A FRACTIONAL COINTEGRATION PANEL MODEL WITH FIXED EFFECT

OLANIRAN SAIDAT FEHINTOLA

UNIVERSITI SAINS MALAYSIA

2023

A FRACTIONAL COINTEGRATION PANEL MODEL WITH FIXED EFFECT

by

OLANIRAN SAIDAT FEHINTOLA

**Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy**

May 2023

ACKNOWLEDGEMENT

I would like to express my profound gratitude to Allah SWT for giving me the opportunity and knowledge to make it possible to achieve this. The Federal Government of Nigeria through Tetfund and Kwara State University are the masterpiece that made this research possible, I appreciate all the aforementioned bodies. A very special gratitude to my darling husband for his moral, financial, physical and emotional support throughout, all of this I won't take for granted In Shaa Allah. My supervisor, Prof Madya Mohd Tahir Ismail deserves all the accolades for his utmost co-operation and full support throughout the research work, the School of Mathematical Sciences, Universiti Sains Malaysia and Institute Postgraduate Studies (IPS) has also tremendously contributed to my success. I would also like to appreciate my children, brothers, friends and other relatives, everyone of them gave me a listening ear and yearned for the success of this research. A special dedication to my late mother who laid the foundation of my success, believed in me and helped tremendously to make my dream come to reality, keep resting mum and be rest assured that Allah took over completely.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

ARDL	Autoregressive distributed lag
CA	Cointegration Analysis
G20	Group of Twenty
GLS	Generalized Least Square
MLE	Maximum Likelihood Estimator
MM2H	Malaysia My 2^{nd} Home
NBLS	Narrow Band Least Square
OECD	Organization for Economic Cooperation and Development
OLS	Ordinary Least Square
PRIMA	Perumahan Rakyat 1 Malaysia
PPP	Purchasing Power Parity
VECM	Vector Error Correction Model

LIST OF SYMBOLS

n Number of cross-sectional units

t Number of time units

T Total of sample size in a panel data

d Fractional integration order of the series

γ Fractional integration order of the residual

$I(0)$ Integration at level

$I(1)$ Integration at first difference

$I(d)$ Integration at fractional order d

$CI(d,b)$ Cointgeration order d,b

x_t Time series process x_t at time t

y_t Time series process y_t at time t

$\boldsymbol{X}_t$ $p-$dimensional time series process at time t

β Cointegrating vector

μ Overall fixed effect

μ_i i^{th} cross-sectional fixed effect

L Lag operator

Δ Differencing operator

Γ Gamma function

$\boldsymbol{G}$ Spectral matrix

$\mathbf{\Lambda}$ Spectral density matrix

λ Fourier frequency

det Determinant of a matrix

log Natural logarithm

e Exponential function

m Bandwidth size

η_m Exponent of bandwidth size

χ^2_{df} Chi-square statistic at df degrees of freedom

$N(0,1)$ Standardized normal distribution

ρ Correlation coefficient

Ψ Characteristics function

$\Re$ Real number

MODEL PANEL KOINTEGRASI PECAHAN DENGAN KESAN TETAP

ABSTRAK

Beberapa penulis telah mengkaji kointegrasi pecahan dalam data siri masa, tetapi sedikit atau tiada pertimbangan telah diperluaskan kepada tetapan data panel. Walau bagaimanapun, ekonomi dan kewangan baru-baru ini seperti pulangan portfolio di seluruh firma, indeks harga dan kadar pertukaran di seluruh negara sering mempamerkan sifat ingatan panjang. Oleh itu, tesis ini bertujuan untuk membangunkan model panel berkointegrasi pecahan dengan andaian kesan tetap. Objektif pertama adalah untuk membandingkan tingkah laku sampel terhingga prosedur ujian siri masa kointegrasi pecahan sedia ada dalam tetapan data panel. Perbandingan ini dilakukan untuk menentukan ujian terbaik yang boleh disesuaikan dengan kointegrasi pecahan dalam tetapan data panel. Khususnya, kajian simulasi dan analisis data kehidupan sebenar telah dilakukan untuk mengkaji perubahan dalam kadar ralat empirikal jenis I dan kuasa enam ujian kointegrasi pecahan semiberparameter dalam tetapan panel. Penemuan analisis menunjukkan bahawa ujian berasaskan reja berguna untuk penyesuaian dalam persekitaran panel. Kedua, dua ujian kointegrasi pecahan siri masa berasaskan reja terbaik yang diperhatikan telah dilaksanakan dalam tetapan panel menggunakan eksperimen simulasi Monte-Carlo. Keputusan eksperimen menunjukkan bahawa salah satu ujian adalah sah untuk tertib kointegrasi pecahan kurang daripada 0.5, yang lain adalah teritlak dan menerima apa-apa tertib kointegrasi pecahan dalam julat $[0,1]$ pada saiz sampel yang berbeza-beza. Akhirnya, pendekatan panel berkointegrasi pecahan telah dibangunkan untuk menguji model Pariti Kuasa Pembelian (PPP) mutlak di kalangan 16 negara Afrika Barat menggunakan data yang merangkumi 49 tahun (1971 - 2019). Keputusan dari ujian kointegrasi pecahan yang baru dibangunkan mengesahkan

kehadiran PPP relatif untuk negara-negara dalam jangka masa panjang, sementara anggaran pintasan biasa dan vektor kointegrasi mengesahkan ketiadaan PPP mutlak untuk negara-negara. Penemuan keseluruhan dalam tesis ini menunjukkan bahawa adalah salah untuk menganggap kewujudan kointegrasi pecahan dalam tetapan panel berdasarkan tertib pecahan siri dan reja seperti yang sering dilakukan dalam kajian sebelumnya. Di samping itu, menggunakan pendekatan panel kointegrasi pecahan negara dan firma adalah lebih bermaklumat dan berintuisi apabila objektifnya adalah untuk menentukan perkembalian min jangka panjang bersama.

A FRACTIONAL COINTEGRATION PANEL MODEL WITH FIXED EFFECT

ABSTRACT

Several authors have studied fractional cointegration in time series data, but little or no consideration has been extended to panel data settings. However, recent economics and financial panel datasets such as portfolio returns across firms, price indices and exchange rates across countries often exhibit long-memory properties. Therefore, this thesis aims to develop a fractional cointegrated panel model with a fixed effect assumption. The first objective was to compare the finite sample behaviour of existing fractional cointegration time-series test procedures in panel data settings. This comparison is performed to determine the best tests that can be adapted to fractional cointegration in panel data settings. Specifically, simulation studies and real-life data analysis were performed to study the changes in the empirical type I error rate and power of six semiparametric fractional cointegration tests in panel settings. The analysis findings revealed that the residual-based tests are useful for adaptation in a panel setting. Secondly, the best two residual-based time series fractional cointegration tests observed were implemented in panel settings using Monte-Carlo simulation experiments. The results of the experiments showed that one of the tests is valid for fractional cointegration order of less than 0.5, the other is generalized and accepts any fractional cointegration order within the range $[0,1]$ at varying sample sizes. Finally, a fractional cointegrated panel approach was developed for testing the absolute Purchasing Power Parity (PPP) model among 16 West African countries using data that spans 49 years $(1971-2019)$. The results from the newly developed fractional cointegration test confirm the presence of relative PPP for the countries in the long run, while the estimation

of common intercepts and cointegration vector confirms the absence of the absolute PPP for the countries. The overall findings in this thesis imply that it is wrong to assume the existence of fractional cointegration in panel settings based on the fractional orders of the series and residuals as often done in previous studies. In addition, using a panel of countries or firms fractional cointegration approach is more informative and intuitive when the objective is to determine the joint long-run mean reversion.

CHAPTER 1

INTRODUCTION

1.1 Background of the study

In econometrics, whether a data is experimental or observational, it can be majorly classified into cross-sectional data, time-series data and panel data whereas the panel data combines the attributes of both the cross-sectional data and the time series. However, in macroeconomics and finance, variables are usually presented in panels to describe the varying characteristics of the different entities such as currencies, assets, countries, income, people, and so on. Since panel data allows for interactions of cross-sections with each other, it leads to a more robust inference when correctly specified.

A panel is a cross-section or group of entities that are surveyed periodically over a given time span. Panel data can also be referred to as longitudinal data or cross-sectional time series data. These longitudinal data contain “observations on the same units across multiple time periods” (Purba and Bimantara, 2020). A panel data set consists of multiple entities, each of which has repeated measurements at various time intervals. Panel data analysis is a method of studying a specific entity within multiple sites that is observed on a regular basis over a set time period. Importantly, panel analysis allows researchers to study the dynamics of change using short time series and can provide a rich and powerful study of a set of subjects by taking into account both the space and time dimensions of the data.

Li et al. (2020) works thoroughly on panel data analysis and concludes that it is an essential method for analyzing longitudinal data and allows for a complex set of regression analyses in both geographical and temporal dimensions. Also, it was added that Panel data analysis might be the only approach to longitudinally analyze data when it comes from multiple sites and the time series is too short for independent time series analysis because even if the series are too short for separate analysis, panel data analysis offers a diverse set of approaches for examining change over time in a certain sort of cross-sectional unit. An illustration of a panel data set is a grouping of five countries that all have the same economic variables, such as personal expenditures, personal and median disposable income, per capita income, population size, unemployment, and employment. With data collected every year for ten years. This pooled data set, also known as time series-cross sectional data, contains $5 * 10 = 50$ observations. In other words, the five countries are followed for ten years and sampled on an annual basis. A panel data's structure confers two dimensions. They have a cross-sectional unit of n observation, which could be a country, as well as a temporal reference, T which could be a year. There are two dimensions to the error term: one for the cross-sectional unit and one for the time period. In some cases, the cross-sections can be nested within time.

Bandi et al. (2021); Phillips and Moon (1999) made a significant contribution to panel literature by thoroughly distinguishing the various limit ideas and clarifying linkages between them. Most crucially, it discovers adequate circumstances for the sequential and joint limits to coincide and these circumstances necessitate some uniformity requirements, as well as limitations on the relative pace at which n and T increases to infinity. There are several advantages of panel data and the first is that Panel data ensures that individual heterogeneity is considered and study the elements of changes. Individual entities are heterogeneous, according to panel data, and failing to account for this heterogeneity leads to serious misspecification. Time-series and cross-section studies that do not account for this heterogeneity risk producing biased results. This is where the strength of panel data analysis lies such that it can control for these state- and time-invariant variables. Lyatuu et al. (2021) stated that, in contrast to cross-sectional data analysis, panel dataset produces information on individual entity changes. Second, unlike time-series studies, which are afflicted by multicollinearity, panel data provide more informative data, more variability, less collinearity among variables, more degrees of freedom, and more efficiency because more data can produce more reliable parameter estimates. Also, panel data analysis accommodates behavioral models and tests when compared to time-series and cross-sectional data analysis.This is because they can uncover and evaluate effects that are simply not observable in pure cross-section or time-series data. In fact, Koop et al. (2001) agrees, claiming that panels are preferable for studying and modelling technical efficiency and quite similarly, Hsiao (2014) claims that in panels on a distributed lag model, less limits can be imposed compared to a pure time series analysis. For other advantages of panel data see Hsiao (2014).

There are also disadvantages of using panel data and the most important one is the design and data collection problems. One of the issues is indequacy in covering the target population. Others include non-response issue whihc arises as a results of lack of cooperation from the respondent and or interviewer error, frequency of interviews, reference period, usage of bounds, measurement error, and time-in-sample bias are examples of design and data collection issues. Non-response occurs in cross-sectional studies as well, but it is more severe in panels because non-response occurs in subsequent waves of the panel Bai et al. (2009).

However, Baltagi (2021) concludes that although panel data has numerous advantages, it is not a cure-all. The ability of panel data to isolate the impacts of individual actions, treatments, or more general policies is highly dependent on the statistical tools' assumptions being compatible with the data collection procedure. Hsiao (2014) added that it may be useful to keep in mind when using panel data ways to increase the efficiency of parameter estimates and the reliability of statistical inference, and how the assumptions underlying the statistical inference procedures and the data-generating process are compatible when choosing a proper method for exploiting the richness and unique properties of the panel. Since panel data combines the strengths and weaknesses of time series data which includes non-stationarity of individual series across panels, series across panels may also be cointegrated. Cointegration occurs when the linear combination of two non-stationary series is stationary (Ahsan et al., 2020; Gjelsvik et al., 2020; Zhang et al., 2021).

The growing availability of panel data with large T dimension (i.e. where the number of time series observations is large) has stimulated a growth in research, both empirical and theoretical, which discusses time series issues in panel data models. Of particular interest are issues relating to nonstationarity and cointegration. The importance of this area of research is evidenced by the increasing tendency for researchers to employ panels of nonstationary processes in empirical studies in macroeconomics and international economics. Baltagi et al. (2021) identifies many areas of application, including purchasing power parity (PPP), growth convergence and international R&D

spillovers. To give one example which illustrates the issues which can be addressed through the use of panel data consider Jacobson et al. (2002). These authors use a multivariate panel cointegration model and demonstrate that, although strong purchasing power parity restrictions are rejected, the location of the cointegrating space is similar for all countries considered. This provides some evidence in support of PPP. There have been a range of methods proposed to obtain inference relating to cointegration in panel data models.

Among many others, we note that residual-based, LM and likelihood based tests have been proposed by McCoskey and Kao (1998), Kao (1999), Larsson et al. (2001a), Groen and Kleibergen (2003), Pedroni (2004) and Rahman and Velayutham (2020). These papers use a variety of estimation methods, ranging from OLS to maximum likelihood and generalized method of moments. In recent years, fractional cointegration has piqued interest in time series econometrics (see, for example, Baltagi (2021)). Cointegrating relationships between non-stationary economic variables can exist without observable processes being unit root $I(1)$ processes or cointegrating errors being $I(0)$ processes, according to fractional cointegration analysis. Although both fractional and standard cointegration were defined at the same time in Engle and Granger (1987), standard cointegration has received more attention. The memory parameter in standard cointegration can only have integer values, and tests for the existence of cointegration rely on unit root theory.Because the memory parameter can take fractional values and be any positive real number, the fractional cointegration framework is more general. Engle and Granger (1987) and Škare et al. (2020) assumed that the cointegrating vector(s) do not change over time in their standard approach. When structural breaks and regime shifts are taken into account, however, the assumption of fixed cointegrating vector(s) becomes quite restrictive.

1.2 Fractional cointegration and cointegrated panels

Fractional cointegration refers to a statistical method used to analyze the long-run relationship between two or more time series data. In traditional cointegration analysis, the time series data is assumed to have a fixed integration order, whereas in fractional cointegration, the integration order is allowed to be a fractional number. Fractional cointegration was first introduced by Granger and Joyeux (1980) and further developed by Hosking (1985) as also reported in Miyandoab et al. (2023). Since then, it has become an important tool for analyzing economic and financial time series data.

The concept of fractional integration was introduced by Granger (1980) and Hosking (1981), who showed that many macroeconomic and financial time series exhibit long memory. This means that the autocorrelation function of the series decays slowly, indicating that the series has a long-term dependence on its past values. Fractional integration models allow for the incorporation of this long-term dependence into the analysis. Fractional cointegration builds on the concept of fractional integration by allowing for the analysis of the long-run relationship between two or more time series that exhibit long memory. The seminal paper on fractional cointegration was written by Engle and Granger (1991), in which they introduced the concept and developed a testing procedure for it. Since then, fractional cointegration has been widely used in various fields, including economics, finance, and engineering. For example, fractional cointegration has been used to analyze the relationship between interest rates and inflation (Ghysels et al., 1996), the relationship between oil prices and stock prices

(Baumeister and Kilian, 2014), and the relationship between exchange rates and stock prices (Chen et al., 2022).

Cointegrated panels, on the other hand, refer to a specific type of cointegration analysis that is used when there are multiple non-stationary time series variables that are observed across multiple individuals or entities (i.e., a panel data set). One of the key challenges in cointegrated panel analysis is that the cointegrating vectors may differ across individuals (Baltagi, 2021). However, fractional cointegrated panels refer to a class of panel data models that allow for non-integer (fractional) integration and cointegration relationships among the variables in the panel. These models have become increasingly popular in econometric research, as they can better capture the long-term dynamics and persistence of economic variables over time (Baltagi, 2021). The fractional cointegrated panel model idea was first formalized by Johansen and Nielsen (2010,1) in the context of a panel of time series with a common factor structure.

One of the most widely used models for fractional cointegrated panels is the common factor model, which was first proposed by Harris and Tzavalis (1999). In this model, the panel data are decomposed into a common factor component and an idiosyncratic component, and the common factor is assumed to be fractionally integrated. This allows for the possibility of fractional cointegration between the variables in the panel. The model has been extended in various ways, such as by including lagged dependent variables (Pesaran et al., 2006) or allowing for individual-specific short-run dynamics (Chudik and Pesaran, 2015). Another important contribution to the literature on fractional cointegrated panels is the work of Breitung and Hassler (2002), who proposed a panel version of the Dickey-Fuller test for fractional integration. This test has become a standard tool for empirical research on fractional cointegration in panel data.

1.3 Statement of problem

Evidence from studies, such as Škare et al. (2020), Yaya et al. (2021) and Kalymbetova et al. (2021) suggested that the equilibrium error may react to shocks more slowly than a stationary $I(0)$ process, making deviations from equilibrium more persistent. Indeed, as demonstrated by the technique suggested by Engle and Granger (1987) there is no requirement that the equilibrium error in a cointegration relationship mimic a $I(0)$ process. Granger and Joyeux (1980) and Hosking (1981) were the first to used the fractional integration in testing the purchasing power parity (PPP) theory. In a similar study by Cheung and Lai (1993), they modeled the equilibrium error as a fractionally-integrated $I(d)$ process.

A two-step testing procedure is often required for all subsequent applications of so-called fractional cointegration, including those cited in Gil-Alana and Hualde (2009),Tan et al. (2021) and Oloko et al. (2021) among others. They performed an OLS estimation of a fractional cointegrating vector based on the assumption that the variables do share common integrated processes (typically $I(1)$), and then check to see if the residuals from the OLS are $I(d)$ and d is a real number less than one. The main variations between those applications are how the fractional difference parameter d is estimated either using the semi-nonparametric method of Geweke and Porter-Hudak (1983) or the conditional sum of squares (CSS) estimator (CSS) of Chung and Baillie (1993) or the maximum likelihood estimator (MLE) of Sowell (1992). However, one major flaw with the empirical studies applications include the impossibility of testing a hypothesis regarding an economic relationship due to the fractional cointegration vector.

The points above show that a lot has been done on fractional cointegration and cointegrated panel data model. However, none of the authors has jointly considered the occurrence of the two problems simultaneously. Also, we are only aware of only one paper that considered fractionally integrated panel data model but not fractional cointegrated. Ergemen (2019); Ergemen and Velasco (2017) considered fractionally integrated panels with fixed effect and cross-sectional dependence. The author considered the fractional integration of individual vector within the panel model and not cointegration of two or more vectors. An empirical illustration of fractionally cointegrated panel model can be deduced by extending the fractional cointegrated model of Cheung and Lai (1993) on modelling foreign and domestic price indices to capture different countries or cross-sectional units. Thus, in this thesis, we present and estimate the fractional cointegration model for modelling panels of price indices.

1.4 Aim and Objectives

This study aims to propose and estimate the fractional cointegration model useful for modelling and testing its existence in panels data. Thus, the following objectives will be considered:

1. Investigate the finite sample behaviour of existing fractional cointegration time-series test procedures in panel data settings.

2. Extend the Chen and Hurvich (2006) and Wang et al. (2015) tests to accomodate

panel data structure.

3. Test the absolute purchasing power parity (PPP) theory using the fractional cointegration panel approach.

1.5 Significance of the study

The models and methods developed in this study are useful in two ways for the time series experts when modelling long memory panel data. Firstly, the new tests developed are useful for testing the existence of fractional cointegration in panel settings. These test procedures are expected to erase the era of arbitrarily assuming the existence of fractional cointegration in panel data based on the estimate of long memory parameters. Secondly, the new estimation method will be handy in modelling bivariate fractionally cointegrated panel data. Furthermore, the developed model is useful for the economist and government in testing the purchasing power parity theory and when comparing individual stock and market share values.

1.6 Scope and limitation

This study is limited in scope to fractional cointegration in panel data with specific application to the fixed-effect panel model. In addition, both residual-based and spectral density based fractional cointegration tests are considered.

1.7 Organization of the thesis

Chapter 1 introduces the problem and gap: testing and modelling fractional cointegration in panel settings. Chapter 2 reviews relevant literature on fractional cointegration tests and models in time series settings. Chapter 3 presents the methods used in this thesis: the fractional cointegration test and model for time series data. Chapter 4 presents the comparative analysis of residual-based and spectral based fractional cointegration for time series and panel data. Chapter 5 presents the modified residual-based fractional cointegration test for panel data developed in this thesis. Chapter 6 presents a generalized residual-based fractional cointegration test for panel data. Chapter 7 presents the estimation of the fractional cointegrated panel model with application in testing the absolute purchasing power parity theory. Chapter 8 present the summary, conclusion, recommendation and policy implication of the thesis.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Recent advancements in the field of panel data, such as nonlinear panels, high-dimensional data, factor models in economics and finance, and pseudo-panels, to name a few, have provided motivation for focusing on, collecting, and critically reviewing publications on fractional panel cointegration because there has been little extensive research on fractional cointegration in panel data, unlike fractional cointegration in time series.

2.2 Cointegration analysis

Cointegration, as defined by Engle and Granger (1987), means that there exists a cointegrating vector $(1-\beta')$, such that the linear combinations $y_t - x_t{}'\beta$ are stationary or are $\boldsymbol{I}(0)$ processes. However, Granger (2004) simplified it by explaining that the difference between two integrated series can be stationary, and it is referred to as "cointegration".

Cointegration analyses are often applied to time series data and panel data. Various cointegration methods have been developed over the years to solve particular problems and used with various models and short and long-run dynamics. There are several cointegration tests, but the popular ones are Engle-Granger, Johansen and Phillips-Ouliaris cointegration test (Khattab, 2021). The novel Johansen cointegration process can be used to estimate many Cointegration/Integration (CI) vectors, but it can also be utilized to generate a test statistic for determining the number of CI vectors.

Some recent research on cointegration analysis, aside from the conventional research that forms its basis, includes the study of Mousavi and Gandomi (2021) to train a Recurrent Neural Network (RNN), a portion of the acquired signals from Variational Mode Decomposition (VMD), as well as a portion of the Johansen Cointegration residuals is utilized as training features and targets, respectively. The remaining portions of the features were examined on both long and short term monitoring tasks and then utilized to forecast future CI residuals using the trained RNN. The suggested method can monitor structures for deterioration even when the Johansen algorithm fails to find a linear CI relationship among the frequency signals.

Cointegration Analysis (CA) is based on the extracted long-term equilibrium relationships. Thus, monitoring relationships based on cointegration analysis can be applied with fewer model updates. However, if the cointegration relationship changes, the previously developed CA model will no longer be able to describe the status of future nonstationary processes effectively. Hence, Yu et al. (2020) proposed a recursive CA-based adaptive monitoring scheme by successfully developing a recursive technique to update the monitoring model. After that, three monitoring statistics were created to indicate the operation status of the industrial process, with representation of both static deviation and dynamic fluctuation. Finally, according to the experimental data from two genuine industrial processes, an adaptive monitoring approach was built based on the suggested recursive CA that may efficiently respond to normal process variations without frequent model updates. However, Pedroni (2019) did a review of the considerable literature on CA starting from strategies for dealing with cross-sectional heterogeneity in cointegration testing and inference, to dealing with heterogeneity in residual based tests for cointegration. They compared residual-based and error correction-based testing to study cointegrating relationships in heterogeneous panels. In addition, they also developed nonparametric direction testing of long-run causality in heterogeneous cointegrated panels.

The Engle-Granger test, Johansen test and Phillips-Ouliaris tests that are mentioned above as traditional cointegration testing have restrictions that cause issues when conducting Cointegration Analysis with mixed ordering of variables. In such situations, researchers may either change the variables into stationary form, obviating the need for cointegration, or eliminate some variables incorrectly. However, in the early millennium, Pesaran in Persyn and Westerlund (2008) and others identified as PSS make some assumptions when creating the limits for an Augmented Autoregressive Distributed Lagged bound test. The Augmented Autoregressive Distributed Lag (ARDL) bounds test for cointegration uses an additional F-test on the lagged levels of the independent variable(s) in the ARDL equation, which does not require the assumption of a $I(1)$ dependent variable.

In addition, the augmented ARDL bounds test was demonstrated in Sam et al. (2019) by first using an empirical study on government spending and taxes to provide both the small sample and asymptotic critical values for simpler test implementation. An application of the Autoregressive Distributed Lag cointegration approach to real life data set is the work of Mostafa (2021) which explores the impact of government initiatives on Malaysian housing prices. The role of the gross domestic product, interest rate, and total population, as emphasized by the Life Cycle and Overlapping Generation Models, is also included in this study. From 1988 through 2017, annual time series data was used in this investigation. Developing a model of housing price determination with a focus on government policy using the autoregressive distributed lag (ARDL) framework. The research concludes that PR1MA has a favorable relationship with house prices and MM2H, on the other hand, is not a major driver of Malaysian housing prices.

As seen in several of the aforementioned literature applications, the extended ARDL bounds test sheds more light on the cointegration state and integration order of the tested variables. In the case of degenerate lagged independent variable(s), the ARDL equation is reduced to the Dickey-Fuller unit root equation, and the dependent variable is represented as $I(0)$; otherwise, it is $I(1)$. If the testing indicates a degenerate dependent variable or non-cointegration, it means the dependent variable is not included in the ARDL equation's cointegrating equation. Therefore, the dependent variable's movement is unresponsive to the movement of the independent variables, showing non-coherence.

2.3 Panel cointegration

In the empirical literature, the use of cointegration techniques to test for the presence of long-run relationships among integrated variables is gaining popularity. Unfortunately, the inherently low power of many of these tests when applied to datasets with both time series and cross-sectional data has been a common source of consternation for practitioners. However, pooling time series has traditionally required a significant amount of sacrifice in terms of the permissible heterogeneity of the individual time series. In ensuring broad applicability for panel cointegration tests, it is necessary to allow for as much heterogeneity among the individual members of the panel.

Persyn and Westerlund (2008) implemented the Westerlund (2007) four error correction based panel cointegration tests. They found that the tests are broad enough to accommodate a wide range of heterogeneity in the long-run cointegrating relationship and short-run dynamics and dependencies within and between cross-sectional units. Droge and Örsal (2009) compared Larsson et al. (2001b) standardized LR-bar statistic in the same line. It was concluded that panel-t and standardized LR-bar statistics had the best size and power properties among the five panel cointegration test statistics assessed. Also, Banerjee and Carrion-i Silvestre (2017) designed a panel cointegration test statistic that takes cross-sectional dependence into account. It demonstrated that consistent estimation of the long-run average parameter is possible when cross-sectional dependency is controlled for using cross-section averages as described in Pesaran et al. (2013).

Kao et al. (1999) is one of the earliest works on panel cointegration. The study provided two approaches for panel cointegration. The first part involves spurious regressions in panel data where the asymptotic properties of the least-squares dummy variable (LSDV) estimator and other conventional statistics are investigated. The LSDV estimator for the time series model differs from the spurious regression because the null distribution of residual-based cointegration tests depends on the estimator. This was found to impact residual-based cointegration tests in panel data significantly. On the other hand, Kao (1999) established a pooled version of the panel Dickey-Fuller test, which is applied to the residuals of a panel data regression estimate. The tests' asymptotic distributions with long-run parameters are derived, and Monte Carlo experiments are performed to evaluate the finite sample properties of the proposed tests.

Similarly, Pedroni (2004) investigated residual-based tests for the null hypothesis of no cointegration for dynamic panels with diverse short-run dynamics and long-run slope coefficients across individual panel members. Individual heterogeneous fixed effects, trend terms, and other variables are also allowed in the tests, considering both the group mean and the pooled within-dimension tests. Their limiting distributions are calculated, which proves that they are valid. The study also presented Monte Carlo data to establish their limiting sample size and power performance and their usage in assessing purchasing power parity for the post–Bretton Woods period.

Furthermore, Westerlund (2005) proposed two new simple residual-based panel data tests for the null of no cointegration. The tests are straightforward because they do not require any correction for the data's temporal dependencies. They are, however, capable of accommodating individual-specific short-run dynamics, intercept, trend terms, and slope parameters. The tests' limiting distributions are derived and shown to be free of nuisance parameters. The Monte Carlo results presented in the research paper indicated that the asymptotic results are well supported even in very small samples. While the Westerlund (2005) tests are in the middle, requiring both parametric modelling and semi-parametric adjustments to account for cross-sectional dependency, Pedroni (2004) test is semi-parametric in terms of the data's temporal dependencies and uses kernel estimation to reduce the nuisance parameters.

Panel cointegration tests were also tried on real-life data sets and for the first time in literature by Narayan (2010) to investigate the Purchasing Power Parity (PPP) evidence for a group of six Asian countries, including Malaysia, Thailand, India, Pakistan, Sri Lanka, and the Philippines. Narayan (2010) used three panel cointegration tests which include Westerlund (2006) which allows for the incorporation of multiple structural breaks, Gregory and Hansen (1996) residual-based test and Pedroni (1999) test without structural breaks. The results provided weak evidence of cointegration between nominal exchange rates vis-a-vis the US dollar and relative prices but when the Lagrange multiplier panel structural break cointegration test was used, strong evidence of panel cointegration, indicating PPP was found.

In the same vein, Pala (2020) empirically explores the relationship between energy use and economic development for a panel of G20 countries. Panel cointegration and vector error correction models were used, and a long-run equilibrium relationship was established. A Fully Modified OLS (FMOLS) and Dynamic OLS (DOLS) are used to assess the long-run relationship. The Panel Granger causality and Vector Error Correction Model (VECM) demonstrated that energy consumption and GDP have a bidirectional link. It shows that the "feedback hypothesis" holds for G20 countries.

Dithmer and Abdulai (2020) further investigated the effect of trade openness on child health using a panel data analysis of 66 countries over the period of 1960 to 2013. To account for the time-series properties of the data and potential cross-country heterogeneity in the impact of trade openness, the study used heterogeneous panel cointegration techniques that are resilient to omitted variables and endogeneity issues. Furthermore, the study showed that trade has a significant long-term potential to lower child mortality rates by decreasing trade openness and child health.

Other related studies on the application of panel cointegration to real-life datasets for various specific countries include; Cuestas and Harrison (2010) that provided information on the patterns of inflation in a panel of Central and Eastern European nations, Zhang (2011) worked on the relationship among inflation persistence, inflation expectations, and monetary policy in China and showed that the structural change is mostly due to better monetary policy conduct and the resulting better anchored inflation expectations using a counterfactual simulation method. Gerlach and Tillmann (2012) targeted inflation in Asia using a median unbiased estimator and bootstrapped confidence bounds by assessing the sum of the coefficients in an autoregressive model. Canarella and Miller (2017) also researched on inflation persistency and structural breaks in targeted 13 OECD countries and the USA. Lovcha and Perez-Laborda (2020) worked on monetary policy and the persistence of inflation in the US by modelling their relationship.

2.4 Fractional cointegration

Standard cointegration has received significantly more attention. As a result, there has been a need to focus heavily on fractional cointegration in recent years due to the paucity of research, particularly for panel data as opposed to time series data. Unlike standard cointegration, that only allows integer values for the memory parameters, the fractional cointegration framework is more general because it enables fractional values for the memory parameter and any positive real number for $d-\gamma$. As indicated by the work of Chen and Hurvich (2006); Marinucci and Robinson (2001); Robinson (2008), fractional cointegration has gotten a lot of attention recently. All of these publications assume that the observed series is either bivariate or that the cointegrating rank is one for time series data.

According to a search on the Science Direct database on 28/02/2022, research on fractional cointegration began around 1998. Hence, for all the cases reviewed in section 2.2 and 2.3, the cointegrated values of the processes are integer-valued and that is why it is standard cointegration. However, we have fractional cointegration when the processes are non-integer value cointegrated. Since fractional cointegration has the same economic ramifications as integer-valued cointegration in the sense that the variables have long-run equilibrium, it is theoretically possible to permit errors with a fractional integration order in a broad setting. This is a crucial generalization. The sole difference is that the convergence rate to equilibrium in the fractional is slower than in the standard cointegration.

One of the earliest works on fractional cointegration that has formed the body of most literature includes the work of Dueker and Startz (1998) which presented a cointegration testing approach based on joint estimations of a cointegrating vector that has a fractional integration order and its initial series. For the case with known fractional cointgeration order, Berg and Lyhagen (1998) presented a method in this case in terms of determining the approximate distribution of the trace test using an Error Correction Model (ECM) that allows fractional order. The assumption that the order of cointegration is known is quite restrictive, and in the case of misspecification, it may have unintended consequences on the test's power, necessitating modification.

Andersson and Gredenhoff (1999) presented the a likelihood ratio test for testing the null of no cointgeration for which the alternatives is fractional. The simulated power of the test showed that a fractional order is indeed present. The usual ML technique for fractional cointegrated systems, on the other hand, produces a significant bias and substantial mean square errors for the n impact matrix. As a result, ignoring fractional cointegration is far more serious than including it when it is not present. Then, Marinucci and Robinson (2001) studied the narrowband frequency domain and least-squares estimate of cointegrating vectors in regression models.

The research of Davidson (2002) concentrated on approaches for employing a parametric bootstrap to evaluate the existence of cointegrating correlations. Breitung and Hassler (2002) proposed a variant of the efficient score test for determining the cointegration rank of fractionally integrated series with fractionally integrated error correction terms. Velasco (2003) presented a consistent semiparametric approach for estimating the memory parameters for a fractionally cointegrated nonstationary series. Marmol and Velasco (2004) suggested Wald statistics for OLS coefficients in testing the null hypothesis of no cointegration in series with unknown fractional order. In a similar study by Gil-Alana and Hualde (2009), an investigation of a two-stage testing approach was presented for fractional cointegration in time-series macroeconomic data. The approach utilized the efficient testing procedure implemented in Robinson (1994).

In the recent time, Nielsen and Frederiksen (2011) worked on estimating a cointegrating relationship of less than half (difference in memory parameters) which is a case of a weak fractional cointegration model just like the stationary fractional cointegrating model by introducing a completely modified Narrow Band Least Square (NBLS) estimator that eliminates bias, and has a faster convergence rate than generic NBLS. Furthermore, it was shown that local Whittle integration error order estimates could be carried out consistently based on NBLS residuals, but the estimator has the same asymptotic distribution as if the errors were only seen under the nonstationary condition. The development of asymptotic distribution theory of Nielsen and Frederiksen (2011) is based upon an alternative representation of spectral density, which is applicable for multivariate fractional integrated processes, compared to other previous research, with the result that lower asymptotic distortions and variations in narrowband estimators are applied. Simulation evidence and several empirical cases were used to demonstrate the practicality and applicability.

However, prior to this asymptotic distribution theory, Nielsen and Shimotsu (2007) studied a stationary, fractionally cointegrated model by proposing the local Whittle quasi maximum likelihood estimator which is a semiparametric that estimates jointly the integration orders, the integration order of the errors and the cointegration vector such that it uses local assumptions on the joint spectral density matrix and errors close to null frequency. Under mild regularity conditions, the proposed estimator for the additional local orthogonality between regressors and cointegration errors was found to be consistent based on evidence from a Monte-Carlo simulation. In the same vein, Hualde and Velasco (2008) used a GLS-type estimator, similar to Robinson (2008) who developed a Wald statistic that is chi-squared distributed under the null hypothesis of no cointegration. Furthermore in $I(1)/I(0)$ cointegration setting, the possibility of fractional parameter are ignored posing an error that often leads to an incorrectly specified likelihood function, potentially implying a significant loss of power for fractional cointegration tests. In many situations, the existing classical analysis of cointegration only considers equilibrium deviations integration of order 0, which often results in a significant loss of power in the fractional case. Hence, for situations where fractional cointegration is allowed under the alternative, Łasak (2010) considered two likelihood ratio tests for the null hypothesis of no cointegration. The test generalizes the maximum eigenvalue and trace tests under the fractional alternative. The tests' power and size observed in the asymptotic distribution's finite sample revealed that the test is optimal when the order of integration is known.

Other related works on fractional cointegration include Nielsen and Frederiksen (2011) model-based inference of a fractionally cointegrated vector autoregressive Gaussian likelihood conditional on initial values with restriction on the parameters the process X_t is of fractional order d and cointegration of order $d-b$. This implies there exist vectors for which X_t is of fractional order $d-b$, and no other fractionality order is possible. The main contributions are the demonstration of the MLE weak convergence of the conditional likelihood as a continuous stochastic process in the parameters when errors are independent and identically distributed with appropriate moment conditions and constrained initial values.

Bayer and Hanck (2013) proposed comprehensive tests that put together information from different cointegration tests. The test employs the distribution of aggregators of the underlying tests it adopts, runs multiple tests, and draws inferences from the most rejecting test with the power properties established using asymptotic and Monte Carlo results. The tests' practicality is confirmed by their application to a broad and current set of research, producing an unambiguous test conclusion in circumstances of conflicting individual tests. Similarly, Boubaker et al. (2017) proposed a new stochastic long memory model that evolves non-linearly according to a Logistic Smooth Transition Autoregressive (LSTAR) specification and contains a time-varying fractional integration parameter. To calculate the fractional integration parameter that varies over time, Gil-Alana et al. (2018) investigated the persistence of income inequality and its major determinants in 26 OECD countries using fractional integration and selected GDP per capita, inflation, and employment as major macroeconomic determinants of income inequality. It finds that income inequality is highly persistent in all the countries examined and that there is a significant long-run equilibrium relationship between GDP growth and income inequality.

2.5 Fractional Cointegration in panel data and other related works

Several works have been done on panel data analysis and cointegration separately or both are considered together i.e., panel cointegration with various residual based tests and spectral based tests derived to solve peculiar problems with different properties like unit roots, cross sectional dependence, heterogeneity and the likes. In fact, Westerlund (2007) reported that in recent years, a great deal of attention had been paid to the problem of unit roots in panel data and the eventual occurrence of cointegration relationships between these variables. Hence, there is a need to look into fractional cointegration in panel data parlance. Existing research focused on fractional cointegration in time series or panel cointegration but not on fractional panel cointegration. Therefore, recent research on fractional cointegration related to this work will be reviewed in this section.