs and probability distribution of variables; therefore, we can draw conclusions about probability distributions. Also, these conclusions are further proved by employing mathematical techniques as indicated in the theoretical paradigms of the graph theory. (Spirits et al. 2000)

But it is noteworthy to reflect on how this technique was not utilized for the time series data. But, in 1997, Swanson and Granger's used this Graph-theoretic approach to understand casual order of SVAR. They assumed that information about the causal ordering of variables of SVAR is actually existed in the covariance matrix of VAR. Hence, after estimating the VAR model, that model's error terms will be treated as the original time series variables, Demiralp and Hoover (2003), (Hoover-Dempsey, Walker et al. 2005). Whereas VAR residuals could bring only current information about cross variable effect.

Let us consider the following VAR model:

$$Y_t = \alpha_1 + \beta_1 X_{t-1} + \beta_2 Y_{t-1} + \varepsilon_{1t} \quad (3.31)$$
$$X_t = \alpha_2 + \beta_3 X_{t-1} + \beta_4 Y_{t-1} + \varepsilon_{2t} \quad (3.32)$$

After estimating this VAR model, we will extract residuals series of both equations (3.31) and (3.32). While keeping in mind that residual series extracted from equation (3.31) only effect of $xt$ could be there, while effect of past values ($X_{t-i} where i$>1) are removed. Thus, it is proved that VAR residuals contain only information about the causal feedback from X to Y and vice versa as well.

Contrary to this scenario, there also exist numbers of univariate methods which are capable of eliminating the non-stationarity even without purging out the effect of past values. However, the power of the residuals extracted from the univariate methods is utilized to determine the causal ordering with the help of PC algorithms of GTA. This research study is intended to modify the original PC algorithm by replacing VAR residuals with univariate models' residuals. And this also transform the name of the algorithm to be Modified PC algorithm. (Haugh (1976); Leong, Rehman et al. (2014)

The FCI approach is a variant of the PC algorithm. (Haugh (1976); Leong, Rehman et al. (2014). One method to causal inference from observational data uses directed acyclic graphs (DAGs) to describe causal linkages and makes assumptions about how the causal DAG structure relates to probability distributions. Finding the "best" DAG or collection of DAGs for a given sample becomes the challenge of establishing causal inferences (Spirtes, Glymour, and Scheines 1993, Heckerman, Meek, and Cooper 1999).

Different causal structures imply different independence relationships at the heart of the constraint-based causal discovery process. The causal connection

A→B→C, for example, suggests that variable A is independent of variable C given B. When A→C←B, on the other hand, A and B are independent (unconditionally), but conditional on C, they become dependent. The "V" structure (also known as collider) is the latter structure, and it has a unique independence connection as compared to other causal interactions. In reality, constraint-based algorithms, such as FCI, seek for it as one of the "primitives." FCI distinguishes out including among constraint-based techniques for its ability to discover latent (unobserved) confounders. This is made feasible by another primitive, the "Y" structure. A "Y" structure is formed when four variables demonstrate the following causal connections: W1 X W2 and X Y. Within the "Y" structure, both W1 and W2 are conditional on X and independent of Y. This conditional independence eliminates the possibility of an unmeasured confusion between X and Y. To put it another way, if FCI finds a "Y" structure in the graph, the causal relationship from X to Y is certain to be un-confounded; otherwise, FCI suggests the presence of unknown confounders. (Spirtes, Glymour et al. 2000)

The FCI algorithm is a fast causal inference technique. FCI creates a causal graph by removing connections that connect conditionally independent variables from a fully connected undirected graph. It orients edges in the second phase by recognizing "V" and "Y" structures and attempting to orient the remaining edges using a set of principles that have been detailed elsewhere.

### 3.1.9. Maximum Entropy Bootstrap Approach

In order to address the problems of mixed results about casual direction in the small samples, a new technique was introduced called the modern maximal entropy bootstrap (meboot) method. It provided better and robust results for large sample sizes (Aqil, Aziz et al. 2014). There were different reasons for using this approach. For instance, in case of social problems null hypothesis is rejected as it can mislead researchers. Similarly, in order to transform the state of data, the researchers often de-trend the original data which can lead to variations in time data. (Yalta (2011); Khan, Ahmed et al. (2019))

Few other drawbacks also include loss/reduction of efficiency, misspecification and inappropriateness with the structural changes and so on that can truly harm the actual condition of the data (Hamilton and Susmel 1994). Moreover, these faults and errors can also lead us directly to the wrong basis of the knowledge (Vinod (2006); Ahmed, Riaz et al. (2015)).

In order to cope with such scenarios, as an alternative, Vinod (2006) proposed the entropy bootstrap (meboot) method.

However, this research study elaborated that the simulation-based meboot method can explain the causal relationship. The constructions of a population of time series that ensembles a high number of times, say _= 999, are among the features of the meboot algorithm, with the use of new computer-intensive techniques for which we see that each potential variable's data series is highly dependent, non-stationary, has spikes and gaps, and has discontinuities or regime shifts. (Ahmed, Riaz, et al. 2015). Henceforth, we can state that each parameter coefficient estimate will have 999 coefficients, from which $(1 - \alpha)$ 100 percent confidence intervals will be derived, and this will be referred to as a high-density zone. (Hyndman, Bashtannyk et al. 1996).

## 3.2. Causality Methods in Cross Sectional

### 3.2.1. Regression Discontinuity Design (1960)

Regression Discontinuity Design (RDD) was introduced by Thistlethwaite and Campbell (1960) in a non-experimental setting to estimate the treatment effects. In a way the effectiveness of a treatment is determined by whether an observable “assignment” variable exceeds a predetermined cutoff threshold.

RDD is a quasi-experimental evaluation method that uses a treatment assignment mechanism to calculate the impact of an intervention. It's based on a continuous eligibility index, which is a continuous distribution variable. And when it's used, the results are consistent to those of a randomized trial in terms of causal direction. It is used for finding out a treatment be it effective or not. The estimates of RDD made by parametric models and also close to a parallel set of nonparametric estimates. For the choice of bandwidth over a wide range, these nonparametric estimates are most of the time come out to be insenstive. (Angrist and Pischke 2014)

There are two styles of RDD one is sharp and the other is fuzzy. The former is basically a selection on-observables story while the latter is an instrumental variable setup. (Angrist and Pischke 2014)

Regression discontinuity (RD) research designs function by exploiting brief knowledge of the rules that can play role in determining the treatment. RD identification primarily functions on the notion that some rules are arbitrary and such rules can be utilized in providing good experiments. The assumption made in this case scenario is that treatment is assigned on the observable variable or index criterion. Moreover, there is also present discontinuity in the probability which can lead to cut off of some variables and the treatment for such index is taken out to be arbitrary (Angrist and Pischke 2014).

Sharp RD is employed when treatment is assigned based on a distinct cut-off point, with all eligible people getting it and all ineligible people not getting it. In order to calculate the effect of the treatment of Sharp RD, we have to simply take comparison of means. If we consider treatment status as a deterministic and discontinuous function of such a covariate, X i, the equation is:

$$D_i = \begin{cases} 1 & if\ x_i \geq x_0 \\ 0 & if\ x_i < x_0 \end{cases}$$

In the above equation, $x_0$ is called a threshold or cutoff.

This mechanism is a deterministic function of a covariate, $X_i$ because we know Di once we know$x_i$. It's a discontinuous function since treatment remains unaltered until $x_i$=$x_0$, regardless of how close $X_i$ comes to $x_0$. However, in the case of Fuzzy RD some qualified assignments do not receive treatment, while others who are ineligible do.

Because of two reasons for it: one is self-selection and the other could be administrative overrides. Practically this type of version is more appropriate. In order to find out the effect of the treatment, the formula for this purpose is:

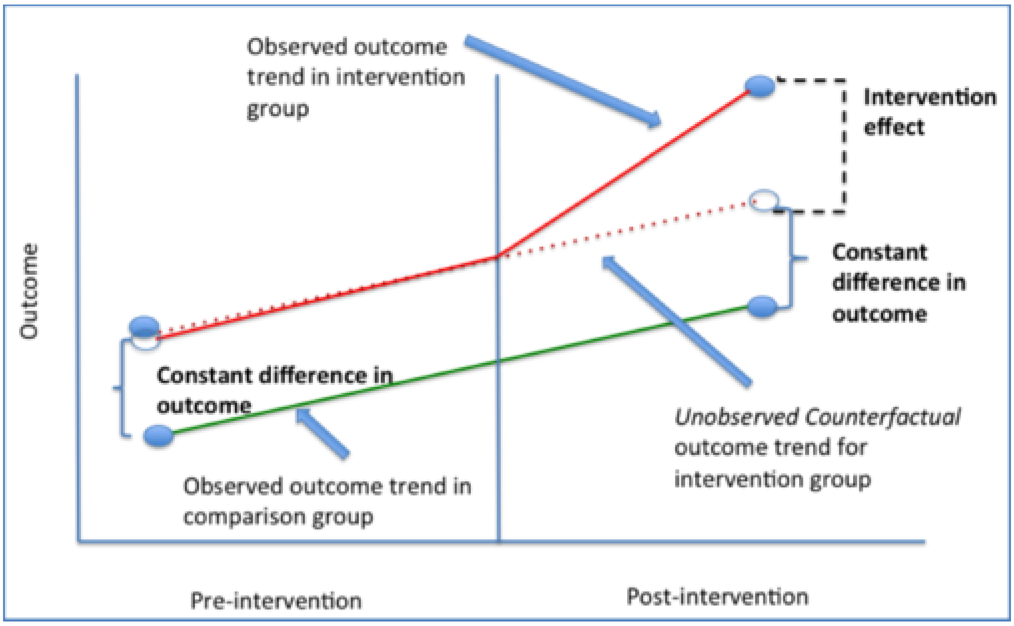
Treatment effect = outcome discontinuity / treatment discontinuity (Angrist and Pischke 2014).

### 3.2.2. Difference in Differences Approach (2005)

Economists utilize one another famous methodology and it called as The Difference in Difference (DID) approach, which was initially developed by John Snowy in 1850's. Later on, it was called controlled before and after study methodology. A quasi-experimental design operates by using control and treatment groups from longitudinal panel data to get casual direction. The technique is specifically designed to find out estimation of the impact of a specific treatment or intervention. (Angrist and Pischke 2008)

This technique is quite interesting as it works on the data from the pre and post treatment or intervention while ensuring that there will not be any biases left in the post interventions. DID is often used to estimate the treatment impact on the treated, but it may also be used to calculate the Average Treatment Effect if robust assumptions are applied. To determine casual effect, three assumptions needed to be taken: the first assumption is positivity; the second assumption is exchangeability, and the last assumption is Stable Unit Treatment Value Assumption (Lechner 2011). After fulfilling these assumptions, it can be continued to identify causal direction through observational data. It is a useful technique for comparison groups may start at changed stages of the result. It is also used for individual and group level data. It is one of the limitations that cannot apply if treatment allocation defined by baseline outcome. If comparisons group contain different conclusion trend. When fluctuations are not stable between pre and post group, it is not applicable. (Abadie 2005).

**Figure 1: Difference in Differences Approach**




Source: (Mastering Matrices, Joshua D.Angrist and Jorn Steffen Pischke, 2015)

DID is employed in observational studies where the treatment and control groups cannot be presumed to be identical. DID is based on a weaker exchangeability assumption, which states that the hidden differences between the control groups and treatment remain the same across time in the absence of treatment. (Bertrand, Duflo et al. 2004).

Though there is no statistical test for this assumption, visual inspection is beneficial when several time points of observation. It's also been suggested that the shorter the time period under examination, the more probable the assumption will hold true. If the parallel trend assumption is violated, the causal effect will be misestimated. (Imbens and Wooldridge 2009).

DID is normally implemented as source of interaction term among time and treatment group dummy variables in a regression line; let us consider the following equation:

$$Y = \beta_0 + \beta_1 * [Time] + \beta_2 * [Intervention] + \beta_3[Time * Intervention] + \beta_4[Covariates] + \varepsilon \qquad (3.33)$$

Numerous things favor the technique: the first benefit is that interpretation is based on the institution, one can easily guess the idea. The second benefit is that if the assumptions are fulfilled, we can easily identify the causal effect from it. The third great thing is that we have choice to use either individual or group level data. The fourth great benefit is that we can chose comparison groups at various levels of the outcome. And the last strength of the method is that we can change factors because of intervention. (Abadie 2005)

Contrary to this, there are different limitations of the method as well. The first limitation is that it demands baseline data and a non-intervention group form. Moreover, if the intervention is baseline outcome, then we cannot use this method. Thirdly, if the comparison groups are having different outcomes, we cannot preferably use this method and lastly if the comparison groups appear to be unstable (Abadie 2005).

### 3.2.3. Cross Correlation Based Methodology

The purpose of correlation analysis is to determine the power of linear relationships among two variables. The method does follow some assumptions. The variables are supposedly random, and observations have to be independent. The same assumptions must be made when evaluating the null hypothesis that the association is zero but both variables should be normally distributed in order to evaluate the correlation coefficient's confidence intervals. However, in the actual world, a nonlinear connection might exist between variables that is inadequately captured by the correlation coefficient and may even go unnoticed. The limitation of the method is that one might get outliers or we can say extreme values in the data, which are not involved in the correlation analysis methodology (White and Peterson 1994).

## 3.3. Causality Methods in Panel Data

### 3.3.1. Bootstrap Panel Granger Causality Test by Konya (2006)

Konya proposed Bootstrap Panel Granger Causality test because it contains slope heterogeneity and cross-sectional dependence. The procedure developed by Kónya (2006) enables the identification of particular countries where the Granger causal relationship exists. The method projected three vital advantages. The first and foremost significance is that it involves seemingly unrelated regression (SUR). The second advantage is that it is based on Wald tests and country-specific bootstrap values critical for the researchers. And the third big advantage is that it does not involve any pretesting for the purpose of co-integration (Kónya 2006).

However, it is mandatory to remember this fact that the results of this method are only focused on short term causality. Before Kónya (2006) approach is briefly presented, we sketch the outline of tests for cross-sectional dependence. The assessment of cross-sectional dependence is of prime importance. The reasons for the presence of the cross-sectional dependence are that in case of Panel data models there are quiet many possibilities of shocks and unobserved data., general residual interdependence, spatial spillovers and so on so forth (Kónya 2006). One reason for this may be connected with the fact that during the previous few decades. Countries and financial institutions have become more integrated economically and financially, which induces strong interdependencies between cross-sectional units. Breitung and

Pesaran (2008) and Bai and Kao (2006) proposed that it has been observed in the case of the default assumption of independence that occur between cross-sections of the cointegration and causality analysis, there is apparent inadequacy.

Cross-sectional dependency is likely to emerge if economic ties between nations are quite strong. As a result, erroneous causal conclusions may result from improper cross-sectional assumptions of independence. As a result, we first opt to test the hypothesis of cross-sectional dependency, considering regularly found cross-sectional dependence in panel data for macroeconomic data. In order to assess the presence of cross-sectional reliance in our data, we will need to use cross-sectional dependency tests established by Pesaran, Schuermann et al. (2004) in the presence of a null hypothesis claiming no cross-sectional dependence. Furthermore, we must express data based on two groups in Kónya (2006) panel causality method frameworks. For example, consider the following equations:

$$y_{1,t} = \alpha_{1,1} + \sum_{l=1}^{mly_1} \beta_{1,1} y_{1,t-l} + \sum_{l=1}^{mlx_1} \delta_{1,1,l} x_{1,t-l} + \sum_{l=1}^{mlz_1} \gamma_{1,1,l} z_{1,t-l} + \sum_{l=1}^{mlv_1} \vartheta_{1,1,j} v_{1,t-l} + \varepsilon_{1,1t} ,$$

$$y_{2,t} = \alpha_{1,2} + \sum_{l=1}^{mly_1} \beta_{1,2} y_{2,t-l} + \sum_{l=1}^{mlx_1} \delta_{1,2,l} x_{2,t-l} + \sum_{l=1}^{mlz_1} \gamma_{1,2,l} z_{2,t-l} + \sum_{l=1}^{mlv_1} \vartheta_{1,2,j} v_{2,t-l} + \varepsilon_{1,2t} \quad (3.34)$$

$$y_{N,t} = \alpha_{1,N} + \sum_{l=1}^{mly_1} \beta_{1,N} y_{N,t-l} + \sum_{l=1}^{mlx_1} \delta_{1,N,l} x_{N,t-l} + \sum_{l=1}^{mlz_1} \gamma_{1,N,l} z_{N,t-l} + \sum_{l=1}^{mlv_1} \vartheta_{1,N,j} v_{N,t-l} + \varepsilon_{1,Nt},$$

$$x_{1,t} = \alpha_{2,1} + \sum_{l=1}^{mly_2} \beta_{2,1} y_{1,t-l} + \sum_{l=1}^{mlx_2} \delta_{2,1,l} x_{1,t-l} + \sum_{l=1}^{mlz_2} \gamma_{2,1,l} z_{1,t-l} + \sum_{l=1}^{mlv_2} \vartheta_{2,1,j} v_{1,t-l} + \varepsilon_{2,1t} ,$$

$$x_{2,t} = \alpha_{2,2} + \sum_{l=1}^{mly_2} \beta_{2,2} y_{2,t-l} + \sum_{l=1}^{mlx_2} \delta_{2,2,l} x_{2,t-l} + \sum_{l=1}^{mlz_2} \gamma_{2,2,l} z_{2,t-l} + \sum_{l=1}^{mlv_2} \vartheta_{2,2,j} v_{2,t-l} + \varepsilon_{2,2t} \quad (3.35)$$

…………

$$x_{N,t} = \alpha_{2,N} + \sum_{l=1}^{mly_2} \beta_{2,N} y_{N,t-l} + \sum_{l=1}^{mlx_2} \delta_{2,N,l} x_{N,t-l} + \sum_{l=1}^{mlz_2} \gamma_{2,N,l} z_{N,t-l} + \sum_{l=1}^{mlv_2} \vartheta_{2,N,j} v_{N,t-l} + \varepsilon_{2,Nt} \text{ ,}$$

In the above equation, we have N that denoted the number of countries in the panel (i=1, 2…, N), t is time period (t=1, 2…, T), and $l$ is the number of lags in equation $\varepsilon_{i,i,t}, \varepsilon_{2,i,t}$ are anticipated correlated concurrently across equations because of the common random shocks. As we can see that this model follows the deterministic trend, unidirectional and bi-directional Granger causality for specifically each country one by one.

The empirical distributions with the help of Wald test and the through this can also received bootstrap critical values. Moreover, we can get regressions from both statistics (Wanat, Papież et al. 2016). In Konya's method, defining the number of lags over all equations is critical. Using Konya's method, we can identify the number of lags in the following equations. We evaluate all equations and apply the Akaike Information Criterion (AIC) to choose the best solution, considering the number of lag is between 1 and 4.

Moreover, in order to get Akaike Information Criterion, use the following equation:

$$AIC_l = ln|W| + \frac{2N^2q}{T}, \qquad (3.36)$$

In the above equation, W represents the estimate residual of the covariance matrix, N represents the number of equations, q represents the number of coefficients of each equation, and T represents the sample size. (Wanat, Papież, et al. 2016)

### 3.3.2. Granger Non-Causality Test in Heterogeneous Panel Data by Hurlin (2012)

Because in today's world, we have increased demand of the macro level panel data, so there is a required new set of techniques that can help the econometricians grow and develop their skills. Hence, a famous approach was introduced by two theorists Dumitrescu and Hurlin (2012) that is primarily used to test for Granger causality in macro panel datasets.

Considering the fast evolution of the literature, practitioners may find it difficult to implement the latest econometric tests. Therefore, the test built by Dumitrescu and Hurlin (2012) summarized. This contribution aims to support the empirical literature using panel causality techniques. Granger non causality test utlizes the Akaike information criteria, Bayesian information criterion, or Hannan–Quinn information criterion to identify the optimal lag length, that is, the length where the mentioned criteria give minimum test statistic value. Finally, to address the empirical problem of cross-sectional dependency, a bootstrap technique was used to obtain p-values and critical values. (Lopez and Weber 2017)

Granger (1969) methodology is primarily established for investigating the causal relations between time series. Let us suppose $x_t$ and $y_t$ are two stationary series.

$$y_t = \alpha + \sum_{k=1}^{K} \gamma_k \, y_{t-k} + \sum_{k=1}^{K} \beta_k \, x_{t-k} + \varepsilon_t \quad with\, t = 1, \dots, T \qquad (3.37)$$

We can say that the model can be utilized to determine whether *x* causes *y*. In other words, if previous values of x are significantly predictive of the present value of y even though previous values of y would be included in the model, then x has a causal effect on y.

While using (3.37), we can easily identify causality that is centered on an *F* test with the following $H_0$:

$$H_0: \beta_1 = \cdots = \beta_k = 0 \qquad (3.38)$$

If $H_0$ is denied, one might infer that there is causation from x to y. Furthermore, the x and y variables may be swapped to test for causation in the opposite direction, and reversible causality can be seen. (Lopez and Weber 2017)

However, the approach followed by Dumitrescu and Hurlin (2012) was an extension designed specifically for detecting the causality in panel data.
We can say the following equation will represent:

$$y_{i.t} = \alpha_i + \sum_{k=1}^{K} \gamma_{ik}\, y_{i,t-k} + \sum_{k=1}^{K} \beta_{ik}\, x_{i,t-k} + \varepsilon_{i,t} \quad with\ \ i = 1, \dots, N\ and\ t = 1, \dots, T \qquad (3.39)$$

In the above equation, we have $x_{it}$ and $y_{it}$ that represent the observations of two stationary variables for individual $i$ in period $t$. Moreover, individual coefficients are permitted to differ, but they are supposed to be time invariant at the same time. It is important to mention here that the lag order $K$ in this scenario is assumed to be equal for all individuals so that the panel should be balanced.

Granger's method is used to establish causality and to assess the substantial impacts of past x values on the current value of y. Hence, we can define then null hypothesis in the following manner:

$$H_0\colon \beta_{i1} = \dots = \beta_{ik} = 0 \qquad \forall\, i = 1, \dots, N \qquad (3.40)$$

Here we can observe for all members of the panel, there is no causation. The reason for this is that in the DH test, we can suppose that causality exists for some people, but not necessary for all. Therefore, in such scenarios, we can assume some alternative hypothesis. For instance, let us consider the following alternative hypothesis:

$$H_1\colon \beta_{i1} = \dots = \beta_{ik} = 0 \qquad \forall\, i = 1, \dots, N1 \qquad (3.41)$$

$$\beta_{i1} \neq 0\ or \dots or\ \beta_{ik} \neq 0 \qquad \forall\, i = N1 + 1, \dots, N \qquad (3.42)$$

Now, we can see that in the above equation, we have $N_1 \in [0, N - 1]$ that is not known. If $N_1 = 0$, all of the individuals in the panel have a causal relationship. $N_1$ must be firmly smaller than $N$; there is no causation for all individuals, and $H_1$ reduces to $H_0$.For such scenarios, Dumitrescu and Hurlin (2012) proposed that we can run the $N$ individual regressions implicitly enclosed in (3.39), perform $F$ tests of the $K$ linear hypotheses $\beta_{i1} = \cdot\cdot\cdot = \beta_{ik} = 0$ to retrieve the individual Wald statistic $W_i$, and finally compute the average Wald statistic $\overline{W}$ ,

$$\overline{W} = \frac{1}{N} \sum_{i=1}^{N} W_i \qquad (3.43)$$

Thus, we can say that this test is primarily focused on detecting the causality at the panel level, and at the same instant, it is designed to reject $H_0$ which will not exclude non-causality for some individuals. However, one can also use Monte Carlo simulations while using $\overline{W}$ asymptotically in order to investigate panel causality (Lopez and Weber 2017). For instance, let us suppose that the assumption for Wald statistics $Wi$ are independently and identically distributed among individuals, it may be expressed as the standardized statistic where first $T \to \infty$ and then $N \to \infty$ (often read as "$T$ should be big relative to $N$") following a standard normal distribution as shown below:

$$\bar{Z} = \sqrt{\frac{N}{2K}} \times (\overline{W} - K) \quad \xrightarrow[T,N \to \infty]{d} \quad N(0,1) \qquad (3.44)$$

Also, for a fixed T dimension with T > 5+3K, the approximated standardized statistic $\tilde{Z}$ follows a standard normal distribution:

$$\tilde{Z} = \sqrt{\frac{N}{2K} \times \frac{T-3K-5}{T-2K-3}} \times \left(\frac{T-3K-3}{T-3K-1} \times \overline{W} - K\right) \quad \xrightarrow[N \to \infty]{d} \quad N(0,1) \qquad (3.45)$$

In this procedure, we can see that the null hypotheses (equation 3.40) is based on $\overline{Z}$ and $\tilde{Z}$. However, if these are pretty larger than the expected standard critical values, then we should reject $H_0$. Moreover, simply, we can conclude that Granger causality exists in this scenario. But it is important to mention here that for large N and T panel data sets, $\overline{Z}$ is a considerable option. $\tilde{Z}$ should be preferred for datasets with large N but relatively small T. Dumitrescu and Hurlin (2012) used Monte Carlo Simulations to show that they exhibit good finite sample characteristics, even when N and T small. The selection lag order (K) is a practical problem that Dumitrescu and Hurlin (2012) do not address. One solution is to use the number of lags based on an information criterion (AIC/BIC/HQIC). (Lopez and Weber 2017)

### 3.3.3. DAG Theory (Spirtes et al. 2000)

Spirtes, Glymour et al. (2000) developed this DAG theory which is impactful and favorable for causality analysis. Much research has been done on the financial markets while using this approach. (Awokuse and Bessler (2003); Bessler and Yang (2003); Yang, Chang et al. (2006); Yang and Bessler (2008)). However, we have witnessed the DAG approach application in energy economics in the recent past. One of the study conducted by Zhou, Li et al. (2014) was focused on the relationship between economic growth, energy consumption, and carbon emissions, and the country focused was India. The successful use of DAG theory supported the researchers' passion who were able to derive the results while using this technique. Cui, Feng et al. (2015) used DAG theory to visualize the dynamics of integration of the international crude oil market. Similarly, both Ji and Fan (2016) conducted another study to see contemporaneous causality between China's oil markets and other commodity markets while utilizing an ECM model along with DAG.

Let us highlight some of the salient features of this approach in detail. First, it is a graph structure that is determined by observed correlations and partial correlations. It represents the causal flow among a set of variables. A directed edge X→Y in the DAG indicates X can cause Y in contemporaneous time. A PC algorithm that helped in computing TETRAD IV software to build DAG was composed of a complete undirected graph and the unconditional correlation matrix between the given variables (Wang and Ji 2017)**.**

Moreover, it is mandatory to emphasize first-order partial correlation that can support removing the edges that are non-static and different from zero. After surviving first order test, the algorithm needs to proceed unless all the edges are removed, or we can also say it like that when an *N*–2 order partial correlation test is finished for *N* variables. The conditional variable(s) on the removed edges is defined as a separate set of the pairwise variables whose edge has been removed. If one edge is removed by unconditional correlation, its separate set is empty. Based on the above two steps, all the remaining edges can be directed using the separate set (Bessler and Yang 2003).

Now, let us suppose that triples of the variables are selected to be directed, considering a triple relation, X—Y—Z, such that X and Y are adjacent as well as Y and Z, but X and Z are not adjacent. Let us suppose that if Y is not in the separate set of X and Z, then X—Y—Z should be directed as X→Y←Z.

Now, on the other hand, we can have only three kinds of results of the orientations here: X→Y→Z, X←Y→Z, or X←Y ←Z. For the sake of determining the correct orientation, we need to have additional information derived from some other identified adjacent linked triples, such as Y→Z←L, and an exogenous restriction, such as X→Y.

These logical algorithms can remove all the remaining edges, which is why they can be directed by confirming DAG (Li, Woodard, et al., 2013). Furthermore, Fisher's $z$ statistic was used to test whether conditional correlations were significantly different from zero:

$$z(\rho(i,j|k),n) = \left[\frac{1}{2}\sqrt{n-|k|-3}\right] ln\left\{\frac{1+\rho(i,j|k)}{1-\rho(i,j|k)}\right\} \tag{3.44}$$

In the above equation, $n$ is the number of observations, $\rho(i,j|k)$ is the population conditional correlation between series $i$ and $j$ conditional on series $k$, and $|k|$ is the number of series in $k$. If series $i$, $j$ and $k$ are normally distributed and $\rho 1(i,j|k)$is the sample conditional correlation of $i$ and $j$ given $k$, then the distribution of $z(\rho(i,j|k),n) - z(\rho 1(i,j|k),n)$ is standard normal (Bessler and Yang 2003); (Ji and Fan 2016).

### 3.3.4. Markov Chain Method (MCM)

A new methodology depends on the Markov Chain method, which is not based on normality and linearity assumptions to find the causal direction. Its representation the state of a system with a random variable that changes over time. In this case, the Markov property proposes that the distribution of a variable based only on the distribution of earlier state. Discrete-Time Markov Chain (DTMC) is a random process which involves the transition from one state to another state-space independently. The working of probability distribution of the next state is dependent only on the current state and the rest of the consequences does not impact it. It has significant applications for different kind of phenomena because of its serial dependence on the formulation of the chain like systems. Where what occurred in future depends only on the present state of system. (Jerrum 1998)

However, certain assumptions need to be fulfilled if we want to use MCM. The first one is that the transition matrix must be stable with respect to space and time Secondly, it is almost impossible to predict the future with certainty. There are numerous other extensions, variations, and generalizations. A discrete-time random process contains a system in a definite state at each step and with the state changing randomly between the steps. The state of a Markov chain at a given point in future, it is generally impossible to predict with certainty because the system transitions randomly. Though, system's future can be predicted by its statistical properties. Normally, there are four possible scenarios in which we can use four different Markov models. Firstly, the Markov Chain Method is considered if the system state is fully observable and autonomous. Secondly, the Hidden Markov model is taken in to account only if state is invisible and the output is observable. Thirdly, Markov's decision process is focused if the system is in control state form and in observable state. Lastly, the partially observable Markov decision process is considered only if the system is in a state of partial control and partial observation. (Harmon and Challenor 1997)

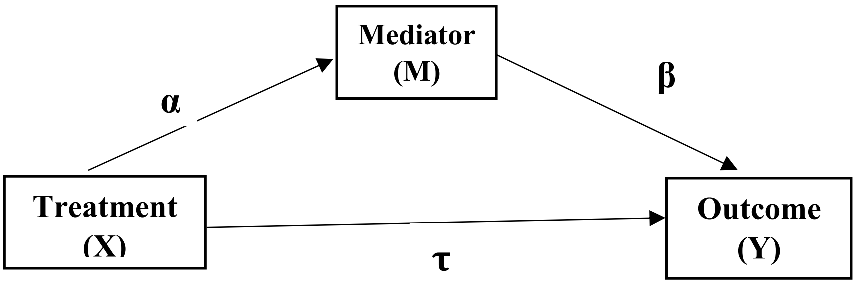
It is significant to highlight here that the Markov chain process/Drunkard's walk has a state space that is basically a transition matrix that represents the transition probabilities and allows for an initial state across the state space. The state system changes are known as Transitions while the probabilities associated with various state changes are said to be called Transition Probabilities. Moreover, there is a need to notice that transition probabilities will only depend on the current position not on its mannerism, to reach the state (Yang and Rannala 1997).

### 3.3.5. Mediation Analysis in Latent Growth Curve Modeling under Structural Equation Framework

In order to examine the intra-individual variations in longitudinal data for the sake of improving statistical inferences, Mediation models are considered to be one of the best choices for economists. Contrary to the three-variable mediation models for cross-sectional data in which only one indirect effect is examined, mediation models for longitudinal data often have multiple indirect effects and even different types of indirect effects. Therefore, it is essential to consider whether and how to summarize these multiple indirect effects. (Cheong, MacKinnon, et al. 2003).

The figure below is an illustration of a single mediator causal model. As we can see that, the theory-based causal variables are basically potential mediating variables (M) that is making intervening relation between the independent variable (X) and the outcome variable (Y). It is observable here that the independent variable (X) is making influence on the outcome directly at one end and at the same instant; it is also indirectly making influence via mediator. Chen (1990) explicitly called it action theory because it is providing link between the treatment program and the mediating variables. He also called it conceptual theory because at the same moment it is linking the mediating variables and the outcome variable (Chen 1990).

**Figure 2: A single Mediator Model**




Direct Effect = $\boldsymbol{\tau}$

Indirect Effect = $\boldsymbol{\alpha\beta}$

Total Effect = $\boldsymbol{\tau + \alpha\beta}$

Interestingly, there is one another method that is called as Product of the coefficient method which is essentially used to find out the point estimates of the mediated effect Aroian (1947); Goodman (1960); MacKinnon and Dwyer (1993); Sobel (1982); MacKinnon, Warsi et al. (1995); MacKinnon, Lockwood et al. (2002).

For the product of coefficients method, we will use the following regression equations in order to estimate the mediated effect:

$$M = \beta_{01} + \alpha X + \varepsilon_1 \ \dots \qquad (3.45)$$

$$Y = \beta_{02} + \beta M + \tau X + \varepsilon_2 \ \dots \qquad (3.46)$$

The above Equation reflects the potential mediator M which is regressed on the regressor X.

While in the Equation 3.46, we can see that the outcome variable Y is regressed on the regressor X and also that the potential mediator M.

Now, we can observe that here Mediation is represented as the indirect effect of (X) on Y. The effect of the regressor on the potential mediator is represented by the coefficient α in equation 3.5. Whereas, the effect of the potential mediator –

(M) on the outcome variable Y is represented by the coefficient β in equation 3.46 after controlling the effect of the regressors. In the similar vein, the effect of the regressor (X) on the outcome variable Y is represented by the coefficient τ in equation 3.46 after controlling for the effect of the mediator

Moreover, we can see that the constants $\beta_{01}$ and $\beta_{02}$ are the regression intercept terms and $\varepsilon_1$ and $\varepsilon_2$ are residuals in the above written two equations.

Moreover, the mediated effect is estimated by the product of the two regression coefficients for α and β. It is to be observed that if the mediated effect is evaluated, the extent of treatment program changes the mediator (α) and the extent to which it changes the outcome (β). Full mediation only takes place if the direct effect is no longer statistically significant after the addition of mediator.

But we will be focusing only on α x β) regardless of the size or statistical significance of τ. Even though there are numerous ways to analyze the product of two coefficients in order to find the standard error (MacKinnon, Lockwood et al. 2002).

The best formula used to find the standard error of the product of the two coefficients is composed on the basis of multivariate delta method (Sobel 1982). Economists often called it as the first-order Taylor series. It is mentioned as follows:

$$\sigma_{\alpha\beta} = \sqrt{\alpha^2\sigma_\beta^2 + \beta^2}\sigma_\alpha^2 \ldots \qquad (3.47)$$

As we can clearly see that in the above written equation, α and $\sigma_\alpha$ are the regression coefficient and its standard error in Equation (3.45) and β and $\sigma_\beta$ are the regression coefficient and its standard error in the Equation 3.46. The observed data in the sample estimates of α, $\sigma_\alpha$, β, and $\sigma_\beta$ which are inserted in the Equation 3.47.

It is important to explain here that the mediated effect is conducted by dividing the estimate of the mediated effect ($\alpha\beta$) by the estimated standard error ($\sigma_{\alpha\beta}$), which is compared to a standard normal distribution. (MacKinnon, Lockwood et al. 2002).

Contrary to the above mentioned method, there is also present an alternative method which is known as the asymmetric CI method (MacKinnon, Lockwood et al. 2002). In this method, both α and β coefficients are converted to z scores (i.e., $z_\alpha$= α/$\sigma_\alpha$ and $z_\beta$ = β/$\sigma_\beta$) and the critical values for the two z scores are found from the tables in Meeker, Cornwell et al. (1981) to construct the CI. This method is workable as it depends on the distribution of the product of αβ which is often asymmetric and also, we are bound to use different values for the upper and lower critical values.

The formula for this purpose is explained as follows:

UCL = αβ + (upper critical value)* $\sigma_{\alpha\beta}$

For, LCL = αβ + (lower critical value)* $\sigma_{\alpha\beta}$

However, if the CI does not include zero, then we say that the mediated effect at given instant is found to be statistically significant. (MacKinnon, Lockwood et al. 2002)

### 3.3.6. A Dynamic Panel Data Approach; Two-step System GMM and Difference GMM

In order to estimate the various parameters in dynamic panel data models, another method is used that is called as Generalized Method of Moments (GMM), given by Arellano and Bover (1995) and Blundell and Bond (1998). In this method, we use a system GMM estimator, which can combine the differenced equation with the level equation. However, the instruments for the level equation are lagged differences of the variables that can only work if the differences are uncorrelated with the individual effects.

While elaborating some of the characteristics of this method, Blundell and Bond (1998) proposed that the system estimator carries superior properties regarding small sample bias and Root Mean Square Error (RMSE), specifically in the case of the persistent series.

It is noticeable here that the practice to use the inverse of the moment matrix of the instruments as the initial weight matrix is a common practice among economists.

In this research study, we will observe the potential efficiency loss from using this weight matrix using the efficiency bounds as derived by (Liu and Neudecker 1997).

Basically, the GMM estimator is a two-step estimator. The first step involves an initial positive semi-definite weight matrix which can be obtained from the consistent estimates of the parameters. Then, after getting estimates, we can further make a weight matrix out of it (Arellano and Bond 1991).

Moreover, the two-step estimated standard errors will have a small sample downward bias, while the one-step estimates will have vigorous standard errors. However, an efficient weight matrix for the differenced model with errors that are homoscedastic and serially non-correlated can easily be attainable.

However, this study will focus on the potential efficiency loss while considering errors to be homoscedastic and serially uncorrelated errors. For this purpose, we will get upper bounds for the efficiency loss by calculating it through Liu and Neudecker's methodology that was primarily based on KI Kantorovich inequality (Liu and Neudecker 1997). It is important to notice here that the upper bounds are indicating the efficiency loss will be severe and variance of the individual unobserved heterogeneity is relatively small.

In order to gain efficiency, let us consider the AR (1) from the panel data specification:

$$y_{it} = \alpha_0 y_{it-1} + \eta_i + \varepsilon_{it} \qquad (3.48)$$

In the above equation, i=1… N, t=2… T, with N large and T fixed.

Now, as we know that the error terms will follow the error components structure like the following equations:

$$E(\eta_i) = 0, \qquad E(\varepsilon_{it}) = 0,$$
$$E(\varepsilon_{it}^2) = \sigma_\varepsilon^2, \qquad E(\eta_{it}^2) = \sigma_\eta^2,$$
$$E(\eta_i \varepsilon_{it}) = 0, \qquad E(\varepsilon_{it}\varepsilon_{is}) = 0, \ \ t \neq s$$

The following observations are made in this process. Firstly, the OLS and within groups estimators of $\alpha_0$ in model (3.48) are biased and inconsistent, as evident from the above equation.

Secondly, there is a consistent estimator for $\alpha_0$ which we are calling a system GMM estimator (Arellano and Bover (1995) and Blundell and Bond (1998). Hence, we can therefore, utilize these conditions (T+1) (T-2)/2 in order to formulate the following equations:

$$E[(\Delta y_{it} - \alpha_0 \Delta y_{it-1})(y_{it-2}, \dots, y_{i1})] = 0, \qquad (3.49)$$
$$E[(y_{it} - \alpha_0 y_{it-1})\Delta y_{it-1}] = 0, \qquad (3.50)$$

In the above equation, as we can observe for t = 3... *T*.

Moreover, moment conditions (3.49) are for the model in first differences as they are utilizing appropriately lagged levels information as instruments. Furthermore, conditions (3.50) are for the model in levels that utilize lagged differences as instruments. Hence, this method is a significantly efficient method of estimator on the basis of mere moments conditions (Blundell and Bond 1998).

# CHAPTER 4

# RESEARCH METHODOLOGY

This chapter will briefly discuss the methodology for assessing causality tests for panel data in simulation research. This study compares the size and power features of each test under consideration. First, we will discuss the data generating process (DGP) and Monte Carlo simulation design used in this research. Second, steps for calculating size and power will be provided. The third section briefly explains the notions and terminologies of panel causality tests, which are utilized in testing and simulations.

## 4.1 Methodology Description

In the present study, we will be examining causal inference methods for testing particular hypothesized causal relations, $H_0$: X causes Y. After that, employing all causality methods in this way requires benefits over the more distinctive application method. First, relatively small numbers of causally related variables are needed. Some variable Z that is causally related in a certain way to X and Y permits rejection $H_0$: irrespective of what other causally related variables may or may not exist. Second, this constricted focus indicates that there are only a limited number of techniques in which latent, related variables might influence the observed variables involved in the test. This permits us to estimate the size and power of the test. Using Monte Carlo Simulation, this study compares the power and size properties of existing and modified panel causality tests. In the existing test, Granger non-causality test, while in modified Sims's and Final Prediction Error method are compared. We modified the Sims test algorithms and the final prediction error method for panel data causality analysis.

We will generate an artificial data-generating process with a known causal relationship among the variables. After that, we will test the methods mentioned above for causality for panel data. This study will show which procedure detects known causal relation or not, employing the appropriate causal method concerning the specific type of data. Further, it will present the results of the Monte Carlo Simulations. We will then analyze the size and power of all causality tests based on the outcome of the Monte Carlo Simulation.

## 4.2 Data Generating Process (DGP):

The objective of the simulation experiment is to find out the size and power properties of methodologies for testing causality. Therefore, we need data with embedded causality (for power) and the data series with no causality (size).

The choice of DGP for comparative studies is critical. Different tests follow different theoretical bases; hence, choosing a data-generating method gives an advantage to some test statistics. Causality methods and tests can be compared in the same framework. The purpose is to eliminate the single test superiority. In such a situation, a simulation study is required. However, we continue with DGP to compare all the tests. For this purpose, we use the following framework. The data for testing properties of causality tests can be generated from a unified framework which is given below:

$$\begin{bmatrix} x_{it} \\ y_{it} \\ z_{it} \end{bmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix} \begin{bmatrix} x_{i,t-1} \\ y_{i,t-1} \\ z_{i,t-1} \end{bmatrix} + \begin{bmatrix} a_{i1} & a_2 \\ b_{i1} & b_2 \\ c_{i1} & c_2 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \begin{bmatrix} \varepsilon_{xit} \\ \varepsilon_{yit} \\ \varepsilon_{zit} \end{bmatrix} \dots (4.1)$$

$$\text{Where, } \begin{bmatrix} \varepsilon_{xit} \\ \varepsilon_{yit} \\ \varepsilon_{zit} \end{bmatrix} \sim N \left[ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho \\ \rho_2 & \rho & 1 \end{pmatrix} \right]$$

This general DGP equation can take various forms by specifying the parameters A, B and Σ. The covariance matrix shows us the contemporaneous causality as they are correlated with each other. The above matrix form equation (4.1) can be written in the following form:

$$X_{it} = A_i X_{i,t-1} + B_i D_t + \gamma w_{it} + \varepsilon_{it} \qquad \varepsilon_{it} \sim N\,[0, \Sigma\,] \dots (4.2)$$

*Where, i=1, 2, 3…N and t=1, 2, 3… T*

$$\text{Where, } X_{it} = \begin{bmatrix} x_{it} \\ y_{it} \\ z_{it} \end{bmatrix}, A_i = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix}, B_i = \begin{bmatrix} a_{i1} & a_2 \\ b_{i1} & b_2 \\ c_{i1} & c_2 \end{bmatrix}, D = \begin{bmatrix} 1 \\ t \end{bmatrix}, \varepsilon_{it} = \begin{bmatrix} \varepsilon_{xit} \\ \varepsilon_{yit} \\ \varepsilon_{zit} \end{bmatrix} \text{ and}$$

$$\Sigma = \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho \\ \rho_2 & \rho & 1 \end{pmatrix}$$

Under the assumption of the Causal relation between X and Y ($H_0$: X causes Y) and under other hypotheses of independence between X and Y ($H_1$: X does not cause Y). As per the definition of Granger causality, Y is caused by X if the lag value of X can be used for predicting Y.

There is introduce a confounding variable $w_{it}$ and we can say that X and Y are not causally correlated, but W indirectly affects them. So it will lead to the result that X is causing Y through confounding variable W. $W_{it}$ is also called strictly exogenous variable, which is highly correlated with regressor but uncorrelated with disturbance term.

The proper specification of the DGP can add confounding variables. In DGP (4.1) suppose $A_{1i} = (\alpha, 0,0), \alpha \in (0,1)$ then $y_{i,t-1}$ and $z_{i,t-1}$ does not appear in the equation $x_{it}$. Therefore, $y_{it}$ and $z_{it}$ doesn't Granger cause $x_{it}$, it means $y_{it}$ and $z_{it}$ have an indirect relationship, which behaves like confounding factors. If $\theta_{12} = \theta_{13} \neq 0$, it means $y_{it}$ and $z_{it}$ Granger cause $x_{it}$. On the other hand, if $A_{1i} = (0, \alpha, 0)$, then $x_{i,t-1}\ and\ z_{i,t-1}$ doesn't appear in the equation $y_{it}$ . Hence, $x_{it}\ and\ z_{it}$ doesn't Granger cause $y_{it}$, it means $x_{it}$ and $z_{it}$ have an indirect relationship which behaves like confounding factors. If $\theta_{21} = \theta_{23} \neq 0$, it means $x_{it}\ and\ z_{it}$ Granger cause $y_{it}$. The same causal direction can be examined if we have a case that $A_{1i} = (0,0,\alpha)$ the above data generating process generates data in different ways by changing the value of parameters or imposing other restrictions.

The series with contemporaneous correlation can also be generated from the data generating process (DGP 1). If $A_i$=0 and $\varepsilon_{it} \sim N\,[0, 1\,]$ then the series generated will be independent of each other. If $\varepsilon_{it} \sim N\,[0, \Sigma\,]$ then the data generating process will generate series with contemporaneous correlation. These cases can be generated by imposing different restrictions on the data generating process. The series with drift and trend can also be developed from equation (4.1) by taking $B_i \neq 0$. We will generate three independent series with drift, without drift, with drift and trend.

First, the data generating process will generate three independent autoregressive series having drift and trend if $\begin{bmatrix} - & \theta_{12} & \theta_{13} \\ \theta_{21} & - & \theta_{23} \\ \theta_{31} & \theta_{32} & - \end{bmatrix}$=0, and $\rho$ = 0. Second, the data generating process will generate three independent autoregressive series having drift if $\begin{bmatrix} - & \theta_{12} & \theta_{13} \\ \theta_{21} & - & \theta_{23} \\ \theta_{31} & \theta_{32} & - \end{bmatrix}$=0, and $\rho$ = 0 and $\begin{bmatrix} - & a_2 \\ - & b_2 \\ - & c_2 \end{bmatrix}$ =0. Third, the data generating process will generate three independent autoregressive series having no drift if $\begin{bmatrix} - & \theta_{12} & \theta_{13} \\ \theta_{21} & - & \theta_{23} \\ \theta_{31} & \theta_{32} & - \end{bmatrix}$=0, and $\rho$ = 0 and $\begin{bmatrix} a_{i1} & - \\ b_{i1} & - \\ c_{i1} & - \end{bmatrix}$=0. Further, we will generate different but correlated series $x_{it}, y_{it}$ , $z_{it}$ with drift, without drift, with drift and trend. 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 $\begin{bmatrix} a_{i1} & - \\ b_{i1} & - \\ c_{i1} & - \end{bmatrix}$=0. Finally, the data generating process will generate three dependent series having no drift if $\begin{bmatrix} - & a_2 \\ - & b_2 \\ - & c_2 \end{bmatrix}$ =0.

The parameter $B_i$ is called “nuisance”. The causality does not depend on the matrix of parameter$B_i$, however the test statistics for coefficient present in “$A_i$” which determine causality is heavily dependent on $B_i$ and incorrect specification of $B_i$ may create bias. So, to avoid biases, we have to include this nuisance term. The first column of $B_i$ shows the individual intercepts; we call this feature “cross-sectional interdependency.”

The generated series are independent of each other, causality cannot be checked, so we will only determine the size of the test. Furthermore, we will generate different but correlated series $x_{it}, y_{it}$ to check the causal ordering through the power of the test.

The term "power of the test" refers to the ability to reject a hypothesis when the alternative hypothesis is true.

$$Power= (rejecting\ H0\ /\ H1\ is\ true)$$

When a valid hypothesis is rejected, an error is created, which is represented by the symbol “α” and determines the size of the test.

$$\alpha = (rejecting\ H0\ /\ H0\ is\ true)$$

Strength is the only number that measures the overall performance of a test. This means that all tests can be compared and ranked based on this measurement. Furthermore, there is a natural and intuitive explanation for the severity of the test. To define stringency- the crucial concept is the power envelope, which is the maximum possible power that can be achieved at a given alternative. A test with stringency zero is a uniformly most powerful test –where power changes and we select the test which offers best power through simulation. It is the most powerful of all the alternatives. If it exists, this test should always be preferred. If a test has a 1% hardness test, it means that the test strength is only 1% less than the most powerful test available on any possible alternative. This test is equally as good as the most powerful test for practical purposes. If 5 to 10 between wiring tests can be found, we 'don't need to search further for functional purposes. If the best available tests have a hardness of 50 or more, we should look for better testing methods. The point is that assessing the severity of the tests provides us with an essential guide for using and comparing tests in practical matters. (Rehman, Zaman, et al. 2017).

### 4.3 Monte Carlo Simulation Design

The study aims to assess all causality tests' performance for panel data by investigating size and power properties. The study mainly focuses on Monte Carlo Simulations, and the optimal procedure will be selected. Finally, through that optimal procedure, causal relations between the intensity of government spending and household spending on education will be explored by taking real panel data.

The following steps are involved in finding the required simulation values;

1. Generate the data for a fixed sample size under null and alternative hypotheses using the given DGP.
2. Testing the Panel Causality Tests
3. Calculating the test's statistics.
4. Size and Power
5. Testing and Simulations
6. Repeat the process a fixed number of times and save the repeated results in a column matrix.

## 4.4 Sample Size Selection

An appropriate sample size is crucial for the data analysis, especially while conducting a panel data analysis. We have annual data sets and selected three different cross-section units for this study. We have categorized them into three other groups to proceed with this simulation study, i.e., 5, 10 and 20; as small cross-section units, medium cross-section units and large cross-section units, respectively. Similarly, three-time series lengths are taken to evaluate the performance of Causality tests; these time series levels are 25, 50, and 200. Similar to categorizing cross-section length into three types, a time series length of 25 indicates small time series, and 50 is assigned as a medium time-series length. At the same time, 200 is allotted as a large time-series length in this study.

There is a problem with the Monte Carlo Sample Size (MCSS); there may be some disruptions when using small MCSS and often face the difficulty to enlarge MCSS. Furthermore, each test's analysis (i.e., simulation) is time-consuming (i.e. may take several days) to produce the same result, which further affects the time required for each refusal. In order to carry out simulations, an MCSS of 10,000 is taken to get the convergence effectively.

## 4.5 Monte Carlo Simulation Design for Size of the Test

Rejecting a true hypothesis is an error denoted by "α". It constitutes the size of the test.

$$\alpha = P\ (rejecting\ H0\ /\ H0\ is\ true)$$

We calculate the test size using the following steps;

1. Generate heterogeneous panel data for a fixed number of times and cross-section dimensions under the null hypothesis $H_0 = 0$. The number of time series and cross-section dimensions considered 25, 50, 200 and 5, 10, 20 respectively.

2. Apply the test statistic and calculate its value.

3. Compute the values of Size and Power of each test.

4. Repeat this process for a fixed Monte Carlo sample size. The percentage of this count is called the size of the test, which is obtained as:

$$Size\ of\ Test = \frac{significant\ count}{MCSS} \times 100$$

### 4.6. Monte Carlo Simulation Design for Power of Tests

The term "power of the test" refers to the ability to reject a hypothesis when the alternative hypothesis is true.

$$Power = (rejecting\ H0\ /\ H1\ is\ true)$$

We compared different panel causality tests used to investigate the causal relations in our analysis. All the tests are compared in a common framework to assess the ability to reject the null hypothesis when it is wrong.

The powers of the tests are calculated following these phases;

1. Generate heterogeneous panel data for a fixed number of time and cross-section dimensions under the whole alternative space $H_1 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$. The number of time series and cross section dimensions have considered are 25, 50, 200 and 5, 10, 20 respectively.

2. Apply test statistics and calculate the test statistic for that particular alternate point.

3. Compute the values of Size and Power of each test.

4. Repeat this process for fixed MCSS; the percentage of these counts is called the power of the test at a specific alternate point.

5. Replication for every point given in alternate space provides us with the power of the test against all possible alternates.

The debate about the DGP conducted in the above section; powers of the test should be considered against each alternate, so we have overall ten alternates revealed in alternate space (AS).

So, the powers of each test against all alternatives will be calculated using 10,000 simulations. These simulations give us powers of the test using three sample sizes; small (5, 25), medium (10, 50) and large (20, 200).

## 4.7 Power Curves

For a specific test, the powers of each test plotted against the elements of alternate space will give us a power curve. For example, we have three-panel causality tests to compare, so three power curves will be obtained for the whole alternative space (0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1), and all the possible causal combinations.

## 4.8 Power Gains and Size Distortions

The size and power of each test are calculated using the theory mentioned above. The empirical analysis is based on two main observations calculated:

$$size\ distortion = Empirical\ size\ of\ the\ test - nominal\ size\ (5\%)$$

$$power\ Gain = Empirical\ power\ of\ the\ test - Empirical\ size\ of\ the\ test$$

## 4.9 Real Data Analysis for Granger Non-Causality Test (2012)

Granger non-causality test detects the correlation and direction of the causality. We claimed this challenge in this study. The two-null hypothesis for Granger non-causality are presented here as;

$$H_0: X_{it} \text{ does not Granger cause } Y_{it}$$

$$H_1: Y_{it} \text{ does not Granger cause } X_{it}$$

In the first hypothesis, it is claimed that the intensity of government spending on education does not Granger Cause the intensity of household spending on education. If this null is rejected, it is claimed as the test's statistical power.

### 4.10 Panel Causality Tests

#### 4.10.1 Granger Non-Causality Test in Heterogeneous Panel Data by Hurlin (2012)

Two theorists, Dumitrescu and Hurlin (2012), introduced a famous approach primarily used to test for Granger causality in macro panel datasets. Because in today's world, we have increased demand for the macro-level panel data, there is a required new set of techniques that can help the econometricians grow and develop their skills.

The Granger non-causality test utilizes the Akaike information criteria, Bayesian information criterion, or Hannan–Quinn information criterion to identify the optimal lag length. The length where the mentioned criteria give minimum test statistic value.

Finally, to address the empirical problem of cross-sectional dependency, a bootstrap technique was used to obtain p-values and critical values. (Lopez and Weber, 2017)

Granger's (1969) methodology is primarily established for investigating the causal relations between time series. Let us suppose $x_t$ and $y_t$ are two stationary series.

$$y_t = \alpha + \sum_{k=1}^{K} \gamma_k \, y_{t-k} + \sum_{k=1}^{K} \beta_k \, x_{t-k} + \varepsilon_t \quad with\; t = 1, \dots, T \qquad (4.3)$$

We can say that the model can be utilized to determine whether *x* causes *y*. In other words, if previous values of x are significantly predictive of the present value of y even though previous values of y would be included in the model, then x has a causal effect on y.

While using (4.3), we can quickly identify causality that is centered on an *F* test with the following $H_0$:

$$H_0: \beta_1 = \cdots = \beta_k = 0 \qquad (4.4)$$

If $H_0$ is denied, one might infer that there is causation from x to y. Furthermore, the x and y variables may be swapped to test for causation in the opposite direction, and reversible causality can be seen. (Lopez and Weber 2017)

However, Dumitrescu and Hurlin followed an extension explicitly designed to detect the causality in panel data.
We can say the following equation will represent:

We can say the following equation will represent the underlying regression:

$$y_{i.t} = \alpha_i + \sum_{k=1}^{K} \gamma_{ik}\, y_{i,t-k} + \sum_{k=1}^{K} \beta_{ik}\, x_{i,t-k} + \varepsilon_{i,t} \quad with\ i = 1, \dots, N\ and\ t = 1, \dots, T \qquad (4.5)$$

In the above equation, we have $x_{it}$ and $y_{it}$ that represents the observations of two stationary variables for individual *i* in period *t*. Moreover, individual coefficients are permitted to differ, but they are supposed to be time-invariant at the same time. It is essential to mention here that the lag order *K* in this scenario is assumed to be equal for all individuals so that the panel should be balanced.

Granger's method is used to establish causality and to assess the substantial impacts of past x values on the current value of y. Hence, we can define the null hypothesis in the following manner:

$$H_0\colon\ \beta_{i1} = \cdots = \beta_{ik} = 0 \qquad \forall\, i = 1, \dots, N \qquad (4.6)$$

Here we can observe that there is no causation for all members of the panel. The reason for this is that in the DH test, we can suppose that causality exists for some people but is unnecessary for all. Therefore, in such scenarios, we can assume some alternative hypotheses. For instance, let us consider the following alternative hypothesis:

$$H_1\colon \beta_{i1} = \cdots = \beta_{ik} = 0 \qquad \forall\, i = 1, \dots, N1 \qquad (4.7)$$

$$\beta_{i1} \neq 0\ or \dots or\ \beta_{ik} \neq 0 \qquad \forall\, i = N1 + 1, \dots, N \qquad (4.8)$$

Now, we can see that in the above equation, we have $N_1 \in [0, N-1]$ that is not known. If $N_1 = 0$, all of the individuals in the panel have a causal relationship. $N_1$ must be firmly smaller than *N*; there is no causation for all individuals, and $H_1$ reduces to $H_0$. For such scenarios, Dumitrescu and Hurlin (2012) proposed that we can run the *N* individual regressions implicitly enclosed in (4.5), perform *F* tests of the *K* linear hypotheses $\beta_{i1} = \cdot\cdot\cdot = \beta_{ik} = 0$ to retrieve the individual Wald statistic $W_i$, and finally compute the average Wald statistic $\overline{W}$ ,

$$\overline{W} = \frac{1}{N}\sum_{i=1}^{N} W_i \qquad (4.9)$$

Thus, we can say that this test is primarily focused on detecting the causality at the panel level, and at the same instant, it is designed to reject $H_0$ which will not exclude non-causality for some individuals.

However, one can also use Monte Carlo simulations while using $\overline{W}$ asymptotically to investigate panel causality. (Lopez and Weber, 2017) For instance, let us suppose that the assumption for Wald statistics $Wi$ are independently and identically distributed among individuals. It may be expressed as the standardized statistic where first $T \rightarrow \infty$ and then $N \rightarrow \infty$ (often read as "$T$ should be big relative to $N$") following a standard normal distribution as shown below:

$$\bar{Z} = \sqrt{\frac{N}{2K}} \times (\overline{W} - K) \quad \xrightarrow[T,N\rightarrow\infty]{d} \quad N(0,1) \qquad (4.10)$$

Also, for a fixed T dimension with T > 5+3K, the approximated standardized statistic $\tilde{Z}$ follows a standard normal distribution:

$$\tilde{Z} = \sqrt{\frac{N}{2K} \times \frac{T-3K-5}{T-2K-3}} \times \left(\frac{T-3K-3}{T-3K-1} \times \overline{W} - K\right) \quad \xrightarrow[N\rightarrow\infty]{d} \quad N(0,1) \qquad (4.11)$$

In this procedure, we can see that the null hypotheses (equation 4.6) is based on $\overline{Z}$ and $\tilde{Z}$. However, if these are pretty larger than the expected standard critical values, then we should reject $H_0$ and simply, we can conclude that Granger causality exists in this scenario. But it is essential to mention here that for large N and T panel data sets, $\overline{Z}$ is a considerable option. $\tilde{Z}$ should be preferred for datasets with large N. Still, relatively small T. Dumitrescu and Hurlin (2012) used Monte Carlo Simulations to show that they exhibits good finite sample characteristics, even when N and T are small. The selection lag order (K) is an empirical problem that Dumitrescu and Hurlin (2012) do not address. One solution is to use the number of lags based on an information criterion (AIC/BIC/HQIC). (Lopez and Weber, 2017)

### 4.10.2 Sims Test (1972)

The GC test (1969) explored that one variable is often regressed on its lags and the lags of another explanatory variable. But it is important to note that it does not include lead values of the explanatory variable during the process. The problem is confronted by Sims (1972). They argued that if one variable regresses on its lags, then the leading values of the explanatory variable will result in causality run from explanatory to regressed variable and all the leading values of the regressor in the regression will not become statistically significant and different from zero as a group. Henceforth, Sims demonstrated that the *"future cannot cause current or past"* (Sims 1972).

Moreover, it assumes that the error term white-noise and variables must be integrated of order zero at a level. However, one variable is non-stationary at level but becomes stationary at the first difference, then in that case variable at first difference should be used preferably (Sims 1972).

Sims (1972) is applicable for only time-series data sets. In order to find the causal relations in panel data, we have modified the Sims test (1972) for panel data. Thus, the application of the modified Sims test will demand an equation for testing panel regression X to Y;

$$Y_{it} = \alpha_i + \sum_{i=1}^{m} \alpha_{1i} X_{it-i} + \sum_{j=1}^{n} b_j X_{it+j} + \nu_{1,it} \qquad (4.12)$$

Here tested the following null hypothesis by using the F-test (eq.4.15):

$$H_0: b_{11} = b_{12} = . \; . \; .= b_{1n} = 0$$

Now, If $H_0$is accepted, then we will say X causes Y; otherwise, we will say that X does not Granger cause Y. Likewise, to assess causality from Y to X, we have to apply the following equation:

$$X_{it} = \alpha_{1i} + \sum_{i=1}^{m} \alpha_{2i} Y_{it-i} + \sum_{j=1}^{n} b_{2j} Y_{it+j} + \nu_{2,it} \qquad (4.13)$$

And for the process of conducting a test for the null hypothesis; we will use the following equation:

$$H_0: b_{21} = b_{22} = . \; . \; .= b_{2n} = 0 \qquad (4.14)$$

In order to see if jointly the coefficients associated with the regressors are statistically significant, conduct an F-test to test the null hypothesis. Now, run the following;

$$F = \frac{(SSR_{restricted} - SSR_{unrestricted})/k}{SSR_{unrestricted}/(T-2k-1)} \sim F_{k,\ T-2k-1} \quad (4.15)$$

Dumitrescu and Hurlin (2012) proposed that we can run the *N* individual regressions implicitly enclosed in (4.12), perform *F* tests of the *K* linear hypotheses $\beta_{i1} = \cdot\cdot\cdot = \beta_{ik} = 0$ to retrieve the individual Wald statistic $W_i$, and finally compute the average Wald statistic $\overline{W}$ ,

$$\overline{W} = \frac{1}{N}\sum_{i=1}^{N} W_i \quad (4.16)$$

Thus, we can say that this test is primarily focused on detecting the causality at the panel level, and at the same instant, it is designed to reject $H_0$ which will not exclude non-causality for some individuals. However, one can also use Monte Carlo simulations while using $\overline{W}$ Asymptotically to investigate panel causality (Lopez and Weber 2017). For instance, let us suppose that the assumption for Wald statistics *Wi* are independently and identically distributed among individuals. It may be expressed as the standardized statistic where first $T \rightarrow \infty$ and then $N \rightarrow \infty$ (often read as "*T* should be big relative to *N*") following a standard normal distribution as shown below:

$$\bar{Z} = \sqrt{\frac{N}{2K}} \times (\overline{W} - K) \quad \xrightarrow[T,N\ \rightarrow\infty]{d} \quad N(0,1) \quad (4.17)$$

### 4.10.3 Final Prediction Error (FPE) Method (Hsiao 1981)

Hsiao introduced one of the relevant and intelligent compositions to handle the limitations of the Grangers Causality test. Hsiao (1981) merged the GC test with Akaike Final Prediction Error (FPE) criterion. The first step of the process involved using the regressed variable through a one-dimensional autoregressive procedure. After regressing a variable only on its lagged, it can calculate the FPE.

In the second step, regress a variable on its lag, plus lags of the explanatory variable, and then calculate its FPE. Now, find out that the FPE of the second step is far less than the first step FPE. This will conclude the causal relationship between the explanatory variable and the explained variable under such a scenario.

The procedure can repeat the same process to examine the GC (1969) among three variables. It is significant to highlight here that both assumptions and methodology will be similar to (Sims 1972). However, the results reflect that FPE will minimize the mean square prediction error, further decreasing the uncertainty at the significance level while using the optimality criterion. Furthermore, Hsiao believed that additional variables are substantially allowed in this method (Hsiao 1981).

We have modified the final prediction error method of Hsiao for panel data. Now represent the Hsiao method in the equations for panel data. As we can see in the first step, we have to estimate the following autoregressive equation having this particular form:

$$Y_{it} = \alpha_i + \sum_{i=1}^{m} \alpha_{1i}\, Y_{it-i} + \nu_{1,it} \quad (4.18)$$

Now here selected “m” to the greatest extent possible. The FPE was then calculated in the following manner for each regression;

$$FPE_{(m)} = \frac{T+m+1}{T-m-1} Q(m)/T \quad (4.19)$$

In the above equation, we have ‘T,’ the number of observations utilized, m' is the lag order ranging from 1 to m, and Q (m) is the related sum of squared residuals. Assume that the precise value of m, say $m^*$, is the optimal lag length, which results in the lowest FPE.

Now in the second stage, treat ‘Y’ as the regressed variable with the optimal lag order set at $m^*$ and ‘X’ is regarded here as a regressor variable with the order of lags ranging from 1 to n. Then, after it, have to run the regression of the following:

$$Y_{i,t} = \alpha_{1,i} + \sum_{i=1}^{m^*} \alpha_{1,i}\, Y_{i,t-i} + \sum_{j=1}^{n} b_{1,j}\, X_{i,t-j} + \nu_{2,i,t} \quad (4.20)$$

As we can see that the corresponding two-dimensional FPE will come out to be:

$$FPE_{(m,n)} = \frac{T+m^*+n+1}{T-m^*-n-1} Q(m,n)/T \quad (4.21)$$

In the above equation, ‘n’ is the order of lags on ‘X.’ We have witnessed here that once again, the optimum ‘n’ say "$n^*$" is picked to reduce FPE (m, n).

Hence, this procedure concludes here that X has Granger causality to Y only if *FPE* $(m^*, n^*)$<*FPE* $(m^*)$.

Furthermore, repeat the same process for the following regression lines if you want to run the GC test (1969) between three variables.

| Restricted equation | | Unrestricted equation |
|---|---|---|
| $Y_{it} = \alpha_1 + lagged\ (Y_{it}) + v_{2,it}$ | (4.22) | $Y_{it} = \alpha_1 + lagged\ \left(Z_{it,}Y_{it}\right) + v_{2,it}$ |
| $Y_{it} = \alpha_1 + lagged\ (Y_{it}) + v_{2,it}$ | (4.23) | $Y_{it} = \alpha_1 + lagged\ \left(X_{it,}Y_{it}\right) + v_{2,it}$ |
| $Z_{it} = \alpha_2 + lagged\ (Z_{it}) + v_{3,it}$ | (4.24) | $Z_{it} = \alpha_2 + lagged\ \left(Z_{it,}Y_{it}\right) + v_{3,it}$ |
| $Z_{it} = \alpha_3 + lagged\ (Z_{it}) + v_{4,it}$ | (4.25) | $Z_{it} = \alpha_3 + lagged\ \left(Z_{it,}, X_{it}\right) + v_{4,it}$ |
| $X_{it} = \alpha_4 + lagged\ (X_{it}) + v_{5,it}$ | (4.26) | $X_{it} = \alpha_4 + lagged\ \left(Z_{it,}, X_{it}\right) + v_{5,it}$ |
| $X_{it} = \alpha_5 + lagged\ (X_{it}) + v_{6,it}$ | (4.27) | $X_{it} = \alpha_5 + lagged\ \left(Y_{it,}, X_{it}\right) + v_{6,it}$ |

As we have written above estimated unrestricted and restricted equations, we found out that the associating minimum FPE for precise values of m and n are present; hence we can appeal conclusions. In all of the regressions mentioned above, errors are white noise, and all variables have been used stationary at their levels.

To see if jointly the coefficients associated with the regressors are statistically significant, conduct an F-test to test the null hypothesis. Now, run the following;

$$F = \frac{(SSR_{restricted} - SSR_{unrestricted})/k}{SSR_{unrestricted}/(T-2k-1)} \sim F_{k,\ T-2k-1} \quad (4.28)$$

Dumitrescu and Hurlin (2012) proposed that we can run the *N* individual regressions implicitly enclosed in (4.20), perform *F* tests of the *K* linear hypotheses

$\beta_{i1}$ = · · · = $\beta_{ik}$ = 0 to retrieve the individual Wald statistic $W_i$, and finally compute the average Wald statistic $\overline{W}$ ,

$$\overline{W} = \frac{1}{N}\sum_{i=1}^{N} W_i \qquad (4.29)$$

Thus, we can say that this test is primarily focused on detecting the causality at the panel level, and at the same instant, it is designed to reject $H_0$ which will not exclude non-causality for some individuals. However, one can also use Monte Carlo simulations while using $\overline{W}$ asymptotically to investigate panel causality (Lopez and Weber 2017). For instance, let us suppose that the assumption for Wald statistics $Wi$ are independently and identically distributed among individuals. It may be expressed as the standardized statistic where first $T \rightarrow \infty$ and then $N \rightarrow \infty$ (often read as "$T$ should be big relative to $N$") following a standard normal distribution as shown below:

$$\bar{Z} = \sqrt{\frac{N}{2K}} \times (\overline{W} - K) \quad \xrightarrow[T,N\ \rightarrow\infty]{d} \quad N(0,1) \qquad (4.30)$$

# CHAPTER 5

# SIZE AND POWER COMPARISON

This chapter briefly discusses the simulation results of panel causality tests with heterogeneous panel DGP under various model specifications. Then, based on Monte Carlo simulation findings, a size and power comparison is performed between the Granger non-causality test, Sims test, and Hsiao's Final prediction error test. Finally, this chapter concludes on which panel causality test is the best performer and which one is the worst based on Monte Carlo simulations.

The study's objective is to evaluate the performance of all causality tests for panel data by investigating size and power properties. To achieve this objective, the study mainly focuses on Monte Carlo Simulations, and the optimal procedure has been selected on its basis.

In view of existing literature for panel data, almost all panel data tests were first developed for the single cross-section with a time series structure, and then expended for more than one cross-section. In the end, an average of all cross-sections is taken to develop a panel data test. Keeping in view the same practice, this study has modified Sim’s time series causality test and the final prediction error method of casualty for time series to panel counterparts to compare with the Granger non-causality test. In the existing test, Granger non-causality test, while in modified Sims and Final Prediction Error are compared. We modified the Sims test algorithms and the final prediction error method for panel data causality analysis.

A comparison of Panel Causality Tests is made through size and power properties. The power of the Granger non-causality Test by Dumitrescu and Hurlin (2012), Sims (1972), and the Final Prediction Error (FPE) method by Hsiao (1981) causal search algorithm is analyzed. The power of any test is defined as the probability of rejecting a null hypothesis when it is false i.e.

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

We analyze the power of Panel causality tests for a variety of situations. The power also depends on several nuisance parameters related to the “deterministic part” and the “stochastic part”. Among the deterministic part are a component of drift and trend. At the same time, among stochastic elements, we have the autoregressive coefficient of the three series (X, Y, and Z), which also determines the series' stationary. This study used three different groups of the sample size, which were categorized into a small sample size, medium sample size, and large sample size for the data generating process under alternative hypotheses to calculate power.

### 5.1. Power Analysis of Panel causality tests with stationary Series

First, we have generated stationary series $x, y,$ and $z$ based on the change in both stochastic and deterministic parts using the data developing process given in equation (5.1). We have attributed different diagonal values in matrix A $\theta_{11} = 0.3$, $\theta_{22} = 0.5$, $\theta_{33} = 0.7$ for making heterogeneous panel data, and the power of panel causality tests is calculated. We have also used heterogeneous drift and trend terms in DGP. The off-diagonal values ($\theta_{12}$ and $\theta_{13}, \theta_{21}$ and $\theta_{23}$ , $\theta_{31}$ and $\theta_{32}$) establish the correlation between x & y and y & z, and its value also changes from 1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, and 0 in matrix A which shows that $x \rightarrow y$ and $y \rightarrow z$ respectively. We have used all six causal combinations of three variables (X, Y, and Z) with different model specifications; only drift, with drift and trend for small, medium, and large panel sample sizes.

**Table 5. 1:** Power Analysis of Panel Causality Tests using stationary series with Drift Only for Small Sample T=25

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel(A) | | X→Y/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho 1$ | $\rho 2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 14.3 | 16.8 | 22.9 | 38.2 | 54.7 | 76.0 | 15.0 | 27.5 | 46.8 |
| 0.1 | 0.9 | 16.2 | 20.7 | 28.0 | 39.1 | 55.8 | 77.3 | 18.6 | 29.8 | 51.1 |
| 0.2 | 0.8 | 23.2 | 33.6 | 47.7 | 40.6 | 56.9 | 78.6 | 20.8 | 33.3 | 58.7 |
| 0.3 | 0.7 | 38.9 | 53.9 | 77.4 | 41.0 | 58.9 | 79.6 | 24.9 | 43.5 | 71.6 |
| 0.4 | 0.6 | 56.4 | 75.8 | 94.5 | 42.1 | 61.2 | 81.1 | 30.9 | 52.4 | 85.3 |
| 0.5 | 0.5 | 73.4 | 92.1 | 99.2 | 42.9 | 62.1 | 83.6 | 41.4 | 66.4 | 92.3 |
| 0.6 | 0.4 | 87.2 | 97.9 | 99.9 | 43.7 | 64.5 | 83.9 | 47.5 | 77.4 | 96.9 |
| 0.7 | 0.3 | 94.7 | 99.8 | 100 | 44.0 | 65.5 | 84.4 | 58.0 | 86.5 | 99.4 |
| 0.8 | 0.2 | 98.1 | 100 | 100 | 44.9 | 67.6 | 85.9 | 71.9 | 92.5 | 99.9 |
| 0.9 | 0.1 | 99.3 | 100 | 100 | 45.3 | 68.3 | 87.2 | 79.8 | 97.8 | 100 |
| 1 | 0 | 99.9 | 100 | 100 | 46.1 | 70.7 | 88.8 | 87.2 | 99.4 | 100 |
| Panel(B) | | X→Z/causality | | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | |
| 0 | 1 | 15.8 | 19.5 | 28.6 | 27.3 | 38.0 | 55.0 | 16.5 | 29.7 | 55.4 |
| 0.1 | 0.9 | 87.8 | 98.5 | 100 | 28.0 | 38.5 | 56.6 | 26.8 | 45.5 | 77.2 |
| 0.2 | 0.8 | 99.2 | 100 | 100 | 28.7 | 39.8 | 57.4 | 31.9 | 50.5 | 83.9 |
| 0.3 | 0.7 | 99.9 | 100 | 100 | 29.3 | 40.1 | 58.0 | 34.2 | 55.7 | 87.6 |
| 0.4 | 0.6 | 100 | 100 | 100 | 30.5 | 41.8 | 59.5 | 37.1 | 63.4 | 89.2 |
| 0.5 | 0.5 | 100 | 100 | 100 | 31.8 | 42.1 | 60.5 | 39.7 | 65.3 | 92.1 |
| 0.6 | 0.4 | 100 | 100 | 100 | 33.6 | 43.4 | 61.9 | 41.8 | 68.7 | 92.7 |
| 0.7 | 0.3 | 100 | 100 | 100 | 34.6 | 44.6 | 62.7 | 44.6 | 69.8 | 96.0 |
| 0.8 | 0.2 | 100 | 100 | 100 | 36.0 | 46.1 | 63.9 | 46.6 | 70.7 | 96.3 |
| 0.9 | 0.1 | 100 | 100 | 100 | 38.7 | 48.5 | 64.7 | 47.7 | 76.7 | 96.9 |
| 1 | 0 | 100 | 100 | 100 | 44.5 | 51.8 | 65.0 | 50.8 | 78.5 | 97.7 |

Panel A of Table 5.1 used stationary series having an autoregressive coefficient value less than one, i.e., 0.3, 0.5, and 0.7 are generated with cross dependence terms $\theta_{12}$ and $\theta_{13}$, keeping deterministic part with only drift term.

The coefficients $\theta_{12}$ and $\theta_{13}$ vary from 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1, in matrix A of DGP given in equation (1). First row of each panel in Table 5.1 shows the size of the test which corresponds to series where $\theta_{12}$ and $\theta_{13} = 0$ at different autoregressive coefficients ($\theta_{11}$ =0.3, $\theta_{22}$ = 0.5, $\theta_{33} = 0.7$). In panels (A) and (B) of Table 5.1, the results indicate the power of the test at each alternative hypothesis when the coefficients of $\theta_{12}$ and $\theta_{13}$ vary from 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1 that treating the power of panel causality tests; Granger non-causality, Sims causality, and FPE method of Hsiao. We find the high power of the Granger non-causality test in each sample size that is small, medium, and large with only drift term. The conventional results in such a scenario indicate that measure of association between x and y ($\theta_{12}$), x and z ($\theta_{13}$)would have very high at each alternative hypothesis. But using the Granger non-causality test and FPE method of Hsiao algorithms in case of $x$ to $y$ ($\theta_{12}$) and $x$ to $z$ ($\theta_{13}$) variable both procedures perform well as compared to Sims causality test.

It is also important to note that actual causal paths go from $x$ to $y$ and $x$ to $z$ in DGP. The power of having these paths is significant and has a regular pattern more in the Granger non-causality test and FPE method of Hsiao.

**Table 5. 2:** Power Analysis of Causality Tests using stationary series with Drift Only for Medium Sample T=50

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | X→Y/causality | | | $\theta_{11} = 0.3$, $\theta_{22} = 0.5$, $\theta_{33} = 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 8.1 | 9.2 | 10.8 | 62.6 | 84.4 | 97.2 | 1.7 | 2.7 | 1.9 |
| 0.1 | 0.9 | 12.7 | 16.2 | 21.1 | 63.8 | 85.7 | 97.9 | 2.3 | 3.5 | 3.8 |
| 0.2 | 0.8 | 29.0 | 44.0 | 63.2 | 65.5 | 86.4 | 98.7 | 4.3 | 6.6 | 13.3 |
| 0.3 | 0.7 | 59.0 | 79.5 | 96.2 | 66.8 | 87.2 | 99.1 | 12.5 | 21.9 | 39.8 |
| 0.4 | 0.6 | 82.4 | 97.0 | 100 | 68.0 | 88.5 | 99.6 | 23.4 | 45.6 | 79.4 |
| 0.5 | 0.5 | 95.7 | 99.8 | 100 | 69.9 | 89.0 | 99.9 | 44.1 | 76.2 | 96.8 |
| 0.6 | 0.4 | 99.3 | 100 | 100 | 71.5 | 90.3 | 100 | 63.3 | 92.9 | 99.5 |
| 0.7 | 0.3 | 99.9 | 100 | 100 | 72.9 | 92.7 | 100 | 85.6 | 99.1 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 73.9 | 94.4 | 100 | 93.9 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 74.0 | 95.4 | 100 | 98.3 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 78.8 | 96.3 | 100 | 99.7 | 100 | 100 |
| Panel (B) | | X→Z/causality | | | $\theta_{11} = 0.3$, $\theta_{22} = 0.5$, $\theta_{33} = 0.7$ | | | | | |
| 0 | 1 | 9.8 | 10.8 | 12.6 | 33.8 | 46.9 | 66.9 | 1.3 | 2.1 | 1.9 |
| 0.1 | 0.9 | 96.1 | 99.8 | 100 | 35.8 | 48.6 | 67.3 | 5.4 | 7.9 | 14.9 |
| 0.2 | 0.8 | 100 | 100 | 100 | 36.1 | 49.8 | 70.0 | 6.8 | 11.5 | 31.6 |
| 0.3 | 0.7 | 100 | 100 | 100 | 37.1 | 50.4 | 74.7 | 13.5 | 21.1 | 42.2 |
| 0.4 | 0.6 | 100 | 100 | 100 | 39.1 | 52.4 | 76.2 | 16.2 | 26.8 | 56.2 |
| 0.5 | 0.5 | 100 | 100 | 100 | 45.3 | 54.1 | 80.2 | 18.4 | 30.8 | 63.5 |
| 0.6 | 0.4 | 100 | 100 | 100 | 46.6 | 56.0 | 83.6 | 21.7 | 39.5 | 73.5 |
| 0.7 | 0.3 | 100 | 100 | 100 | 49.2 | 58.2 | 84.0 | 23.4 | 49.2 | 83.9 |
| 0.8 | 0.2 | 100 | 100 | 100 | 51.6 | 60.1 | 86.8 | 29.1 | 60.5 | 89.3 |
| 0.9 | 0.1 | 100 | 100 | 100 | 56.6 | 69.7 | 89.1 | 35.6 | 63.6 | 93.8 |
| 1 | 0 | 100 | 100 | 100 | 60.1 | 79.9 | 96.8 | 40.6 | 75.9 | 96.3 |

**Table 5. 3**: *Power Analysis of Panel Causality Tests using stationary series with Drift Only for Large Sample T=200*

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | X→Y/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 5.0 | 5.3 | 5.6 | 9.9 | 10.8 | 12.9 | 0.57 | 1.63 | 2.19 |
| 0.1 | 0.9 | 27.9 | 40.1 | 59.2 | 49.9 | 54.2 | 76.1 | 1.28 | 2.89 | 4.69 |
| 0.2 | 0.8 | 86.0 | 98.0 | 100 | 64.0 | 100 | 100 | 18.7 | 35.4 | 72.1 |
| 0.3 | 0.7 | 99.8 | 100 | 100 | 100 | 100 | 100 | 71.7 | 96.5 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 100 | 100 | 100 | 98.2 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Panel (B) | | X→Z/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| 0 | 1 | 5.2 | 5.3 | 5.7 | 82.7 | 96.5 | 98.9 | 0.2 | 0.4 | 0.9 |
| 0.1 | 0.9 | 99.9 | 100 | 100 | 83.3 | 97.5 | 100 | 1.5 | 2.0 | 4.3 |
| 0.2 | 0.8 | 100 | 100 | 100 | 84.3 | 98.4 | 100 | 8.2 | 21.8 | 48.7 |
| 0.3 | 0.7 | 100 | 100 | 100 | 86.9 | 99.1 | 100 | 23.8 | 49.9 | 84.9 |
| 0.4 | 0.6 | 100 | 100 | 100 | 88.8 | 99.9 | 100 | 40.6 | 75.9 | 97.6 |
| 0.5 | 0.5 | 100 | 100 | 100 | 89.9 | 100 | 100 | 54.8 | 88.5 | 99.6 |
| 0.6 | 0.4 | 100 | 100 | 100 | 90.7 | 100 | 100 | 70.8 | 95.9 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 93.6 | 100 | 100 | 82.6 | 99.1 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 95.8 | 100 | 100 | 92.6 | 99.4 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 96.3 | 100 | 100 | 96.5 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 98.8 | 100 | 100 | 99.9 | 100 | 100 |

The same procedure is used in Table 5.2 and Table 5.3, but the only difference is that the panel series have different sample sizes; the medium sample size (i.e. T=50) in Table 5.2 and the large sample size (i.e. T=200) in Table 5.3. In Table 5.2 and Table 5.3, stationary series are generated with cross dependence terms $\theta_{12}$ and $\theta_{13}$, keeping deterministic part only drift term present. We found about the same results as shown in Table 5.1. As the number of time lengths increases, each Panel Causality Test (PCT) gains much power at each stage of increasing cross-section units (i.e. N=5, N=10, and N=20) if DGP and test equation has drift term only. However, the Sims causality test has 78% maximum empirical power at N=5, T=50 when the causal combination is X causes Y in the Panel (A) of Table 5.2. Still, as the cross-section unit increases to 10 and 20, then the empirical power of the SIM test statistic has been observed as 100% at 0.6 alternatives. At large time length, T=200 in Table

5.3, Granger non-causality test, Sims Causality and FPE Method of Hsiao have gained 100% empirical power at 0.1 alternatives even at small cross-section unit.

In comparing size, the GC test has the least size distortion compared to size distortion of SIM and FPE causality tests at small, medium, and large cross-section units. All three tests archive increasing power pattern as a parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium, or large.

**Table 5. 4:** Power Analysis of Panel Causality Tests using stationary series with Drift Only for Small Sample T=25

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Y→X/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} =\ 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 11.8 | 14.1 | 19.5 | 20.0 | 27.3 | 40.8 | 15.8 | 25.3 | 44.8 |
| 0.1 | 0.9 | 29.7 | 37.7 | 57.4 | 21.7 | 30.6 | 43.3 | 19.8 | 32.9 | 57.9 |
| 0.2 | 0.8 | 66.8 | 89.7 | 98.3 | 24.4 | 33.4 | 48.2 | 35.2 | 61.3 | 87.5 |
| 0.3 | 0.7 | 94.5 | 99.8 | 100 | 27.7 | 36.6 | 52.5 | 63.2 | 87.7 | 99.3 |
| 0.4 | 0.6 | 100 | 100 | 100 | 29.2 | 38.4 | 55.3 | 85.4 | 99.9 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 34.3 | 40.1 | 57.9 | 95.6 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 36.8 | 43.3 | 58.4 | 99.5 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 38.0 | 45.1 | 60.1 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 40.8 | 54.2 | 64.6 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 43.5 | 56.2 | 67.3 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 46.0 | 58.0 | 73.0 | 100 | 100 | 100 |
| Panel (B) | | | Y→Z/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} =$ 0.7 | | | | |
| 0 | 1 | 20.4 | 27.8 | 39.4 | 34.5 | 49.0 | 70.3 | 22.7 | 34.9 | 62.2 |
| 0.1 | 0.9 | 95.7 | 99.9 | 100 | 38.9 | 50.1 | 73.6 | 32.7 | 53.7 | 83.1 |
| 0.2 | 0.8 | 100 | 100 | 100 | 40.8 | 54.9 | 76.0 | 38.7 | 61.1 | 89.4 |
| 0.3 | 0.7 | 100 | 100 | 100 | 42.3 | 58.8 | 79.7 | 42.9 | 70.7 | 93.9 |
| 0.4 | 0.6 | 100 | 100 | 100 | 45.7 | 60.7 | 81.5 | 44.8 | 75.1 | 96.8 |
| 0.5 | 0.5 | 100 | 100 | 100 | 47.8 | 64.0 | 85.2 | 53.2 | 80.5 | 97.7 |
| 0.6 | 0.4 | 100 | 100 | 100 | 48.7 | 67.2 | 88.3 | 55.6 | 82.2 | 98.4 |
| 0.7 | 0.3 | 100 | 100 | 100 | 52.9 | 70.7 | 89.0 | 58.3 | 85.5 | 98.9 |
| 0.8 | 0.2 | 100 | 100 | 100 | 55.7 | 73.0 | 93.2 | 60.8 | 86.9 | 99.6 |
| 0.9 | 0.1 | 100 | 100 | 100 | 58.4 | 76.1 | 95.0 | 63.9 | 88.8 | 99.8 |
| 1 | 0 | 100 | 100 | 100 | 60.4 | 80.3 | 97.5 | 66.3 | 91.8 | 99.9 |

Table 5.4 shows the size and power attainment of panel causality tests for small-time dimensions and small, medium and large cross-section dimensions when both DGP and test equation have drift term only. Panel (A) of Table 5.4 explains the results for the situation when Y Granger causes X. In compression of size, the GC test has the least size distortion compared to size distortion of SIM and FPE tests at small, medium, and large cross-section units. All three tests archive increasing power pattern as a parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium or large. This test archives 100% power at 0.4 alternative corresponding to N=5 and at 0.2 for a large cross-section unit and is recognized as the best performer compared to the other two tests. Among Sims and FPE tests, the former gains the least power at all alternatives compared to the latter, which corresponds to small, medium and large cross-section units and is thus identified as the worst performer.

A similar pattern has been observed for almost all tests in Table 5.5 and Table 5.6 when the causal combinations are Y causes X ($\theta_{21}$) and Y causes Z ($\theta_{23}$) in heterogeneous panel data in case of only drift term. At the medium-time length T=50, Panel (A) of Table 5.5 explains the results for the causal combination of Y causes X. In a comparison of size, the GC test has the least size distortion compared to the size distortion of SIM and FPE causality tests at small, medium, and large cross-section units. The (B) panel of Table 5.5 with a causal combination $\theta_{23}$SIM causality test has taken maximum empirical power of 60%, 80%, and 97% at N=5, 10, and 20, respectively. In contrast, GC and FPE tests for both causal combinations have attained maximum empirical power equal to 100%.

All three tests archive increasing power pattern as a parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions. GC test archives 100% power at 0.2 alternative corresponding to N=5 and at 0.1 for large cross-section unit and recognized the best performer compared to the other two tests. Among SIM and FPE tests, the former gains least power at all alternatives compared to the later one corresponding to small, medium and large cross-section units and thus identified as the worst performer.

Further, at T=200, almost all test have attained their highest empirical power equal to 100% in Table 5.6. As the time series and cross-section length increase, the power of PCT tests also increases. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium, or large.

**Table 5. 5:** Power Analysis of Panel Causality Tests using stationary series with Drift Only for Medium Sample T=50

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Y→X/causality | | | $\theta_{11}=0.3,\ \theta_{22}=0.5,\ \theta_{33}=\ 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 8.3 | 9.2 | 11.3 | 15.0 | 20.1 | 28.3 | 1.7 | 2.7 | 4.5 |
| 0.1 | 0.9 | 34.9 | 54.7 | 81.5 | 19.4 | 24.9 | 29.0 | 6.3 | 9.8 | 18.9 |
| 0.2 | 0.8 | 93.6 | 99.6 | 100 | 23.2 | 28.2 | 30.2 | 34.5 | 64.8 | 92.9 |
| 0.3 | 0.7 | 100 | 100 | 100 | 26.5 | 32.5 | 36.2 | 86.7 | 99.9 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 31.6 | 35.1 | 39.2 | 99.5 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 33.2 | 37.5 | 42.8 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 38.7 | 40.8 | 48.4 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 40.8 | 44.2 | 56.7 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 43.2 | 47.9 | 59.0 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 47.2 | 49.5 | 66.3 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 50.3 | 57.4 | 74.5 | 100 | 100 | 100 |
| Panel (B) | | | Y→Z/causality | | | $\theta_{11}=0.3,\ \theta_{22}=0.5,\ \theta_{33}=\ 0.7$ | | | | |
| 0 | 1 | 10.1 | 13.4 | 16.0 | 27.1 | 39.7 | 56.4 | 1.7 | 2.5 | 3.4 |
| 0.1 | 0.9 | 99.1 | 100 | 100 | 28.4 | 40.2 | 57.7 | 7.37 | 9.6 | 22.5 |
| 0.2 | 0.8 | 100 | 100 | 100 | 30.2 | 43.4 | 59.5 | 10.2 | 22.7 | 40.5 |
| 0.3 | 0.7 | 100 | 100 | 100 | 33.3 | 47.6 | 61.8 | 16.8 | 29.9 | 62.2 |
| 0.4 | 0.6 | 100 | 100 | 100 | 36.0 | 49.6 | 64.0 | 24.3 | 48.8 | 84.7 |
| 0.5 | 0.5 | 100 | 100 | 100 | 38.8 | 52.1 | 66.8 | 31.9 | 61.3 | 92.3 |
| 0.6 | 0.4 | 100 | 100 | 100 | 40.3 | 56.9 | 68.9 | 42.7 | 74.9 | 97.3 |
| 0.7 | 0.3 | 100 | 100 | 100 | 43.3 | 59.0 | 74.0 | 50.5 | 81.3 | 98.9 |
| 0.8 | 0.2 | 100 | 100 | 100 | 46.8 | 64.9 | 78.1 | 59.6 | 88.9 | 99.5 |
| 0.9 | 0.1 | 100 | 100 | 100 | 49.8 | 69.9 | 84.4 | 67.5 | 94.2 | 99.9 |
| 1 | 0 | 100 | 100 | 100 | 58.5 | 73.8 | 88.1 | 74.7 | 96.2 | 100 |

**Table 5. 6:** Power Analysis of Panel Causality Tests using stationary series with Drift Only for Large Sample T=200

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Y→X/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho 1$ | $\rho 2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 5.6 | 7.9 | 8.1 | 35.4 | 50.5 | 70.9 | 0.3 | 1.4 | 2.1 |
| 0.1 | 0.9 | 94.4 | 99.8 | 100 | 38.9 | 54.9 | 78.1 | 30.2 | 57.7 | 93.2 |
| 0.2 | 0.8 | 100 | 100 | 100 | 42.9 | 63.1 | 83.7 | 99.9 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 48.3 | 68.5 | 89.0 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 52.3 | 75.0 | 93.8 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 59.1 | 81.5 | 96.3 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 65.8 | 86.9 | 98.4 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 71.5 | 91.3 | 99.4 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 78.0 | 94.0 | 99.7 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 80.5 | 96.1 | 99.9 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 84.2 | 97.3 | 100 | 100 | 100 | 100 |
| Panel (B) | | Y→Z/causality | | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | |
| 0 | 1 | 5.4 | 6.4 | 7.0 | 57.6 | 77.3 | 94.9 | 0.3 | 2.9 | 4.1 |
| 0.1 | 0.9 | 100 | 100 | 100 | 61.2 | 81.4 | 96.6 | 3.2 | 5.5 | 15.8 |
| 0.2 | 0.8 | 100 | 100 | 100 | 66.1 | 86.6 | 97.7 | 17.1 | 41.8 | 75.5 |
| 0.3 | 0.7 | 100 | 100 | 100 | 70.4 | 88.9 | 98.8 | 43.6 | 74.8 | 98.4 |
| 0.4 | 0.6 | 100 | 100 | 100 | 73.2 | 92.0 | 99.5 | 69.5 | 95.7 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 77.0 | 93.9 | 99.7 | 90.5 | 99.5 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 81.0 | 95.7 | 99.8 | 96.7 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 82.7 | 96.8 | 99.9 | 99.8 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 83.9 | 97.1 | 100 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 84.8 | 98.9 | 100 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 86.0 | 99.4 | 100 | 100 | 100 | 100 |

**Table 5. 7:** Power Analysis of Panel Causality Tests using stationary series with Drift Only for Small Sample T=25

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Z→X/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 10.8 | 13.6 | 15.9 | 27.4 | 38.7 | 57.7 | 16.8 | 25.4 | 48.7 |
| 0.1 | 0.9 | 99.9 | 100 | 100 | 33.0 | 39.4 | 58.0 | 94.6 | 99.8 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 36.2 | 40.6 | 60.3 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 39.1 | 43.6 | 62.2 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 40.4 | 46.3 | 64.0 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 43.0 | 48.0 | 66.3 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 46.2 | 50.5 | 68.3 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 48.6 | 54.4 | 70.8 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 52.9 | 58.4 | 73.1 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 55.0 | 60.0 | 75.5 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 63.1 | 68.4 | 79.7 | 100 | 100 | 100 |
| Panel (B) $\theta_{33} = 0.7$ | | | | | Z→Y/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,$ | | |
| 0 | 1 | 10.3 | 12.8 | 16.6 | 40.1 | 60.1 | 80.6 | 14.7 | 25.3 | 52.3 |
| 0.1 | 0.9 | 88.3 | 98.3 | 100 | 46.2 | 58.2 | 84.7 | 92.3 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 49.5 | 59.5 | 88.9 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 50.9 | 63.9 | 90.9 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 54.3 | 65.3 | 93.1 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 58.4 | 68.4 | 97.3 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 62.6 | 70.6 | 98.6 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 65.4 | 73.4 | 99.4 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 67.1 | 75.1 | 99.6 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 74.2 | 77.9 | 99.9 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 78.5 | 85.6 | 100 | 100 | 100 | 100 |

Tables 5.7, Table 5.8, and Table 5.9 show the size and power attainment of panel causality tests when the causal combinations are Z cause X ($\theta_{31}$) and Z causes Y ($\theta_{32}$) in heterogeneous panel data in case of only drift term. At the small-time length T=25, the Panel (A) of Table 5.7 with a causal combination of Z causes X ($\theta_{31}$), SIMS causality test has taken maximum empirical power of 79.7% at N=20, respectively. In compression of size, the GC test has the least size distortion from the nominal size of 5% compared to size distortion of SIM and FPE tests at small, medium, and large cross-section units. All three tests archive increasing power pattern as the parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives,

whether the cross-sectional length is small, medium, or large. This test archives 100% power at 0.2 alternative corresponding to N=5 and at 0.1 for large cross-section unit for both causal combinations and recognized the best performer compared to the other two tests. Among SIM and FPE tests, the former gains least power at all alternatives compared to the latter, which corresponds to small, medium and large cross-section units and is thus identified as the worst performer.

**Table 5. 8:** Power Analysis of Panel Causality Tests using stationary series with Drift Only for Medium Sample T=50

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Z→X/causality | | | $\theta_{11}=0.3,\ \theta_{22}=0.5,\ \theta_{33}=\ 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 6.9 | 7.1 | 8.3 | 21.6 | 29.3 | 41.5 | 0.8 | 1.3 | 1.8 |
| 0.1 | 0.9 | 100 | 100 | 100 | 23.3 | 33.0 | 43.9 | 100 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 25.9 | 36.3 | 47.7 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 27.6 | 37.0 | 48.0 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 29.2 | 39.4 | 49.9 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 32.7 | 40.1 | 50.3 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 34.8 | 44.5 | 52.8 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 39.5 | 48.4 | 55.5 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 41.7 | 49.8 | 59.5 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 45.6 | 53.1 | 61.6 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 55.7 | 60.1 | 65.4 | 100 | 100 | 100 |
| Panel (B) | | | Z→Y/causality | | | $\theta_{11}=0.3,\ \theta_{22}=0.5,\ \theta_{33}=\ 0.7$ | | | | |
| 0 | 1 | 7.3 | 8.0 | 8.5 | 28.1 | 40.6 | 60.8 | 0.3 | 1.7 | 2.3 |
| 0.1 | 0.9 | 99.9 | 100 | 100 | 29.9 | 44.6 | 62.9 | 100 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 33.3 | 47.5 | 64.8 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 35.5 | 50.8 | 68.1 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 38.9 | 52.1 | 69.6 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 40.2 | 54.8 | 70.2 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 42.5 | 56.8 | 74.2 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 45.2 | 58.4 | 77.2 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 48.4 | 60.5 | 79.3 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 52.2 | 64.2 | 80.7 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 58.6 | 70.4 | 83.6 | 100 | 100 | 100 |

The same procedure is used in Table 5.8 and Table 5.9, but the only difference is that the panel series have different sample sizes; the medium sample size (i.e., T=50) in Table 5.8 and the large sample size (i.e. T=200) in Table 5.9. A similar pattern has been observed for almost all tests in Tables 5.8 and 5.9. GC and SIM test archive 100% power at 0.1 alternative corresponding to N=5 and recognized the best performer compared to another test. SIM test gains least power at all alternatives compared to others one corresponding to small, medium, and large cross-section units and thus identified as the worst performer. Further, at the large sample size of T=200, the GC and SIM tests have attained their highest empirical power, equal to 100% in Table 5.9. At the same time, the SIM test has gained a maximum observed power of 83% at T=200 and N=20. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium or large.

**Table 5. 9:** Power Analysis of Panel Causality Tests using stationary series with Drift Only for Large Sample T=200

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Z→X/causality | | | $\theta_{11}=0.3,\ \theta_{22}=0.5,\ \theta_{33}=\ 0.7$ | | | | | |
| $\rho 1$ | $\rho 2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 4.2 | 4.6 | 5.1 | 16.2 | 21.4 | 29.0 | 0.6 | 2.1 | 3.2 |
| 0.1 | 0.9 | 100 | 100 | 100 | 18.1 | 22.5 | 30.4 | 100 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 20.2 | 24.3 | 33.8 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 23.7 | 27.8 | 35.8 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 27.0 | 30.2 | 38.6 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 30.1 | 34.4 | 40.2 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 32.0 | 37.3 | 43.5 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 34.2 | 38.0 | 47.5 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 36.2 | 40.2 | 51.4 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 38.3 | 45.0 | 55.5 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 42.1 | 50.8 | 60.2 | 100 | 100 | 100 |
| Panel (B) | | Z→Y/causality | | | | $\theta_{11}=0.3,\ \theta_{22}=0.5,\ \theta_{33}=$ 0.7 | | | | |
| 0 | 1 | 5.0 | 5.3 | 6.0 | 21.7 | 33.2 | 46.1 | 0.2 | 0.6 | 1.2 |
| 0.1 | 0.9 | 100 | 100 | 100 | 22.4 | 35.1 | 47.5 | 100 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 25.2 | 36.4 | 49.2 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 28.6 | 40.1 | 50.8 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 30.8 | 42.3 | 54.4 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 35.3 | 44.9 | 57.2 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 37.3 | 45.9 | 58.1 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 41.1 | 49.9 | 62.2 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 44.6 | 54.3 | 66.2 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 48.8 | 64.5 | 72.1 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 53.1 | 70.1 | 83.5 | 100 | 100 | 100 |

**Table 5. 10:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for Small Sample T=25

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | X→Y/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 17.8 | 22.5 | 32.1 | 26.3 | 38.0 | 56.1 | 20.7 | 32.6 | 57.4 |
| 0.1 | 0.9 | 18.4 | 25.3 | 34.9 | 30.2 | 41.2 | 59.4 | 23.0 | 35.9 | 57.8 |
| 0.2 | 0.8 | 26.9 | 36.8 | 53.6 | 30.6 | 43.5 | 62.2 | 25.7 | 42.4 | 66.8 |
| 0.3 | 0.7 | 40.7 | 58.5 | 79.5 | 32.1 | 46.9 | 65.9 | 29.2 | 47.2 | 76.8 |
| 0.4 | 0.6 | 59.1 | 79.4 | 95.8 | 34.3 | 48.8 | 69.5 | 34.1 | 59.6 | 86.7 |
| 0.5 | 0.5 | 74.1 | 92.6 | 99.6 | 36.6 | 51.8 | 72.8 | 39.0 | 69.5 | 92.9 |
| 0.6 | 0.4 | 87.7 | 98.2 | 100 | 37.8 | 55.0 | 75.1 | 49.1 | 81.3 | 98.1 |
| 0.7 | 0.3 | 94.7 | 99.8 | 100 | 38.4 | 57.0 | 78.5 | 61.3 | 87.1 | 99.5 |
| 0.8 | 0.2 | 98.3 | 99.9 | 100 | 39.9 | 58.4 | 79.3 | 71.3 | 94.6 | 100 |
| 0.9 | 0.1 | 99.4 | 100 | 100 | 43.0 | 61.8 | 83.3 | 80.7 | 97.0 | 100 |
| 1 | 0 | 99.9 | 100 | 100 | 44.3 | 64.5 | 84.1 | 86.8 | 99.2 | 100 |
| Panel (B) | | X→Z/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| 0 | 1 | 23.6 | 32.0 | 47.0 | 25.7 | 35.9 | 53.1 | 22.6 | 39.5 | 68.2 |
| 0.1 | 0.9 | 49.8 | 70.4 | 90.6 | 27.4 | 36.6 | 54.1 | 26.9 | 46.3 | 75.4 |
| 0.2 | 0.8 | 82.4 | 96.6 | 99.9 | 26.6 | 38.1 | 56.8 | 34.2 | 54.7 | 86.4 |
| 0.3 | 0.7 | 97.9 | 100 | 100 | 27.3 | 40.8 | 58.3 | 40.2 | 67.6 | 93.4 |
| 0.4 | 0.6 | 99.9 | 100 | 100 | 29.1 | 43.2 | 60.5 | 47.6 | 75.7 | 97.6 |
| 0.5 | 0.5 | 100 | 100 | 100 | 30.0 | 47.6 | 64.1 | 51.8 | 83.1 | 98.7 |
| 0.6 | 0.4 | 100 | 100 | 100 | 34.1 | 49.4 | 67.2 | 58.4 | 88.4 | 99.2 |
| 0.7 | 0.3 | 100 | 100 | 100 | 36.2 | 50.1 | 70.2 | 60.7 | 89.7 | 99.6 |
| 0.8 | 0.2 | 100 | 100 | 100 | 37.5 | 53.4 | 73.3 | 69.6 | 93.8 | 99.8 |
| 0.9 | 0.1 | 100 | 100 | 100 | 38.0 | 54.5 | 75.1 | 69.7 | 94.4 | 100 |
| 1 | 0 | 100 | 100 | 100 | 40.4 | 55.5 | 81.8 | 71.6 | 94.8 | 100 |

Table 5.10 shows the size and power analysis of the panel causality test with the causal combinations are X causes Y ($\theta_{12}$) and X causes Z ($\theta_{13}$) in heterogeneous panel data in case of keeping both deterministic parts with drift trend term.

At the small-time length T=25, the Panel (A) of Table 10 with a causal combination of X causes Y ($\theta_{12}$). In compression of size, the GC test has the least size distortion compared to size distortion of SIM and FPE tests at small, medium, and large cross-section units. All three tests archive increasing power pattern as the parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium or large. This test archives 100% power at 0.5 alternative corresponding to N=5 and at 0.3 for large cross-section unit for causal combination X causes Z ($\theta_{13}$) in the Panel (B) and recognized the best performer compared to the other two tests. SIM causality test has taken maximum empirical power of 84.1% at N=20, respectively. Among SIM and FPE tests, the former gains the least power at all alternatives compared to the later one corresponding to small, medium and large cross-section units and thus identified as the worst performer.

**Table 5. 11:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for Medium Sample T=50

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | X→Y/causality | | | $\theta_{11}=0.3,\ \theta_{22}=0.5,\ \theta_{33}=0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 10.8 | 11.6 | 14.7 | 54.1 | 74.5 | 94.0 | 1.1 | 2.5 | 3.3 |
| 0.1 | 0.9 | 12.4 | 16.9 | 22.6 | 55.9 | 77.1 | 94.2 | 3.3 | 3.5 | 5.3 |
| 0.2 | 0.8 | 31.0 | 45.0 | 65.3 | 56.5 | 79.8 | 95.0 | 3.7 | 7.1 | 12.8 |
| 0.3 | 0.7 | 58.4 | 79.7 | 96.0 | 59.3 | 80.8 | 96.4 | 10.6 | 19.6 | 39.9 |
| 0.4 | 0.6 | 83.6 | 96.7 | 99.9 | 62.2 | 82.8 | 96.5 | 24.6 | 47.2 | 79.6 |
| 0.5 | 0.5 | 96.4 | 99.8 | 100 | 61.9 | 84.5 | 97.4 | 46.2 | 75.9 | 97.6 |
| 0.6 | 0.4 | 99.2 | 100 | 100 | 63.3 | 84.8 | 98.0 | 66.8 | 92.9 | 99.9 |
| 0.7 | 0.3 | 99.9 | 100 | 100 | 64.6 | 85.9 | 98.5 | 83.8 | 98.5 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 65.7 | 86.8 | 99.4 | 95.4 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 66.7 | 88.0 | 100 | 98.9 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 67.8 | 89.1 | 100 | 99.3 | 100 | 100 |
| Panel (B) | | X→Z/causality | | | $\theta_{11}=0.3,\ \theta_{22}=0.5,\ \theta_{33}=0.7$ | | | | | |
| 0 | 1 | 11.8 | 14.4 | 19.4 | 28.2 | 40.6 | 58.7 | 2.5 | 3.2 | 4.6 |
| 0.1 | 0.9 | 47.2 | 68.2 | 89.2 | 29.0 | 41.3 | 60.6 | 2.8 | 4.8 | 8.7 |
| 0.2 | 0.8 | 93.6 | 99.5 | 100 | 30.2 | 42.2 | 61.9 | 6.1 | 12.6 | 21.6 |
| 0.3 | 0.7 | 99.9 | 100 | 100 | 33.9 | 43.7 | 63.7 | 11.4 | 21.8 | 43.5 |
| 0.4 | 0.6 | 100 | 100 | 100 | 34.7 | 45.4 | 66.3 | 19.7 | 35.6 | 71.6 |
| 0.5 | 0.5 | 100 | 100 | 100 | 36.6 | 46.9 | 66.8 | 28.3 | 52.7 | 85.3 |
| 0.6 | 0.4 | 100 | 100 | 100 | 37.6 | 47.7 | 68.7 | 32.1 | 62.6 | 94.1 |
| 0.7 | 0.3 | 100 | 100 | 100 | 38.9 | 48.8 | 70.5 | 40.2 | 72.4 | 97.2 |
| 0.8 | 0.2 | 100 | 100 | 100 | 39.4 | 49.9 | 71.9 | 50.1 | 81.3 | 98.6 |
| 0.9 | 0.1 | 100 | 100 | 100 | 40.0 | 50.4 | 72.6 | 56.3 | 87.9 | 99.7 |
| 1 | 0 | 100 | 100 | 100 | 42.5 | 52.0 | 74.0 | 61.7 | 91.5 | 99.9 |

**Table 5. 12:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for large Sample T=200

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | X→Y/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho 1$ | $\rho 2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 5.5 | 5.8 | 6.1 | 38.2 | 54.7 | 76.0 | 0.5 | 1.50 | 2.2 |
| 0.1 | 0.9 | 26.4 | 38.0 | 58.5 | 39.1 | 55.8 | 77.3 | 1.5 | 2.19 | 3.8 |
| 0.2 | 0.8 | 85.1 | 97.6 | 100 | 40.6 | 56.9 | 78.6 | 16.6 | 38.9 | 74.4 |
| 0.3 | 0.7 | 99.9 | 100 | 100 | 41.0 | 57.9 | 79.6 | 70.7 | 95.3 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 42.1 | 58.2 | 80.1 | 97.4 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 43.1 | 59.1 | 80.6 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 44.7 | 60.5 | 81.9 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 46.8 | 61.5 | 83.0 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 48.4 | 63.6 | 83.9 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 49.3 | 64.3 | 84.2 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 50.1 | 66.7 | 85.8 | 100 | 100 | 100 |
| Panel (B) | | X→Z/causality | | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | |
| 0 | 1 | 5.6 | 5.9 | 7.1 | 78.9 | 95.4 | 99.4 | 0.1 | 0.4 | 0.9 |
| 0.1 | 0.9 | 55.9 | 76.2 | 94.9 | 80.0 | 95.5 | 99.7 | 0.8 | 1.3 | 2.1 |
| 0.2 | 0.8 | 99.6 | 100 | 100 | 82.4 | 95.9 | 99.9 | 2.7 | 3.5 | 6.9 |
| 0.3 | 0.7 | 100 | 100 | 100 | 83.4 | 96.4 | 100 | 9.3 | 20.5 | 42.5 |
| 0.4 | 0.6 | 100 | 100 | 100 | 84.7 | 97.2 | 100 | 27.6 | 55.7 | 88.6 |
| 0.5 | 0.5 | 100 | 100 | 100 | 85.8 | 98.4 | 100 | 52.6 | 86.3 | 99.9 |
| 0.6 | 0.4 | 100 | 100 | 100 | 86.1 | 98.8 | 100 | 74.3 | 97.4 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 87.0 | 99.6 | 100 | 88.1 | 99.9 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 88.9 | 99.9 | 100 | 96.6 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 89.6 | 100 | 100 | 98.7 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 90.4 | 100 | 100 | 99.9 | 100 | 100 |

The same procedure is used in Table 5.11 and Table 5.12, but the only difference is that the panel series have different sample sizes; the medium sample size (i.e., T=50) in Table 5.11 and the large sample size (i.e. T=200) in Table 5.12. A similar pattern has been observed for almost all tests in Tables 5.11 and 5.12. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium, or large. This test archives 100% power at 0.4 alternative corresponding to N=5 and at 0.2 for a large cross-section unit in the Panel (A) and recognized the best performer compared to the other two tests. Among SIM and FPE tests, the former gains the least power at all alternatives compared to the latter, which corresponds to small, medium and large cross-section units and is thus identified as the worst performer.

**Table 5. 13:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for Small Sample T=25

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Y→X/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho 1$ | $\rho 2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 18.3 | 24.3 | 35.9 | 21.1 | 33.7 | 47.6 | 17.1 | 29.3 | 54.1 |
| 0.1 | 0.9 | 48.5 | 68.3 | 88.6 | 24.1 | 35.1 | 48.7 | 23.2 | 38.8 | 66.1 |
| 0.2 | 0.8 | 87.8 | 98.3 | 100 | 26.9 | 37.1 | 49.2 | 39.5 | 66.9 | 88.8 |
| 0.3 | 0.7 | 99.4 | 100 | 100 | 28.5 | 38.2 | 51.5 | 63.5 | 90.2 | 99.6 |
| 0.4 | 0.6 | 100 | 100 | 100 | 30.2 | 40.4 | 54.9 | 86.7 | 98.6 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 35.2 | 43.4 | 56.3 | 96.6 | 99.9 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 36.8 | 45.1 | 57.3 | 99.5 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 39.7 | 48.1 | 59.8 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 40.2 | 50.4 | 61.1 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 42.4 | 53.5 | 64.1 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 45.4 | 55.5 | 65.2 | 100 | 100 | 100 |
| Panel (B) | | Y→Z/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| 0 | 1 | 32.9 | 47.8 | 66.4 | 37.0 | 53.9 | 74.7 | 29.3 | 47.3 | 75.8 |
| 0.1 | 0.9 | 71.9 | 90.9 | 99.1 | 39.6 | 55.8 | 76.1 | 37.5 | 58.3 | 85.8 |
| 0.2 | 0.8 | 97.1 | 99.9 | 100 | 40.6 | 58.1 | 79.1 | 46.8 | 75.7 | 95.9 |
| 0.3 | 0.7 | 99.8 | 100 | 100 | 46.2 | 61.9 | 81.6 | 60.8 | 89.4 | 99.1 |
| 0.4 | 0.6 | 100 | 100 | 100 | 47.0 | 63.0 | 83.1 | 72.3 | 94.6 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 49.5 | 67.5 | 84.0 | 79.4 | 98.1 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 50.4 | 69.9 | 86.2 | 82.9 | 98.6 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 51.1 | 70.7 | 91.5 | 86.2 | 99.3 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 53.7 | 72.0 | 94.5 | 89.6 | 99.9 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 54.4 | 76.5 | 95.0 | 89.8 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 55.9 | 81.3 | 98.5 | 91.5 | 100 | 100 |

The Panel (A) of Table 5.13, the causal combinations are Y causes X ($\theta_{21}$) and Y causes Z ($\theta_{23}$) in heterogeneous panel data to keep both deterministic parts with drift and trend term.

At the small-time length T=25, the Panel (A) of Table 5.13 with a causal combination of Y causes X ($\theta_{21}$). In compression of size, the GC test has the least size distortion compared to size distortion of SIM and FPE tests at small, medium, and large cross-section units. All three tests archive increasing power pattern as a parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium or large. This test archives 100% power at 0.4 alternative corresponding to N=5 and at 0.2 for a large cross-section unit for both causal combinations and recognizes the best performer compared to the other two tests. SIM causality test has taken maximum empirical power of 65.2% and 98.5% at N=20, for Y causes X and Y causes Z, respectively. Among Sims and FPE tests, the former gains the least power at all alternatives compared to the latter corresponding to small, medium and large cross-section units and thus identified as the worst performer.

**Table 5. 14:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for Medium Sample T=50

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Y→X/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 9.8 | 11.0 | 13.6 | 14.8 | 19.6 | 22.5 | 1.8 | 2.7 | 2.9 |
| 0.1 | 0.9 | 55.4 | 76.6 | 94.4 | 16.9 | 22.5 | 25.2 | 6.9 | 11.1 | 20.3 |
| 0.2 | 0.8 | 98.8 | 100 | 100 | 19.1 | 25.9 | 29.1 | 34.2 | 66.3 | 92.7 |
| 0.3 | 0.7 | 100 | 100 | 100 | 24.4 | 29.7 | 33.2 | 86.1 | 99.3 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 25.8 | 32.8 | 38.9 | 99.6 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 27.7 | 37.5 | 42.8 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 29.4 | 39.5 | 45.6 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 30.5 | 43.3 | 48.6 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 31.9 | 45.7 | 54.2 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 35.4 | 48.7 | 57.8 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 45.9 | 50.8 | 65.1 | 100 | 100 | 100 |
| Panel (B) | | | Y→Z/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | |
| 0 | 1 | 14.7 | 20.3 | 28.0 | 25.3 | 36.4 | 52.4 | 2.9 | 4.3 | 5.7 |
| 0.1 | 0.9 | 81.6 | 96.2 | 99.9 | 28.5 | 38.3 | 55.3 | 5.2 | 8.1 | 19.5 |
| 0.2 | 0.8 | 100 | 100 | 100 | 32.4 | 43.7 | 58.7 | 15.9 | 34.8 | 59.8 |
| 0.3 | 0.7 | 100 | 100 | 100 | 35.5 | 46.6 | 62.8 | 34.5 | 62.8 | 92.9 |
| 0.4 | 0.6 | 100 | 100 | 100 | 36.6 | 46.7 | 64.0 | 53.2 | 86.8 | 99.4 |
| 0.5 | 0.5 | 100 | 100 | 100 | 38.1 | 48.4 | 65.9 | 67.8 | 95.9 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 49.5 | 51.7 | 69.2 | 82.3 | 98.4 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 50.5 | 55.1 | 74.2 | 89.1 | 99.5 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 54.9 | 59.3 | 75.2 | 93.1 | 99.8 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 56.9 | 64.3 | 79.8 | 95.7 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 61.2 | 68.8 | 81.1 | 96.4 | 100 | 100 |

**Table 5. 15:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for Large Sample T=200

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Y→X/causality | | | $\theta_{11} = 0.3, \theta_{22} = 0.5, \ \theta_{33} = \ 0.7$ | | | | | |
| $\rho 1$ | $\rho 2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 5.3 | 5.4 | 6.0 | 30.7 | 46.6 | 68.6 | 0.3 | 0.5 | 1.2 |
| 0.1 | 0.9 | 97.0 | 99.9 | 100 | 100 | 100 | 74.1 | 30.4 | 58.6 | 92.9 |
| 0.2 | 0.8 | 100 | 100 | 100 | 100 | 100 | 81.1 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 100 | 100 | 86.4 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 100 | 100 | 90.8 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 100 | 100 | 95.4 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 100 | 100 | 97.6 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 100 | 100 | 98.6 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 100 | 100 | 99.7 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 81.6 | 97.8 | 100 | 100 | 100 | 100 |
| Panel (B) | | Y→Z/causality | | | $\theta_{11} = 0.3, \theta_{22} = 0.5, \ \theta_{33} = \ 0.7$ | | | | | |
| 0 | 1 | 5.8 | 7.1 | 7.9 | 46.3 | 65.1 | 86.0 | 0.2 | 0.6 | 0.9 |
| 0.1 | 0.9 | 94.5 | 99.7 | 100 | 49.9 | 71.7 | 90.3 | 0.9 | 1.10 | 2.8 |
| 0.2 | 0.8 | 100 | 100 | 100 | 55.0 | 75.8 | 93.5 | 14.1 | 25.8 | 57.2 |
| 0.3 | 0.7 | 100 | 100 | 100 | 60.3 | 81.2 | 96.4 | 50.2 | 83.3 | 99.6 |
| 0.4 | 0.6 | 100 | 100 | 100 | 65.2 | 85.7 | 97.7 | 86.2 | 99.7 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 69.8 | 89.5 | 98.7 | 98.1 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 74.8 | 92.8 | 99.5 | 99.9 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 78.6 | 95.3 | 99.8 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 82.3 | 96.2 | 99.9 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 83.4 | 96.7 | 100 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 84.7 | 97.2 | 100 | 100 | 100 | 100 |

The same procedure is used in Table 5.14 and Table 5.15, but the only difference is that the panel series have different sample sizes; the medium sample size (i.e., T=50) in Table 5.14 and the large sample size (i.e., T=200) in Table 5.15. A similar pattern has been observed for almost all tests in the abovementioned tables. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium or large. This test archives 100% power at 0.2 alternative corresponding to N=5 and at 0.1 for a large cross-section unit and recognized the best performer compared to the other two tests. Among SIM and FPE tests, the former gains the least power at all alternatives compared to the latter, which corresponds to small, medium and large cross-section units and is thus identified as the worst performer.

**Table 5. 16:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for Small Sample T=25

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Z→X/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 12.9 | 15.8 | 20.6 | 36.3 | 51.8 | 73.8 | 17.9 | 29.8 | 52.6 |
| 0.1 | 0.9 | 100 | 100 | 100 | 38.4 | 52.5 | 75.5 | 93.1 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 43.9 | 55.7 | 81.0 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 44.9 | 58.7 | 83.2 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 46.3 | 61.3 | 88.1 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 48.7 | 67.0 | 89.7 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 52.9 | 68.6 | 90.3 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 54.2 | 71.6 | 93.5 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 57.6 | 72.2 | 97.6 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 64.9 | 77.7 | 98.5 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 67.8 | 84.3 | 99.6 | 100 | 100 | 100 |
| Panel (B) | | Z→Y/causality | | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | |
| 0 | 1 | 12.9 | 15.8 | 21.5 | 76.9 | 94.0 | 99.6 | 18.8 | 34.1 | 61.9 |
| 0.1 | 0.9 | 87.2 | 97.9 | 100 | 100 | 100 | 100 | 65.3 | 90.4 | 99.6 |
| 0.2 | 0.8 | 100 | 100 | 100 | 100 | 100 | 100 | 99.5 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table 5.16, Table 5.17, and Table 5.18 show the size and power attainment of panel causality tests when the causal combinations are Z cause X ($\theta_{31}$) and Z causes Y ($\theta_{32}$) in heterogeneous panel data in case of with drift and trend term. At the small-time length T=25, the Panel (A) of Table 5.16 with a causal combination of Z causes X ($\theta_{31}$), SIM causality test has taken maximum empirical power of 99.6% at N=20. In compression of size, the GC test has the least size distortion compared to size distortion of SIM and FPE tests at small, medium, and large cross-section units. All three tests archive increasing power pattern as a parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions. However, the power attainment of GC and FPE tests is much better than SIM tests at all alternatives, whether the cross-sectional length is small, medium, or large. GC and

FPE test archive 100% power at 0.2 alternative corresponding to N=5 and at 0.1 for large cross-section unit for both causal combinations and recognized the best performer compared to tests. Among SIM and FPE tests, the former gains the least power at all alternatives compared to the latter, which corresponds to small, medium and large cross-section units and is thus identified as the worst performer.

**Table 5. 17:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for Medium Sample T=50

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Z→X/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho1$ | $\rho2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 7.7 | 8.3 | 9.6 | 24.9 | 36.3 | 51.8 | 1.2 | 2.7 | 2.8 |
| 0.1 | 0.9 | 100 | 100 | 100 | 25.8 | 38.0 | 53.7 | 100 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 27.5 | 40.2 | 55.0 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 34.4 | 43.4 | 58.9 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 35.7 | 45.8 | 60.5 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 38.6 | 51.7 | 63.3 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 40.3 | 52.9 | 65.3 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 43.1 | 53.2 | 68.5 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 45.8 | 56.0 | 70.8 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 47.0 | 59.4 | 72.1 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 56.9 | 65.6 | 74.3 | 100 | 100 | 100 |
| Panel (B) | | Z→Y/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| 0 | 1 | 7.9 | 8.4 | 9.5 | 50.5 | 70.8 | 91.5 | 1.7 | 2.6 | 3.3 |
| 0.1 | 0.9 | 99.8 | 100 | 100 | 53.0 | 73.9 | 93.2 | 88.6 | 99.6 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 59.6 | 76.1 | 95.3 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 61.4 | 78.0 | 97.7 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 69.9 | 81.1 | 99.5 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 70.5 | 86.9 | 99.9 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 72.5 | 88.0 | 100 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 75.8 | 89.3 | 100 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 81.8 | 90.3 | 100 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 85.8 | 94.7 | 100 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 90.0 | 97.5 | 100 | 100 | 100 | 100 |

**Table 5. 18:** Power Analysis of Panel Causality Tests using stationary series with Drift and Trend for Large Sample T=200

| | | Granger Non-causality | | | Sims Causality | | | FPE Hsiao Causality | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel (A) | | Z→X/causality | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | | |
| $\rho 1$ | $\rho 2$ | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* | *N=5* | *N=10* | *N=20* |
| 0 | 1 | 5.0 | 5.6 | 6.1 | 17.9 | 22.9 | 32.4 | 0.4 | 0.7 | 0.9 |
| 0.1 | 0.9 | 100 | 100 | 100 | 18.2 | 25.5 | 36.0 | 100 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 20.8 | 28.0 | 38.5 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 21.1 | 29.5 | 41.0 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 24.9 | 31.1 | 43.7 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 27.6 | 34.0 | 46.8 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 28.2 | 35.5 | 51.1 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 30.2 | 38.1 | 54.1 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 33.5 | 41.9 | 56.8 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 36.0 | 43.6 | 59.7 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 39.7 | 47.0 | 64.2 | 100 | 100 | 100 |
| Panel (B) | | Z→Y/causality | | | | $\theta_{11} = 0.3,\ \theta_{22} = 0.5,\ \theta_{33} = 0.7$ | | | | |
| 0 | 1 | 4.6 | 4.8 | 5.6 | 24.1 | 34.4 | 50.4 | 0.2 | 0.5 | 0.8 |
| 0.1 | 0.9 | 100 | 100 | 100 | 27.9 | 37.4 | 64.3 | 100 | 100 | 100 |
| 0.2 | 0.8 | 100 | 100 | 100 | 32.9 | 38.9 | 67.7 | 100 | 100 | 100 |
| 0.3 | 0.7 | 100 | 100 | 100 | 34.9 | 41.9 | 69.4 | 100 | 100 | 100 |
| 0.4 | 0.6 | 100 | 100 | 100 | 36.4 | 47.1 | 70.8 | 100 | 100 | 100 |
| 0.5 | 0.5 | 100 | 100 | 100 | 39.2 | 53.7 | 71.1 | 100 | 100 | 100 |
| 0.6 | 0.4 | 100 | 100 | 100 | 43.6 | 55.7 | 76.7 | 100 | 100 | 100 |
| 0.7 | 0.3 | 100 | 100 | 100 | 45.6 | 57.0 | 79.4 | 100 | 100 | 100 |
| 0.8 | 0.2 | 100 | 100 | 100 | 46.7 | 59.9 | 84.4 | 100 | 100 | 100 |
| 0.9 | 0.1 | 100 | 100 | 100 | 49.5 | 64.8 | 85.0 | 100 | 100 | 100 |
| 1 | 0 | 100 | 100 | 100 | 51.2 | 70.9 | 88.5 | 100 | 100 | 100 |

The same procedure is used in Table 5.17 and Table 5.18, but the only difference is that the panel series have different sample sizes; the medium sample size (i.e., T=50) in Table 5.17 and the large sample size (i.e., T=200) in Table 5.18. A similar pattern has been observed for almost all tests in the abovementioned tables. However, the power attainment of GC and FPE tests is much better than SIM tests at all alternatives, whether the cross-sectional length is small, medium, or large. GC and FPE test archive 100% power at 0.1 alternative corresponding to N=5 and for large cross-section unit for both causal combinations and recognized the best performer in comparison of tests. Among SIM and FPE tests, the former gains the least power at all alternatives compared to the latter, which corresponds to small, medium and large cross-section units and is thus identified as the worst performer. Further, at T=200, almost all test has attained their highest empirical power equal to 100% in Table 5.18. As the time series and cross-section length increase the power of PCT tests also increases. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium or large.

***Figure 5. 1***: Power Analysis of Panel Causality Tests for X→Y and X→Z with Drift Only, T=25

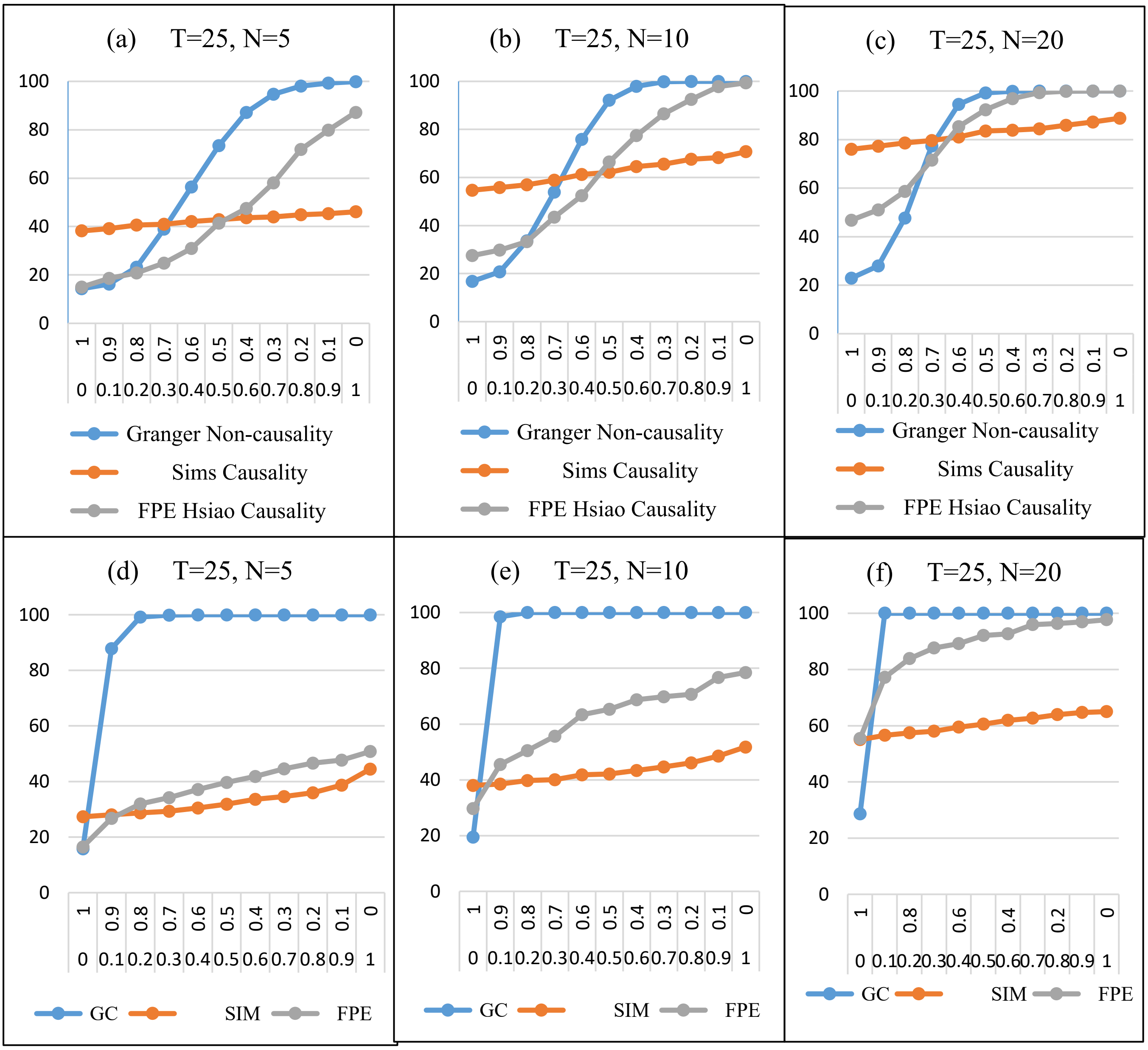


*Figure 5. 2*: Power Analysis of Panel Causality Tests for X→Y and X→Z with Drift Only, T=50

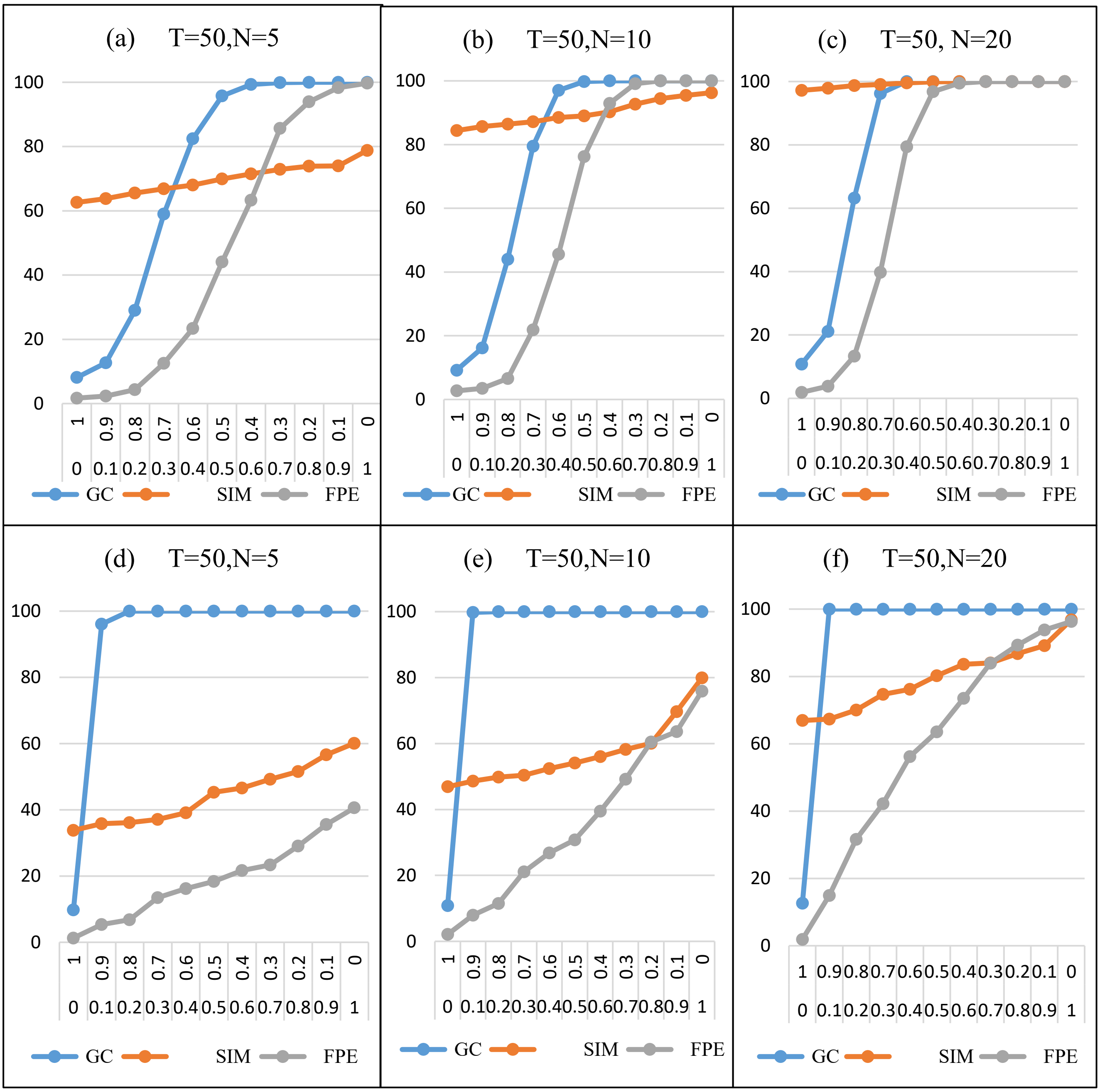
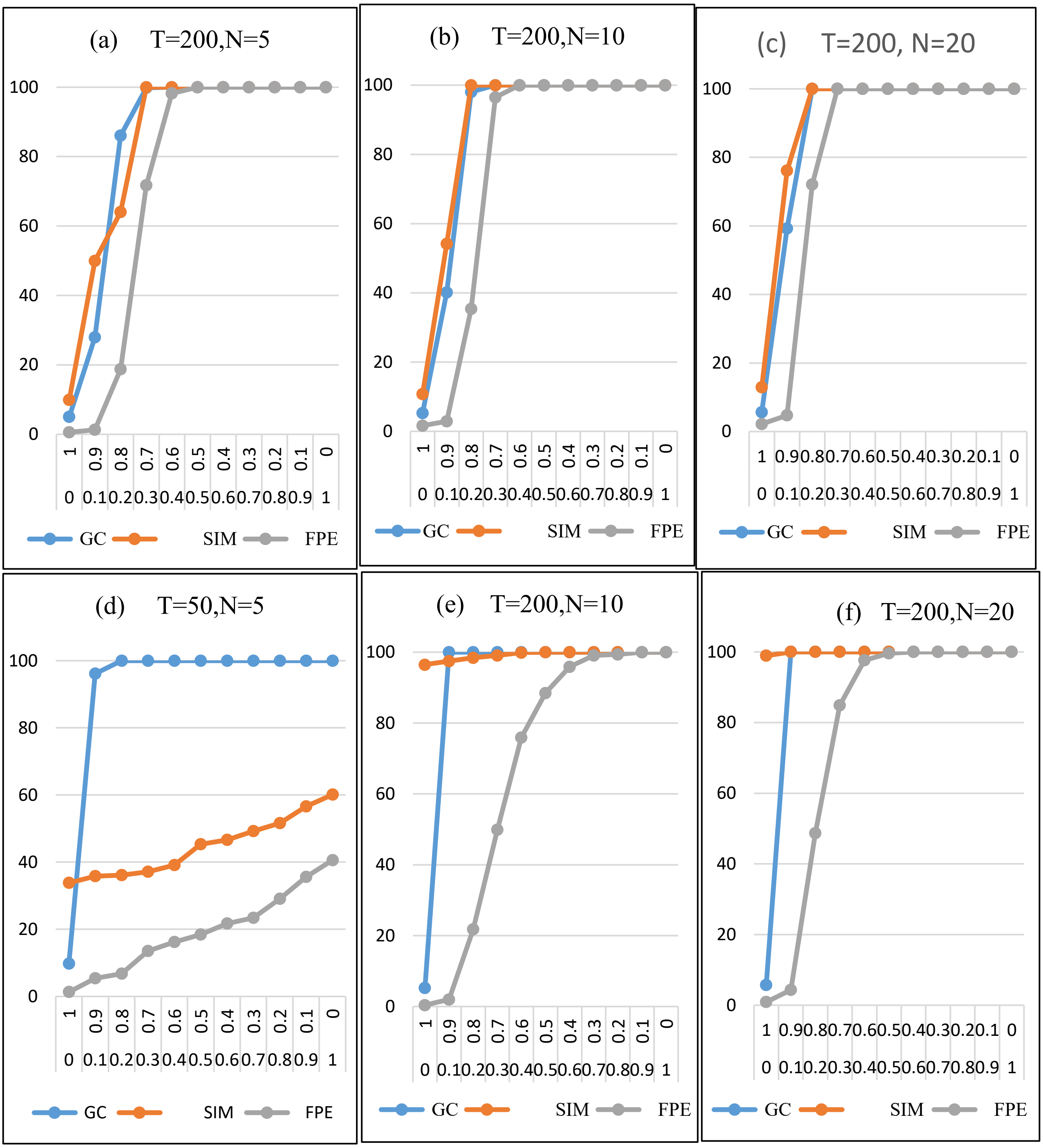


***Figure 5. 3:*** Power Analysis of Panel Causality Tests for X→Y and X→Z with Drift Only, T=200



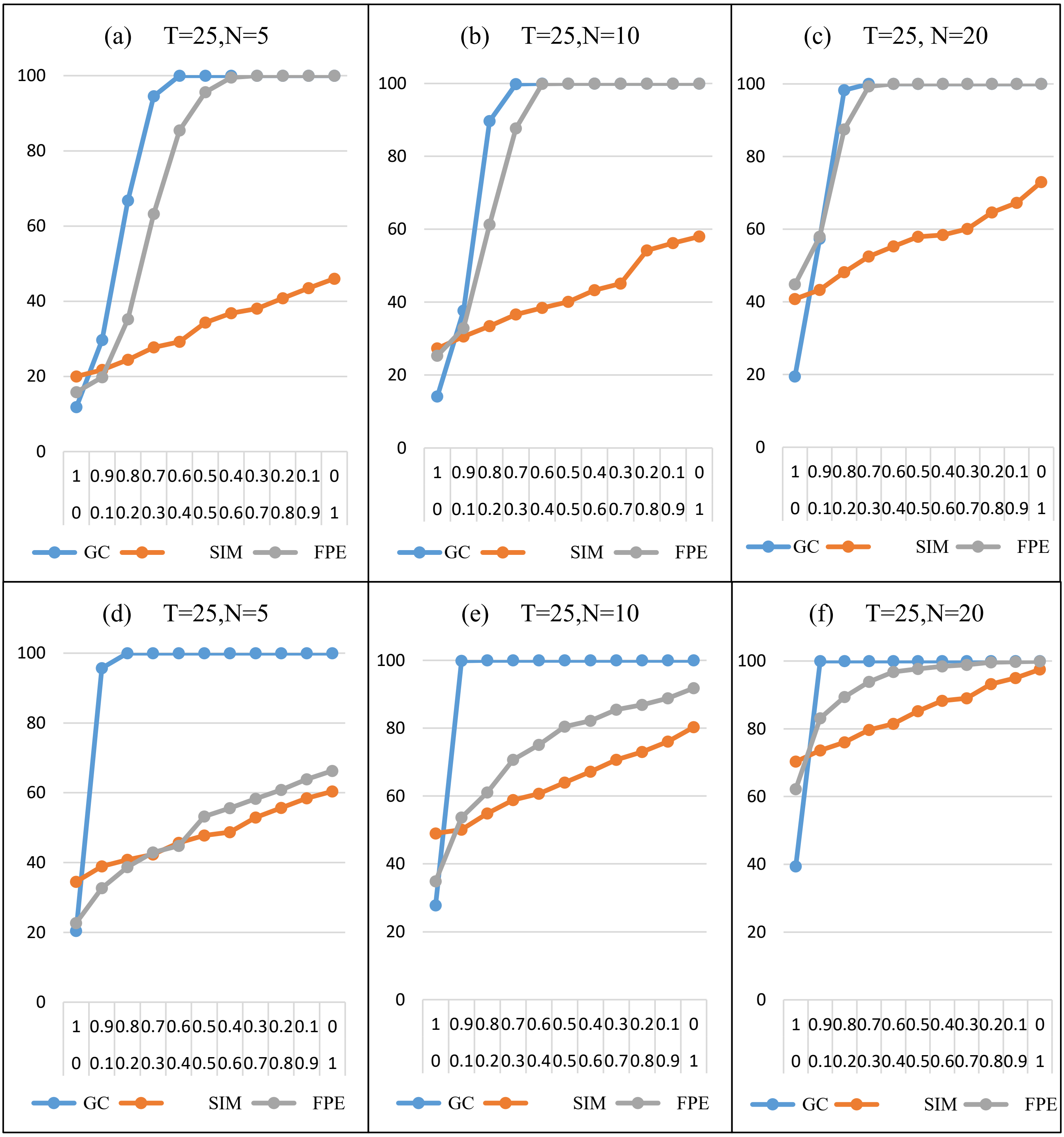



***Figure 5. 4*:**Power Analysis of Panel Causality Tests for Y→X and Y→Z with Drift Only, T=25

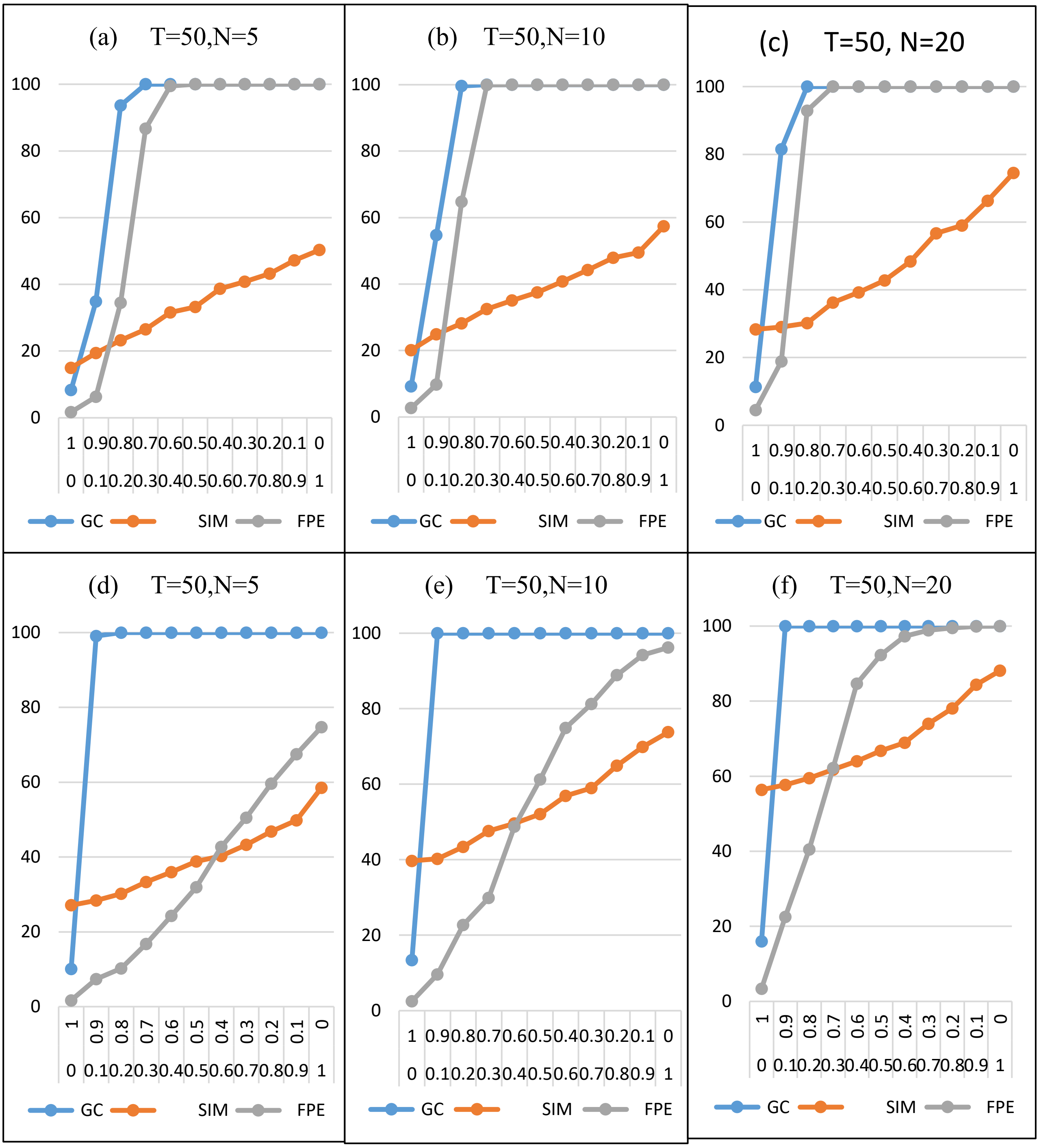


***Figure 5. 5*:** Power Analysis of Panel Causality Tests for Y→X and Y→Z with Drift Only, T=50

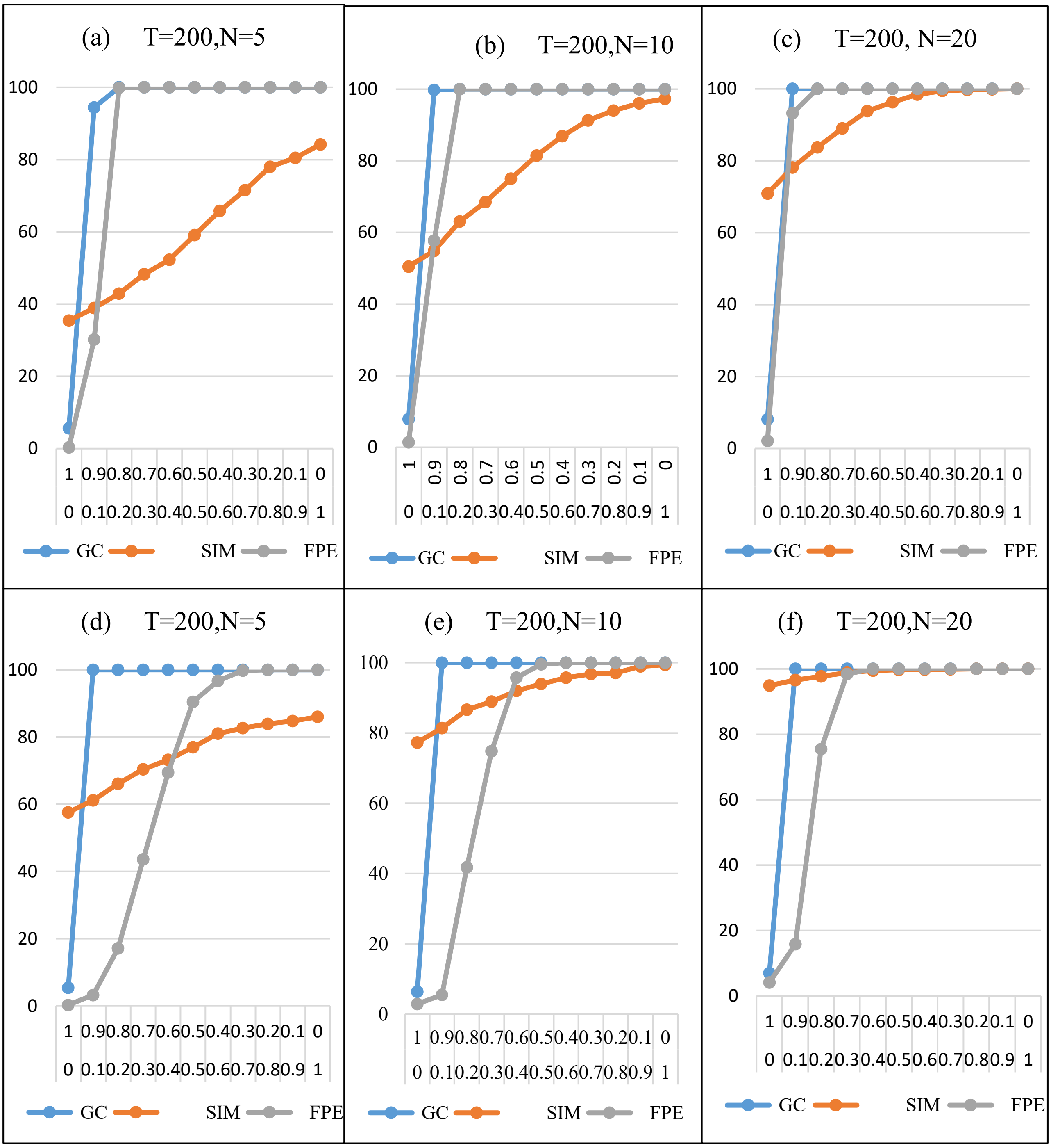



***Figure 5. 6***: Power Analysis of Panel Causality Tests for Y→X and Y→Z with Drift Only, T=200

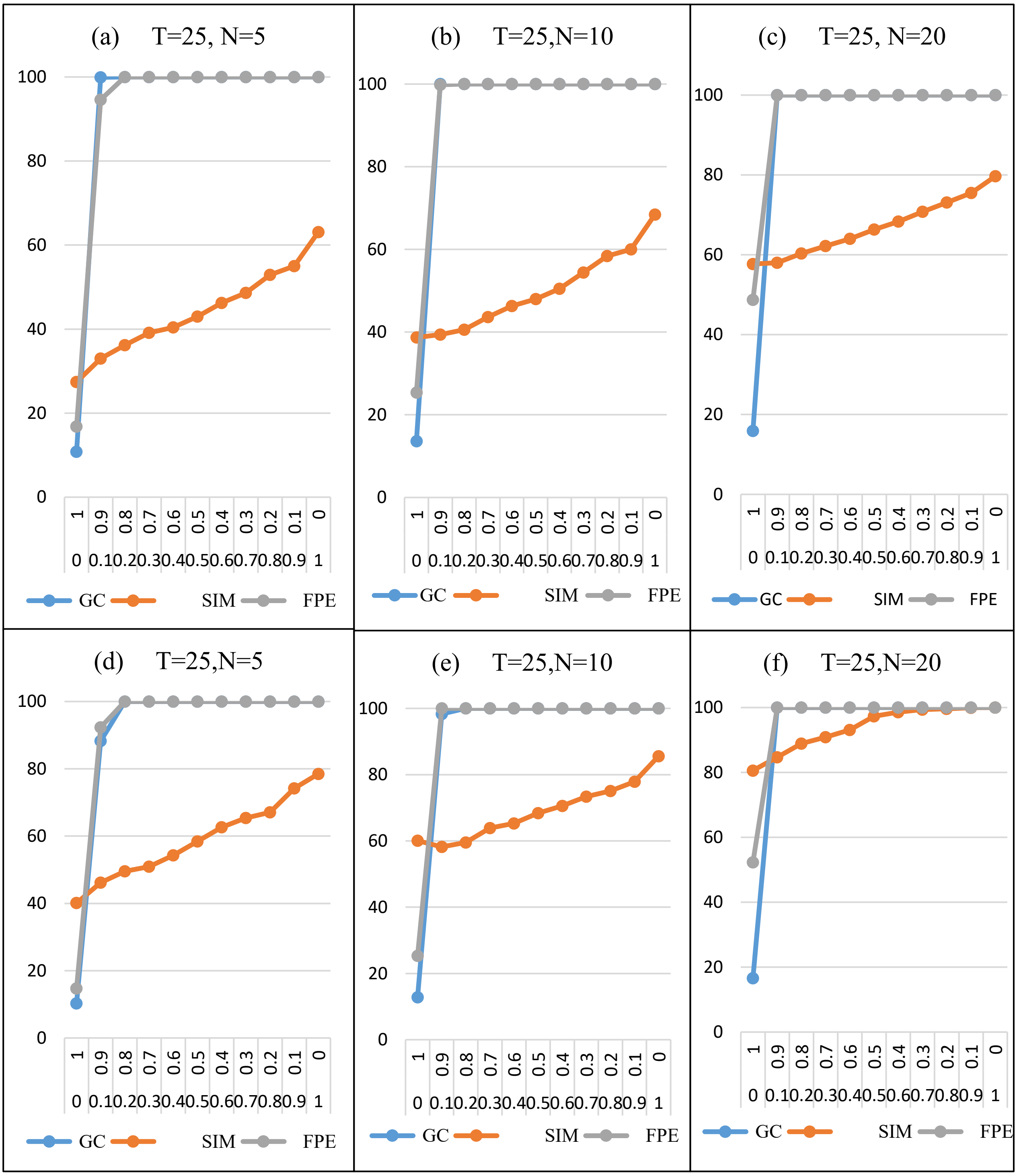



***Figure 5. 7*:** Power Analysis of Panel Causality Tests for Z→X and Z→Y with Drift Only, T=25

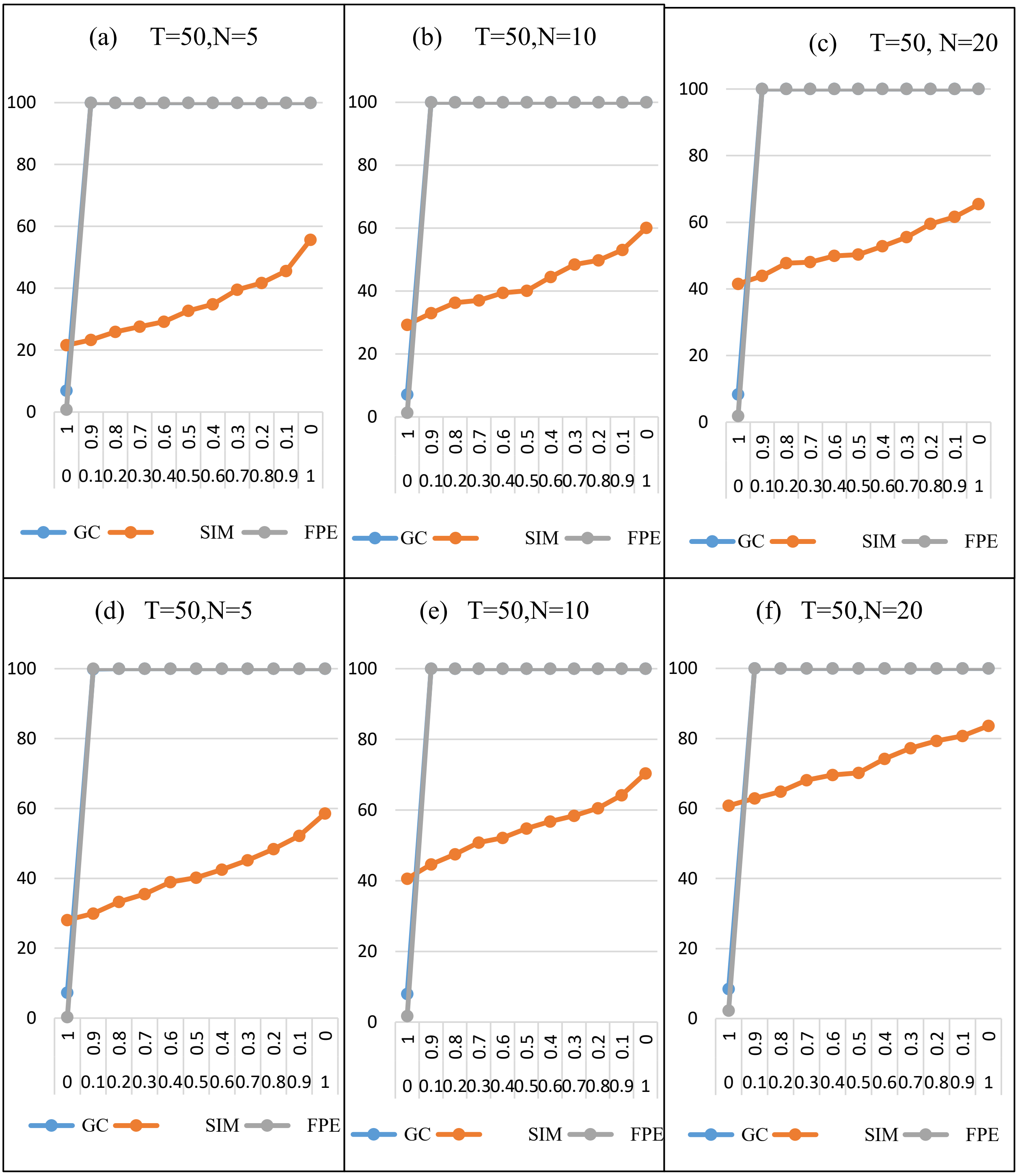



***Figure 5. 8*:** Power Analysis of Panel Causality Tests for Z→X and Z→Y with Drift Only, T=50

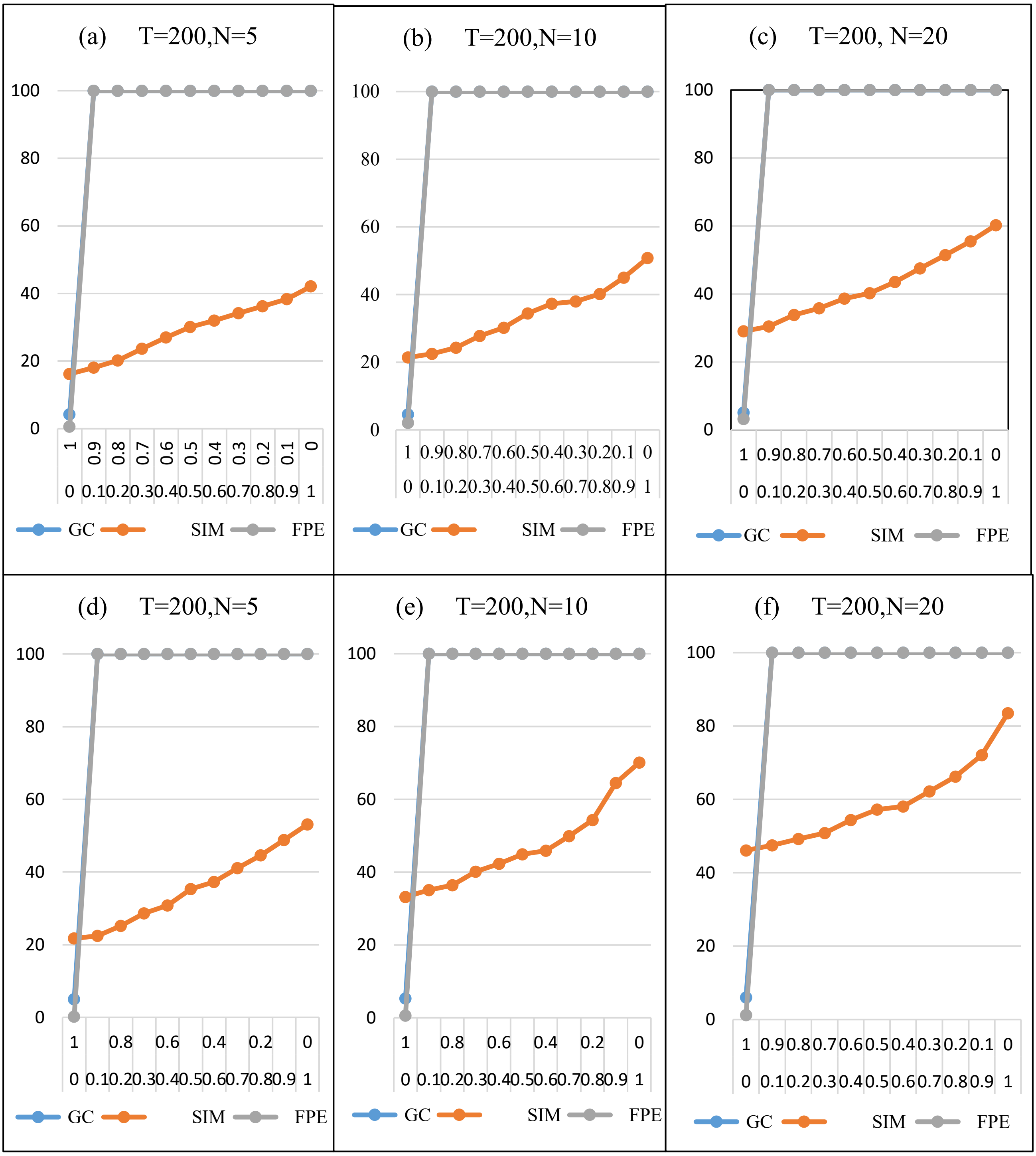



***Figure 5. 9***: Power Analysis of Panel Causality Tests for Z→X and Z→Y with Drift Only, T=200

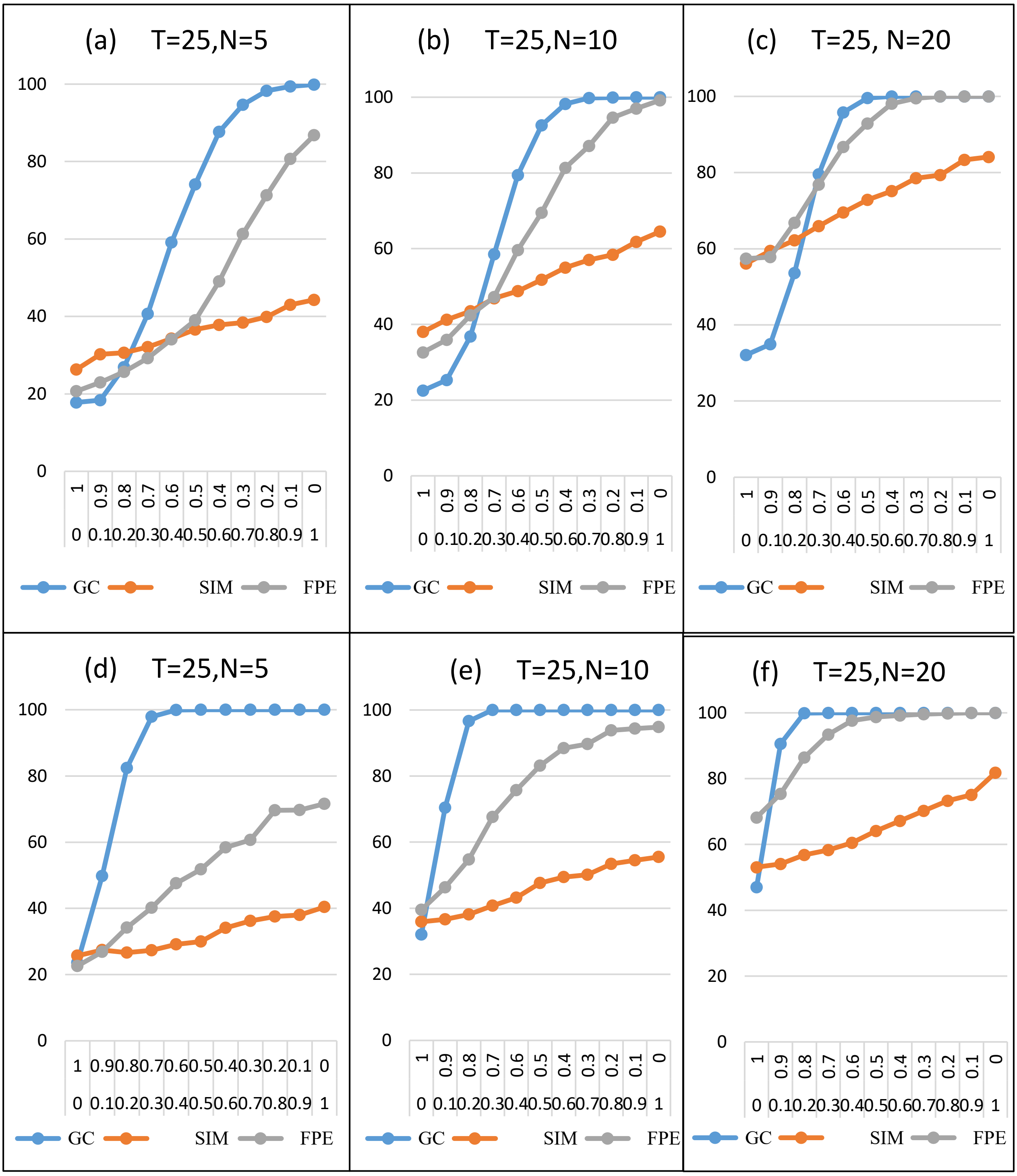



**Figure 5. 10:** Power Analysis of Panel Causality Tests for X→Y and X→Z with Drift and Trend, T=25

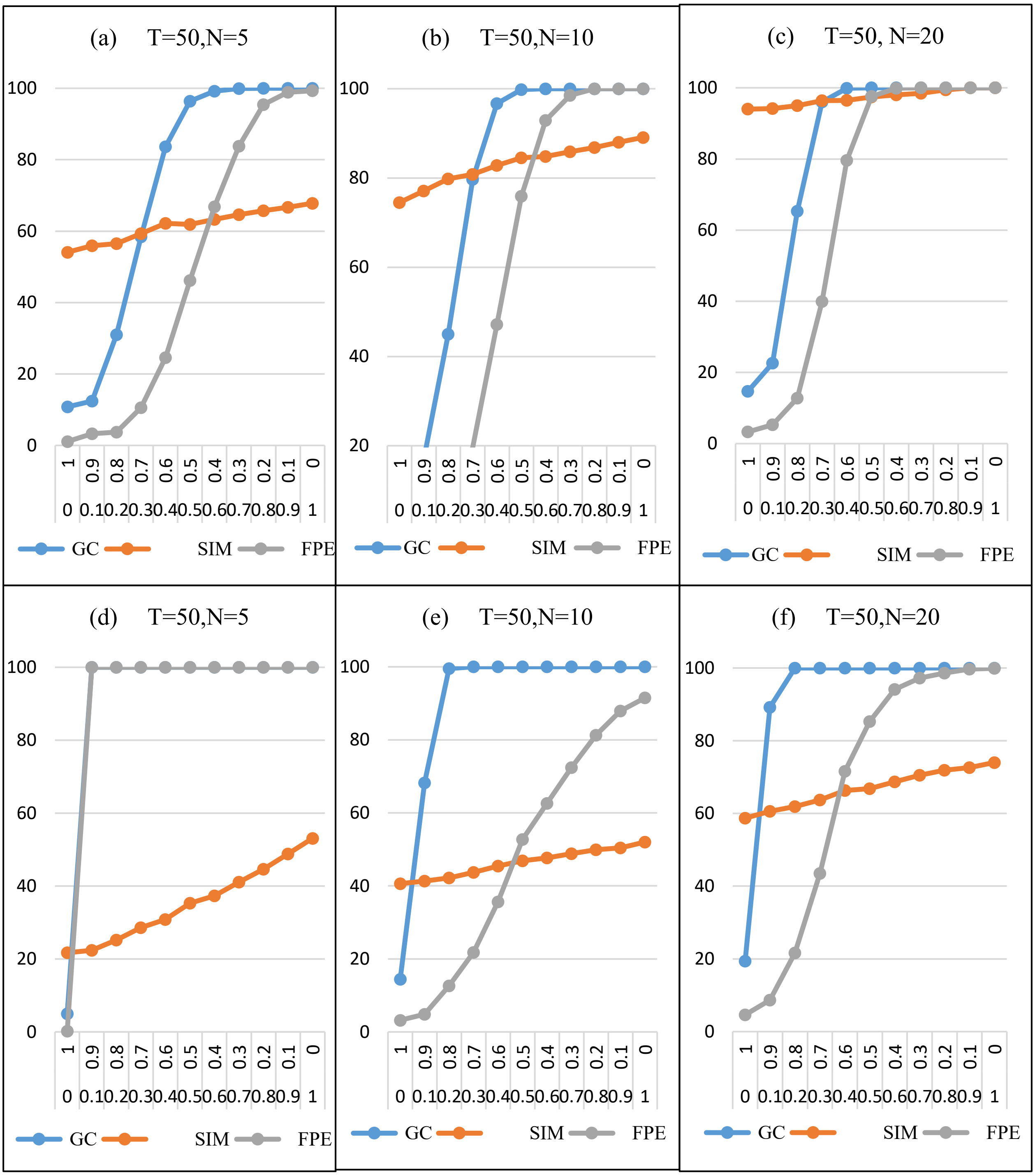



***Figure 5. 11:*** Power Analysis of Panel Causality Tests for X→Y and X→Z with Drift and Trend, T=50

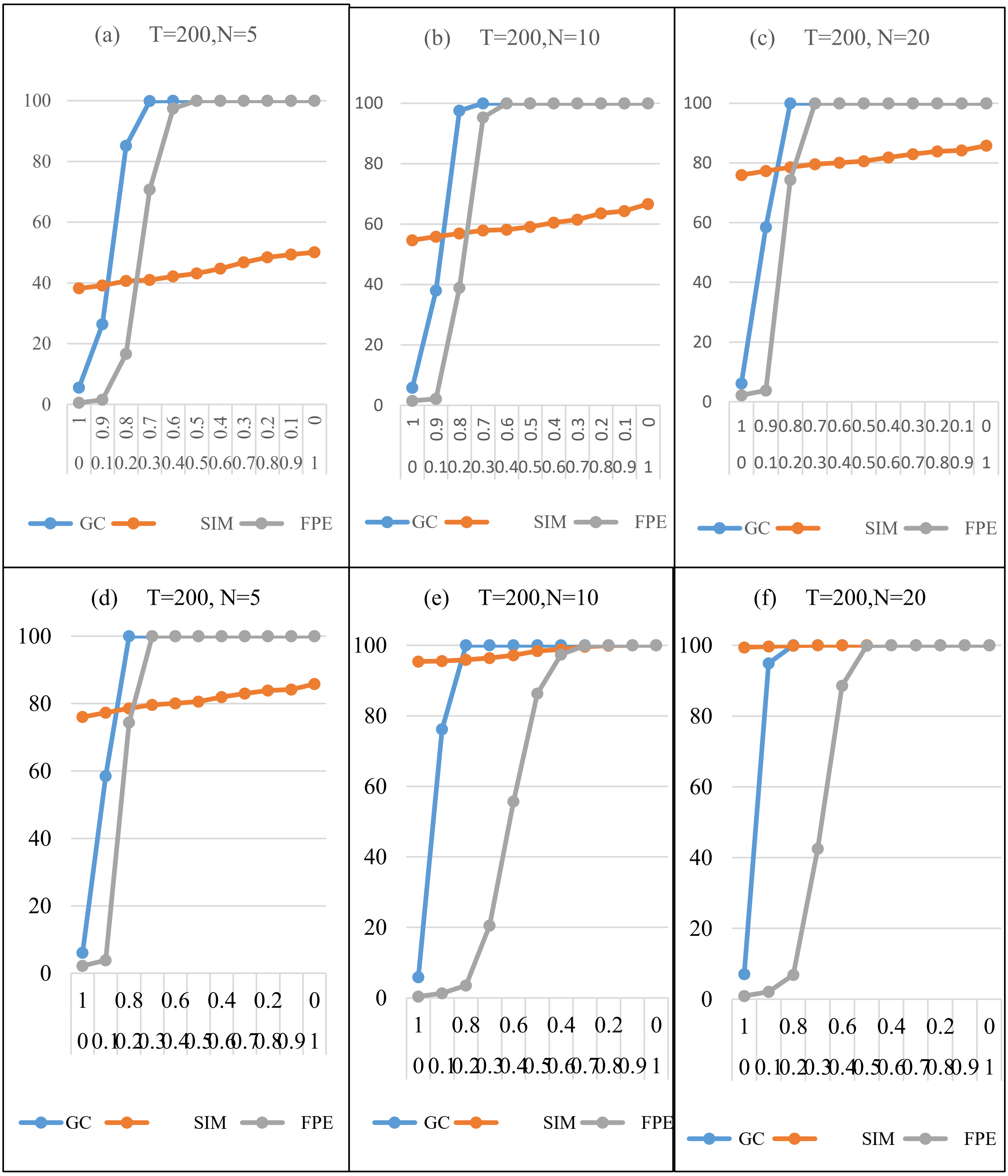



***Figure 5. 12:*** Power Analysis of Panel Causality Tests for X→Y and X→Z with Drift and Trend, T=200

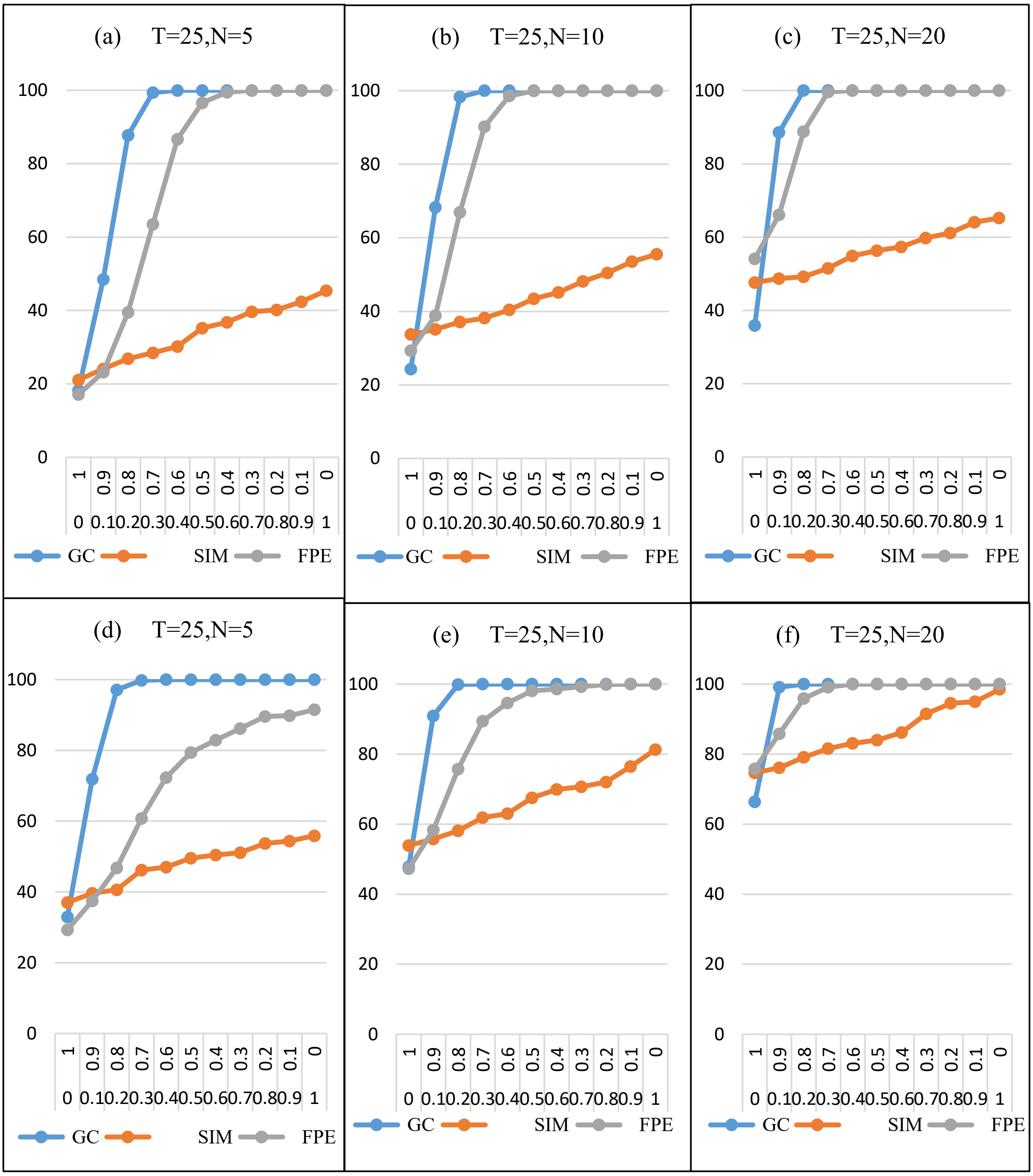



***Figure 5. 13***: Power Analysis of Panel Causality Tests for Y→X and Y→Z with Drift and Trend, T=25

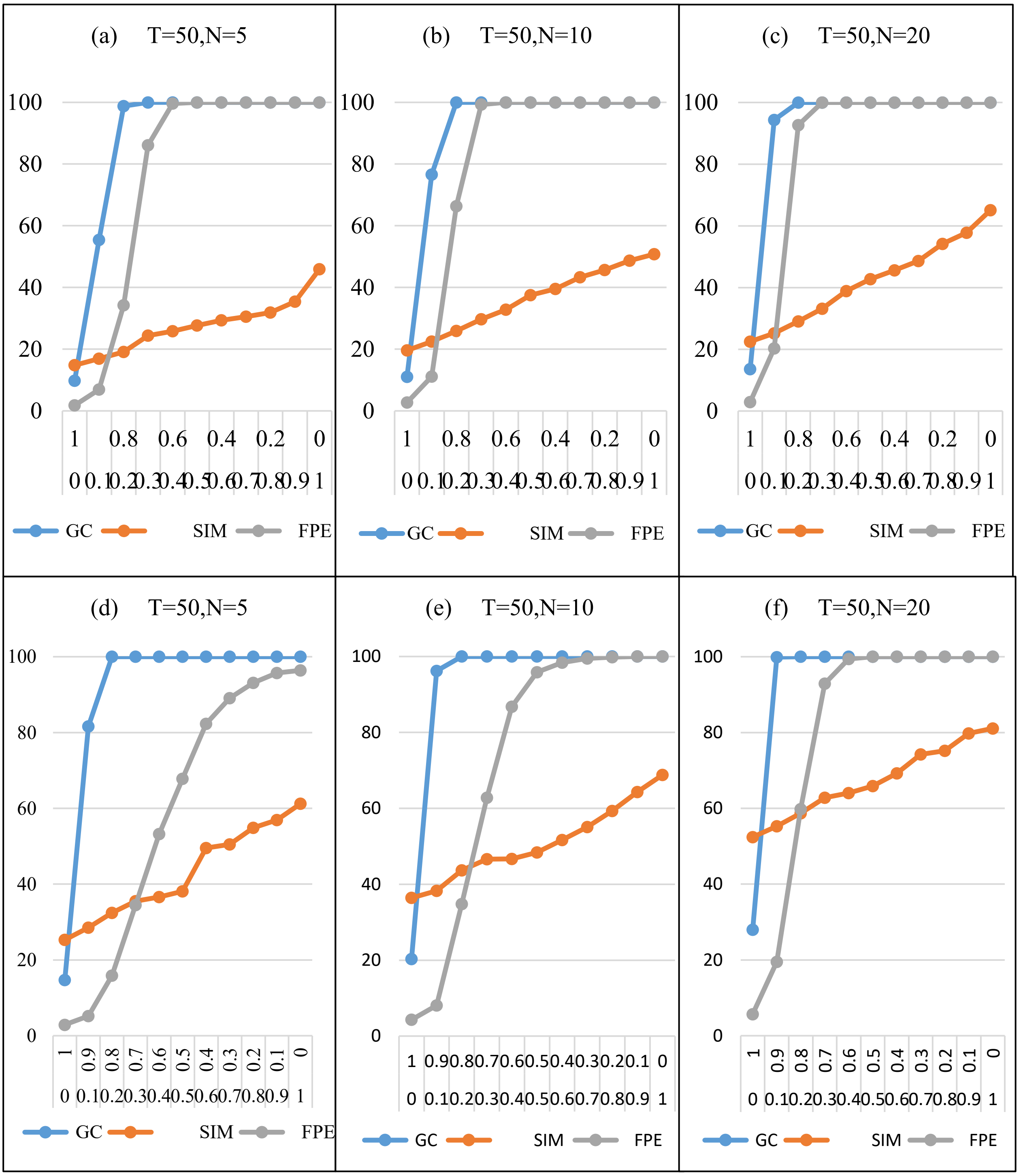



***Figure 5. 14***: Power Analysis of Panel Causality Tests for Y→X and Y→Z with Drift and Trend, T=50

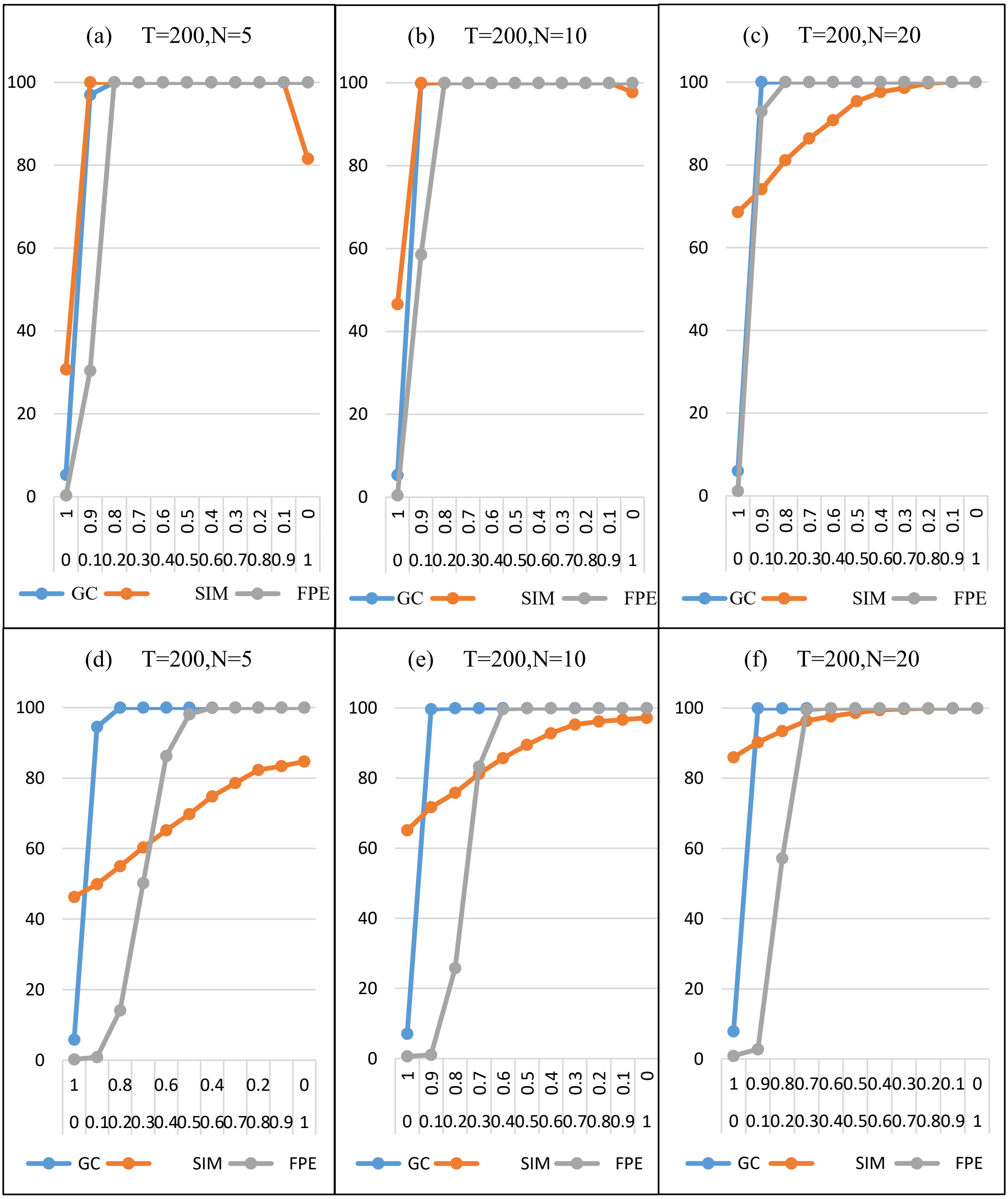



***Figure 5. 15***: Power Analysis of Panel Causality Tests for Y→X and Y→Z with Drift and Trend, T=200

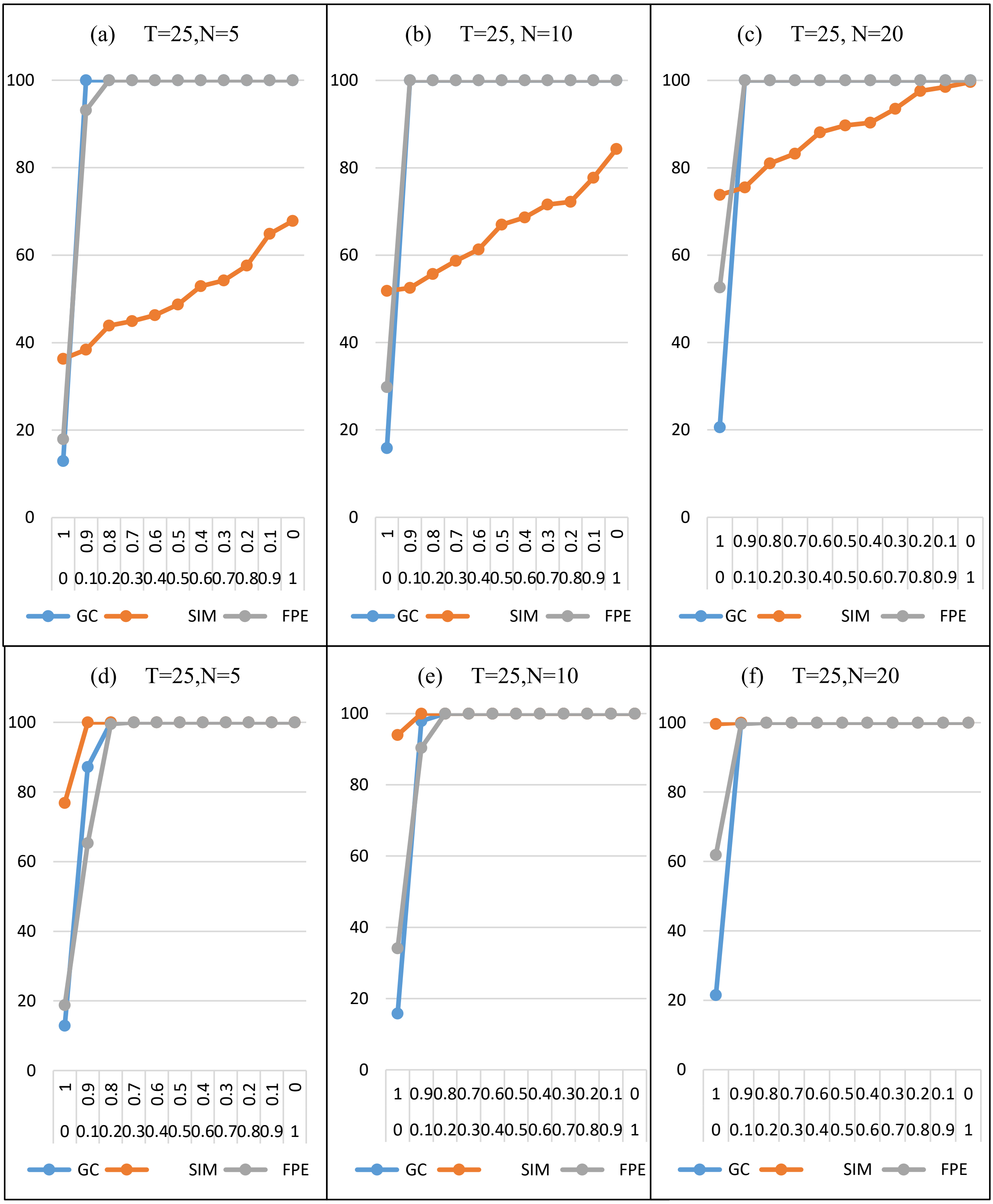



**Figure 5. 16**: Power Analysis of Panel Causality Tests for Z→X and Z→Y with Drift and Trend, T=25

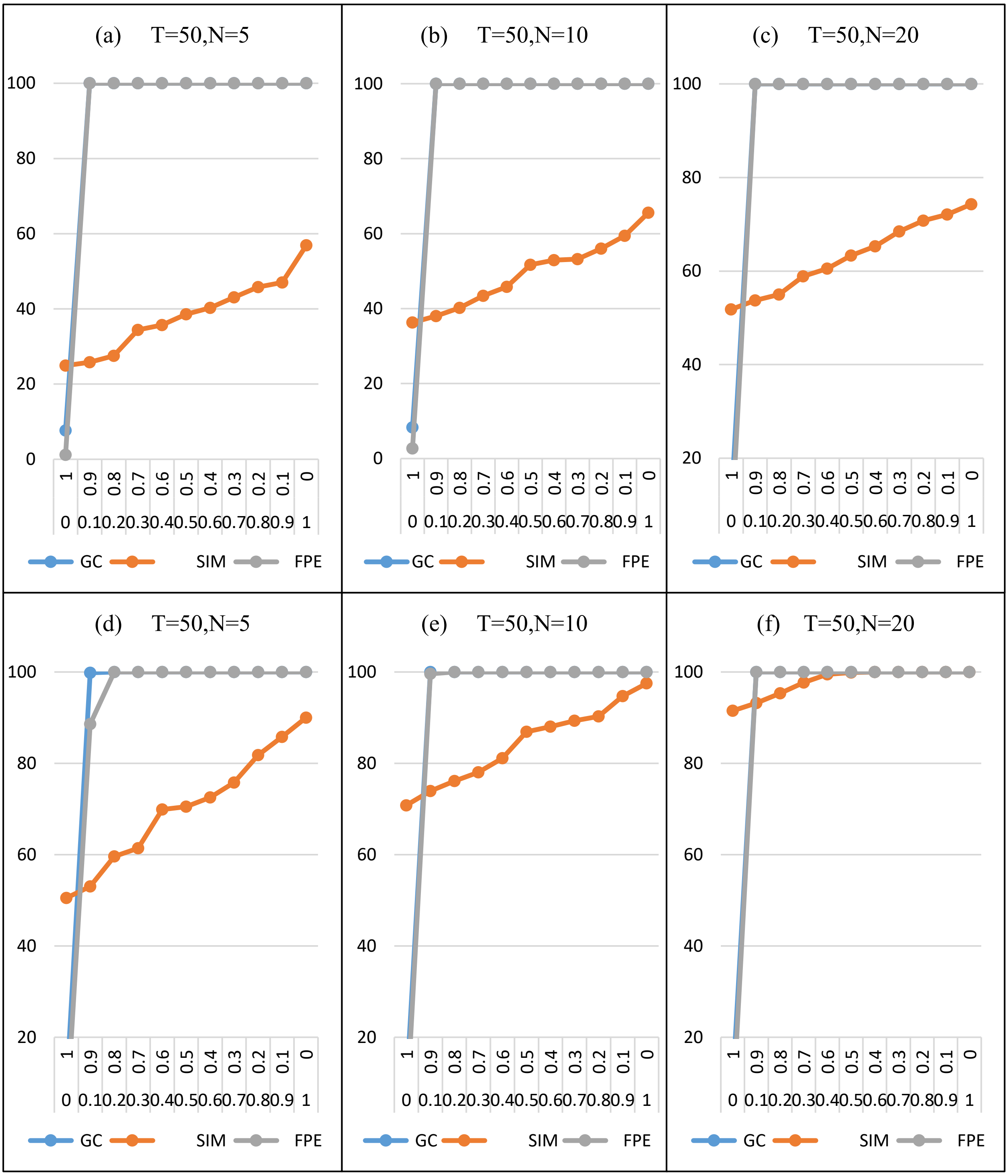



***Figure 5. 17***: Power Analysis of Panel Causality Tests for Z→X and Z→Y with Drift and Trend, T=50

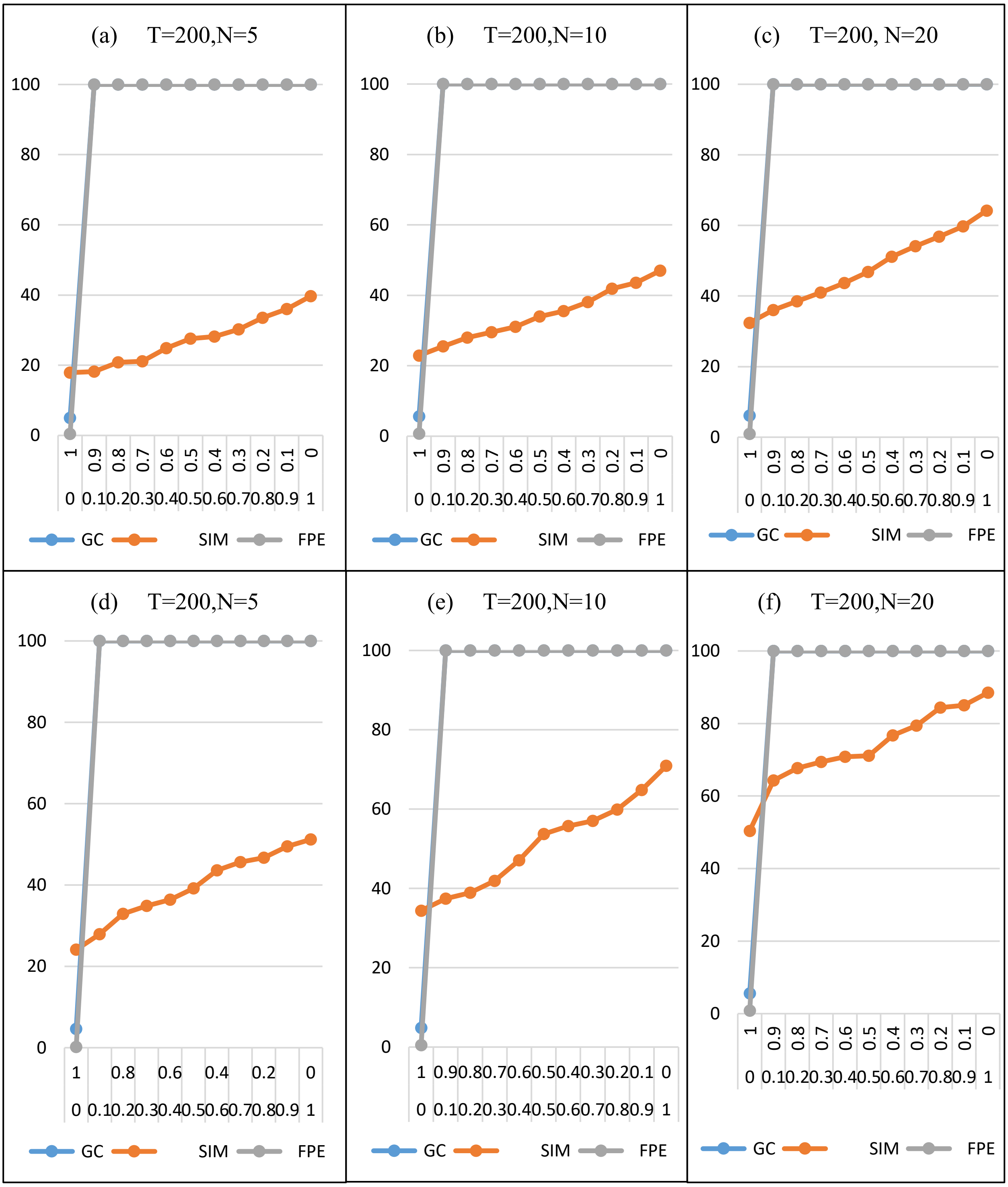



***Figure 5. 18***: Power Analysis of Panel Causality Tests for Z→X and Z→Y with Drift and Trend, T=200

## 5.2. Concluding Remarks

The study's objective is to evaluate the performance of all causality tests for panel data by investigating size and power properties. To achieve these objectives, the study mainly focuses on Monte Carlo Simulations, and the optimal procedure has been selected on its basis.

In this chapter, a comparison of Panel Causality Tests is made through size and power properties. The power of the Granger non-causality Test (GC) by Dumitrescu and Hurlin (2012), Sims (SIM) test (1972), and the Final Prediction Error (FPE) method by Hsiao (1981) causal search algorithm is analyzed. The power of any test is defined as the probability of rejecting a null hypothesis when it is false i.e.

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

We analyze the power of Panel causality tests for a variety of situations. We know that the power also depends on several nuisance parameters related to the "deterministic part" and the "stochastic part". Among the deterministic part are a component of drift and trend, while among stochastic parts, we have the autoregressive coefficient of the three series(x,y,z), which also determine the stationary of the series. This study used three different groups of the sample size, which were categorized into a small sample size, medium sample size, and large sample size for the data generating process under alternative hypotheses to calculate power.

Table 1 to Table 18 show the size, and power analysis of panel causality tests with the all six causal combinations i.e., X causes Y ($\theta_{12}$) and X causes Z ($\theta_{13}$) , Y causes X ($\theta_{21}$) and Y causes Z ($\theta_{32}$), Z causes X ($\theta_{31}$) and Z causes Y ($\theta_{32}$) in heterogeneous panel data in case of keeping both deterministic part with drift only and with drift and trend.

In compression of size, GC test has the least size distortion from the nominal size of 5% compared to size distortion of SIM and FPE tests at small, medium and large cross-section units. All three tests archive increasing power pattern as a parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions. However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium or large. This test archives 100% power at 0.3/0.2 alternative corresponding to N=5 and at 0.1 for large cross-section unit for all causal combinations and recognized the best performer compared to the other two tests. Among SIM and FPE tests, the former gains the least power at all alternatives compared to the latter, which corresponds to small, medium and large cross-section units and is thus identified as the worst performer. A similar pattern has been observed for almost all tests at different sample sizes; medium sample size (i.e., T=50) and large sample size (i.e., T=200). Based on the comparison of size and power analysis of the panel causality tests, this study concludes that the GC test is a point optimal. Overall research shows that the GC test performs better at all causal combinations and panel dimensions, whether drift only or both drift and trend take into account. On the other hand, the SIMS test with its lowest power gain at all causal combinations and panel dimensions is the worst performer test. However, the FPE test having a power curve between the better and worst performer test is graded as the average performer test. Theoretically, the GC test has more power than the AR (1) structure of DGP against all alternative hypotheses, which supports the GC test method. SIM's algorithm's lead values are not supported by the DGP, which might be one of the reasons for the algorithm's low performance.

# CHAPTER 6

# ANALYSIS OF THE CAUSALITY BETWEEN INTENSITY OF GOVERNMENT SPENDING AND THE INTENSITY OF HOUSEHOLD SPENDING ON EDUCATION WITH THE ROLE OF CREDIT CONSTRAINTS: PANEL DATA

This chapter is organized as follows; the first section describes the research topic's brief introduction, objectives, and significance. The existing literature on Government spending on education and household spending on education is reviewed in Section 2. Section 3 presents the data, variables and econometric methodology. Section 4 describes the results and discussions the empirical results. Finally, Section 5 offers a conclusion.

The simulation results reported in chapter 5 show that Granger non-causality test by Dumitrescu & Hurlin (2012) is the best performing test among its counterparts. Therefore, the study utilizes the said technique to analyze the causal relationship between the intensity of government spending and the intensity of household spending on education with the role of credit constraint. Alongside another econometric model for panel data is used. This research aims to find the causal relationship between government and household spending on education.

## 6.1. Introduction

Causality is the most important concept which is tested frequently in social sciences. Unfortunately, it is not easily detected from observational studies. In the natural sciences, causality can be determined through controlled experiments, whereas controlled experiments are difficult to be carried out in social sciences. Furthermore, experimental and observational studies have different statistical tools that can be explained using descriptive analyses. Therefore, one has to investigate the causal analysis for observational data. However, in observational data, the causal inferences are among the most difficult and have many issues.

The concept of Causality is essential but, unfortunately, cannot be detected easily. The exploration of data intensively is necessary for the detection of causality.

In some cases, the association between two variables exists due to the other confounding variables. Therefore, researchers were not able to identify which method should be preferred for testing causality in the presence of confounding variables. Freedman, Hall et al. (1995) states that “Indeed, causal inference requires many skills, intelligence and hard work. Natural variations need to be identified, and data must be collected. “Confounders need to be considered; Alternative explanations have to be tested”, Theory must support finding true causes than statistical analysis”.

Economic theory is mainly concerned with causal relationships among social and economic variables, for example, government expenditures on education, Household expenditure on education with the role of credit constraints (nonperforming bank loans), and other control variables are GDP per capita of growth, Population density (annual %), Unemployment (total % of the labour force), and Inflation of these entities with time.

Pakistan, a developing country, is struggling with economic growth and development. Through the mutual collaboration of federal and provincial institutions, the country spends a considerable amount every year on its education sector. In FY2020-21, the government allocated Rs.83 billion for the education sector. However, there is a need to spend more on education in Pakistan to ensure productivity and economic growth by developing human capital (PBS, 2020).

It is no doubt that in any developing country, only education can ensure the probability of better opportunities in the future because of three significant outcomes. The first role of education is to prepare highly trained, skilled, productive human resources that can actively strengthen economic growth and stability. Another part is to facilitate trained people to advance their knowledge through experimentation and learning in pure sciences and applied sciences. Moreover, the last role is to bring conscious understanding and awareness of the surrounding environment's social, physical, economic, and political functionality (Downes, 2001).

Contextualizing Barbados' economic growth and productivity, Downes observed in his research a solid and positive relationship between the public expenditure on education and human capital. He elaborates further by establishing the impact of education on the quality of labourers and their cognitive and critical skills after getting proper education and learning traditional skills in various fields of knowledge. In the specific context of Pakistan, it has been established through past research that only education can help minimize and alleviate poverty. Therefore, this research will narrow the need for more public expenditure on the education sector. The independent variables of the current study include: Household expenditure on education, Bank nonperforming loans, GDP per capita of growth, Population density (annual %), Unemployment (total % of the labor force), Inflation, and the dependent variable is Government expenditure on education.

The empirical analyses worldwide reflect the fast-track stimulation of economic growth after spending on the education sector. The most striking outcome of educated and trained human resources is that they begin to contribute to constructing sustainable economic accomplishments. Through skills and education, the chances of economic and social progress increased manifolds. Moreover, the pertinent need for efficient government expenditure on education can transform the economic landscape of the country (Kelly 1997).

In the current scenario, the significance of the allocated budget for education by the government each year is insufficient to meet up the challenges and demands of education for developing countries in modern times. Therefore, the size of public allocation of the education budget concerning the household expenditure on education, bank nonperforming loans, GDP per capita of growth, population density, inflation, and unemployment is an exciting study that will reflect the problematic factors of budget allocation.

Panel data offer both opportunities and challenges for causal inference. One key advantage of panel data over cross-sectional data is that it allows researchers to better handle the effects from unobserved time-invariant factors. At the same time, a fundamental problem in analyzing panel data is to account for possible serial correlations in the error terms for each individual. The underlying study aims to present models and methods for analyzing panel data, with particular attention to examining how the various models and methods handle causality issues.

The main idea of the research question is the following: Is the tendency of households to spend on education a consequence of the government's lack of education provision, or is it the case that the government itself responds to the households' lack of spending on education? What is the role of credit constraints?

In other words, is it households' spending on education that affects the government's spending on education, or is it the government's provision of education that affects household spending on education? It could also be that the level of credit provision affects household spending on education, affecting government spending.

The theoretical studies on the effect of government expenditures on education show that If households are credit constrained, their spending on education is low (they may not be able to borrow to spend on education). So the government invests in education, making up for households' lack of spending on education. Therefore, there might be a link between consumer credit provision and government spending on education via household spending. However, is this link statistically significant, or is it the case that households simply respond to government spending? In other words, if the government provides public education, then there is no need for households to spend on education. What is the correct causality?

We have found no functional, effective comparison of different causality methods in the literature. As it is known that other causality techniques are applicable in different scenarios. Therefore, it is necessary to determine which statistical technique/test gives us better statistical properties in a particular scenario. Hence, it is required to explore which conventional causality method is more appropriate to determine the true causal relationship between government expenditures on education and household expenditures on education with credit constraints.

This chapter contributes to the existing literature by examining the causal relationship between government spending on education and household spending on education with other confounding variables by using all conventional causality methods and tests for panel data.

### 6.2. Literature Review

In 1998, a comprehensive study conducted by Judson examined data of 138 countries to examine the impact of budget allocation on generating human capital resources and the relation of economic productivity with the spending on the education sector. The results revealed that the size and quality of budget allocation dominate the chances of economic growth compared to countries that were spending less on people's education sector and social welfare. Hence, education stimulates and sustains the different areas of economic growth and prosperity (Judson 1998).

The chances of educated labour income increase with time and experience, which automatically contributes to improving the financial institution of individuals in any country. For instance, in 1992, Fraumeni and Jorgenson conducted a study about how economic growth is linked with the level of education of labourers. Their findings confirmed that labourers who had attained better education and were skilled had better and higher lifetime incomes. The tremendous significance of investment in education can guarantee better human capital production and utilization (Jorgenson and Fraumeni 1992). In the like vein, Hutchinson and Schumacher (1997) observed that education is given the status of merit good due to the high impact factor. Government investment in an individual’s education can become a tool to secure future positive economic performance. The fiscal policy design of Latin American and Caribbean countries is evidently reflective of significant budget allocation for education expenditure.

One of the problems that many employees in the job market face are that the income is decided on the worker's merit or level of education. Individuals who had better exposure and were more educated had high-income rates, while considerable income inequality was observed among less qualified individuals. In 2002, Sylwester proposed a solution for income inequalities by increasing the total allocated budget for education. He suggested that individuals can overcome income inequalities through the availability of getting an education. This study was conducted in OECD and non-OECD countries from 1970 to 1990. The income disparity solution lay only in public education expenditure (Sylwester 2002).

In 2003, a study was conducted in Tanzania and Zambia about the relationship between education and poverty alleviation to find a path for economic growth. The study results ensured that when the government focused on education expenditure, the scope of labor expertise increased, which became a source of economic development of the country; through resourcing human capital with education, the potential of alleviating poverty increases (Jung and Thorbecke 2003).

Similarly, a study conducted in Jamaica and Guyana from 1970-to 2004 identified that after allocating the budget in the social sector transformation that is related to the education and skill-building capacity of labours, a positive trend of economic productivity had been observed in the social sector, which allows us to witness the immediate effect of education on individual’s chances towards prosperity and stability. (Conrad 2011)

To find out the relationship between public education expenditure and GDP/Capita, research was conducted in Pakistan in 2014 by Aqil et al., which revealed that education is a kinetic factor in building human resources for better economic performance in a robust manner. The overall relationship between education expenditure and GDP/capita is positive, therefore providing a vital opportunity for economists and researchers to witness the country's education sector (Aqil, Aziz, et al. 2014).

The core significance of education is that it allows individuals to attain knowledge of various fields. Moreover, knowledge facilitates technological development and advancement, and its utilization becomes a source of economic activity.  A study conducted between the European Union and BRICS in 2015 reaffirmed a positive correlation between education expenditure and economic growth. Hence, the significance of public education expenditure can never be undermined (Tomić 2015).

Another significant study conducted in Malaysia by Ramli et.al; identified that there is a direct relationship between investment on education sector and economic growth. The analysis of the study concluded that education is the driving force for generating skilled labor for market demand, by educating more people, the ability to secure market jobs around the globe increases, contributing to the country's economic growth. Therefore, the annual spending on education expenditure significantly impacts the country's prosperity. Another benefit is that all the investment in the skills of the labourers can provide them with long-term earning opportunities and benefits, which will directly strengthen the country's employment sector, making more people stable and independent. The demand for technology-equipped labor increases worldwide due to digitalization and globalization. Therefore, only through education is there the possibility to produce skilled labor that can compete in the comparative digitalized contemporary demands. For that purpose, the investment in education expenditure is justified (Ramli, Hashim, et al., 2016).

Frank (2018) analyzed through the long run growth accounting model that there is a positive effect on economic growth through spending on the education sector. He took data from 179 countries from 1970-to 2014. The study results once again affirmed the relationship between education expenditure on economic development and the country's well-being. However, this particular research highlighted significant room for economic development by prioritizing technological innovation and advancement. As technology is demanded in all international and national markets worldwide, it can become a continuous source of collecting revenue. Therefore, the relationship between education expenditure and economic growth is multidimensional. It helps alleviate poverty and build capacity for new employment and space for new investments in the country market.

The studies mentioned earlier refer to the fact that the education budget allocation can directly contribute to the country's long-term growth and economic development. Moreover, Pakistan is still struggling to achieve the status of a developed country. The only practical way to achieve this position is through spending on the education sector, contributing to long-term economic growth.

## 6.3. Covariates of Public Education Expenditure

In this section, there is a discussion about different covariates/determinants of public education expenditure. Much research has been conducted in the area. Still, none has been done to examine the immediate need to increase total education expenditure as a strategy to overcome the economic crisis. The current study will examine the various determinants of public education expenditure, keeping in mind the socio-economic and socio-political conditions of the country. It is evident through past research that education is the sole contributor to generating human capital. For this purpose, government allocation of the education budget is crucial in the long-run economic development plan. By drawing upon past studies, this section will discuss covariates of public education expenditure in Pakistan.

A study conducted in the United States revealed five factors that control per capita education expenditure. These factors include population size, demographic variables like age and structure of the population, physical infrastructure including size and form of buildings, the role of government, and the economic features/status of the country (Hirsch 1960).

However, a study conducted through cross-section analysis by McMahon in 1970 highlighted that the expenditure on public education is directly related to the demand, cost of production and tax behaviour of individuals. The study also highlighted several indicators that contribute to the education sector's overall expenditure. These include the number of pupils per teacher, school-age population of children, and substitutions of public schools. The study confirmed the population of school-age children as a significant factor. Another considerable point highlighted in the study is that the state of employment and the state aids are two factors that contribute to designing the allocated budget for education expenditure (McMahon 1970).

Another determinant is the state of the running economy of the country. If the government is going through an economic crisis, then the treatment of education expenditure is different, especially in Pakistan. Research from the world also ensured this factor that economies prefer to benefit from instantly from their investments compared to long-term investment plans. A research study conducted by Tilak (1989) reflected that the country's economic state directly participates in allocating the size of the budget. Therefore, the precarious economic state can contribute to less allocation of the total funding in the education sector compared to a stable economic state. The study further concluded that the most common vulnerable budget area is education in the recession.

Tilak (1989) gave various reasons for that purpose. Firstly, the investment in the education sector is seen as a long-term investment by the government. So, it is overlooked in developing countries where poverty rates are high, and employment rates are low. Secondly, during the critical inflation time, the allocation of the education sector seems impractical to most policymakers. Thirdly, the intangible benefits of education are not viewed as productive for the economic growth activity. The state of the country's economic activity is a significant contributor to determining the size of the allocated budget (Tilak 1989).

A study highlighted that in OECD countries since 1960, it has been surveyed that the influence of demographic factors and economic state fluctuations control the total budget allocation for education expenditure (Castles 1989). In a similar vein, a study conducted by Falch and Falch and Rattsø (1997) proposed that the role of a total number of older people out of the total population, public debt, unemployment and inflation contribute to fluctuating allocation of the education budget. He further explained that macroeconomic economy fluctuations intensify the chances of vulnerability for the education sector, especially in developing countries (Falch and Rattsø 1997).

On the other hand, while analyzing the impact of demographic factors on the political functionality of the spending budget in the education sector, Poterba commented that the elderly population, compared to the school-age children population, hinders the massive investment in the education sector. It is because the government has to place subsidies for older people because of their passive role in the development of economic growth. At the same time, the population of school-age children is a controllable factor in the budget (Poterba 1997).

Another critical finding from the study of Verbina and Chowdhury (2004) reflected that population density negatively influences the allocation of the education expenditure budget. Due to the constantly increasing population rate, the budget allocated for education expenditure becomes more vulnerable because of its relationship to poverty and unemployment. Additionally, the study also discussed the increase in per capita revenue of the country has a positive influence on the size of the allocating budget and its distribution (Verbina and Chowdhury 2004).

Two researchers in Switzerland confirmed almost the same findings. Their study revealed that demographic factors create an intensively competitive environment for allocating and distributing funds during the budget planning process. As a result, the financial constraints negatively delimit the budget's size for education spending. Furthermore, through empirical findings of the study, the results reflected that between young adults and older people, the financial expenditure and distribution become problematic (Grob and Wolter 2007).

The study conducted in 2007 through time series analysis revealed similar findings. The population sample was 21 OECD countries, and the data was taken from 1980-to 2001. Results, however, again confirmed that the level of economic development is directly linked with the demographic factors of the country. This implied that positive GDP growth is a product of the spending budget on labourers' skills and education (Busemeyer 2007).

Another research on African countries revealed that education had become one of the most successful investments to alleviate poverty and positive economic growth to exercise public welfare policies and strategies. The study also reflected that because the population was under 14 years of age. So, the investment in the education sector became more productive as it created skilled and educated labourers and professionals. Moreover, positive GDP growth has been observed after spending on the education sector (Akanbi and Schoeman 2010). Similarly, a study conducted in Thailand about the determinants of education expenditure revealed that the education expenditure budget was not dependent on demographic factors. The government of Thai made it compulsory to allocate a specific budget for education irrespective of the country's economic status. Moreover, education is not compromised because of inflation or recession. The priority was education over the demographical needs of the homeland (Sagarik 2013).

In a study conducted in India, it has been found that when the government gives more aid and help to educational institutions, the rate of economic growth increases. Moreover, the total revenue collected from taxation allowed the government to allocate a higher budget for education expenditure, which positively influences the overall GDP growth. However, it has also been found that due to the high population of school-age children in the community, the spending of education budget gets lower as the total financial cost of education expenditure becomes far less than the demand present (Chatterji, Mohan et al., 2014).

A study conducted in Cameroon, Chad, and the Central African Republic estimated a negative relationship between inflation and the probability of allocating a high budget for the education sector. The study explained that the chances of employment decrease simultaneously at the time of inflation, and the total budget allocated for different social sectors is greatly disturbed. Therefore, the efficiency of education expenditure to the economic instability is greatly hindered by uncertainty. Additionally, during the stable economic condition, the efficiency and the distribution of allocated education expenditure are enhanced. The relationship between the two variables mentioned above confirmed that the state's economic condition could directly influence the country's GDP (Fonchamnyo and Sama 2016).

Bischoff and Prasetyia (2015) analyzed the education spending determinants that involved 398 Indonesian panel data from the years 2005 to 2012 by using random effects and fixed effects models in their research study. Their results exposed that the public education expenditure increases due to the larger share of children population in any community. Moreover, the study further highlighted that the citizens were found unwilling to support high shares of public expenditure invested in the form of tax contributed by them because it created financial distress for them to pay for communal welfare.

Contradictory to the above studies, a study conducted by Kurban, Gallagher et al. (2015) confirmed a different perspective. The study's findings refuted the earlier studies' results confirming the elderly population as a source of a hindrance to the growth and prosperity of the per-pupil education expenditure of the total population of the children. By conducting a re-inspection on the United States data, it has been found that older adults do not control or create any competition on the resources allocated for the education of the children who school are going age (Kurban, Gallagher et al. 2015).

The study conducted by Imana (2017) found that some significant economic factors, such as real GDP per capita, budget deficits, and the education lagged expenditure, contributed negatively to controlling the size of the allocation budget. She further suggested that urbanization has become another factor in managing the education expenditure budget, and more specifically, primary education expenditure is directly suffered from it. Moreover, due to the domestic debt, low-income rates, and high inflation, the children's chances of secondary school education become scarce. She highlighted various reasons for it. Firstly, the role of government is crucial; more borrowing and loans led to high taxes and to meet the revenue demand, the prices became high, and the affordability factor of services area became challenging to attain. Secondly, due to inflation, the necessities are challenging to fulfill, and the long term plans of investing in the future become worthless (Imana 2017).

A critical assessment of the determinants of education expenditures within Malaysian society was carried out by Abdul Jabbar and Selvaratnam (2017)and Yun and Yusoff (2018). Both studies pointed out that the economic demographics and political factors played a vital role in determining the Malaysian public education expenditure from 1990 to 2015. The study results confirmed that revenue had been positively significant for education expenditure, while the results were negatively significant in the budget deficit. Additionally, the unemployment rate had an inverse and insignificant impact on the overall education expenditures of the country. The same studies reflected that the policymakers did not consider economic and political indicators as the decision on the education expenditure was given priority to achieve more important future goals. Both studies emphasized that the long term sustainable economic growth is dependent on the quality education expenditure, and the short-term variable like poverty, unemployment, inflation, and budget deficits were handled with other economic strategies while securing the allocated budget for the education sector.

Smith (1776) represented the idea of labour specialization that can contribute to the total world economy. He found out that the number of annual products of any nation depends on the amount of labor employed in the production and the productivity of the labor. He explained that the first element is of lesser importance than the second element because the quality of the productivity of the laborers will decide their payments. So, by creating more educated and skilled laborers, the amplitude of financial worth will increase and positively impact economic growth. The demand and the need of the time are crucial with much technological advancement in today’s world. However, the investment in the education sector is not up to the demand.

Adolph Wagner first witnessed the positive correlation between economic growth and government activity, and active growth. As well pointed out by Henrekson (1993), Wagner saw three significant reasons for assessing the state's role in maintaining the pace of economic activity. Firstly, due to rapid industrialization and modernization in contemporary times, it is estimated that there would be a definite increase in private sector development and economic activity. Therefore, the total expenditures allocated for the law-and-order situation and the contractual enforcement will increase. Secondly, the increase in real income due to taxation will enhance and expand the income elastic earmarked for the budget of cultural and social welfare projects and their expenses. Lastly, Wagner identified that the government is a better controller and executor of the expenditure only in the education and cultural sectors.

Sinha (1998) observed that the government should deal with natural monopolies such as railroads because private companies cannot plot these monopolies. For private companies, it is impossible to arrange huge finances for mere losses. It is an interesting phenomenon that the government can prefer to spend the budget on building failed and unproductive industrial significance projects for years. Still, it cannot see the possibility of spending on the education sector for the long-term benefits.

The selected research area is undoubtedly significant in the contemporary world. Still, the problems lie in understanding the practical worth of the relationship between education expenditure and economic prosperity in the long run. Several empirical studies are conducted to assess the relationship between human capital investment and the country's economic growth. Researchers such as Ghosh Dastidar, Mohan et al. (2013) elaborated the relationship is tested positive in their seminal work. In Tanzania and Zambia, a study was conducted by Jung and Thornback and in Nigeria by Oguibe and Adeniyi.  And in India, a survey conducted by Chandra revealed a positive relationship between education expenditure and the economic growth rate. Hence, all these papers concluded a positive relationship between the spending on education and the GDP growth rate.

Fiszbein and Psacharopoulos carried out a study to analyze the impact of educational investment in Venezuela. They identified that investments in primary education could largely boost economic growth. On the other hand, the total investments in the higher education category yield the lowest results among the three distinctive levels of the educational institution of the society.

In a study conducted by Oluwatobi and Ogunrinola (2011) for Nigeria, it has been noted that there is a positive relationship between the increased growth of expenditures in the education and the economic sector by using an expanded model of economic reproduction, in which they also included the impact of costs of education and health care in designing the size of the budget. Similarly, Sinha (1998) carried out research in Malaysia, which revealed a significant long-term relationship between the costs spent on education and economic growth; however, there is no mutual relationship between the increase in the total cost of education and the economic growth. This analysis is conflicting yet present in the research analysis.

Idrees and Siddiqi (2013) conducted research based on panel analysis. They identified a significant positive relationship present between the rising cost of education and economic growth due to different reasons. Firstly, the investments spent on the education sector ensure the availability and access to learning various skills and techniques that can generate income. Secondly, there is a vast competition present in contemporary times regarding the quality of education and the research-oriented studies that can effectively help other people and communities. Thirdly, the total amount spent on the education sector increases due to the digitalization and commercialization of the education field and the increased taxes related to the services industries. Therefore, the market worth of well-educated and well-equipped labour is increased.

One of the most typical examples given in this regard is often India. Chandra (2010) observed that the sole reason for India’s boom in the software industry in the twentieth century is the result of the enormous investments in the 1950s and 1960s in the technological and education sector. The 40-year-old investment is paying immensely in multiple employment spaces and the revenue through the taxation system. Moreover, it is also generating competition in the global job market. The fruitful investment in the long run in the education sector is evident. Therefore, the relationship of economic growth and education is vividly positive.

It is essential to understand that education in Pakistan is related to the provincial governments. According to the population growth rate, the National Finance Commission Award reveals that the provinces receive the considerate amount of funds from the federal divisible pool. Then these provinces prioritize their funding according to the needs and requirements of the services sector. As education is one of the country's sectors, its expenditure is decided according to the total budget. Through past research, it has been observed that Punjab allocated 30 percent of the funds for education. It has been observed further that KPK allocates the second big budget. On the other hand, it has been observed that Sindh and Baluchistan spend less budget on the education sector.

Since (2003), according to the districts' devolution plan, the districts have received education-development funds from the respective provincial governments. In addition to their resources and the presence of the allocated funds, the districts later distribute funds across various sectors, including the Education sector. The low priority accorded to the education sector is observed primarily in the development expenditures. It is one of the reasons that we keep many variations in the literacy levels among the various districts of the same province. Some researchers believe that there is a possibility that expenditures on education and literacy levels are interdependent with each other. For instance, the study conducted by Husain, Qasim et al. (2003) shows that large-scale disparities are present among Urban and rural Punjab at one side and on the other side between urban and rural Sindh on the other side in terms of literacy rates. The problem goes on because these specific districts that include Rajanpur, Muzaffargarh, Lodhran, D. G. Khan, etc., in Punjab and Mithi, Thatta, Badin, etc. in Sindh that are highly illiterate on record are also found in allocating fewer budgets for education expenditure purpose (Husain, Qasim et al. 2003).

According to the different studies, it has been identified that the public sector expenditure on education does not equally benefit all the discreet groups of the population. Other factors can influence the benefits of spending. These factors include income, age, gender and region. Husain, Qasim et al. (2003) noted that the expenditure on the state of art coronary care services is selectively beneficial for rich people because of the affordability factor, while financing in the unemployment insurance company will help empower the poor.

While it is observed that the competition between populations of different ages may vary the benefit of expenditures, it has been observed that a higher allocation of funds for pension is beneficial to only older people because they cannot earn at this point in time. On the other hand, funding a school meal is only significant to the young ones (Husain, Qasim, et al., 2003).

Another perspective raised in this regard is that public spending on education can be progressive or regressive. For instance, studies like Gupta, Hanges et al. (2002), and SPDC Report (2004) reflected that in the countries like Columbia, Ecuador, Malaysia, Philippines, and Pakistan, there is a progressive benefit linked with the expenditure on primary and secondary education, health care, public transport and infrastructure. According to Rasmus et al. (2001), this is possible due to the following reasons. Firstly, the spending on primary education, especially in rural areas and less developed sectors, will be progressive.

Also, it is observed that the access to publicly provided education services may vary because of race, gender, caste, region, and religious factors. Government expenditure will not reach such groups since the chances of service utilization rates in such cases are lower than in the privileged groups. For example, research by Al-Samarrai and Zaman (2007) in Malawi; Sabir, Ahmed et al. (2003); Shahin (2001) in Côte d'Ivoire and Selden and Wasylenko (1992) in the Peru States have been able to establish the fact that females of the school-age group are less privileged in terms of having fewer benefits as compare to men in getting an education.

Cama, Jorge et al. (2016) has observed the same results that describe that the females are less prioritized by their parents to accomplish their educational goals. The expenditure spent on females in Pakistan is far less than that spent on males.

In a cross-sectional study of 56 countries studied by Gupta, Hanges et al. (2002), it has been observed that the increase in total government expenditure on the education sector is directly associated with the improvement in the accessibility and attainment of education in the schools. It is significant to highlight that the government expenditure on education can be utilized in the long run to reduce poverty. Analyzing the fiscal policy in Thailand, a study conducted by Hyun in 2006 explained that the government subsidies the tax exemption o for the poor and the subsidies in the education services would help in getting education significantly it will benefit the less privileged and the marginalized communities.

The studies in the same vein conducted by researchers like Sahn and Younger (2000), Demery and Verghis (1994) and Younger, Warrington et al. (1999) have critically examined the education expenditure at different levels of education. The study's findings have revealed that primary education is the most progressive than secondary education, public universities, and private sector universities. (Hakro 2007) Different researches have explained the different determinants of public education expenditure worldwide. However, this literature review is limited to the selected variables used for this study purpose. Nevertheless, with the help of different studies, a theoretical ground has been established that explains the relationship of public education expenditure with inflation, unemployment, population density and GDP per capita growth.

### 6.4. Data and Econometric Methodology

The study's annual data from 2004 to 2018 for all countries are available data for variables. This panel data set is explored by *UNESCO*, the data bank of the World Bank *World development indicators,* and the data bank of *IMF's International Financial Statistics* online database. Note that $HEX_{it}$ is the Initial household funding of secondary education, (% of GDP) and $GEX_{it}$ is Government expenditure on education, total (of % of GDP). Measuring credit constraints is a bit tricky, especially for panel data. Some researchers use non-performing loans; others construct specific indexes. Here, $\mathrm{BNL}_{it}$ is bank nonperforming loans to total gross loans (%) as a proxy of Consumer credit constraints. The problem with these approaches (mentioned below) is that such data may not be available for a long time series and an adequate number of countries. Therefore, we examined a full sample with a mixture of countries. Further, this study also employs other four potential determinants of public expenditures on education Consumer price index (2010 = 100), Population density, GDP per capita (constant 2010 US$), Unemployment, total (% of the total labor force) (modelled ILO estimate). The summary statistics of all variables are reported in Table 6.1. The figures in the table below show the means and standard deviations of variables.

**Table 6. 1:** Descriptive Statistics

| | Mean | Median | Maximum | Minimum | Std.Dev. |
|---|---|---|---|---|---|
| $LnHEX_{it}$ | -1.182 | -0.966 | 1.5247 | -6.778 | 1.397 |
| $LnGEX_{it}$ | 1.443 | 1.498 | 2.147 | 0.412 | 0.349 |
| $LnNPL_{it}$ | 1.524 | 1.405 | 4.090 | -0.581 | 0.8163 |
| $LnCPI_{it}$ | 4.629 | 4.637 | 5.947 | 3.809 | 0.2412 |
| $LnPOP.Den_{it}$ | 4.437 | 4.717 | 7.322 | 0.963 | 1.277 |
| $UNP_{it}$ | 6.529 | 5.890 | 26.091 | 0.130 | 3.945 |
| Observations | 600 | 600 | 600 | 600 | 600 |

Note: Annual data for the period 2004-2018; whole sample, 40 countries.

### 6.4.1 Econometric Methodology

This paper investigates the causality between household spending and government spending on education. In doing so, we consider the ability of households to borrow as it may affect the causality both directly and indirectly. Approximately, the ability to borrow by credit risk as non-performing loans express it. Thus, the latter is incorporated as a mediator in estimating the causality. This paper has applied Panel Granger non-causality test proposed by Dumitrescu and Hurlin (2012), and the Panel corrected standard error (PCSEs) model to tackle the problem of heteroscedasticity Serial correlation of AR (1), and Cross-sectional dependence. For testing stationarity of the variables, the second-generation panel unit root test is Im-Pesaran and Shin (IPS) Test at level and difference. Pedroni's cointegration test and Panel ARDL model, i.e., Pooled Mean Group (PMG) estimation for heterogeneous panel data is used for cointegration analysis. This model (PMG) is used for long-run, short-run causality, and the error correction term (ECT). Dynamic panel data estimation two-step system GMM is also used to handle the problem of endogeneity. The endogeneity problem means that the dependent variable in each regression is correlated with the regression's error term. Therefore, least squares cannot be used to estimate the model. Alternatively, we have estimated our model with a two-step system GMM, using lagged values of the regressors as instruments. As a robustness test estimated a VAR (3) and computed the Impulse response function using Cholesky decomposition and a 95% confidence interval.

### 6.4.2 Panel Granger non-causality test

Panel Granger non-causality test was proposed by Dumitrescu and Hurlin (2012) is applied to investigate the causal relationships between Government expenditure on education and household expenditures on education with credit constraints. As a preliminary overview, present tests of causalities (whole sample, low-income subsample, high-income subsample) using the Dumitrescu and Hurlin (2012) test. Table 6.2 displays the p-values for both $\bar{Z}$ and $\tilde{Z}$.

**Table 6. 2:** Granger non-causality test – Dumitrescu & Hurlin (2012)

| $H_0$ | whole sample | | high income | | low income | |
|---|---|---|---|---|---|---|
| | p-values | | | | | |
| | $\bar{Z}$ | $\tilde{Z}$ | $\bar{Z}$ | $\tilde{Z}$ | $\bar{Z}$ | $\tilde{Z}$ |
| $HEX_{it} \nrightarrow GEX_{it}$ | 0.00*** | 0.15 | 0.00*** | 0.09* | 0.01*** | 0.17 |
| $NPL_{it} \nrightarrow GEX_{it}$ | 0.00*** | 0.03** | 0.00*** | 0.15 | 0.02*** | 0.21 |
| $GEX_{it} \nrightarrow HEX_{it}$ | 0.00*** | 0.02** | 0.00*** | 0.26 | 0.00*** | 0.00*** |
| $NPL_{it} \nrightarrow HEX_{it}$ | 0.33 | 0.92 | 0.00*** | 0.45 | 0.30 | 0.29 |
| $GEX_{it} \nrightarrow NPL_{it}$ | 0.05** | 0.80 | 0.00*** | 0.90 | 0.01*** | 0.96 |
| $HEX_{it} \nrightarrow NPL_{it}$ | 0.00*** | 0.14 | 0.00*** | 0.01** | 0.00*** | 0.00*** |

*** 1% significance level; ** 5% significance level; * 10% significance level.
Annual data for 2004-2018; whole sample, 40 countries; high income, 29 countries; low income, 11 countries.
The lags for each case were chosen using BIC; the minimum number of lags is found to be 1 and the maximum 3.

Table 6.2 displays results from the procedure proposed by Dumitrescu & Hurlin (2012) –referred to hereafter as the DH approach- an extension of Granger's bivariate framework to test stationarity in panel data. The null hypothesis is that *x* does not Granger cause *y*, and the table presents p-values for the standardized statistics $\bar{Z}$ and $\tilde{Z}$, for the whole sample, a subsample contains only the low-income countries and a subsample that comprises only the high-income countries.[4] The p-values for the $\bar{Z}$ Statistic suggests that there is bivariate causality between $HEX$, $GEX$ and $NPL$ that runs both ways in all samples apart from causality $NPL \rightarrow HEX$, which does not hold for the whole sample and the low-income sample.

The $\bar{Z}$ statistic, however, is more appropriate when both *T* and *N* are relatively large, and *T* is large relative to *N*. The $\tilde{Z}$ statistic, on the other hand, corresponds to the case where *N* is large relative to $T$ and $T > 5 + 3K$, where *K* denotes the number of lags for *x*. Since $T = 18$, $N = 40$ and the maximum *K* is found to be 3 using BIC, the $\tilde{Z}$ Statistics appear to be the most appropriate measure for our samples. The results based on the $\tilde{Z}$ statistic firstly indicates that for the low-income and whole sample, the causality only runs from $GEX$ to $HEX$ at the 1% and 5% significance levels, respectively. For the high-income sample, it runs only from $HEX$ to $GEX$ at the 10% significance level. This leads us to conclude that the weakening of the statistical significance of $GEX \rightarrow HEX$ in the whole sample relative to that of the low-income sample is due to the reversal of the causality in the high-income sample.

[4] The second-generation panel unit root test of Im, Pesaran and Shin (1997) indicates that $GEX_{it}$, $HEX_{it}$ and $NPL_{it}$ are all level stationary at 5% significance levels, without including a trend.

Secondly, the causalities are only direct as no secondary statistically significant causalities support indirect effects.[5] It is also worth highlighting two more aspects of those results. First, the fact that causality $NPL \rightarrow GEX$ is statistically supported in the whole sample but none of the other two subsamples causes doubts about the validity of this causality. Since $NPL \rightarrow GEX$ holds for the entire sample, one would naturally expect it to hold in at least one subsample. Second, while causality $HEX \rightarrow NPL$ is statistically insignificant in the whole sample, it appears to be highly statistically significant in the two subsamples. These results are consistent with one another, as long as the effects of $HEX$ on $NPL$ in the two subsamples go in opposite directions in a way that they cancel one another. Our estimates suggest that this is indeed the case.

The above Table (6.2) establishes the causalities based on past values of one of the variables. Next, we have quantified the causal relationships by controlling for the possible dependence of other variables that may act as mediators. Finally, in examining the causality between *HEX* and *GEX*, we consider the role of credit market risk, which is approximated by non-performing loans ($NPL$). In particular, we believe the following model:

$$\boldsymbol{A}\boldsymbol{Y}_t = \boldsymbol{A}_0 + \sum_{i=1}^{N} \boldsymbol{A}_j \boldsymbol{Y}_{t-j} + \boldsymbol{\varepsilon}_t \qquad (6.1)$$

Where $\boldsymbol{Y}_t$ includes the three variables of interest, $HEX_t$, $GEX_t$ and $NPL_t$ as well as other variables which may interact and feedback the three variables of interest. Specifically,

$$\boldsymbol{Y}_t \equiv [HEX_t \quad GEX_t \quad NPL_t \quad CPI_t \quad POP_t \quad UNP_t]',$$

All variables are stationary and expressed in logarithm, apart from the unemployment rate, and thus coefficients in the regressions refer to elasticities.

Matrix $\boldsymbol{A}$ is 6 x 6 with all diagonal elements equal to 1, and the error vector $\boldsymbol{\varepsilon}_t$ has zero mean and a diagonal variance-covariance matrix. To estimate the sign and magnitude of the causality between $HEX_t$ and $GEX_t$ requires not only the estimation of $\{\boldsymbol{A}_j\}_{j=1}^{N}$ but also $\boldsymbol{A}$. The fact that there are infinite combinations of different values

[5] Even though causalities $NPL \rightarrow GEX$ for the whole sample and $HEX \rightarrow NPL$ for the two subsamples are statistically significant, neither of them supports indirect effects of GEX on HEX and HEX on GEX respectively. In the first case, the effect is not triggered by GEX while in the second case, although the effect is triggered by HEX, NPL has no effect on GEX.

for $\{\boldsymbol{A}_j\}_{j=1}^{N}$ And $\boldsymbol{A}$ that implies the same probability distribution for the observed data leads to a well-known fundamental problem in estimation—in particular, estimating equation (6.1) as a structural VAR requires identifying restrictions on parameters.

The problem is that the theory is not informative, let alone conclusive, regarding possible restrictions that would be adequate to identify the first two rows of the equation (6.1). Therefore, since firm theoretical foundations are absent, a priori restrictions will be somewhat arbitrary and thus not well justified. An alternative approach is to find six exogenous variables which would be used as instruments to help with identification. For each equation in (6.1), the corresponding exogenous variable should be highly correlated with the dependent variable but uncorrelated with the contemporaneous, future, or past realizations of the structural shocks of the other equations.

Finding such distinct exogenous variables is a significant challenge. Sims (1980) hinted that even though many of these variables are treated as exogenous by default, there are no good reasons to believe they are strictly exogenous. Thus, the choice of exogenous variables that can be used as instruments is quite controversial. Within our framework, where structural errors are assumed to be *independent and identically distributed*, prominent instruments are lagged values of the regressors. As explained below, we adopt a version of the latter approach. Furthermore, estimating equation (1) as an SVAR restricts us from using a third dimension of the data: the country cross-sectional dimension. The latter is significant as time series data for several countries is relatively short, while many countries have data availability, including developing economies.

In this analysis, we adopt a rather heuristic approach to the problem by estimating directly and separately each equation included in equation (6.1) using cross-country data. Although we have several control variables, we also add fixed effects in the regressions to mitigate further bias concerns due to omitted variables. The latter enables us to control for the impact of time-invariant unobserved heterogeneity. The regressions are specified as follows.

$$y_{i,t} = \alpha^y + \boldsymbol{\beta}^y(L)\boldsymbol{Y}_{i,t} + \gamma^y f_i^y + \varepsilon_{i,t}^y, \qquad (6.2)$$

For $y = HEX, GEX, NPL, CPI, POPDEN, UNP$

where, $\boldsymbol{Y}_{i,t} \equiv [HEX_{i,t} \quad GEX_{i,t} \quad NPL_{i,t} \quad CPI_{i,t} \quad POP_{i,t} \quad UNP_{i,t}]'$ and $\boldsymbol{\beta}^{y}(L)$ is a 1 x 6 vector of polynomials in the lag operator, $L$, with elements $\beta_s^y(L) = \sum_{j=0}^{N} \beta_{s,j}^{y} L^j$ for the coefficients that do not correspond to the variable in $\boldsymbol{Y}_{i,t}$ which coincides with $y_{i,t}$, and $\beta_s^y(L) = \sum_{j=0}^{N-1} \beta_{s,j+1}^{y} L^{j+1}$ for the coefficient that corresponds to the variable in $\boldsymbol{Y}_{i,t}$ which coincides with $y_{i,t}$, for s = 1, 2, 3, …, 6.

For instance, if $y = HEX$, then $\beta_1^{HEX}(L) = \sum_{j=0}^{N-1} \beta_{1,j+1}^{HEX} L^{j+1}$ while for s = 2, 3, 4, 5, 6. $\beta_s^{GEX}(L) = \sum_{j=0}^{N} \beta_{s,j}^{GEX} L^j$. Likewise, if $y = GEX$, then $\beta_2^{GEX}(L) = \sum_{j=0}^{N-1} \beta_{2,j+1}^{GEX} L^{j+1}$ while $\beta_s^{GEX}(L) = \sum_{j=0}^{N} \beta_{s,j}^{GEX} L^j$ for s = 1, 3, 4, 5, 6. Note that whether there is a direct or indirect effect depends on the statistical significance of the estimated coefficients. Noticeably, there is an endogeneity problem, which means that the dependent variable in each regression is correlated with the regression's error term. Therefore, least squares cannot be used to estimate the model. Alternatively, we can estimate those equations with GMM, using lagged values of the regressors as instruments. Then, after obtaining estimates for the three equations and assuming $N = 3$ (according to the BIC criterion), the dynamic components of those equations can be written as a VAR (3) process:

$$\hat{B}_0 \mathrm{Y}_t = \sum_{j=1}^{3} \hat{B}_j \, \mathrm{Y}_{t-j} + \epsilon_t \qquad (6.3)$$

Where $\mathrm{Y}_t = [HEX_t \quad GEX_t \quad NPL_t \quad CPI_t \quad POP_t \quad UNP_t]'$ is country-invariant, $\hat{B}_0, \hat{B}_1, \ldots, \hat{B}_3$ are matrices that include the estimated parameters of the previous regressions:

$$\hat{B}_0 = \begin{bmatrix} 1 & \hat{\beta}_{2,0}^{HEX} & \hat{\beta}_{3,0}^{HEX} & \hat{\beta}_{4,0}^{HEX} & \hat{\beta}_{5,0}^{HEX} & \hat{\beta}_{6,0}^{HEX} \\ \hat{\beta}_{1,0}^{GEX} & 1 & \hat{\beta}_{3,0}^{GEX} & \hat{\beta}_{4,0}^{GEX} & \hat{\beta}_{5,0}^{GEX} & \hat{\beta}_{6,0}^{GEX} \\ \hat{\beta}_{1,0}^{NPL} & \hat{\beta}_{2,0}^{NPL} & 1 & \hat{\beta}_{4,0}^{NPL} & \hat{\beta}_{5,0}^{NPL} & \hat{\beta}_{6,0}^{NPL} \\ \hat{\beta}_{1,0}^{CPI} & \hat{\beta}_{2,0}^{CPI} & \hat{\beta}_{3,0}^{CPI} & 1 & \hat{\beta}_{5,0}^{CPI} & \hat{\beta}_{6,0}^{CPI} \\ \hat{\beta}_{1,0}^{POP} & \hat{\beta}_{2,0}^{POP} & \hat{\beta}_{3,0}^{POP} & \hat{\beta}_{4,0}^{POP} & 1 & \hat{\beta}_{6,0}^{POP} \\ \hat{\beta}_{1,0}^{UNP} & \hat{\beta}_{2,0}^{UNP} & \hat{\beta}_{3,0}^{UNP} & \hat{\beta}_{4,0}^{UNP} & \hat{\beta}_{5,0}^{UNP} & 1 \end{bmatrix},$$

$$\hat{B}_j = \begin{bmatrix} \hat{\beta}_{1,j}^{HEX} & \hat{\beta}_{2,j}^{HEX} & \hat{\beta}_{3,j}^{HEX} & \hat{\beta}_{4,j}^{HEX} & \hat{\beta}_{5,j}^{HEX} & \hat{\beta}_{6,j}^{HEX} \\ \hat{\beta}_{1,j}^{GEX} & \hat{\beta}_{2,j}^{GEX} & \hat{\beta}_{3,j}^{GEX} & \hat{\beta}_{4,j}^{GEX} & \hat{\beta}_{5,j}^{GEX} & \hat{\beta}_{6,j}^{GEX} \\ \hat{\beta}_{1,j}^{NPL} & \hat{\beta}_{2,j}^{NPL} & \hat{\beta}_{3,j}^{NPL} & \hat{\beta}_{4,j}^{NPL} & \hat{\beta}_{5,j}^{NPL} & \hat{\beta}_{6,j}^{NPL} \\ \hat{\beta}_{1,j}^{CPI} & \hat{\beta}_{2,j}^{CPI} & \hat{\beta}_{3,j}^{CPI} & \hat{\beta}_{4,j}^{CPI} & \hat{\beta}_{5,j}^{CPI} & \hat{\beta}_{6,j}^{CPI} \\ \hat{\beta}_{1,j}^{POP} & \hat{\beta}_{2,j}^{POP} & \hat{\beta}_{3,j}^{POP} & \hat{\beta}_{4,j}^{POP} & \hat{\beta}_{5,j}^{POP} & \hat{\beta}_{6,j}^{POP} \\ \hat{\beta}_{1,j}^{UNP} & \hat{\beta}_{2,j}^{UNP} & \hat{\beta}_{3,j}^{UNP} & \hat{\beta}_{4,j}^{UNP} & \hat{\beta}_{5,j}^{UNP} & \hat{\beta}_{6,j}^{UNP} \end{bmatrix}, \text{ for } j = 1, 2, \dots, N.$$

As long as the inverse of the matrix $\hat{B}_0$ exists, the reduced form of (6.3) is given by

$$\mathrm{Y}_t = \textstyle\sum_{j=1}^{3} \hat{\Gamma}_j \, \mathrm{Y}_{t-j} + u_t, \qquad (6.4)$$

Where $\hat{\Gamma}_j = \hat{B}_0^{-1}\hat{B}_j$ and $u_t = \hat{B}_0^{-1}\epsilon_t$. (4) Can be further written as

$$X_t = \Pi X_{t-1} + \lambda_t \qquad (6.5)$$

Where $X_t = [\mathrm{Y}_t \quad \mathrm{Y}_{t-1} \quad \mathrm{Y}_{t-2}]'$, $\lambda_t = [u_t \quad 0_6^1 \quad 0_6^1]'$ and

$$\Pi = \begin{bmatrix} \hat{\Gamma}_1 & \hat{\Gamma}_2 & \hat{\Gamma}_3 \\ I_6 & 0_6 & 0_6 \\ 0_6 & I_6 & 0_6 \end{bmatrix},$$

Where $0_6^1$ 6 x 1 vector of zeros is, $0_6$ is a 6 x 6 matrix of zeros and $I_6$ is a 6 x 6 identity matrix. As long as all eigenvalues of $\Pi$ are strictly less than one in modulus, (6.5) can be written as a moving average process;

$$X_t = [I_{18} - \Pi \mathrm{L}]^{-1}\lambda_t = \lambda_t + \Pi\lambda_{t-1} + \Pi^2\lambda_{t-2} + \cdots + \Pi^j\lambda_{t-j} + \cdots \quad (6.6)$$

It follows that the percentage dynamic response of HEX to a 1% innovation in GEX is given by the first element of $\Pi^j\lambda$, where $\lambda = [\hat{B}_0^{-1}\epsilon^2 \quad 0_6^1 \quad 0_6^1]'$ and $\epsilon^2 = [0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0]'$. In other words, we can plot the first element of $\Pi^j\lambda$ as a function of j. Likewise, the percentage dynamic response of GEX to a 1% innovation in HEX is given by the second element of $\Pi^j\lambda$, where $\lambda = [\hat{B}_0^{-1}\epsilon^1 \quad 0_6^1 \quad 0_6^1]'$ and $\epsilon^1 = [1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]'$. These dynamic responses will quantify the causality $GEX \to HEX$ and the causality $HEX \to GEX$, respectively. To quantify the indirect effect of GEX on HEX via NPL, we can set $\hat{\beta}_{2,j}^{NPL} = 0$ for j = 0, 1, 2, 3 and recompute the response and then compare them with the previous, likewise, for the indirect effect of HEX and GEX via NPL.

## 6.5. Empirical Results and Discussions

Table 6.3 represents the estimators of Pooled OLS, Fixed Effect Model, and Random Effect Model with their standard errors in parentheses. According to our panel data set, the Hausman test to check which model (Fixed effect or Random effect) is appropriate. The p-value of the Hausman test is 0.0000, statistically significant, which concludes that the fixed effect model is suitable for this panel data set. Therefore, again regressed the fixed effect model. After it, checked the diagnostic test for Cross-sectional dependence is the Pesaran, Schuermann et al. (2004). This diagnostic test examined whether the residuals are correlated across entities. Here, the p-value of 0.000 rejects the null hypothesis and concludes that cross-sectional dependence exists across the Panel members. Therefore, the Panel Corrected Standard error model has been used to tackle the problems of heteroscedasticity, serial correlation of AR (1), and cross-sectional dependence.

**Table 6. 3:** Pooled, Fixed Effect, Random Effect, and Panel Corrected Standard Error (PCSE) Model

| | Coefficient-values | | | | | |
|---|---|---|---|---|---|---|
| Models | Pooled | Fixed Effect | Random Effect | PCSE | [95% Conf. Interval] PCSE | |
| | $LnGEX_{it}$ | $LnGEX_{it}$ | $LnGEX_{it}$ | $LnGEX_{it}$ | | |
| $LnHEX_{it}$ | -0.188** | -0.203 | -0.243 | -0.027* | -0.058 | 0.003 |
| $LnNPL_{it}$ | 0.043*** | -0.012** | -0.008 | -0.052*** | -0.089 | -0.014 |
| $LnCPI_{it}$ | -0.078 | -0.339** | -0.145 | 0.063 | -0.074 | 0.201 |
| $LnPOP.Den_{it}$ | 0.001*** | 0.005*** | 0.002*** | 0.012 | -0.013 | 0.038 |
| $UNP_{it}$ | -0.038** | -0.009 | -0.019* | 0.000 | -0.009 | 0.009 |
| Constant | 0.897* | -0.447* | 1.395 | 1.133*** | 0.485 | 1.781 |
| R Squared | 0.275 | 0.174 | 0.215 | 0.491 | Rho | 0.8661 |

*** 1% significance level; ** 5% significance level; * 10% significance level.

Annual data for 2004-2018; whole sample, 40 countries. The number of observations is 600.

### 6.5.1 Panel Unit Root Test

The second-generation panel unit root test was used to test the stationarity of the variables, which is the Im-Pesaran and Shin Test (1997) at level and difference. This test (IPS) concludes that$LnGEX_{it}$, $LnHEX_{it}$, $\mathrm{LnBNL_{it}}$ and $LnCPI_{it}$ are stationary at level and difference as well. While other control variables $LnPOP.Den_{it}$, and $UNP_{it}$ are non-stationary at level then become stationary at the first difference of the variables.

### 6.5.2 Panel Cointegration Test

Pedroni's panel cointegration test is displayed in table 6.4. The results of this test contain seven statistics, i.e., panel and group statistics. Most of them are statistically significant to accomplish the presence of cointegrated relationships among $LnGEX_{it}$,$LnHEX_{it}$, $LnBNL_{it}$, $LnCPI_{it}$, $LnPOP.Den_{it}$, and $UNP_{it}$.

### 6.5.3 Pooled Mean Group (PMG) Estimation for heterogeneous panel data.

This model (PMG) is used for long-run and short-run coefficients, implying long-run and short-run causality. These coefficients and the error correction term both show strong causality among variables. The PMG model assumed that long-run coefficients are the same across all countries in the panel. Here, the most important thing is the long-run coefficients of the PMG model presented in table 4. This shows long-run coefficients are statistically significant at a 1% level to indicate long-run causal relationships exist among variables $LnGEX_{it}$, $LnHEX_{it}$, $\mathrm{LnBNL}_{\mathrm{it}}$, $LnCPI_{it}$, $LnPOP.Den_{it}$and $UNP_{it}$. The PMG also shows short-run coefficients at the difference of independent variables and the error correction term (ECT).

The PMG model also assumed that short-run coefficients and ECT are not the same for each country in the panel. ECT is negative -0.6770 and statistically significant at a 1 % level of significance, which shows a cointegration relationship among panel variables and indicates that any deviation from long-run equilibrium is corrected at 67% speed of adjustment. ECT gives a joint causal effect among the variables.

### 6.5.4 Dynamic Panel Data Estimation, Two-step System GMM

This study applies a two-step system GMM model to investigate the true causal relationship between the intensity of household spending on education (HEX) and government spending on education (GEX) with the role of credit constraints. This model requires the number of instruments should be less than the number of groups (countries), and the overall validity of the instruments is determined by the values of

AR (1), AR (2), Sargan and Hansen Test. These statistic values should be higher to conclude that can't reject the null hypothesis, here null hypothesis is desirable i.e., Instruments as a group are exogenous.

**Table 6. 4:** Pedroni's Cointegration Test and Pooled Mean Group (PMG) Estimation

| | | | | |
|---|---|---|---|---|
| Pedroni's Cointegration Test | | | | |
| | Panel v-statistic | -3.676*** | Group rho-statistic | 7.539** |
| | Panel rho-statistic | 5.238** | Group t-statistic | -17.55*** |
| | Panel t-statistic | -13.06*** | Group ADF-statistic | 3.973*** |
| | Panel ADF-statistic | 2.951 | | |
| Pooled Mean Group Estimation (PMG) | | | | |
| Dependent variable | $dlnGEX_{it}$ | | | |
| Long-run | Coefficients | Short-run | Coefficients | |
| $LnHEX_{it}$ | -0.079*** | ECT | -0.677*** | |
| $LnNPL_{it}$ | 0.001 | $dlnHEX_{it}$ | -0.061 | |
| $LnCPI_{it}$ | -0.149*** | $dlnNPL_{it}$ | -0.0323 | |
| $LnPOP.Den_{it}$ | 0.524*** | $dlnCPI_{it}$ | -0.973** | |
| $UNP_{it}$ | -0.017*** | $dlnPOP.L$ | -0.633 | |
| | | $dUNP_{it}$ | -0.016 | |

Note: *, **, *** indicate statistical significance at the 10%、5%, 1% level, respectively.

The Two-step System GMM model shows higher AR (1), AR (2), Sargan and Hansen Test, which indicates the overall validity of the instruments. And the number of instruments is also less than the number of groups that lead to being a good model. All the diagnostic tests are satisfied the criterion at the 5 % level of significance. Table 6.5 displays the estimated coefficients of the six-panel regressions. In summary, table 6.5 confirms the main result of the DH bivariate test for the whole and low-income sample under $\tilde{Z}$, That is, only *GEX* causes *HEX*, while the causality is only direct. The latter is inferred that we do not find statistically significant mediators that would support an indirect relationship among our set of variables.

**Table 6. 5:** Dynamic Panel Data Estimation: Two-step System GMM

| | dependent variables | | | | | |
|---|---|---|---|---|---|---|
| regressors | $\widetilde{HEX}_{i,t}$ | $\widetilde{GEX}_{i,t}$ | $\widetilde{NPL}_{i,t}$ | $\widetilde{CPI}_{i,t}$ | $\widetilde{POP}_{i,t}$ | $\widetilde{UNP}_{i,t}$ |
| $HEX_{i,t}$ | | -0.223 | -0.911 | 0.576 | 0.057$^{*}$ | 9.319 |
| | | (0.49) | (1.72) | (0.68) | (0.03) | (7.35) |
| $GEX_{i,t}$ | 3.192$^{**}$ | | 1.941 | 0.351 | 0.188 | -0.004 |
| | (1.63) | | (1.95) | (0.93) | (0.16) | (4.04) |
| $NPL_{i,t}$ | 0.121 | -0.223 | | 0.062 | 0.082 | 21.80 |
| | (0.657) | (0.17) | | (0.24) | (0.51) | (19.06) |
| $CPI_{i,t}$ | 3.435 | 1.485 | -7.868$^{*}$ | | -0.056$^{**}$ | -1.474 |
| | (4.54) | (2.13) | (4.55) | | (0.02) | (2.02) |
| $POP_{i,t}$ | -1.333 | 0.040 | 0.606 | 33.38$^{**}$ | | 8.568 |
| | (9.87) | (2.71) | (2.99) | (17.19) | | (25.02) |
| $UNP_{i,t}$ | 0.073 | 0.015 | .2103$^{*}$ | 0.003 | 0.004 | |
| | (0.18) | (0.05) | (0.12) | (0.09) | (0.01) | |
| $HEX_{i,t-1}$ | 0.733$^{***}$ | 0.158 | -1.313 | -0.284 | -0.045$^{*}$ | -6.128 |
| | (0.13) | (0.33) | (1.10) | (0.63) | (0.02) | (2.81) |
| $GEX_{i,t-1}$ | -1.98$^{**}$ | 0.664$^{***}$ | 0.428 | -0.385 | -0.095 | 0.281 |
| | (1.09) | (0.13) | (1.091) | (0.52) | (0.10) | (2.38) |
| $NPL_{i,t-1}$ | -0.163 | 0.176 | 1.078$^{***}$ | -0.062 | 0.051$^{***}$ | 1.436 |
| | (0.56) | (0.16) | (0.26) | (0.24) | (0.02) | (26.07) |
| $CPI_{i,t-1}$ | -4.485 | -2.434 | 10.266 | 1.809$^{***}$ | -0.052 | -28.659 |
| | (6.35) | (3.21) | (6.31) | (0.52) | (0.71) | (3.45) |
| $POP_{i,t-1}$ | 1.391 | -23.18 | 19.043 | -69.16$^{**}$ | 1.639$^{***}$ | -3.876 |
| | (0.89) | (2.23) | (1.89) | (34.88) | (0.64) | (0.20) |
| $UNP_{i,t-1}$ | 0.009 | 0.002 | -0.144 | 0.042 | 0.004 | 0.402$^{**}$ |
| | (.06) | (0.024) | (0.10) | (0.04) | (0.01) | (2.12) |
| $HEX_{i,t-2}$ | 0.099 | -0.017 | -0.139 | 0.021 | -0.014 | -2.195 |
| | (0.16) | (0.12) | (0.42) | (0.20) | (0.01) | (0.82) |
| $GEX_{i,t-2}$ | -0.252 | 0.008 | -0.047 | -0.134 | -0.032 | 0.487 |
| | (0.25) | (0.10) | (0.44) | (0.12) | (0.02) | (1.08) |
| $NPL_{i,t-2}$ | 0.066 | -0.005 | -0.169 | 0.013 | 0.007 | -0.601 |
| | (0.14) | (0.06) | (0.31) | (0.13) | (0.01) | (5.18) |
| $CPI_{i,t-2}$ | 0.963 | 1.194 | -1.691 | -0.685 | 0.027 | 3.302 |
| | (1.93) | (0.97) | (2.26) | (0.89) | (0.18) | (79.55) |
| $POP_{i,t-2}$ | 4.617 | -1.191 | -2.936 | 42.254$^{*}$ | -0.392 | -26.794 |
| | (27.7) | (7.21) | (8.75) | (23.98) | (1.28) | (0.11) |
| $UNP_{i,t-2}$ | 0.036 | -0.008 | 0.058 | 0.005 | 0.002 | 0.038 |
| | (0.05) | (0.01) | (0.05) | (0.04) | (0.01) | (1.17) |
| $HEX_{i,t-3}$ | 0.147 | 0.069 | -0.424 | -0.077 | 0.003 | -0.915 |
| | (0.11) | (0.11) | (0.42) | (0.25) | (0.01) | (0.92) |
| $GEX_{i,t-3}$ | -0.767 | 0.265$^{**}$ | 0.405 | -0.019 | -0.058 | -0.120 |
| | (0.51) | (0.11) | (0.46) | (0.17) | (0.03) | (0.36) |
| $NPL_{i,t-3}$ | -0.000 | -0.001 | 0.055 | -0.001 | -0.013 | 0.277 |
| | (0.13) | (0.04) | (0.13) | (0.13) | (0.01) | (0.34) |
| $CPI_{i,t-3}$ | 0.341 | -0.259 | -1.047$^{*}$ | -0.640$^{**}$ | -0.204$^{***}$ | 5.533 |
| | (0.63) | (0.53) | (0.66) | (0.32) | (0.07) | (4.51) |
| $POP_{i,t-3}$ | -3.283 | 1.156 | 2.344 | -6.477 | -0.246 | 18.148 |
| | (17.98) | (4.58) | (6.64) | (8.66) | (0.65) | (54.78) |
| $UNP_{i,t-3}$ | 0.000 | 0.017 | -0.046 | 0.002 | 0.000 | -0.220 |
| | (0.03) | (0.01) | (.05) | (0.02) | (0.003) | (0.14) |
| intercept | -1.267 | -1.26$^{**}$ | -1.92 | .067$^{***}$ | -2.602$^{***}$ | .067$^{**}$ |
| | (1.22) | (0.66) | (1.71) | (0.57) | (0.19) | (4.50) |

$^{***}$ 1% significance level; $^{**}$ 5% significance level; $^{*}$ 10% significance level.
Numbers in parenthesis correspond to the standard deviations of the estimates.
According to the BIC criterion, we choose 3 lags for each regression.

Moreover, our estimates suggest that an increase of *GEX* by 1% affects *HEX* both in the current year as well as in the subsequent year, inducing an overall increase in *a roughly equal percentage of HEX*. Specifically, an increase in the intensity of government education spending by 1% increases *HEX* on impact by about 3% and decreases it the following year by about 2%.

Our findings indicate that an increase in the intensity of government spending on education encourages households to increase the intensity of their spending on education in the same year. They do so, but the percentage increase is three times higher than the increase in the intensity of government spending. For instance, an increase in government investment in infrastructure, say via an investment in school premises and computer labs, induces households to increase their spending significantly more by enrolling in private classes and purchasing new equipment. The overreaction of households to the increase in the intensity of government spending on education is followed by a "*correction*" the year after. The "*correction*" in relative expenditure that occurs in the subsequent year brings the overall percentage increase of the intensity of household spending to the same level as the initial percentage increase of the intensity of government spending. Among others, we do not find any evidence that credit constraints, proxied by the share of non-performing loans, affect directly or indirectly the intensity of spending on education.

To confirm the validity of the instruments, we perform further tests displayed in table 6.6. Given that table 6.5 displays no evidence for mediator variables that channel indirect relationships between GEX and HEX, we report results only for the first two main regressions. First, we check for serial correlation in the residuals of the system GMM by employing the Arellano-Bond test. The results for the main regressions are displayed in the first two columns. While the first order is statistically significant for both regressions with p-values of 5.2% and 6.5%, respectively, the second-order is clearly statistically insignificant as the p-values are as high as 82% and 66%, respectively. First-order serial correlation is not surprising since residuals in first differences correlate by construction.

On the other hand, the absence of second-order serial correlation implies that residuals are uncorrelated in levels, suggesting that the instruments are strictly exogenous. As shown in table 6.6, statistics from the Sargan and Hansen tests of over-identifying restrictions are in line with the Arellano and Bond test results. Specifically, the Sargan-Hansen result suggests that the instruments are jointly uncorrelated with the error term as the null hypothesis of over-identifying restrictions cannot be rejected.

**Table 6. 6:** Two-step System GMM: test results

| | dependent variables | | | |
|---|---|---|---|---|
| | $\overline{HEX}_{i,t}$ | $\overline{GEX}_{i,t}$ | $HEX_{i,t}$ | $GEX_{i,t}$ |
| regressors | $GEX_{i,j}$ | $HEX_{i,j}$ | $\overline{GEX}_j$ | $\overline{HEX}_j$ |
| $j = t, t-1, t-2, t-3$ | $t^{**}, t-1^{**}$ | | | |
| CD Pesaran test | 0.105 | 0.143 | 0.341 | 0.464 |
| Arellano-Bond, $AR(1)$ | 0.052 | 0.065 | 0.895 | 0.500 |
| Arellano-Bond, $AR(2)$ | 0.824 | 0.658 | 0.870 | 0.812 |
| Sargan test | 0.992 | 0.273 | 0.110 | 0.413 |
| Hansen test | 0.940 | 0.243 | 0.206 | 0.384 |

The number of instruments is 33; $\overline{GEX}_t \equiv \left(\sum_{j=1}^{N} GEX_{j,t}\right)/N$ and; $\overline{HEX}_t \equiv \left(\sum_{j=1}^{N} HEX_{j,t}\right)/N$; $^{**}$ 5% significance level.

To examine whether there are differences in the causality across -low and high-income countries, similarly to what the bivariate tests suggest, we extend (6.3) by introducing income-level dummies that enable us to capture possible differentiated effects. Nonetheless, the dummies are statistically insignificant, which indicates that, on average, income levels are irrelevant to the causality between the intensities of government and household spending on education.[6] This finding refutes the differentiated responses across the two subsamples implied by DH bivariate tests. The discrepancies between the multivariate and bivariate models results could be attributed to the aforementioned missing aspects of the bivariate test. Finally, to examine whether $GEX_{i,t}$ and $HEX_{i,t}$ respond to country invariant components of $HEX$ and $GEX$, respectively, we replace regressors at time $t$ with $\overline{HEX}_t$ and $\overline{GEX}_t$ and in all corresponding lags, and then re-estimate the models. We find that all regressors involving the country invariant factors are statistically insignificant both contemporaneously and in lags. This result indicates that household spending intensities on education respond only to country-specific changes in corresponding government intensities. We do not report the complete set of available estimates upon request to save on space. However, we report the various tests for the two regressions in the last two columns of table 4 to demonstrate that the models are well-specified.

[6] To save on space, we do not report the estimates with dummies which are available upon request.

## 6.6. Conclusion

The central research question of this study is to evaluate the tendency of household educational spending vis-à-vis government spending on education, given the household's credit constraints. This analysis used the annual panel data from 2004 to 2018 for all countries with available data. All variables are expressed in logarithm, apart from the unemployment rate, and thus, coefficients in the regressions refer to elasticities. The study took into account the ability of households to borrow as it may affect the causality both directly and indirectly—the ability to borrow by credit risk as expressed by non-performing loans approximately. Thus, the latter is incorporated as a mediator in estimating the causality.

As a preliminary overview, the Panel Granger non-causality test by Dumitrescu and Hurlin (2012) is applied to investigate the above causal relationships. The best performing test, the Granger non-causality test by Dumitrescu & Hurlin (2012), has been used to examine the causal relationship between the intensity of government spending and the intensity of household spending on education with the role of credit constraint, based on simulation results obtained in Chapter 5. In addition, panel data econometric models are applied. After that, the Panel corrected standard error (PCSEs) model is applied to tackle heteroscedasticity, Serial correlation of AR (1), and Cross-sectional dependence. For testing stationarity of the variables used second-generation panel unit root test that is the Im-Pesaran and Shin (IPS) Test at level and difference conclude that $LnGEX_{it}$, $LnHEX_{it}$, $\mathrm{LnBNL_{it}}$ and $LnCPI_{it}$ are stationary at level and difference as well. While other control variables $LnPOP.Den_{it}$, and $UNP_{it}$ are non-stationary at level then become stationary at the first difference of the variables.

In this paper, we examine the causality between the intensities of government and household spending on education. Using data from a cross-country panel, we show that appropriate bivariate causality tests suggest that the intensity of government spending on education causes the intensity of household spending on education, while the reverse does not hold. Although we find a relatively weak reversal of the causality for high-income countries, we demonstrate that this reversal, and thus the differentiated responses among low- and high-income countries, disappears when we consider a multivariate model that also controls for contemporaneous relationships, cross-country dependence, homoskedasticity and autocorrelation within countries as well as country fixed effects. The result is fascinating and useful for policymakers. It further shows that the causality clearly runs from the intensity of government spending on education to the corresponding household intensity, but the effect is only direct.

Our findings suggest that households tend to overreact when the government increases its spending intensity by increasing its intensity three times more. In the following year, however, they *correct* their response by decreasing their spending intensity so that there is an overall one-to-one relationship between government and household spending intensities on education. Interestingly, when we approximate credit market tightness with the percentage of non-performing loans, we find no evidence that the latter affects the intensity of household spending on education or government spending on education.

# CHAPTER 7

# CONCLUSION AND FUTURE RESEARCH DIRECTION

## 7.1. Summary

Causality is the most important concept which is tested frequently in social sciences. Unfortunately, it is not easily detected from observational studies[7]. In the natural sciences, causality can be determined through controlled experiments, whereas controlled experiments are difficult to be carried out in social sciences. Experimental and observational studies have different statistical tools, which can be explained with various descriptive analyses. Therefore, one has to investigate the causal analysis for observational data. But in observational data, the causal inferences are among the most challenging inferences and have several issues. The first and the most critical point is that causality is not directly observable in the non-experimental data. Second, one cannot control primary confounding factors in observational data. Third, statistical measures of relationship are symmetric and don't directly in the form of causality. Hence, it's difficult for the researcher to differentiate between cause and effect.

The main objective of this study is to analyze the effective functional comparison of all the causality methods regarding the type of data, i.e., cross-section, time series, and panel data. Furthermore, this study will show which methodology detects true causal relation by employing the appropriate method for specific data types.

We have found no effective functional comparison of different causality methods in the literature as it is known that different causality techniques are applicable in different scenarios. Therefore, it is necessary to determine which statistical technique/test gives better statistical properties in a particular scenario. Econometric analyses like cointegration, error-correction model and Granger or Sims causality tests are applied to examine this issue. These existing regression methodologies rely on normality and linearity assumptions not supported by the data

[7] Observational study draws inferences from a sample to a population where the independent variable is not under the control of the researcher because of ethical concerns or logistical constraints.

used in the analysis and may lead to inconclusive results. These techniques sometimes may not be able to provide proper/true results. Therefore, it becomes impossible to formulate policies based on the studies' findings following conventional methods. Hence, exploring which traditional method of causality is more appropriate to find out true causal relationships is required. Therefore, the present study has analyzed the performance of the panel causality test.

Further, it investigated which causality test is the best performer based on the size and power properties. This study modified the time series causality test for panel data, i.e., the Sims test and Hsiao's final prediction error method. Monte Carlo simulations are used to calculate these methods' size and power characteristics, and an optimum causal approach is proposed. The objective of the simulation experiment is to find out the size and power properties of methodologies for testing causality. Therefore, we need data with embedded causality (for power) and the data series with no causality (size).

The selection of DGP for simulation study is crucial in most comparative analyses. Different tests have a diverse theoretical base, so selecting one DGP can benefit some test statistics. Causality methods and tests can be compared in a homogeneous framework to conclude the superiority of a single test or show the strength and weaknesses of the tests under some alternative. In this regard, a simulation study should be needed; however, we proceed with our research by selecting a feasible DGP to compare all the tests. Under the Null hypothesis of a causal relation between X and Y ($H_0$: X causes Y) and under alternative hypotheses of independence between X and Y ($H_1$: X doesn't cause Y). As per the definition of Granger causality, Y is caused by X if the lag value of X can be used for predicting Y. The generated series are independent of each other, causality cannot be checked, and so we will only determine the size of the test. Furthermore, we will develop different but correlated series $x_{it}, y_{it}$ To check the causal ordering through the power of the test.

Power of the test means rejecting a hypothesis when the alternative hypothesis is true.

$$Power = \text{Prob}\ (rejecting\ H0\ /\ H1\ is\ true)$$

When a true hypothesis is rejected, this is an error probability denoted by "α" and establishes the size of the test.

$$\alpha = \text{Prob}\ (rejecting\ H0\ /\ H0\ is\ true)$$

Selection of sample size is critical in analyzing any data, especially for the panel data set. We have annual data sets and selected three cross-section units for this study. We have categorized them into three groups for this simulation study, i.e., 5, 10, and 20, as small cross-section units, medium cross-section units, and large cross-section units, respectively. Similarly, three-time series lengths are taken to evaluate the performance of Causality tests; these time series levels are 25, 50, and 200. Similar to categorizing cross-section length into three types, a time series length of 25 indicates a small-time series, and 50 is assigned as a medium time-series length. At the same time, 200 is allotted as a large time-series length in this study. To carry out simulations and MCSS of 10,000 is taken to get the convergence effectively.

The study also includes a real-world application. The best causality test algorithms are used for real-world data analysis to discover the causal drivers of government expenditure on education intensity. The present investigation also uncovers answers that provide researchers with a high confidence level. The result is fascinating and useful for policymakers. It further shows that the causality clearly runs from the intensity of government spending on education to the corresponding household intensity, but the effect is only direct.

## 7.2. Conclusion

In compression of size, GC test has the least size distortion from the nominal size of 5% compared to size distortion of SIM and FPE tests at small, medium, and large cross-section units. All three tests archive increasing power pattern as the parameter of interest ($\rho_1$ and $\rho_2$) moves away from the null hypothesis corresponding to all cross-section dimensions.

However, the power attainment of the GC test is much better than the other two tests (SIM and FPE) at all alternatives, whether the cross-sectional length is small, medium, or large. This test archives 100% power at 0.3/0.2 alternative corresponding to N=5 and at 0.1 for a large cross-section unit for all causal combinations and recognizes the best performer compared to the other two tests. Among SIM and FPE tests, the former gains least power at all alternatives compared to the latter, which corresponds to small, medium and large cross-section units and is thus identified as the worst performer. A similar pattern has been observed for almost all tests at different sample sizes; medium sample size (i.e. T=50) and large sample size (i.e. T=200).

Based on the comparison of size and power analysis of the panel causality tests, this study concludes that the GC test is a point optimal. Overall analysis shows that GC test performs better at all causal combinations and panel dimensions, whether drift only or both drift and trend take into account. On the other hand, the Sim's test, with its lowest power gain at all causal combinations and panel dimensions, is the worst performer test. However, the FPE test with a power curve between the better and worst performer is graded as the average performer test. The GC test's theoretical reasoning has better power because of the AR (1) structure of DGP against all alternative hypotheses, which supports the algorithm of GC test. On the other hand, the DGP does not keep the lead values for Sim's algorithm, which might be one of the potential reasons for Sim's poor performance.

### 7.3. Limitations and the Direction for Future Research

This comparative simulation research studied the performance of panel causality tests using a stationary scenario with varied DGPs of causal combinations for three-panel series. Model specifications have been employed with drift only and with drift & trend for different sample sizes, small, medium, and large. The current study used DGP as AR (1) process only; the autoregressive structure can be converted to a distributed lag structure so that one can use distributed lag structure in the future.

The optimal point tests under the stringency criteria must be developed for future research studies. For example, this study employed three-panel causality tests because of a shortage of time and resources in the future to investigate at least six to eight other panel causality tests and evaluate the size and power performance of these tests. Then, create the optimal point test and determine the most stringent test.

This research observed less convergence and the size distortion for small-time series and cross-section dimensions, especially for SIM and FPE tests. Therefore, a panel causality test (PCT) with adequate maximum power at small time-series and cross-section length must be developed to close the gap.

# REFERENCES


Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies, 72*(1), 1-19.

Abdul Jabbar, N., & Selvaratnam, D. P. (2017). Analysis of the determinants of education expenditures in Malaysia. *Journal of International Business, Economics and Entrepreneurship (JIBE), 2*(1), 1-10.

Adämmer, P. (2019). lpirfs: An R package to estimate impulse response functions by local projections. *The R Journal (2019), 11*(2), 421-438.

Ahmed, M., Riaz, K., Khan, A. M., & Bibi, S. (2015). Energy consumption–economic growth nexus for Pakistan: Taming the untamed. *Renewable and Sustainable Energy Reviews, 52*, 890-896.

Akanbi, O. A., & Schoeman, N. J. (2010). The determinants of public expenditure and expenditure on education in particular in a selection of African countries: economics. *South African Journal of Economic and Management Sciences, 13*(1), 50-61.

Al-Samarrai, S., & Zaman, H. (2007). Abolishing school fees in Malawi: The impact on education access and equity. *Education Economics, 15*(3), 359-375.

An, W., & Winship, C. (2017). Causal inference in panel data with application to estimating race-of-interviewer effects in the general social survey. *Sociological Methods & Research, 46*(1), 68-102.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics*: Princeton university press.

Angrist, J. D., & Pischke, J.-S. (2014). *Mastering'metrics: The path from cause to effect*: Princeton university press.

Aqil, M., Aziz, S. F., Dilshad, M., & Qadeer, S. (2014). Relationship between public education expenditures and economic growth of Pakistan. *IOSR Journal of Humanities and Social Science, 19*(3), 153-155.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies, 58*(2), 277-297.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics, 68*(1), 29-51.

Aroian, L. A. (1947). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics*, 265-271.

Asghar, Z. (2008a). Energy-GDP relationship: a causal analysis for the five countries of South Asia. *Applied Econometrics and International Development, 8*(1).

Asghar, Z. (2008b). Simulation evidence on Granger causality in presence of a confounding variable. *International Journal of Applied Econometrics and Quantitative Studies, 5*(2), 71-86.

Asghar, Z., & Abid, I. (2007). Performance of lag length selection criteria in three different situations.

Attiaoui, I., Toumi, H., Ammouri, B., & Gargouri, I. (2017). Causality links among renewable energy consumption, CO 2 emissions, and economic growth in Africa: evidence from a panel ARDL-PMG approach. *Environmental science and pollution research, 24*(14), 13036-13048.

Awokuse, T. O., & Bessler, D. A. (2003). Vector autoregressions, policy analysis, and directed acyclic graphs: an application to the US economy. *Journal of Applied Economics, 6*(1), 1-24.

Bai, J., & Kao, C. (2006). On the estimation and inference of a panel cointegration model with cross-sectional dependence. *Contributions to economic analysis, 274*, 3-30.

Baltagi, B. H. (2008). *Econometric analysis of panel data* (Vol. 4): Springer.

Bayraktar-Sağlam, B. (2016). The stages of human capital and economic growth: Does the direction of causality matter for the rich and the poor? *Social Indicators Research, 127*(1), 243-302.

Bayramoğlu, A. T., & Öztürk, Z. (2018). Assessing the twin and triple deficit hypotheses in developing economies: A panel causality analysis. In *Global Approaches in Financial Economics, Banking, and Finance* (pp. 209-225): Springer.

Bedir, S., & Yilmaz, V. M. (2016). CO 2 emissions and human development in OECD countries: Granger causality analysis with a panel data approach. *Eurasian Economic Review, 6*(1), 97-110.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics, 119*(1), 249-275.

Bessler, D. A., & Yang, J. (2003). The structure of interdependence in international stock markets. *Journal of international money and finance, 22*(2), 261-287.

Bhattacharjee, B., Schwer, D. A., Barton, P. I., & Green, W. H. (2003). Optimally-reduced kinetic models: reaction elimination in large-scale kinetic mechanisms. *Combustion and Flame, 135*(3), 191-208.

Bilal, K., Malik, S. U. R., Khalid, O., Hameed, A., Alvarez, E., Wijaysekara, V., . . . Ali, M. (2014). A taxonomy and survey on green data center networks. *Future Generation Computer Systems, 36*, 189-208.

Bischoff, I., & Prasetyia, F. (2015). *Determinants of local public expenditures on education: empirical evidence for Indonesian districts between 2005 and 2012*. Retrieved from

Blaug, M., & Mark, B. (1992). *The methodology of economics: Or, how economists explain*: Cambridge University Press.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics, 87*(1), 115-143.

Boţa-Avram, C., Groşanu, A., Răchişan, P.-R., & Gavriletea, M. D. (2018). The bidirectional causality between country-level governance, economic growth and sustainable development: A cross-country data analysis. *Sustainability, 10*(2), 502.

Breitung, J. (2001). *The local power of some unit root tests for panel data*: Emerald Group Publishing Limited.

Breitung, J., & Pesaran, M. H. (2008). Unit roots and cointegration in panels. In *The econometrics of panel data* (pp. 279-322): Springer.

Bryant, H. L., Bessler, D. A., & Haigh, M. S. (2009). Disproving causal relationships using observational data. *Oxford Bulletin of Economics and Statistics, 71*(3), 357-374.

Busemeyer, M. R. (2007). Determinants of public education spending in 21 OECD democracies, 1980–2001. *Journal of European Public Policy, 14*(4), 582-610.

Cama, M. G., Jorge, M. L., & Peña, F. J. A. (2016). Gender differences between faculty members in higher education: A literature review of selected higher education journals. *Educational Research Review, 18*, 58-69.

Cartwright, N. (2003). Counterfactuals in economics: a commentary.

Castles, F. G. (1989). Explaining public education expenditure in OECD nations. *European Journal of Political Research, 17*(4), 431-448.

Chandra, A. (2010). *Does government expenditure on education promote economic growth? An econometric analysis*. Retrieved from

Chang, T., Chu, H.-P., Deale, F. W., & Gupta, R. (2016). The causal relationship between happiness and smoking: a bootstrap panel causality test. *Journal of Happiness Studies, 17*(3), 1327-1336.

Chang, T., Chu, H.-P., Deale, F. W., Gupta, R., & Miller, S. M. (2017). The relationship between population growth and standard-of-living growth over

1870–2013: evidence from a bootstrapped panel Granger causality test. *Empirica, 44*(1), 175-201.

Chatterji, M., Mohan, S., & Dastidar, S. G. (2014). Determinants of public education expenditure: Evidence from Indian states.

Chen, H. T. (1990). *Theory-driven evaluations*: Sage publications.

Cheong, J., MacKinnon, D. P., & Khoo, S. T. (2003). Investigation of mediational processes using parallel process latent growth curve modeling. *Structural equation modeling, 10*(2), 238-262.

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *Journal of abnormal psychology, 112*(4), 558.

Conrad, D. A. (2011). The level-specific effects of education on economic growth: Evidence from four Caribbean countries. *The journal of developing areas, 45*(1), 279-290.

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

Cui, B., Feng, Q., Wang, H., Wang, M., Peng, Z., Li, P., . . . Fan, Z. (2015). Fecal microbiota transplantation through mid-gut for refractory C Crohn's disease: Safety, feasibility, and efficacy trial results. *Journal of gastroenterology and hepatology, 30*(1), 51-58.

De Hoyos, R. E., & Sarafidis, V. (2006). Testing for cross-sectional dependence in panel-data models. *The stata journal, 6*(4), 482-496.

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

Dietrich, A. (2012). Does growth cause structural change, or is it the other way around? A dynamic panel data analysis for seven OECD countries. *Empirical economics, 43*(3), 915-944.

Dumitrescu, E.-I., & Hurlin, C. (2012). Testing for Granger non-causality in heterogeneous panels. *Economic modelling, 29*(4), 1450-1460.

Eberhardt, M. (2011). *Panel time-series modeling: New tools for analyzing xt data.* Paper presented at the 2011 UK Stata Users Group meeting.

Ender, E. (1995). *Sexing the mind: Nineteenth-century fictions of hysteria*: Cornell University Press.

Engle, R. F., Hendry, D. F., & Richard, J.-F. (1983). Exogeneity. *Econometrica: Journal of the Econometric Society*, 277-304.

Falch, T., & Rattsø, J. (1997). Political economic determinants of school spending in federal states: Theory and time-series evidence. *European Journal of Political Economy, 13*(2), 299-314.

Ferreira, C. (2016). Debt and economic growth in the European Union: A panel granger causality approach. *International Advances in Economic Research, 22*(2), 131-149.

Fonchamnyo, D. C., & Sama, M. C. (2016). Determinants of public spending efficiency in education and health: Evidence from selected CEMAC countries. *Journal of Economics and Finance, 40*(1), 199-210.

Frank, N. (2018). Making the grade: The contribution of education expenditure to economic growth. *Undergraduate Economic Review, 14*(1), 11.

Freedman, D. A. (1991). Statistical models and shoe leather. *Sociological methodology*, 291-313.

Freedman, R., Hall, M., Adler, L. E., & Leonard, S. (1995). Evidence in postmortem brain tissue for decreased numbers of hippocampal nicotinic receptors in schizophrenia. *Biological psychiatry, 38*(1), 22-33.

Geweke, J. (1982). "Measurement of linear dependence and feedback between multiple time series." Journal of the American statistical association **77**(378): 304-313.

Ghosh Dastidar, S., Mohan, S., & Chatterji, M. (2013). The relationship between public education expenditure and economic growth: The case of India.

Glymour, C. N., & Cooper, G. F. (1999). *Computation, causation, and discovery*: Aaai Press.

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child development*, 80-92.

Goodman, L. A. (1960). On the exact variance of products. *Journal of the American statistical association, 55*(292), 708-713.

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

Grob, U., & Wolter, S. C. (2007). Demographic change and public education spending: A conflict between young and old? *Education Economics, 15*(3), 277-292.

Guloglu, B., & Tekin, R. B. (2012). A panel causality analysis of the relationship among research and development, innovation, and economic growth in high-income OECD countries. *Eurasian Economic Review, 2*(1), 32-47.

Gupta, V., Hanges, P. J., & Dorfman, P. (2002). Cultural clusters: Methodology and findings. *Journal of world business, 37*(1), 11-15.

Haavelmo, T. (1944). The probability approach in econometrics. *Econometrica: Journal of the Econometric Society*, iii-115.

Hakro, A. N. (2007). The Incidence of Government Expenditures on Education and Health: Microeconomic Evidence from Pakistan Ahmed Nawaz Hakro and Muhammed Akram. *Lahore Journal of Economics, 12*(2), 27-48.

Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of econometrics, 64*(1-2), 307-333.

Hanck, C. (2011). Joshua D. Angrist and Jorn-Steffen Pischke (2009): Mostly Harmless Econometrics: An Empiricist's Companion. *Statistical Papers, 52*(2), 503.

Harmon, R., & Challenor, P. (1997). A Markov chain Monte Carlo method for estimation and assimilation into models. *Ecological modelling, 101*(1), 41-59.

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), 378-385.

Hausman, J. A. (1983). Specification and estimation of simultaneous equation models. *Handbook of econometrics, 1*, 391-448.

Heckman, J. J. (2000). Causal parameters and policy analysis in economics: A twentieth century retrospective. *The Quarterly Journal of Economics, 115*(1), 45-97.

Hendry, D. F., & Mizon, G. E. (1998). Exogeneity, causality, and co-breaking in economic policy analysis of a small econometric model of money in the UK. *Empirical Economics, 23*(3), 267-294.

Henrekson, M. (1993). Wagner's law-a spurious relationship? *Public finance, 46*(3).

Hirsch, W. Z. (1960). Determinants of public education expenditures. *National Tax Journal, 13*(1), 29-40.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica: Journal of the econometric society*, 1371-1395.

Hoover-Dempsey, K. V., Walker, J. M., Sandler, H. M., Whetsel, D., Green, C. L., Wilkins, A. S., & Closson, K. (2005). Why do parents become involved? Research findings and implications. *The elementary school journal, 106*(2), 105-130.

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

Hoover, K. D., Demiralp, S., & Perez, S. J. (2006). *A bootstrap method for identifying and evaluating a structural vector autoregression*. Retrieved from

Hsiao, C. (1981). Autoregressive modelling and money-income causality detection. *Journal of Monetary economics, 7*(1), 85-106.

Hume, D. (1752). *Political discourses*: R. Fleming.

Hurlin, C., & Dumitrescu, E. (2008). Testing for Granger non-causality in heterogeneous panels. *Department of Economics, University of Orleans Working Paper*, 1-32.

Husain, F., Qasim, M. A., & Sheikh, K. H. (2003). An analysis of public expenditure on education in Pakistan. *The Pakistan development review, 42*(4), 771-780.

Hutchinson, G. A., & Schumacher, L. (1997). Fiscal expenditure policy and economic growth: Evidence from Latin America and the Caribbean. *Social and Economic Studies*, 1-16.

Hyndman, R. J., Bashtannyk, D. M., & Grunwald, G. K. (1996). Estimating and visualizing conditional densities. *Journal of Computational and Graphical Statistics, 5*(4), 315-336.

Idrees, A. S., & Siddiqi, M. W. (2013). Does public education expenditure cause economic growth? Comparison of developed and developing countries. *Pakistan Journal of Commerce and Social Sciences (PJCSS), 7*(1), 174-183.

Imana, D. K. (2017). The determinants of public education expenditures: an empirical analysis of changing patterns and growth of public expenditure on education in Kenya. *Education, 7*(4).

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of economic literature, 47*(1), 5-86.

Inglesi-Lotz, R., Chang, T., & Gupta, R. (2015). Causality between research output and economic growth in BRICS. *Quality & Quantity, 49*(1), 167-176.

Jerrum, M. (1998). Mathematical foundations of the Markov chain Monte Carlo method. In *Probabilistic methods for algorithmic discrete mathematics* (pp. 116-165): Springer.

Ji, Q., & Fan, Y. (2016). How do China's oil markets affect other commodity markets both domestically and internationally? *Finance Research Letters, 19*, 247-254.

Jorgenson, D. W., & Fraumeni, B. M. (1992). Investment in education and US economic growth. *The Scandinavian Journal of Economics*, S51-S70.

Judson, R. (1998). Economic growth and investment in education: how allocation matters. *Journal of Economic Growth, 3*(4), 337-359.

Jung, H.-S., & Thorbecke, E. (2003). The impact of public education expenditure on human capital, growth, and poverty in Tanzania and Zambia: a general equilibrium approach. *Journal of Policy Modeling, 25*(8), 701-725.

Kalim, R., Arshed, N., & Shaheen, S. (2019). Does competitiveness moderate inclusive growth: a panel study of low-income countries. *Competitiveness Review: An International Business Journal*.

Kandil, M., Shahbaz, M., & Nasreen, S. (2015). The interaction between globalization and financial development: new evidence from panel cointegration and causality analysis. *Empirical Economics, 49*(4), 1317-1339.

Kar, M., Nazlıoğlu, Ş., & Ağır, H. (2011). Financial development and economic growth nexus in the MENA countries: Bootstrap panel granger causality analysis. *Economic modelling, 28*(1-2), 685-693.

Kelly, T. (1997). Public expenditures and growth. *The Journal of Development Studies, 34*(1), 60-84.

Khan, A., Ahmed, M., & Bibi, S. (2019). Financial development and economic growth nexus for Pakistan: a revisit using maximum entropy bootstrap approach. *Empirical Economics, 57*(4), 1157-1169.

Khan, H. N., Khan, M. A., Razli, R. B., Shehzada, G., Krebs, K. L., & Sarvghad, N. (2016). Health care expenditure and economic growth in SAARC countries (1995–2012): a panel causality analysis. *Applied Research in Quality of Life, 11*(3), 639-661.

Khoo, S. T. (2001). Assessing program effects in the presence of treatment—baseline interactions: A latent curve approach. *Psychological Methods, 6*(3), 234.

Kim, Y., & Steiner, P. M. (2021). Causal graphical views of fixed effects and random effects models. *British Journal of Mathematical and Statistical Psychology, 74*(2), 165-183.

Kirmizi-Alsan, E., Bayraktaroglu, Z., Gurvit, H., Keskin, Y. H., Emre, M., & Demiralp, T. (2006). Comparative analysis of event-related potentials during Go/NoGo and CPT: decomposition of electrophysiological markers of response inhibition and sustained attention. *Brain Research, 1104*(1), 114-128.

Kónya, L. (2006). Exports and growth: Granger causality analysis on OECD countries with a panel data approach. *Economic Modelling, 23*(6), 978-992.

Kurban, H., Gallagher, R. M., & Persky, J. J. (2015). Demographic changes and education expenditures: A reinterpretation. *Economics of Education Review, 45*, 103-108.

Leamer, E. E. (1985). *Vector autoregressions for causal inference?* Paper presented at the Carnegie-rochester conference series on Public Policy.

Lechner, M. (2011). *The estimation of causal effects by difference-in-difference methods*: Now Hanover, MA.

Leong, A., Rehman, W., Dastani, Z., Greenwood, C., Timpson, N., Langsetmo, L., . . . Wong, B. Y. (2014). The causal effect of vitamin D binding protein (DBP) levels on calcemic and cardiometabolic diseases: a Mendelian randomization study. *PLoS medicine, 11*(10), e1001751.

Li, Y., Woodard, J. D., & Leatham, D. J. (2013). Causality among foreign direct investment and economic growth: A directed acyclic graph approach. *Journal of Agricultural and Applied Economics, 45*(4), 617-637.

Liu, S., & Neudecker, H. (1997). Experiments with mixtures: optimal allocations for Becker's models. *Metrika, 45*(1), 53-66.

Lopez, L., & Weber, S. (2017). Testing for Granger causality in panel data. *The Stata Journal, 17*(4), 972-984.

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation review, 17*(2), 144-158.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological methods, 7*(1), 83.

MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate behavioral research, 30*(1), 41-62.

Mathur, S. K., Arora, R., Ghoshal, I., & Singh, S. (2016). Domestic Energy Consumption and Country's Income Growth: A Quantitative Analysis of Developing and Developed Countries Using Panel Causality, Panel VECM, Panel Cointegration and SURE. *Journal of Quantitative Economics, 14*(1), 87-116.

McMahon, W. W. (1970). An economic analysis of major determinants of expenditures on public education. *The Review of Economics and Statistics*, 242-252.

Meeker, W. Q., Cornwell, L. W., & Aroian, L. A. (1981). *The product of two normally distributed random variables*: American Mathematical Soc.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*(1), 107-122.

Mill, J. S. (1884). *Principles of political economy*: D. Appleton.

Moneta, A., Chlaß, N., Entner, D., & Hoyer, P. (2011). *Causal search in structural vector autoregressive models.* Paper presented at the NIPS Mini-Symposium on Causality in Time Series.

Mundial, B. (1995). *Priorities and strategies for education: a World Bank review*: Banco Mundial.

Mutascu, M. (2015). A bootstrap panel Granger causality analysis of government revenues and expenditures in the PIIGS countries.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52*(3), 431-462.

Norris, R., Nichols, S., Ransom, G., Webb, A., Stewardson, M., Liston, P., & Mugodo, J. (2008). Causal Criteria Analysis Methods Manual. *Canberra: eWater*.

Oluwatobi, S., & Ogunrinola, I. O. (2011). Government expenditure on human capital development: Implications for economic growth in Nigeria. *Journal of sustainable development, 4*(3).

Omri, A. (2014). An international literature survey on energy-economic growth nexus: Evidence from country-specific studies. *Renewable and Sustainable Energy Reviews, 38*, 951-959.

Ozturk, I. (2010). A literature survey on energy–growth nexus. *Energy Policy, 38*(1), 340-349.

Parés, J. M., Pérez-González, A., Rosas, A., Benito, A., de Castro, J. B., Carbonell, E., & Huguet, R. (2006). Matuyama-age lithic tools from the Sima del Elefante site, Atapuerca (northern Spain). *Journal of Human Evolution, 50*(2), 163-169.

Paulusma, C. C., Bosma, P. J., Zaman, G. J., Bakker, C. T., Otter, M., Scheffer, G. L., . . . Elferink, R. P. O. (1996). Congenital jaundice in rats with a mutation in a multidrug resistance-associated protein gene. *Science, 271*(5252), 1126-1128.

Payne, J. E. (2010). A survey of the electricity consumption-growth literature. *Applied energy, 87*(3), 723-731.

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

Pearl, J. (2000). Causal inference without counterfactuals: Comment. *Journal of the American Statistical Association, 95*(450), 428-431.

Pedroni, P. (2013). Structural Panel VARs. Econometrics, 1 (2), 180-206. In.

Pesaran, M. H., Schuermann, T., & Weiner, S. M. (2004). Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business & Economic Statistics, 22*(2), 129-162.

Phillips, P. C. (1986). Understanding spurious regressions in econometrics. *Journal of econometrics, 33*(3), 311-340.

Pierce, D. A., & Haugh, L. D. (1977). Causality in temporal systems: Characterization and a survey. *Journal of econometrics, 5*(3), 265-293.

Polomé, P. P. Dynamic Panel Data: Intro 2016-17.

Poterba, J. M. (1997). Demographic structure and the political economy of public education. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management, 16*(1), 48-66.

Pradhan, R. P., Arvin, M. B., & Hall, J. H. (2019). The nexus between economic growth, stock market depth, trade openness, and foreign direct investment: the case of ASEAN countries. *The Singapore Economic Review, 64*(03), 461-493.

Ramli, N., Hashim, E., & Marikan, D. A. A. (2016). Relationship between education expenditure, capital, labor force and economic growth in Malaysia. *International Journal of Academic Research in Business and Social Sciences, 6*(12), 459-468.

Raudenbush, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual review of psychology, 52*(1), 501-525.

Rebonato, R. (2016). Mostly Harmless Econometrics: An Empiricist's Companion; Mastering 'Metrics: The Path from Cause to Effect. In: Taylor & Francis.

Rehman, A., Zaman, A., & Zaman, A. (2017). The concept of stringency for test comparison: the case of a Cauchy location parameter. *International Econometric Review, 9*(1), 1-20.

Reichenbach, H. (1956). *The direction of time* (Vol. 65): Univ of California Press.

Requena, G., Garcés, G., Asghar, Z., Marks, E., Staron, P., & Cloetens, P. (2011). The Effect of the Connectivity of Rigid Phases on Strength of Al-Si Alloys. *Advanced Engineering Materials, 13*(8), 674-684.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The stata journal, 9*(1), 86-136.

Sabir, S., Ahmed, S., & Lodhi, N. (2003). Morphological and biochemical variation in Sea buckthorn Hippophae rhamnoides ssp. turkestanica, a multipurpose plant for fragile mountains of Pakistan. *South African journal of botany, 69*(4), 587-592.

Sagarik, D. (2013). An Analysis of the Determinants of Education Expenditure in T hailand. *Asian Politics & Policy, 5*(2), 227-248.

Sahn, D. E., & Younger, S. D. (2000). Expenditure incidence in Africa: microeconomic evidence. *Fiscal Studies, 21*(3), 329-347.

Selden, T. M., & Wasylenko, M. J. (1992). *Benefit incidence analysis in developing countries* (Vol. 1015): World Bank Publications.

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in human development, 6*(2-3), 144-164.

Selowsky, M. (1979). *Who benefits from government expenditure?* : Published for the World Bank [by] Oxford University Press.

Shawe-Taylor, J., & Cristianini, N. (2000). *An introduction to support vector machines and other kernel-based learning methods* (Vol. 204): Volume.

Shimizu, S., Hoyer, P. O., Hyvärinen, A., Kerminen, A., & Jordan, M. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research, 7*(10).

Sigmund, M., & Ferstl, R. (2019). Panel vector autoregression in R with the package panelvar. *The Quarterly Review of Economics and Finance*.

Simon, E. J., & Shemin, D. (1953). The preparation of S-succinyl coenzyme A. *Journal of the American Chemical Society, 75*(10), 2520-2520.

Simon, H. A. (1953). Notes on the observation and measurement of political power. *The Journal of Politics, 15*(4), 500-516.

Sims, C. A. (1972). Money, income, and causality. *The American economic review, 62*(4), 540-552.

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

Sinha, D. (1998). Government expenditure and economic growth in Malaysia. *Journal of Economic Development, 23*(2), 71-80.

Smith, A. (1776). An inquiry into the nature and causes of the wealth of nations: Volume One. In: London: printed for W. Strahan; and T. Cadell, 1776.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology, 13*, 290-312.

Spirtes, P., Glymour, C., & Scheines, R. (1993). Discovery algorithms for causally sufficient structures. In *Causation, prediction, and search* (pp. 103-162): Springer.

Spirtes, P., Glymour, C. N., Scheines, R., & Heckerman, D. (2000). *Causation, prediction, and search*: MIT press.

Steel, Z., Chey, T., Silove, D., Marnane, C., Bryant, R. A., & Van Ommeren, M. (2009). Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement: a systematic review and meta-analysis. *Jama, 302*(5), 537-549.

Stokes, P. A., & Purdon, P. L. (2017). A study of problems encountered in Granger causality analysis from a neuroscience perspective. *Proceedings of the national academy of sciences, 114*(34), E7063-E7072.

Suppes, P. (1970). A Theory of Probabilistic Causality. In: Amsterdam: North-Holland Publishing Company.

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), 357-367.

Sylwester, K. (2002). Can education expenditures reduce income inequality? *Economics of education review, 21*(1), 43-52.

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology, 51*(6), 309.

Tilak, J. B. (1989). The recession and public investment in education in Latin America. *Journal of interamerican studies and world affairs, 31*(1-2), 125-146.

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

Tomić, Z. (2015). Analysis of the impact of public education expenditure on economic growth of European Union and BRICS. *Economic analysis, 48*(1-2), 19-38.

Verbina, I., & Chowdhury, A. (2004). What determines public education expenditures in Russia? 1. *Economics of Transition, 12*(3), 489-508.

Vinod, H. D. (2006). Maximum entropy ensembles for time series inference in economics. *Journal of Asian Economics, 17*(6), 955-978.

Wanat, S., Papież, M., & Śmiech, S. (2016). Insurance Market Development and Economic Growth in Transition Countries: Some new evidence based on bootstrap panel Granger causality test.

Wang, K., & Ji, Q. (2017). *Real time eye gaze tracking with 3d deformable eye-face model.* Paper presented at the Proceedings of the IEEE International Conference on Computer Vision.

White, R. J., & Peterson, B. M. (1994). Comments on cross-correlation methodology in variability studies of active galactic nuclei. *Publications of the Astronomical Society of the Pacific, 106*(702), 879.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116*(2), 363.

Windmeijer, F. (2000). Efficiency comparisons for a system GMM estimator in dynamic panel data models. In *Innovations in Multivariate Statistical Analysis* (pp. 175-184): Springer.

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

Woodward, J. (2003). Scientific explanation. *The Stanford*.

Yalta, A. T. (2011). Analyzing energy consumption and GDP nexus using maximum entropy bootstrap: the case of Turkey. *Energy Economics, 33*(3), 453-460.

Yang, J., & Bessler, D. A. (2008). Contagion around the October 1987 stock market crash. *European Journal of Operational Research, 184*(1), 291-310.

Yang, R.-Y., Chang, L.-C., Hsu, J.-C., Weng, B. B., Palada, M. C., Chadha, M., & Levasseur, V. (2006). Nutritional and functional properties of Moringa leaves– From germplasm, to plant, to food, to health. *Moringa leaves: Strategies, standards and markets for a better impact on nutrition in Africa. Moringanews, CDE, CTA, GFU. Paris*, 1-9.

Yang, Z., & Rannala, B. (1997). Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. *Molecular biology and evolution, 14*(7), 717-724.

Younger, M., Warrington, M., & Williams, J. (1999). The gender gap and classroom interactions: reality and rhetoric? *British Journal of Sociology of Education, 20*(3), 325-341.

Yun, W. S., & Yusoff, R. (2018). The determinants of public education expenditure in Malaysia. *Jurnal Ekonomi Malaysia, 52*(2), 109-122.

Yun, W. S., & Yusoff, R. (2019). Determinants of Public Education Expenditure: A Review. *SOUTHEAST ASIAN JOURNAL OF ECONOMICS, 7*(2), 127-142.

Zapata, H. O., & Rambaldi, A. N. (1997). Monte Carlo evidence on cointegration and causation. *Oxford Bulletin of Economics and Statistics, 59*(2), 285-298.

Zellner, A. (1979). *Causality and econometrics.* Paper presented at the Carnegie-Rochester conference series on public policy.

Zhang, X., Wang, C., Tan, Y., & Zhang, F. (2011). *Causality between energy consumption and economic growth in Beijing: evidence from cross-industry panel data.* Paper presented at the International Conference on Computer Science, Environment, Ecoinformatics, and Education.

Zheng, W., & Sakellariou, R. (2013). Stochastic DAG scheduling using a Monte Carlo approach. *Journal of Parallel and Distributed Computing, 73*(12), 1673-1689.

Zhou, W., Li, W., Wang, J.-Q., Qu, Y., Yang, Y., Xie, Y., . . . Zhao, D. (2014). Ordered mesoporous black TiO2 as highly efficient hydrogen evolution photocatalyst. *Journal of the American Chemical Society, 136*(26), 9280-9283.