Bachelor's Thesis

Statistical models for the evaluation of factors affecting CO_2 emission

Department of Statistics
Ludwig-Maximilians-Universität München

Author: Weiyun Hong
Supervisor: Prof. Dr. Helmut Küchenhoff
Submission date: July 17th, 2023

Abstract

This thesis delves into the analytical frameworks and statistical models that have been employed in the studies investigating the factors influencing CO_2 emissions. The electronic databases Web of Science, SCOPUS, and ResearchGate were used to identify citation records to discuss this topic. Subsequently, 30 journal articles met the selection criteria and were retained for discussion. A detailed literature review was conducted to understand the multifaceted factors influencing CO_2 emissions, with particular emphasis on socio-economic, energy-related and policy-related aspects. Varied modelling approaches were identified in these studies, but the Autoregressive Distributive Lagged Model, Difference-in-Differences Model and Panel Quantile Regression Model were the main models used by authors. By investigating various statistical methods used in the selected studies, a well-rounded understanding of how determinants of CO_2 emissions are identified and quantified is provided.

Contents

List of Abbreviations

ADF	Augmented Dickey–Fuller Test
AR(1)	Autoregressive Model of Order 1
ARDL	Autoregressive Distributive Lagged
DID	Difference-in-Differences Model
FEM	Fixed Effects Model
FGLS	Feasible Generalized Least Squares
GMM	Generalized Method of Moments
IPS	Im-Pesaran-Shin Test
IV	Instrumental Variable
MMQR	Method of Moment Quantile-Regression
OLS	Ordinary Least Squares
PCSE	Panel-corrected Standard Errors
PQR	Panel Quantile Regression
REM	Random Effects Model
STIRPAT	Stochastic impacts by regression on population, affluence, and technology
VAR	Vector Autoregressive Model
WLS	Weighted Least Squares
2FE	Two-way Fixed Effects Model

1 Introduction

Global climate change, driven mainly by anthropogenic carbon dioxide (CO_2) emissions, has become a critical issue nowadays. As CO_2 continues to accumulate in the atmosphere, causing substantial shifts in global weather patterns, rising sea levels, and increased frequency of extreme weather events. To combat these terrible consequences, countries worldwide are taking initiatives to reduce CO_2 emissions, including measures like promoting renewable energy, fostering technology innovation and implementing policy regulations. Therefore, it is imperative to understand the various factors contributing to CO_2 emissions to mitigate their impacts effectively.

Researchers worldwide have dedicated their efforts to identifying and analyzing main driving factors of CO_2 emissions, many of which have relied on the Environmental Kuznets Curve (EKC) as a theoretical framework. The EKC proposes an inverted-U relationship between environmental degradation (in this case, CO_2 emissions) and per capita income, suggesting that economic development initially leads to increased CO_2 emissions, but as a certain income threshold is attained, further development mitigates the emissions. While the EKC offers valuable perspectives into the relationship between income and emissions, it doesn't encompass all of the crucial factors influencing CO_2 emissions. Consequently, recent empirical studies have extended this framework involving socio-economic, demographic, and energy-related factors, among others.

Several studies have incorporated variables such as population size, energy consumption patterns and technological advancements, into their research frameworks based on various research contexts. These factors are multifaceted and complex, making their relationships with CO_2 emissions challenging to quantify. To handle this complexity, various statistical models have been employed in the studies, providing crucial insights into the relationship between variables in the system.

In this thesis, it is aimed to discuss and analyze the statistical models applied in the existing research evaluating the factors affecting CO_2 emissions. I'll explore the strengths and limitations of each method, along with their unique features and applicability. In addition, the findings drawn from these methods will be presented, contributing to a comprehensive understanding of the determinants of CO_2 emissions.

The outline of this thesis is as follows: Section 2 presents a literature review relating to the focus and statistical models. Section 3 describes the main statistical methods employed to evaluate the factors influencing CO_2 emissions, illustrating their application in selected studies. Section 4 explores alternative estimation methods for analyzing the influencing factors. Section 5 discusses the statistical methods described above. A summary and outlook in section 6 conclude the thesis.

2 Literature Review

2.1 Criteria for Selection of Papers

This thesis aims to provide the most relevant insights into the factors affecting CO_2 emissions and the statistical models employed in this research domain. The papers are sourced primarily from electronic databases Web of Science, SCOPUS, and ResearchGate, using the search terms: "CO_2", "Carbon emissions" combined with "statistical analysis", "modelling", or "statistical model" to ensure relevance to the research domain. Throughout the process of literature selection, well-defined exclusion criteria were followed.

Firstly, only studies employing statistical methods to assess CO_2 emissions were selected, priority was given to studies published after 2020, as they provide the most recent work and up-to-date insights into this evolving field. It is essential to reflect the current state of research and the latest methods used in the study. I also limited the search to indexed journal articles written in English to maintain consistency in interpretation.

Besides, the geographical focus is considered, with priority given to studies conducted in regions with high CO_2 emissions, including China, the US, Europe, etc. These regions, known as major CO_2 emitters, provide the most critical case studies for understanding the determinants of CO_2 emissions. The selected literature for this thesis mainly includes research conducted in these key regions.

It is also important to note that CO_2 emissions are a crucial component of carbon emissions. Carbon emissions can also include the emissions of other greenhouse gases such as methane, fluorocarbons, and nitrogen oxides, among others. However, due to the prevalence and importance of CO_2 as a greenhouse gas, discussions on carbon emissions often specifically focus on CO_2 emissions. Consequently, studies analyzing carbon emissions are also included. On the other hand, papers examining the factors influencing carbon efficiency are not included in this thesis, as each paper might calculate the efficiency differently, the varying methods make it challenging to ensure consistency and could complicate the analysis.

By adhering to these criteria, this thesis identifies 30 relevant journal articles that provide a current investigation of the factors contributing to CO_2 emissions, as well as the statistical methods employed in their research.

2.2 Summary of the Reviewed Papers

The 30 journal articles selected in this thesis can be found in Table 1, the key information and research content of each paper is clearly shown in this table.

Authors, Year	Regions(s)	Time Span	Focus	Main Methods	Main Findings
Azimi and Bian, 2023	China	2006-2017	Carbon neutraility policy	Panel-corrected standard errors, Feasible generalized least squares	Carbon neutrality policy has a significant negative impact on per capita CO_2 emissions. Improving energy efficiency and renewable energy power generation can reduce CO_2 emissions.
S. Wen and H. Liu, 2022	China	2003–2017	Low-carbon city pilot policy	Time-varying Difference-in-Differences model	The implementation of Low-carbon city pilot policy led to a significant decrease in energy intensity and carbon emission intensity of the pilot cities.
X. Liu et al., 2022	China	2005-2016	Low-carbon city pilot policy	Time-varying Difference-in-Differences model	The Low-carbon city pilot policy has effectively promoted the reduction of CO_2 emissions; Energy structure, industrial structure, and innovation level have a significant impact on the effect of low-carbon city pilot policies.
Q. Yirong, 2022	China, USA, India, Russia, and Japan	1990-2019	Environmental policy stringency	Non-linear panel ARDL model	An increase in environmental policy stringency improves the environmental quality by reducing CO_2 emissions in the long run.
L.A.Attílio et al., 2023	U.S., U.K., Japan, and the Eurozone	1990-2018	Monetary policy	Global Vector Auto-regressive, Variance decompositions	Monetary policy does not seem to reduce short-run emissions in the U.K., or long-run emissions in the Eurozone, but it has a strong effect in Japan.
A. Xu, W. Wang and Y. Zhu, 2023	China	2009-2018	Smart city pilot policy	Time-varying Difference-in-Differences model, Mediating effect model	Smart city pilot policy can curb CO_2 emissions by strengthening the intensity of environmental regulation and promoting green technological innovation. Mediating effect of resource allocation efficiency only works in low-carbon industries, not high-carbon industries.

Table 1: Summary of selected studies (continued on next page)

Authors, Year	Regions(s)	Time Span	Focus	Main Methods	Main Findings
Y. KEHO, 2020	45 countries	1980-2011	Socio-economic factors	Mean-based regression methods, Panel quantile regression	Energy consumption and financial development increase CO_2 emissions, Industrialization increases CO_2 emissions especially in countries with higher level of CO_2.
A.O. Acheampong et al. ,2020	83 countries	1980–2015	Financial development	Instrumental variable Generalized Method of Moment	For standalone financial economies, the overall financial market development and its sub-indicators have no direct impact on carbon emissions.
O.K. Essandoh et al., 2020	52 countries	1991- 2014	International trade and foreign direct investment	Panel pooled mean group-ARDL model, Granger causality	For developed countries, there exists a negative long-run relationship between CO_2 emissions and trade. For developing countries, foreign direct investment inflows exhibit a positive long-run relationship with CO_2 emissions.
M. K. Anser et al., 2020	8 countries in South Asian	1994-2013	Human and economic factors	Fixed effect regression model	Population growth has increased energy use and contributed to carbon emissions.
A. Awan, K.R. Abbasi, S. Rej et al., 2022	10 emerging countries	1996-2015	Foreign direct investment	MMQR, Fixed effects-OLS model with robust Driscoll-Kraay standard errors	Foreign direct investment's effect on CO_2 emissions is significant and positive at 0.05th-0.50th quantiles, it becomes insignificant at higher quantile levels. Urbanization enhances while renewable energy mitigates CO_2 emissions at all quantile levels.
M. Shahbaz et al., 2020	China	1984-2018	Investment in energy sector	Bootstrapping ARDL	Public-private partnerships investment in energy increased CO_2 emissions. Exports are positively linked with CO_2 emissions.
D. Xu et al., 2022	G7-Countries	1986-2019	Financial development, Renewable Energy	Non-linear ARDL, Two-stage least squares	An increase in Finance development can decrease carbon emissions. Renewable energy exerts a negative and significant effect on CO_2 emission.

Table 1: Summary of selected studies (continued on next page)

Authors, Year	Regions(s)	Time Span	Focus	Main Methods	Main Findings
G.N. Ike et al., 2020	15 oil-producing countries	1980–2010	Fossil fuel energy production	MMQR with fixed effects, Modified Ordinary Least Squares	Electricity production increase CO_2 emissions while trade condenses CO_2 emissions across all the quantiles. Impacts of oil production and electricity production on CO_2 emissions are positive.
Q. Jing et al.,	China	2010–2019	Public Transportation development	Two-way fixed-effect model, Mediating-effect model	Public transport development level and CO_2 emissions are negatively correlated, showing an "Inverted U-shaped" curve relationship. Energy consumption is the transmission path of the carbon emissions reduction effect of public transport development level.
C.-W. Su et al., 2020	USA	1990-2017	International trade and Technological Innovation	Autoregressive Distributive Lagged	Exports and consumption-based carbon emissions are negatively associated, and technological innovation helps reducing the adverse effect of CO_2 growth.
R. Wu et al., 2021	18 developed countries	2005-2016	Renewable Energy	Two-way fixed effects models, Granger causality	The transition from fossil fuels to renewable energy and variations in energy intensity and fossil CO_2 intensity were the primary factors contributed to the decrease in CO_2 emissions.
A. Azam et al., 2021	10 leading CO_2 emitter countries	2000-2016	Renewable Energy, Technology Innovation	Granger causality, Pooled regression, Fixed effect model	Renewable energy, Information and communication technologies-trade contribute to eliminating CO2 emissions. Innovation and green energy eliminate the CO2 emissions.
M.M. Rahman, K. Alam and E. Velayutham, 2022	22 developed countries	1990-2018	Renewable energy, Export quality	Non-linear ARDL, Pooled mean group estimation	Renewable energy and export quality are found as contributory factors for the reduction of CO2 emissions. Bidirectional causality is found between renewable energy and CO2 emissions.

Table 1: Summary of selected studies (continued on next page)

Authors, Year	Regions(s)	Time Span	Focus	Main Methods	Main Findings
I. Khan et al., 2021	USA	1971-2016	Energy consumption, Population growth	Generalized Method of Moments, Generalized linear model, Robust least-squares, Granger causality	Bidirectional causality runs between natural resources and CO_2 emissions. Negative relationships between renewable energy and CO2 emissions. Non-renewable energy consumption, population growth have a positive relationship with CO_2 emissions.
Z. Wang, Y. Zhu, 2020	China	2001-2017	Renewable Energy	Panel ordinary least squares model	Renewable energy technology innovation could control carbon emissions, while fossil energy technology innovation might boost carbon emissions.
C.W. Su, F. Liu, P. Stefea et al., 2023	USA	2010-2021	Technological Innovation	Rolling-window causality test, Granger causality	Technological innovation has positive and negative influences in mitigating greenhouse gas emissions. The negative impact confirms that Technological innovation is a dominant factor in mitigating carbon emissions.
C. Cheng et al., 2021	35 OECD countries	1996-2015	Technological Innovation	Panel quantile regression with non-additive fixed effects	Technological innovation directly reduces CO_2 emissions through Research and Development investment and education expenditure; Technological innovation can offset the positive impacts of economic growth on CO_2 emissions by reducing energy intensity.
B. Wang et al., 2022	China	2007-2017	High-quality energy development	System Generalized Method of Moments	The overall effect of High-quality energy development on CO_2 emissions is negative, it not only curbs CO_2 emissions directly, but also mitigates CO_2 emissions indirectly by reducing total energy consumption.

Table 1: Summary of selected studies (continued on next page)

Authors, Year	Regions(s)	Time Span	Focus	Main Methods	Main Findings
Y. Chen, C.-C. Lee, 2020	96 countries	1996-2018	Technological Innovation	Generalized Nesting Spatial model, Two-way fixed effect model	Technological innovation has no significant mitigation effect on CO_2 emissions globally, but Technological innovation in high-income, high-technology, and high-CO_2 countries can significantly reduce CO_2 in neighboring countries
H. Jeon, 2022	48 U.S. states	1997–2017	Renewable Energy	Panel fixed-effects, Two-step Generalised Method of Moments, MMQR with fixed effects	Renewable energy consumption, electricity prices, and primary energy prices have negative impact on CO_2 emissions whereas Heating Degree Days have a positive impact on CO_2 emissions.
C.K. Lau et al., 2023	36 OECD countries	1970-2021	Green quality of energy mix	System Generalized Method of Moments, Feasible generalized least squares	Per capita income, institutional quality, and technology increase CO_2 emissions. The Green quality of energy mix is negatively related to the level of CO_2 emissions.
Chen et al., 2022	China	2011-2017	Artificial Intelligence	Two-way fixed effects model	Artificial Intelligence has a significant inhibitory effect on carbon emissions; Artificial Intelligence reduces carbon emissions through optimizing industrial structure, enhancing information infrastructure.
J. Liu et al., 2022	China	2005-2016	Artificial Intelligence	Feasible generalized least squares	Artificial Intelligence significantly reduces carbon emissions.
Wang et al., 2023	China	2008-2019	Industrial Robots	Fixed-effects model, Moderating effect model	Although using industrial robots reduces carbon emissions, it also leads to an energy rebound effect.

Table 1: Summary of selected studies (continued)

In general, retained studies involve multiple developed and developing countries as well as global-level analysis. A survey of the literature reveals a shared focus across most studies on socio-economic, demographic, and energy-related aspects. These appear to be the predominant factors of interest in current research regarding CO_2 emissions. Besides these critical research areas, it is worth noting that many studies also explore the role of policy-related factors in determining carbon emissions.

Economic development and rapid urbanization are critical processes boosting fossil fuel consumption, and consequently, CO_2 emissions worldwide (Debone *et al.*, 2021). Therefore, it is essential to account for the influence of socio-economic factors, including international trade, foreign direct investment, and population size, when investigating the determinants of carbon emissions. At the same time, with the rapid advancement of technology, the impacts of emerging technologies such as renewable energy development and artificial intelligence have been incorporated by scholars into the analytical framework for CO_2 emissions, integrating sustainable manufacturing into environmental governance.

2.3 Comments and Evaluation of Strategies

2.3.1 Strategy Overview

Current research investigating factors influencing CO_2 emissions mainly employs panel data analysis. This allows the consideration of data that includes both cross-sectional and time-series aspects, providing a more detailed regional examination of their relationships. Different statistical methods were found in the selected papers. Granger causality, Panel quantile regression models, Two-way fixed-effect model, Autoregressive Distributed Lag (ARDL) model, and Difference-in-differences model are the main modelling approaches used by the authors. The most common estimation approaches used by researchers include Feasible Generalized Least Squares (FGLS), Generalized Method of Moments (GMM) and Method of Moment Quantile-Regression (MMQR).

Meanwhile, a majority of the investigated studies employ the Stochastic Impacts by Regression on Population, Affluence and Technology (STIRPAT) model. This environmental impact assessment model serves as an analytical framework for their investigations. The STIRPAT model can comprehensively consider the impact of social, economic, and technological driving factors on CO_2 emissions (Su and Lee, 2020). Hence, it has been widely used in the research of influencing factors of CO_2 emissions. This detailed model will be further reviewed in section 3.

In the review of the studies, it is observed that the ARDL is the most frequently employed modelling approach. The ARDL, NARDL (Nonlinear Autoregressive Distributed Lag), and BARDL (Bootstrap Autoregressive Distributed Lag) models are all variants of the Autoregressive Distributed Lag model, which is widely used in statistical time series analysis.

ARDL is a general specification taking into account the lag structure. (Ghouse and Khan, 2018). Pesaran and Shin (1995) revealed that asymptotically valid inference on short-run and long-run parameters could be made by employing ordinary least square estimations of the ARDL model. The ARDL model can handle variables that are stationary, nonstationary, or a mixture of both, making the model particularly suitable for modelling complex economic and social phenomena in which the effects may not be immediate but rather distributed over time.

A NARDL model is an extension of the ARDL model that allows for asymmetric effects of changes in the explanatory variables. This means that increases and decreases in the explanatory variables might affect the dependent variable differently. Jareño *et al.* (2020) state that the NARDL approach has some advantages over other estimation techniques. The NARDL approach is not inclined to omit lag bias, and is suitable irrespective of the stationary properties of the variables. This approach simultaneously produces estimates of short and long-run nonlinearities through the positive and negative partial sum of the decomposition of the regression.

Rahman *et al.* (2022) explored the asymmetrical influence of changes in technological innovation, specifically research and development (R&D), on carbon emissions. They employed the NARDL model to examine the asymmetrical linkage between the selected variables in the studied countries. The study indicates that negative shocks to technological innovation (declines in R&D expenditure) have a significant and positive long-term effect on carbon emissions. However, the influence of positive technological innovation shocks (increases in R&D expenditure) presents an insignificantly negative association with carbon emissions.

The BARDL model is another extension of the ARDL model. This method employs bootstrap resampling, a statistical technique that can be used to estimate the sampling distribution of an estimator by resampling with a replacement from the original sample. The bootstrap approach can improve the accuracy of hypothesis tests and confidence intervals in small samples. The BARDL bounds testing approach works better than the traditional ARDL model, as this approach considers the joint F-test on all lagged level variables, t-test on the lagged level of dependent variable and the regressors, which will help with respect to cointegration equilibrium between the sample variables (McNown *et al.*, 2018). This methodology has a main advantage in addressing the problem of weak size and power properties encountered in the traditional ARDL method.

Shahbaz *et al.* (2020) applied the BARDL cointegration technique and identified the presence of five cointegrating vectors between carbon emissions and their determinants, suggesting a long-run cointegration relationship between these variables. The presence of cointegration between the variables led researchers to examine the long-run and short-run effects of influencing factors on carbon emissions. They found that public-private partnerships investment in energy sector has a positive and significant effect on carbon emissions, and the relationship between technological innovations and carbon emissions is negative and significant.

In summary, based on the ARDL model, these models are used to examine the relationship between variables over time, and differ in their specific capabilities and applications. The ARDL model is the most basic, the NARDL model includes the capacity to handle asymmetric effects, and the BARDL model uses bootstrap methods to improve accuracy in small samples.

Pedroni's Fully Modified Ordinary Least Squares (FMOLS) model, the Dynamic Ordinary Least Squares (DOLS) estimator, and the Fixed Effects Ordinary Least Squares (FE-OLS) technique are the prominent estimation methods used in the selected studies. For example, Yirong (2022) found that an increase in environmental policy stringency reduces CO_2 emissions in the long run with the use of panel ARDL and NARDL methods, and the results were confirmed using FMOLS and DOLS.

The central reasons for concern in estimating dynamic cointegrated panels as highlighted by Pedroni (2004) are heterogeneity issues with differences in means between cross-sections and differences in cross-sectional adjustment to the cointegrating equilibrium. Pedroni's Fully Modified Ordinary Least Squares (FMOLS) model includes individual-specific intercepts and allows for heterogeneous serial correlation properties of the error processes across individual members of the panel, and thus, deals with these issues accordingly.

The DOLS estimator is another extension of OLS to the panel data setting based on the results of Monte Carlo simulations and was found to be unbiased compared to both OLS and the FMOLS estimators in finite samples. The DOLS estimator also accounts for possible endogeneity of the regressors through the augmentation of lead and lagged differences to suppress the endogenous feedback (Kao and Chiang, 2001). In addition, the FE-OLS technique is augmented with Driscoll and Kraay standard errors, which are robust to general forms of cross-sectional dependence and autocorrelation up to a certain lag. Apart from the above estimators, other estimation techniques employed in the selected studies will be discussed in detail in Section 4.

2.3.2 Diagnostic Test

Before conducting regression analysis, researchers perform a series of tests on the data to ensure reliable results. An essential step in this process, particularly when dealing with time series analysis, is the stationarity test. The presence of a stochastic trend in a variable set can give rise to misleading and spurious regression results. As such, it is important to validate the stationarity of a series through unit root tests. The augmented Dicky-Fuller (ADF) test, the principle of which is to verify the existence of unit roots in time series data, is one of the most generally employed unit root tests.

If the result shows that the time series has a unit root, then the data needs to be differentiated to make it stationary, avoiding the pseudo-regression of the variables. When all data is non-stationary at the same level and tends to stabilize after the first difference, a cointegration analysis will be applied to verify that the time series shows long-term cointegration. The cointegration tests, including Pedroni, IPS and Kao test, are performed to show long-term connections between non-stationary sequences. These tests will be described in detail in section 3.1.

Cross-sectional Dependence (CD) tests are used to confirm cross-sectional dependence within the panel data. Cross-sectional panel dependence may arise when, for instance, countries respond to common shocks or spatial diffusion processes are present (such as shocks from the financial crisis). If cross-sectional dependence is present, these results, at least, are in the inefficiency of the estimators and invalid inference when using standard estimation techniques (Ali, 2017). Investigating the presence of cross-sectional dependency is crucial in panel data sets, as regions being studied may share similar economic and institutional attributes that can cause cross-sectional dependency.

Panel data models have two main structures: Fixed Effects Models (FEM) and Random Effects Models (REM). The Hausman test (Hausman, 1978) is commonly used in the investigated studies to select the most appropriate model for the data structure. The null hypothesis is that the preferred model is REM, while the alternative is that the preferred model is FEM.

The recurring preference for the FEM (Wang *et al.*, 2023; Azam *et al.*, 2021 etc.), as supported by the Hausman test, can be attributed to its ability to control for time-invariant unobserved individual characteristics. This trait of the FEM is especially beneficial in the context of panel studies, where unobserved region-specific characteristics can significantly influence CO_2 emissions. By accounting for such unobserved heterogeneity, the FEM can provide more accurate and reliable results than the REM, assuming that individual effects are uncorrelated with the explanatory variables.

In order to ensure the robustness of their results, researchers often employ various robustness checks and methods. Robustness checks are typically used to assess the sensitivity and robustness of a model to outliers, model assumption violations, and other confounders that may occur in the data.

A common approach to check the robustness in the selected papers is Placebo test, which is essential to determine if the observed effect in the primary analysis is not due to chance or spurious correlation. To conduct the process, the primary analysis must be reproduced by substituting the key variable of interest with another variable known as the "placebo". This placebo variable should be irrelevant or unrelated to the model and is expected to have no impact on the outcome. It could be a variable from a control group that was not exposed to the treatment, or a variable that represents a period before the actual treatment occurred. For instance, in a study (Xu *et al.*, 2023) assessing the carbon reduction effect of Smart City Pilot Programs (SCPP), the authors identified "false smart cities" randomly as the treatment group and set the policy implementation time at random. Then they estimated the coefficients using these new treatment and control groups, effectively completing a Placebo test. This process was repeated 500 times to establish a distribution of estimated coefficients, which confirmed the reliability of the baseline results. The results of the placebo test reveal that, the coefficient estimates show an approximate normal distribution trend with a mean value close to 0, which is significantly different from the coefficient of SCPP in the baseline regression. The above analysis fully illustrates that the negative effect of the SCPP on the CO2 emissions of industrial firms is robust (Xu *et al.*, 2023).

Endogeneity tests are employed to assess whether an explanatory variable in a regression model is correlated with the error term, thus violating the fundamental assumption of regression analysis. Durbin-Wu-Hausman test is one common test for endogeneity. Durbin-Wu-Hausman test compares the results of an OLS regression with those of an instrumental variable regression, where the instrumental variable (IV) is correlated with the endogenous explanatory variable but not with the error term. The test suggests that endogeneity is present if the coefficients are significantly different (Greene and William, 2012).

Endogeneity can arise due to simultaneity (e.g., economic growth can lead to increased CO_2 emissions due to increased industrial activity, simultaneously, higher CO_2 emissions can impact economic growth due to health-related costs), omitted variables (e.g., if a study only includes economic growth and omits technology advances in the model, the coefficient on economic growth could be biased as it might also be capturing the effect of technological progress), or measurement error (Johnston, 1972). Apparently, endogeneity is often unavoidable in studies investigating factors impacting CO_2 emissions. For instance, Liu *et al.* (2022) recognized that the relationship between Artificial Intelligence (AI) application and CO_2 emissions could be mutually influencing, the authors considered AI as potentially endogenous. In response, they introduced a lagged version of AI as an IV in the empirical model, using the Two-Stage Least

Squares method to estimate parameters. The validity and strength of this IV were tested through several statistical measures. The Anderson Canonical Correlation Lagrange Multiplier statistic rejected the null hypothesis that the IV is not identifiable. The Cragg-Donald Wald F statistic rejected the weak IV hypothesis, indicating that the IV is strongly correlated with the endogenous predictor. The Anderson-Rubin Wald test further confirmed the correlation between the IV and AI (Liu *et al.*, 2022).

Moreover, alternative models for further robustness checks are applied in the selected papers. For instance, Su *et al.* (2020) adopted ARDL to analyze the relationship between carbon emissions and trade, GDP and technological innovation, and subsequently implemented dynamic ordinary least squares (DOLS) to check the robustness of the results. Wang *et al.* (2022) also conducted a robustness test by substituting the CO_2 emission variable with per capita CO_2 emissions and then reestimated the parameters applying the FGLS, and SYS-GMM techniques simultaneously.

These different robustness checks and methodologies not only provide a solid foundation for the research but also significantly enhance the validity of the results. Overall, the analyzed methods could be considered efficient in investigating the relationships between the studied variables.

3 Main Strategies

3.1 Time Series Methods

Dealing with a time series with unit roots is essential because many statistical techniques assume stationarity. Thus, the stationarity of a series must be tested through unit root tests before constructing the statistical model to avoid misleading regression results. Once it is determined that individual series are non-stationary, cointegration tests can be applied to investigate the long-term relationships among non-stationary series. After extensive empirical tests, the Granger causality tests further offer insights into the direction of influence among variables.

3.1.1 Unit Root Test

In the context of time series analysis, stationarity is a critical concept. A stationary time series has properties that do not depend on the time at which the series is observed. This implies that the mean, variance, and autocorrelation structures do not change over time. Stationarity can be divided into two types: strictly stationary and weakly stationary. A time series $\{y_t\}$ is weakly stationary if for all t:

$\mu(t) = E(y_t) = \mu$, and

$\gamma(t+h,t) = \mathrm{Cov}(y_{t+h}, y_t) = E[(y_{t+h} - \mu(t))(y_t - \mu(t))] = \gamma(h)$.

This means that the time series $\{y_t\}$ moves in a similar way as the "shifted" time series $\{y_{t+h}\}$ for all h. A process is called strictly stationary if the joint distribution of $\{y_1, \ldots, y_k\}$ is identical to that of $\{y_{1+t}, \ldots, y_{k+t}\}$ (Wohlrabe, 2023). Strict stationarity implies that the process always maintains an invariant distribution function. All strictly stationary processes satisfy the property of weak stationarity, but the converse is not valid. In practice, strict-sense stationarity is too restrictive for many applications. Therefore, the concept of weak stationarity is usually applied. In general, the term "stationarity" is used to refer to the weak form of stationarity.

A unit root test is used to determine whether a time series is non-stationary and possesses a unit root. A unit root implies that the series is non-stationary and exhibits a stochastic trend. Consider the AR(1) case without drift,

$$y_t = \rho y_{t-1} + \epsilon_t, \quad \epsilon_t \overset{iid}{\sim} N(0, \sigma^2), \quad \sigma^2 < \infty \tag{1}$$

to test for the presence of a unit root is to test the hypothesis

$$H_0 : \rho = 1 \quad \text{and} \quad H_0 : \rho < 1.$$

The test is formulated such that the presence of unit root is the null hypothesis. A major difficulty arises from the fact that, under the null hypothesis, the OLS estimate of ρ is not normally distributed, consequently, the t–statistic is not described by a t–distribution.

Dickey and Fuller (1979) investigated this problem after subtracting y_{t-1} from both sides of Eq.(1), yielding

$$\Delta y_t = (\rho - 1)y_{t-1} + \epsilon_t \tag{2}$$

so that the test is

$$H_0 : \rho - 1 = 0 \quad \text{and} \quad H_1 : \rho - 1 < 0$$

Dickey and Fuller used Monte-Carlo simulations to generate the critical values for the non-standard t-distribution arising under the null. This test procedure is called the Dickey-Fuller (DF) test. Dickey and Fuller considered a more general model

$$\Delta y_t = f(t) + (\rho - 1)y_{t-1} + \epsilon_t$$

where $f(t)$ is a deterministic function of time, it turns out that the critical value varies with the choice of $f(t)$. The deterministic trend specification assumes stationary deviations ϵ_t around a deterministic trend (t). They provide critical values for the three cases:

$$f(t) = \begin{cases} 0 & \text{Case 1: random walk} \\ c & \text{Case 2: random walk with drift} \\ c_0 + c_1 t & \text{Case 3: random walk with deterministic trend} \end{cases}$$

If, in Case 2, there is a unit root, i.e., $y_t = c + y_{t-1} + \epsilon_t$, the constant c acts like a linear trend. In each period, the level of y_t shifts, on average, by the amount c. A process of this nature is said to have a stochastic trend and to be difference stationary. In Case 3 with a unit root, i.e., $y_t = c_0 + c_1 t + y_{t-1} + \epsilon_t$, we refer to a deterministic trend; and the process becomes trend stationary after differencing.

A problem with the AR(1)-based unit-root test is that the $\hat{\epsilon}_t$ obtained from Eq.(2) tend to be autocorrelated. To circumvent this, one can add sufficiently many lagged Δy_{t-i}s on the right-hand side of Eq.(2)until the residuals appear to be white noise. If the unit-root test described above is based on the OLS-estimated coefficient of y_{t-1} in

$$\Delta y_t = f(t) + (\rho - 1)y_{t-1} + \sum_{i=1}^{p-1} \alpha_i \Delta y_{t-i} + \epsilon_t$$

this refers to the augmented Dickey-Fuller (ADF) test. By including lags of order p, the ADF formulation allows higher-order autoregressive processes (Imam, 2016). The unit root test is then carried out under the null hypothesis $\rho - 1 = 0$ against the alternative hypothesis of $\rho - 1 < 0$. The test statistic is $t = (\hat{\rho} - 1)/\hat{\sigma}_\rho$, where $\hat{\sigma}_\rho$ is the standard error of $\hat{\rho}$. The ADF test, which is the most widely used unit-root test, relies on the same critical values as the DF test (Wohlrabe, 2023). If the calculated test statistic is significantly less than the critical value, then the null hypothesis is rejected, and no unit root is present.

While the ADF test is robust and commonly used, it is important to note that it primarily analyzes single time series data. When handling panel data, the panel unit root tests could be more suitable, such as the Im-Pesaran-Shin (IPS) test proposed by Im, Pesaran and Shin (2003).

The IPS test evaluates the t-test for unit roots in heterogeneous panels and has the added advantage of allowing for cross-sectional dependence. Consider a sample of N cross sections (regions or countries) observed over T time periods. We suppose that the stochastic process $y_{i,t}$, is generated by the first-order autoregressive process:

$$y_{i,t} = \rho_i y_{i,t-1} + \gamma z_{i,t} + \epsilon_{i,t}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T$$

also can be expressed as: $\Delta y_{i,t} = \delta_i y_{i,t-1} + \alpha_i + \beta_i t + \epsilon_{i,t}$,
where $\delta_i = \rho_i - 1$, $z_{i,t}$ is the exogenous term that exists in a constant form with fixed effects and time trends, $\epsilon_{i,t}$ is the error term. And assume that there are n independent individuals in the panel data series.

First, an ADF statistic $t_i = \frac{\hat{\rho}_i - 1}{\hat{\sigma}_{\rho_i}}$ is constructed for each independent individual.

Then, t_i of these individuals are averaged to calculate $\bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i$.

Thus, deriving the ADF statistic for the panel data:

$$Z_{\bar{t}} = \frac{\bar{t} - E(\bar{t})}{\sqrt{\frac{Var(\bar{t})}{n}}} \xrightarrow{d} N(0, 1).$$

$$\mathrm{H}_0 : \delta_1 = \delta_2 = \cdots = \delta_N = 0$$
$$\mathrm{H}_1 : \delta_i < 0, i = 1, \ldots, N; \delta_j < 0, j = 1, \ldots, N$$

That is, we consider that under the alternative hypothesis, a part of the individual time series in the panel data is stationary, while the remaining part of the individual time series is non-stationary.

Overall, the IPS test proposes a standardized $\bar{t}$-bar test statistic based on the ADF statistics averaged across the groups. Under a general setting, this statistic is shown to converge in probability to a standard normal variate sequentially (Barbieri, 2005).

3.1.2 Cointegration Test

Non-stationary data, as is often the case in many socio-economic variables, can be transformed into a stationary sequence to permit further analysis. Integrated processes represent a class of non-stationary time series that can be made stationary through differencing. A time series is said to be integrated of order 'd' (denoted as I(d)) if it must be differenced 'd' times to become stationary. The number of differences required to achieve stationarity represents the order of integration (Harris and Sollis, 2003). For example, if a time series is non-stationary and one differencing step is required to make it stationary, we denote this series as I(1) - integrated of order one. In the multivariate case, there may exist a linear combination of non-stationary time series, which is stationary without taking differences. This property is called cointegration.

Cointegration theory holds that although some economic variables are non-stationary, their linear combination may be a stationary sequence. Engle and Granger (1987) were the first to formalize this concept: The components of a $(k \times 1)$ vector X_t are said to be cointegrated of order (d, b), denoted by $\mathrm{CI}(d, b)$, if each component individually taken is $\mathrm{I}(d)$, and a vector $\beta = (\beta_1, \ldots, \beta_k) \neq 0$ exists such that the linear combination $Z_t := \beta^\top X_t$ is $\mathrm{I}(d - b)$.

The linear combination is called a cointegration relation and the vector β is called a cointegration vector. Cointegration means that two or more time series are connected and form a long-term equilibrium represented by the linear combination $\beta^\top X_t$. Although the individual components may contain stochastic trends, they move closely together over time and show only short-term deviations from their equilibrium (Harris and Sollis, 2003).

The cointegration test is a statistical method used to determine the long-term equilibrium relationship between two or more non-stationary time series variables. It helps to identify whether there is a stable long-term relationship among the variables, despite them individually exhibiting non-stationary behavior. One popular cointegration test is the Engle-Granger two-step method. It is used to test the null hypothesis of no cointegration between two time series.

Assume that x_t and y_t are non-stationary time series and both are integrated of order one. The first step of the testing procedure is to estimate the long-run equilibrium relationship in terms of ordinary least squares (OLS) regression:

$$y_t = \alpha_0 + \alpha_1 x_t + u_t,$$

where α_1 is the coefficient of cointegration and u_t is an error term. Note that the OLS regression estimation results are reliable only when both time series are cointegrated. In this case, the OLS estimator of α_1 is considered super-consistent, meaning that it converges to the true parameter much faster than in the standard case with I(0) variables. If no cointegration is present, this technique leads to the problem of spurious regression and can produce misleading results. The estimated cointegrating regression yields the residual series $\hat{u}_t = y_t - \hat{\alpha}_0 - \hat{\alpha}_1 x_t$.

In the second step of the testing procedure, a unit root test is applied to the residuals $\hat{u}_t$ to determine whether they are stationary or not. For this purpose, an ADF test is usually performed on the following model:

$$\Delta\hat{u}_t = \psi^*\hat{u}_{t-1} + \sum_{j=1}^{p-1} \psi_i \Delta\hat{u}_{t-j} + \varepsilon_t$$

where ε_t is assumed to be white noise, with univariate unit root test

$$\mathrm{H}_0 : \psi^* = 1 \text{ and } \mathrm{H}_1 : -1 < \psi^* < 1.$$

If the null hypothesis of non-stationarity of the residuals $\hat{u}_t$ can be rejected, then they can be considered stationary, which in turn means that the investigated time series x_t and y_t are cointegrated.

Given the potential limitations of the Engle-Granger test, particularly when dealing with more complex scenarios involving more than two time series and multiple cointegration relationships, researchers often turn to panel cointegration tests such as the Kao panel cointegration test. Assuming a panel data model:

$$y_{i,t} = \alpha_i + \beta x_{i,t} + \varepsilon_{i,t}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T$$

Kao (1999) proposed a panel cointegration test generalized from the Augmented Dickey-Fuller test, and the test model is as follows:

$$\hat{\varepsilon}_{i,t} = \rho\hat{\varepsilon}_{i,t-1} - \sum_{j=1}^{p} \varphi\Delta\hat{\varepsilon}_{i,t-j} + v_{i,t}$$

Based on the Engle-Granger two-step method, if $\mathrm{H}_0 : \rho = 1$ can be rejected, there is a long-term relationship between the variables.

Currently, the Kao test is the most widely used panel cointegration test (Lin and Xu, 2018). Many investigated papers used the Kao test to examine whether there is a cointegration relationship between the variables in their studies.

3.1.3 Granger Causality

If the results of the cointegration test show that two time series are cointegrated, the Granger causality analysis (Granger, 1980) could be applied to investigate the causal relationship between these variables. Granger causality is based on linear regression and its time series must be stationary, otherwise, spurious regression problems may occur (He and Maekawa, 2001).

Granger causality test is a method used to test whether one variable X is the cause of a change in another variable. If variable X helps predict variable Y, the relationship is defined as Granger causality. This test can only be applied to the test of a time series data model with stationarity. The procedure for testing whether X is the Granger cause of a change in Y is as follows:

Let Y_t and X_t be stationary time series. To test the null hypothesis that X_t does not Granger-cause Y_t, first, find the appropriate lagged values Y_{t-i} to include in a univariate autoregression of Y:

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \epsilon_t \quad \text{(restricted regression)}$$

Next, the autoregression is augmented by including lagged values of X:

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \beta_i X_{t-i} + \epsilon_t \quad \text{(unrestricted regression)}$$

where α_0 denotes the constant term, p and q are the longest lag length of the variables Y_t and X_t, respectively, which can usually be assumed to be somewhat larger, and ϵ_t is the white noise. The residual sum of squares of the two regression models, RSS_U and RSS_r, were then used to construct the F-statistic:

$$F = \frac{\frac{RSS_r - RSS_U}{q}}{\frac{RSS_U}{n-p-q-1}} \sim F(q, n-p-q-1)$$

where n is the sample size. If $F > F_\alpha(q, n-p-q-1)$, then $\beta_1, \ldots, \beta_q$ are significantly different from zero, and the null hypothesis that X_t does not Granger-cause Y_t must be rejected. Otherwise, the null hypothesis is accepted if and only if no lagged values of X are retained in the regression.

Dumitrescu and Hurlin panel causality test is an extension of the Granger causality test that adapted for the context of panel data. The test accounts for cross-sectional independence and is applicable to large panel data situations (both N and T are large). The following model detects the causality in panel data:

$$Y_{i,t} = \alpha_i + \sum_{j=1}^{J} \lambda_i^J Y_{i,t-J} + \sum_{j=1}^{J} \beta_i^J X_{i,t-J} + \epsilon_{i,t}$$

where $X_{i,t-J}$ and $Y_{i,t-J}$ denote the observations of two stationary variables for the individual i in period t. j shows the lag length, λ_i^J represents the autoregressive parameter, while β_i^J is the regression coefficient that varies within the groups. It is assumed that the lag order J is the same for all individuals in a balanced panel. This test is a fixed type of test that yields a fixed coefficient model. It allows for heterogeneity and maintains a normal distribution.

The null hypothesis of no causal relationship and the alternative hypothesis for testing a causal relationship are as follows:

$$H_0 : \beta_i = 0, \forall i = 1, \dots, N$$
$$H_1 : \beta_i = 0, \forall i = 1, \dots, N_1;$$
$$\beta_i \neq 0, \forall i = N_1 + 1, N_1 + 2, \dots, N$$

where N_1 is unknown but satisfies the condition $0 \leq N_1/N < 1$. The ratio N_1/N is necessarily less than one since, if $N_1 = N$, there is no causality for any of the individuals in the panel, which is equivalent to the null hypothesis. Conversely, when $N_1 = 0$, there is causality for all the individuals in the panel. (Dumitrescu and Hurlin, 2012)

In general, the Granger causality is widely used to confirm the effect of each variable on the others, effectively exposing whether a uni-directional or bi-directional relationship exists among the variables in the studies.

3.1.4 Application in Selected Studies

The study conducted by Xu *et al.* (2022) analyzed the nonlinear effects of financial development (FD), renewable energy (REC), and human capital (HC) on CO_2 emissions across G7 countries, which include Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. The researchers initiated the study by examining the stationarity of all the variables using unit root tests.

Variables	Test	Canada	France	Germany	Italy	Japan	UK	USA
CO2	ADF	I(0)	I(1)	I(0)	I(1)	I(0)	I(0)	I(0)
	PP	I(0)	I(1)	I(0)	I(1)	I(0)	I(0)	I(0)
FD	ADF	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)
	PP	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)
REC	ADF	I(0)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)
	PP	I(0)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)
HC	ADF	I(0)	I(1)	I(0)	I(0)	I(1)	I(0)	I(1)
	PP	I(0)	I(1)	I(0)	I(0)	I(1)	I(0)	I(1)

Table 2: ADF and PP unit root test for time series data (Xu *et al.*, 2022)

Note: ADF and PP represent Augmented Dickey-Fuller, Phillips, and Perron. FD: Financial Development, REC: Renewable Energy, HC: Human Capital.

Table 2 demonstrates that CO_2, FD, REC, and HC are stationary at either level I(0) or the first difference I(1), and the results from the ADF and PP unit root tests are consistent with each other. I(0) means the variable is stationary at its original level, whereas I(1) means it is stationary at the first difference level.

Similarly, Table 3 shows the results of the second-generation (CIPS and CADF) unit root test. CIPS and CADF represent Cross-sectional Augmented Im-Pesaran-Shin and Cross-sectional Augmented Dicky-fuller Statistic.

Variables	Level	First difference	Order	Level	First difference	Order
CO2	-2.343***	-	I(0)	-2.343***	-	I(0)
FD	-3.078**	-	I(0)	-2.987	-2.342***	I(1)
REC	-2.654**	-	I(0)	-2.346*	-	I(0)
HC	-2.876***	-	I(0)	-2.098***	-	I(0)

Table 3: CIPS and CADF unit root test for panel data (Xu *et al.*, 2022)

Note: FD: Financial Development, REC: Renewable Energy, HC: Human Capital. Moreover, ***, ** and * indicate the level of significance at 1%, 5%, and 10%.

It can be observed from Table 3 that variables CO_2, REC, and HC have significant CADF and CIPS test statistics at the original level, indicating that these variables are stationary without any differencing. Consider the variable "FD", in the CIPS test, it has a t-value of -3.078 at the original level, which is significant at the 5% significance level, indicating that the "FD" data is stationary at its original level, hence it is I(0). However, in the CADF test, "FD" has a t-value of -2.987 at the original level, which is not significant, but its t-value at the first difference level is -2.342, which is significant at the 1% significance level, suggesting that the "FD" data becomes stationary after first differencing, hence it is I(1). Results confirm that all the variables are stationary at the level and the first difference. This is a desirable property for further analysis.

To confirm the long-run relationship among CO_2 emission, FD, REC, and HC, this study employs the bound test with the optimal lag order. The basic idea behind the bound test is to test the joint significance of the lagged levels' coefficients in an unrestricted equilibrium correction model, when the calculated F-statistic is above the upper bound value, the null hypothesis of no cointegration can be rejected, indicating the presence of a long-run relationship between the variables. The results are displayed in Table 4,

Table 4 reveals the existence of the long-run relationship between CO_2 emissions, FD, REC, and HC, as all G7 countries reject the null hypothesis of no long-run cointegration at 1%. For example, for Canada, the F-statistics value is 3.39, which is significant at 1%, suggesting that the cointegration exists. The F-statistics value is above the I(0) bound at all significance levels and also above the I(1) bound at the 10% level, providing evidence of cointegration. The last column shows the chosen ARDL model for each country and the numbers in parentheses represent the lag order selected for

Countries	F-Statistics	Outcome	Model Selection
Canada	3.39***	Co-integration	ARDL(1,1,0,0,0,1,0)
France	9.32***	Co-integration	ARDL(1,3,3,3,1,3,2)
Germany	3.31***	Co-integration	ARDL(7, 2, 2, 2, 2, 2, 2)
Italy	16.64***	Co-integration	ARDL(1,3,2,2,2,3,3)
Japan	25.14***	Co-integration	ARDL(3,3,3,2,3,3,0)
UK	6.99***	Co-integration	ARDL(2,2,2,1,1,1,1)
USA	7.86***	Co-integration	ARDL(2,3,2,3,1,0,2)
			Critical values (k = 6)
Sig		I0 Bound	I1 Bound
10%		2.12	3.23
5%		2.45	3.61
2.50%		2.75	3.99
1%		3.15	4.43

Table 4: Bound test for asymmetric cointegration (Xu *et al.*, 2022)

the variables in the model.

The Dumitrescu-Hurlin Granger causality test is employed as the final test before regression analysis to estimate the panel causality effect between CO_2, FD, REC, and HC. The estimated results are presented in Table 5.

Null Hypothesis:	W-Stat.	Zbar-Stat.	P-value
CO2 → FD	3.600	1.628	0.1034
FD → CO2	4.031*	2.115	0.0344
CO2 → RE	2.719	0.632	0.5273
RE → CO2	4.277*	2.393	0.0167
CO2 → HC	1.817	−0.387	0.6987
HC → CO2	6.278***	4.656	0.0000

Table 5: Dumitrescu-Hurlin (D-H) panel causality test (Xu *et al.*, 2022)

Note: → represent unidirectional causality; FD: Financial Development, REC: Renewable Energy, HC: Human Capital.

Table 5 presents diverse findings from the causality tests. In the case of financial development and CO_2, a unidirectional causal relation running from financial development to CO_2 emission is observed. This reveals that a change in financial development significantly affects CO_2 emission levels in G7 countries. Similarly, a unidirectional causal relationship exists from renewable energy to CO_2, revealing the impact of renewable energy on CO_2.

3.2 Regression Analysis

After determining the characteristics of the data through a series of time series tests, the regression model can be implemented. Several modelling approaches are used to specify the functional form of the relationship between variables. Panel data models are commonly employed due to their ability to control for both individual-specific and time-specific effects. Various techniques exist that are suitable for different research scenarios. In this section, the application of the Two-Way Fixed Effect model, Difference-in-differences model, and the Panel Quantile Regression model in the context of CO_2 emissions studies is explored.

3.2.1 Two-Way Fixed-Effect Model

In panel data analysis, we often observe data for multiple individuals or entities over several time periods. In order to consider both the time period fixed effect and the individual fixed effect, a Two-Way Fixed Effect Model (2FE) was constructed based on the general panel data model. Two-way fixed effects refer to a statistical modelling technique that incorporates fixed effects for both the individual (cross-sectional) units and the time periods.

Suppose we have a panel data set of N individuals and T time periods. Let $X_{i,t}$ and $Y_{i,t}$ represent the variables for individual i at time t, respectively. We consider the following two-way linear fixed effects regression model,

$$Y_{i,t} = \alpha_i + \gamma_t + \beta X_{i,t} + \varepsilon_{i,t}$$

for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$ where α_i represents the individual fixed effects and γ_t represents the time fixed effects (lmai and kim, 2020).

Including individual fixed effects α_i allows researchers to control for unobserved individual-level heterogeneity that remains constant over time. Similarly, the inclusion of time fixed effects γ_t helps capture time-specific factors that affect all individuals equally, such as macroeconomic shocks or policy changes.

By incorporating both individual and time-fixed effects, the two-way fixed effects model provides a robust approach to mitigate omitted variable bias and account for unobserved heterogeneity in panel data analysis. It allows researchers to analyze the effects of variables of interest while controlling for individual and time-specific factors that may confound the relationship.

3.2.1.1 Application in Selected Studies

Specifically, Chen *et al.* (2022) investigates the impact of artificial intelligence on carbon emissions using the two-way fixed effects model. The specific regression equation is constructed as follows:

$$CE_{i,t} = \alpha + \beta AI_{i,t} + \lambda X_{i,t} + u_i + \gamma_t + \epsilon_{i,t}$$

where $CE_{i,t}$ denotes the carbon emission intensity (represents the amount of total CO_2 emissions per unit of GDP) of city i in year t; $AI_{i,t}$ measures the artificial intelligence (AI) development level; $X_{i,t}$ is a set of control variables including fixed asset investment, financial development level, urbanization ratio, and government intervention; u_i and γ_t denote the area fixed effect and time fixed effect; $\epsilon_{i,t}$ is the error term. The coefficient β represents the net effect of the level of AI development on carbon emissions. A significant negative β indicates a reduction in carbon emissions due to improved AI development. The results of the two-way fixed effects model are shown in Table 6.

Variables	**(1)**	**(2)**
$\ln ETR$	−0.2720 ***	−0.2389 ***
	(0.0306)	(0.0336)
fin		0.0852 **
		(0.0373)
$invest$		0.0937
		(0.0708)
$urban$		−0.9216 ***
		(0.0804)
$expenditure$		2.0483 ***
		(0.3446)
Constant	0.6761 ***	0.5837 ***
	(0.0739)	(0.1004)
Observations	1890	1614
Adjusted R-squared	0.6586	0.6815
F statistics	79.20	43.35
Year FE	YES	YES
Province FE	YES	YES

Table 6: Results of two-way fixed effects model in China (Chen *et al.*, 2022)

Note: $\ln ETR$: Exposure to robot, *fin*: Financial development, *invest*: Fixed asset investment, *urban*: Urbanization Ratio, *expenditure*: government intervention.

Column (1) displays the regression results with the level of AI development as the single explanatory variable, whereas column (2) displays the results of regression with the incorporation of control variables. As observed in Table 6, each 1% improvement in AI development results in a 0.0027% decrease in carbon emissions. The estimated coefficients become smaller when control variables are included but remain significant.

This demonstrates that the development of artificial intelligence has a significant effect on reducing carbon emissions.

3.2.2 Difference-in-Differences Model

The Difference-in-Differences (DID) method is one of the most popular methods in the social sciences for estimating causal effects in non-experimental settings (Roth *et al.*, 2022), as it allows for the control of pre-existing differences in the objects of study. DID estimates the causal effect of an intervention by comparing the average change over time in the outcome variable for the treatment group with the average change for the control group. This helps in isolating factors that might influence the empirical results.

Assume that there are $i \in \{1, \ldots, n\}$ individuals, and $t \in \{1, \ldots, T\}$ time periods. We aim to evaluate the impact of a treatment on an outcome $Y_{i,t}$ over a population of individuals. Two groups are indexed by treatment status 0 and 1, where $Y_{i,t}(0)$ indicates individuals who do not receive treatment, i.e., the control group, and $Y_{i,t}(1)$ indicates individuals who do receive treatment, i.e., the treatment group.

The key assumption is that the treatment is independent of time t, observed covariates U_i, and the identity of the observation $X_{i,t}$. Assume that time and linear effects are constant:

$$E[Y_{i,t}(0)|U_i, X_{i,t}, t] = \beta_0 + U_i'\gamma + X_{i,t}'\beta.$$

Then assume that the causal effect is constant and has a linear functional form:

$$E[Y_{i,t}(1)|U_i, X_{i,t}, t] = E[Y_{i,t}(0)|U_i, X_{i,t}, t] + \tau.$$

This implies: $E[Y_{i,t}|U_i, X_{i,t}, t, D_{i,t}] = \alpha_i + \delta_t + \tau D_{i,t} + U_i'\gamma + X_{i,t}'\beta$. (Lechner, 2010)

A typical DID estimator is based on a regression model with fixed effect terms for individual (e.g., city) and year:

$$Y_{i,t} = \alpha_i + \delta_t + \tau D_{i,t} + X_{i,t}'\beta + \epsilon_{i,t},$$

where $D_{i,t}$ is a dummy variable for whether the individuals received the treatment in year t, α_i is the city-related fixed effect, which controls all the city-related factors that do not change with time, and δ_t is the time-related fixed effect, which controls all the time-related factors that do not change with city changes. This research focuses on the positive and negative direction of coefficient τ and its significance. If the estimated value $\hat{\tau}$ is different from zero, it suggests that the treatment's effect is significant.

3.2.2.1 Application in Selected Studies

The standard DID model requires individuals to be impacted by the treatment at the same point in time. However, this is not always the case, such as with the phased implementation of the Low-carbon city pilot policy in 2010, 2012, and 2017. In this case, a time-varying DID model is more applicable.

In some investigated studies exploring the impact of policies on CO_2 (Wen and Liu, 2022; Liu *et al.*, 2022), the time-varying DID model analysis are employed to compare the differences in carbon emissions between pilot and non-pilot cities before and after policy implementation. This requires that the parallel trend assumption is met – that is, carbon emissions in the treatment and control groups follow a similar trend before the start of the Policy. In other words, if the low-carbon city pilot policy is not implemented, the variation trend of carbon emission in pilot cities and non-pilot cities should be the same. Consequently, parallel trend test and dynamic effect analysis are required to be performed, namely:

$$Y_{i,t} = \beta_0 + \sum_{l=t-p}^{\min\{t,0\}} \beta_{i,t} D_{i,t}^l + \beta_1 \times \text{Control}_{i,t} + \theta_i + \mu_t + \epsilon_{i,t}$$

in which $D_{i,t}^l$ is the dummy variable. Provided that the year when city i became a low-carbon city pilot is p, then set $l = t - p$. When l is negative, if the studied year t is smaller than the year when the policy is implemented, then we set $D_{i,t}^l = 1$; otherwise, we set $D_{i,t}^l = 0$. When l is non-negative, if t is larger than the year when the policy is implemented, then we set $D_{i,t}^l = 1$; otherwise, we set $D_{i,t}^l = 0$.

In the study from Liu *et al.* (2022), considering China's low-carbon city pilot was launched in 2010, and the sample data for this study were selected from 2006 to 2017, covering the four years before and seven years after the implementation, therefore the parallel trend test would be:

$$Y_{i,t} = \beta_0 + \sum_{l=-4}^{7} \beta_{it} D_{it}^l + \beta_1 \times \text{Control}_{i,t} + \theta_i + \mu_t + \epsilon_{i,t}$$

The estimated coefficient β_{it} indicates the differences in carbon emissions between the experimental group and the control group in the year after the implementation of the policy. If the trend of β_{it} is relatively flat while l is negative, it is proven that the parallel trend hypothesis is true. On the contrary, it indicates that the two groups have significant differences before the implementation of the policy and the parallel trend hypothesis is false. Figure 1 shows the estimation results of β_{it} under the 95% confidence intervals for the indexes of carbon emissions.

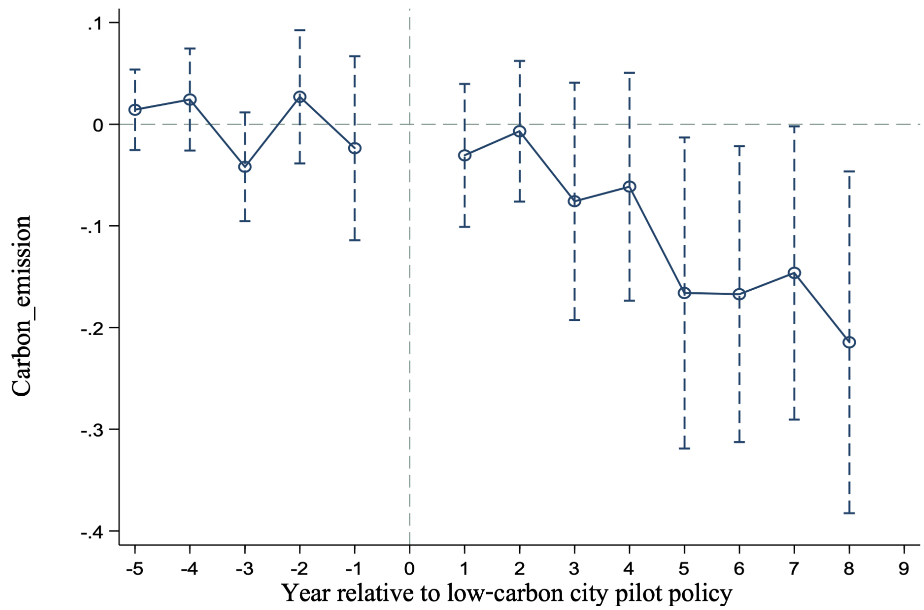

Figure 1: Dynamic effect of low-carbon city pilot policy on carbon emissions in China (Liu *et al.*, 2022)

β_{it} was not significant in the first four years after the implementation of the emission trading policy. In other words, before the pilot implementation, there was little difference between the treatment group and the control group, so the hypothesis of the parallel trend was established. Besides, β_{it} showed a decreasing trend year by year, and the carbon emission decreased significantly from the fifth year. This shows that the effect of the low-carbon city pilot policy needs time to accumulate.

The main reason is probably that enterprises mainly improve environmental efficiency by new technology development and industrial structure adjustment, which need a large investment and a long time to complete. Therefore, there is a lag in the policy effect of the emission trading system.

The parallel trends assumption cannot be directly tested for the period after the treatment is implemented, as we do not observe the counterfactual outcome for the treated group. Therefore, we can only test it for the pre-treatment period and assume that it holds for the post-treatment period as well. This method helps mitigate potential biases in the estimated treatment effect that could be due to permanent differences between those who receive the intervention and those who do not, or due to trends over time.

Overall, time-varying DID provides a flexible framework for estimating causal effects that can vary over time, allowing researchers to capture the dynamic nature of treatment effects in observational studies.

3.2.3 Panel Quantile Regression Model

The data of investigated variables are probably characterized by high kurtosis, thick tails, and heteroscedasticity, therefore, it is worth analyzing the entire data distribution instead of just the mean level. This is more informative because the way in which factors affect CO_2 could be different in high-pollution and low-pollution economies.

Panel Qantile Regression models (PQR) allow the researchers to account for unobserved heterogeneity and heterogeneous covariates effects, while the availability of panel data potentially allows the researcher to include fixed effects to control for some unobserved covariates (Canay, 2011). A simplified panel quantile regression model can be formulated as follows:

$$Y_{it} = X_{it}\theta_\tau + \epsilon_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T$$

where θ_τ are the $k \times 1$ regression coefficients at the τ-th quantile of the dependent variable:

$$P(Y_{it} \leq X_{it}\theta_\tau + \epsilon_{it}|X_{it}) = \tau$$

In contrast to OLS, which is based on minimizing the sum of squared residuals, the τ-th quantile regression estimator of θ minimizes a weighted sum of absolute errors:

$$\min_\theta \left[\sum_{(Y_{it} \geq X_{it}\theta_\tau)} \tau|Y_{it} - X_{it}\theta_\tau| + \sum_{(Y_{it} \leq X_{it}\theta_\tau)} (1-\tau)|Y_{it} - X_{it}\theta_\tau| \right]$$

The weighting scheme with τ and $1 - \tau$ is based on the Check Loss Function, which is designed to penalize deviations of predictions from the true values. Loss function is integral to the calculation of quantile regression estimator, it provides a mechanism to apply different weights to the residuals. These weights signify that underestimation ($Y_{it} \geq X_{it}\theta_\tau$), for the given quantile τ, are given τ times as weight in the overall loss function. In the symmetric case of absolute value loss, it is well known to yield the median. When the loss is linear and asymmetric, we prefer a point estimate more likely to leave us on the flatter of the two branches of marginal loss (Koenker and Machado, 1999). Hence, for higher quantiles (larger τ), the underestimations are given more weight. Compared to OLS regression, quantile regression minimizes the τ-weighted sum of absolute residuals, thus providing a more robust estimation in the presence of outliers.

In addition, quantile regression is less restrictive than the OLS approach as it allows the slope coefficient to vary across the quantiles of the dependent variable. This makes PQR particularly useful when dealing with heterogeneous effects or when the distribution of the dependent variable is skewed or contains outliers.

3.2.3.1 Application in Selected Studies

The investigated studies (Cheng *et al.*, 2021; Jeon, 2022; Keho, 2020) utilize panel quantile regression to estimate their models, as this method offers a systematic strategy for examining how the driven factors influence CO_2 emissions in countries across the entire conditional distribution of CO_2 emissions. For instance, Keho (2020) has modeled in his study:

$$Q(CO_{2it}|\Omega_{it}) = \theta_0 + \theta_1 E_{it} + \theta_2 y_{it} + \theta_3 y_{it}^2 + \theta_4 Ind_{it} + \theta_5 F_{it} + \theta_6 T_{it} + \theta_7 U_{it} + \epsilon_{it}$$

where $Q(\mathrm{CO}_{2it}|\Omega_{it})$ is the conditional quantile of CO_2 emissions and Ω_t contains the available information known at time t for country i. As for the independent variables: E is per capita energy consumption, y indicates per capita real GDP, y^2 is the square of per capita real GDP, Ind is an industrial value added as a share of GDP, F is financial development indicator, T is trade openness, and U is urbanization rate. Figure 2 shows Part of the results.

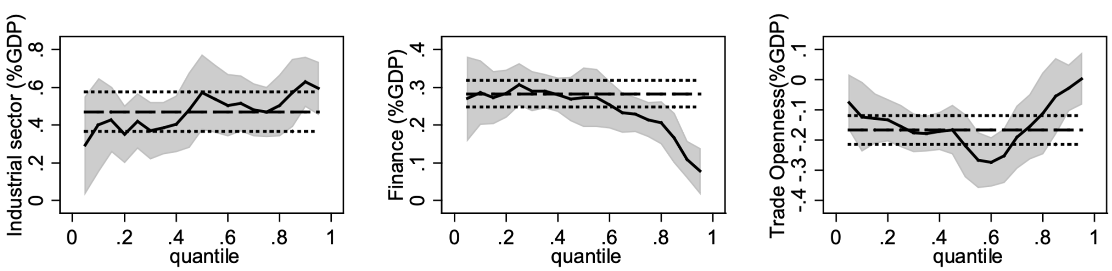

Figure 2: The parameter estimates of quantile and OLS regressions and their confidence intervals across 45 countries (Keho, 2020)

Note: The x-axis represents the conditional quantile of CO_2 emissions. The horizontal dashed line represents the OLS estimates. The two dotted lines depict the 95 percent confidence intervals for the OLS estimates. The solid line represents the quantile regression estimates; and the shaded grey area plots the 95 percent confidence band for the quantile regression estimates.

The regression results suggest some important differences across different quantiles in the conditional distribution of CO_2 emissions. The impact of the industrial sector is positive and larger at the top tail of the distribution, suggesting that industrialization increases pollution, especially in countries with higher pollution levels. Financial deepening, measured as bank credit to the private sector, contributes to increase CO_2, and its impact is larger in countries with lower CO_2 emissions.

With respect to trade openness, the effect is negative and shows a U-shaped relation with quantiles, suggesting that openness to trade reduces CO_2 with a larger reducing effect in countries in the middle part of the pollution distribution. The impact of trade is not significant at the top tail of the emission distribution (i.e., 0.80 quantile and higher), this suggests that in regions with high level of CO_2, the trade openness has no significant impact on CO_2 emissions.

3.3 Specific Framework - STIRPAT

After exploring various regression models used for investigating the determinants of CO_2 emissions, one might wonder how to select the most relevant factors in the models that precisely reflect the complexities of carbon emissions. To address this issue, STIRPAT (Stochastic Impacts by Regression on Population, Affluence, and Technology) model has been widely employed as a theoretical framework for examining the potential determinants of CO_2 emissions.

The IPAT model considers environmental impacts resulting from three drivers: population, affluence, and technology. The IPAT model could be expressed as $I = P \times A \times T$, where I denotes the human impact on the environment, P stands for population size, A represents average affluence, and T is technological level. Taking CO_2 emissions as an example, the equation will be:

$$CE = POP \times \frac{GDP}{POP} \times \frac{CE}{GDP}$$

where CE stands for the total CO_2 emissions level, denoted as CE; P represents the population, denoted as POP; A stands for GDP per capita, denoted as GDP/POP; and T is carbon intensity which is calculated by CO_2 emissions per unit of GDP, denoted as CE/GDP. The IPAT framework was subsequently improved by way of the ImPACT model, which decomposes CO_2 emissions per unit of GDP (T) into energy consumption per unit of GDP (C) and CO_2 emissions per unit of consumption (T) (Su *et al.*, 2020).

However, the ImPACT and IPAT models remain extremely limited in their application because they do not allow the influencing factors to change non-monotonically and non-proportionally. For example, the "Environmental Kuznets Hypothesis" points out that the link between environmental quality and per capita income is not a linear, monotonous relationship, but rather in the form of an inverted U-shaped curve (Grossman and Krueger, 1995). For addressing this issue, Dietz and Wa (1994) developed the IPAT identity into a stochastic form to create the STIRPAT model. This stochastic model has been popularly applied to examine the influence of driving forces on environmental changes. The STIRPAT model can be written as:

$$I_i = a \cdot P_i^b \cdot A_i^c \cdot T_i^d \cdot e_i \tag{3}$$

where a denotes the model's constant coefficient; b, c, and d stand for the estimated parameters, and e stands for the error term; when the values of a, b, c, d, e are all equal to 1, the STIRPAT model will change to the IPAT identity. After taking logarithm, Eq.(3) can be transformed into:

$$\ln(I_i) = \ln(a) + b\ln(P_i) + \ln(A_i) + \ln(T_i) + \ln(e_i)$$

Overall, IPAT/STIRPAT is a coordinated research program dedicated to understanding the dynamic coupling between human systems and the ecosystems on which they depend. The STIRPAT model not only allows each coefficient as a parameter to estimate, but also allows the proper decomposition of each factor, which means new influencing factors can be added to the STIRPAT model framework according to the characteristics of each study.

3.3.1 Application in Selected Studies

Wu *et al.* (2021) extended the STIRPAT theoretical model to meet the design requirements regarding previous literature. The first extension was made in acknowledgement that CO_2 emissions are affected by fossil energy efficiency. Fossil energy intensity (FEI) is an expression of efficiency that was thus considered in their model.

Secondly, renewable energy consumption share (RES) in the energy mix is considered, as increases in the renewable energy consumption share can decrease the direct use of fossil energy, thus decreasing CO_2 emissions. Third, the model also extends to include fossil fuel CO_2 intensity (FCI), which denotes the carbon content of the fossil fuel mix. Fourthly, the industrial structure (IS), expressed by the share of value added of industry (including construction) of GDP, was also included, as a decrease in the proportion of secondary industries signals the gradual replacement of high-energy-consuming industries by low-carbon industries and thus a gradual reduction in carbon emissions. In addition, this study also introduced a time trend variable (TT), which can explain exogenous changes in CO_2 emissions that are not explained by other independent variables.

Finally, since CO_2 emissions have path-dependent inertial characteristics and are subject to continuous dynamic adjustment, the emissions may experience lagged effects, making the introduction of a dynamic model lagged term (CE_{it-1}) necessary. Taking all of this into account, the study proposes the following dynamic panel data model:

$$\begin{aligned}\ln(CE_{it}) = &\ln(CE_{it-1}) + \beta_1 \ln(POP_{it}) + \beta_2 \ln(GDPPC_{it}) + \beta_3 FEI_{it} \\ &+ \beta_4 RES_{it} + \beta_5 IS_{it} + \beta_6 \ln(FCI_{it}) + TT_{it} + \mu_i + \varepsilon_i\end{aligned}$$

where CE_{it} represents the total CO_2 emissions of country i in year t, where i denotes the country and t is the year; CE_{it-1} denotes the lagged term of CE_{it} (the CO_2 emissions of the following year after CE_{it}); POP stands for the population; $GDPPC$ is GDP per capita; FEI represents fossil energy intensity (final energy consumption / GDP); RES stands for the share of renewable energy use (renewable energy consumption / total final energy); IS represents the industrial structure (the value added by industry, including construction, within GDP); FCI represents fossil CO_2 intensity (CO_2 / fossil energy consumption); TT stands for time trend variable; β_1-β_6 are estimated parameters; μ_i is unobservable individual effect; and ε_i is a random disturbance term.

This extended STIRPAT model provides a comprehensive framework to dissect the intricate relationship between CO_2 emissions and a range of economic, energy, and structural factors. The inclusion of variables such as fossil energy intensity, renewable energy consumption share, industrial structure, and lagged variables, adds refinement and specificity to the understanding of CO_2 emissions.

Generally, emissions were dissected into several contributing factors using an extended STIRPAT model. This model allowed for an estimation of their effects on CO_2 emissions from both historical and prospective perspectives. Following this, various estimation methods could then be used to determine the parameters in the model.

4 Other Estimation Methods

4.1 Feasible Generalized Least Squares

Ordinary least squares is a technique for estimating unknown parameters in a linear regression model. OLS yields the maximum likelihood in a vector β, assuming the parameters have equal variance and are uncorrelated, in a noise ε homoscedastic:

$$y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I)$$

where the response values are placed in a vector $y = (y_1, \ldots, y_n)^\top$, and the predictor values are placed in the design matrix $X = (x_1, \ldots, x_n)^\top$, where $x_i = (1, x_{i2}, \ldots, x_{it})$ is a vector of the ith individual at time t (including a constant), and the error term $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^\top$.

Generalized least squares (GLS) proposed by Aitken (1935) allows this approach to be generalized to give the maximum likelihood estimate β when the model assumes the conditional variance of the error term given X is a known nonsingular covariance matrix Ω:

$$E[\varepsilon|X] = 0, \quad \mathrm{Cov}[\varepsilon|X] = \Omega, \quad \mathrm{Cov}(\varepsilon\varepsilon^\top) = \sigma^2\Omega$$

Typically, this leads to the treatment that presents as follows:

$$y = X\beta + \eta, \quad \eta \sim N(0, \sigma^2\Omega)$$

In an ideal situation, if we precisely know the form of the covariance matrix of the error term, we can employ the GLS method to attain more efficient and unbiased estimation results (Baltagi, 2008).

$$\Omega = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix}$$

Since Ω is a real symmetric matrix, it can be decomposed as $\Omega = DD^\top$. By left multiplying both sides of $y = X\beta + \eta$ by D^{-1}, we have:

$$D^{-1}y = D^{-1}X\beta + D^{-1}\eta,$$
$$Y^* = X^*\beta + \eta^*$$
$$Cov(\eta^*\eta^{*\top}) = E(\eta^*\eta^{*\top}) = E(D^{-1}\eta\eta^\top D^{-\top})$$
$$= D^{-1}E(\eta\eta^\top)D^{-\top} = D^{-1}\sigma^2\Omega D^{-\top} = \sigma^2 D^{-1}DD^\top D^{-\top} = \sigma^2 I$$

where Y^*, X^* and η^* are the transformed vectors. Therefore, the least squares method can be used to obtain the estimator:

$$\hat{\beta} = \left(X^{*\top} X^*\right)^{-1} X^{*\top} Y^* = \left(X^\top D^{-\top} D^{-1} X\right)^{-1} X^\top D^{-\top} D^{-1} y = \left(X^\top \Omega^{-1} X\right)^{-1} X^\top \Omega^{-1} y$$

GLS is equivalent to applying OLS to a linearly transformed data version. The GLS estimator is unbiased, consistent, efficient, and asymptotically normal with $E[\hat{\beta}|X] = \beta$ and $Cov[\hat{\beta}|X] = \left(X^\top \Omega^{-1} X\right)^{-1}$.

However, in practical applications, we often do not know the exact form of the covariance matrix of the error term. One can obtain a consistent estimate of Ω, denoted as $\hat{\Omega}$, using an implementable version of GLS known as the Feasible Generalized Least Squares (FGLS) estimator.

FGLS is a two-step process. First, it uses OLS or other methods to preliminarily estimate the covariance structure of the error term $\hat{u}_j = (y - X\hat{\beta}_{OLS})_j$. For simplicity, consider the model for heteroscedastic and not-autocorrelated errors. Assume that the variance-covariance matrix Ω of the error vector is diagonal, or equivalently that errors from distinct observations are uncorrelated. Then each diagonal entry may be estimated by the fitted residuals $\hat{u}_j$, so $\hat{\Omega}_{OLS}$ may be constructed by $\hat{\Omega}_{OLS} = diag(\hat{\sigma}_1^2, \hat{\sigma}_2^2, \ldots, \hat{\sigma}_n^2)$. Then, estimate $\hat{\beta}_{FGLS}$ using $\hat{\Omega}_{OLS}$:

$$\hat{\beta}_{FGLS} = \left(X^\top \left(\hat{\Omega}_{OLS}^{-1}\right) X\right)^{-1} X^\top \left(\hat{\Omega}_{OLS}^{-1}\right) y,$$

along with its covariance matrix $\hat{\Sigma}_{FGLS} = \hat{\sigma}_{FGLS}^2 \left(X^\top \left(\hat{\Omega}_{OLS}^{-1}\right) X\right)^{-1}$, where

$$\hat{\sigma}_{FGLS}^2 = \frac{y^\top \left[\hat{\Omega}_{OLS}^{-1} - \hat{\Omega}_{OLS}^{-1} X \left(X^\top \hat{\Omega}_{OLS}^{-1} X\right)^{-1} X^\top \hat{\Omega}_{OLS}^{-1}\right] y}{T - p}.$$

This process can be iterated. The first iteration is given by

$$\hat{u}_{FGLS1}^2 = Y - X\hat{\beta}_{FGLS1},$$

$$\hat{\Omega}_{FGLS1} = \mathrm{diag}(\hat{\sigma}_{FGLS1,1}^2, \hat{\sigma}_{FGLS1,2}^2, \ldots, \hat{\sigma}_{FGLS1,n}^2),$$

$$\hat{\beta}_{FGLS2} = \left(X^\top \left(\hat{\Omega}_{FGLS1}^{-1}\right) X\right)^{-1} X^\top \left(\hat{\Omega}_{FGLS1}^{-1}\right) y.$$

This estimation can be iterated to convergence. (Greene, 2003)

In summary, FGLS is an extension or a practical improvement of the GLS method. It allows us to employ the GLS method for more efficient parameter estimation in the absence of precise knowledge about the covariance structure of the error term. However, the FGLS estimator is not always consistent. One case in which FGLS might be inconsistent is if there are individual specific fixed effects (Hansen, 2007). This is because the FGLS method attempts to transform the model based on an estimated covariance matrix of the errors such that the transformed errors become uncorrelated. If the errors are not independent and their correlation structure is unknown, the estimate of the error covariance matrix used to perform this transformation may be incorrect, leading to an inappropriate transformation and subsequently to biased and inefficient estimates.

In such cases, it is recommended to use Weighted Least Squares (WLS). WLS is a variant of OLS that weights the observations differently based on the heteroscedasticity of the errors. Observations with a higher variance have less weight than observations with a smaller variance (Fahrmeir *et al.*, 2013). The basic idea is that observations with larger variances are considered less reliable. WLS directly accounts for heteroscedasticity in the model, this can make it a more robust method in certain circumstances, such as when the errors are heteroscedastic and the fixed effects are present.

4.1.1 Application in Selected Papers

Azimi and Bian (2023) examined the nexus between carbon neutral policies and CO_2 emissions using the FGLS. They took a comprehensive approach in assessing the carbon neutral policy, examining it from both source control and sink increase perspectives. Source control refers to the process of managing and limiting the production of carbon dioxide at its origin. Sink increase, on the other hand, pertains to the enhancement and expansion of carbon sinks, which are systems that absorb and store carbon dioxide from the atmosphere. The analytical model is as follows:

$$\begin{aligned}\ln\left(\mathrm{CO}_{2it}\right) = \alpha_0 &+ \alpha_1 \ln\left(\mathrm{R\&D}_{it}\right) + \alpha_2 \ln\left(\mathrm{GREEN}_{it}\right) + \alpha_3 \ln\left(\mathrm{EE}_{it}\right) + \alpha_4 \ln\left(\mathrm{REPG}_{it}\right) \\ &+ \alpha_5 \ln\left(\mathrm{GDP}_{it}\right) + \alpha_6 \ln\left(\mathrm{GDP}^2_{it}\right) + \alpha_7 \ln\left(\mathrm{URB}_{it}\right) + \varepsilon_{it}\end{aligned}$$

where the subscript it refers to the region i and the year t. Variable definitions:

CO_2 : Per capita CO_2 emissions

R&D : CCUS (carbon capture, utilisation, and storage) development measured by R&D (Research and development) investment per unit of GDP

GREEN : Green space development measured by the green-covered area in cities

EE : Energy efficiency measured by industrial output value per unit of energy consumption

REPG : Renewable energy power generation rate measured by the percentage of power generation from renewable sources

GDP : Gross domestic product per capita

URB : Percentage of urban population from the total population

The carbon neutrality policy includes R&D and GREEN as a sink increase perspective and EE and REPG as a source control perspective. Each of these variables represents dynamic processes that can change over time, leading to potential autocorrelation and heteroscedasticity. FGLS used to address these issues by estimating a feasible version of the variance-covariance matrix of the errors. To explore the long-term relationships among the variables, the authors ran the FGLS estimators. Applying the FGLS estimator

$$\hat{\beta}_{\text{FGLS}} = \left(X^T \left(\hat{\Omega}_{\text{OLS}}^{-1}\right) X\right)^{-1} X^T \left(\hat{\Omega}_{\text{OLS}}^{-1}\right) y$$

Where: $y = (\ln \mathrm{CO}_{2it}^{\top})$
$X = (1, \ln \mathrm{R\&D}_{it}, \ln \mathrm{GREEN}_{it}, \ln \mathrm{EE}_{it}, \ln \mathrm{REPG}_{it}, \ln \mathrm{GDP}_{it}, \ln \mathrm{GDP}^2_{it}, \ln \mathrm{URB}_{it})$
$\beta = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7)$
$\hat{\Omega}_{\text{OLS}}$: Estimated covariance matrix obtained from the OLS residuals.

The result $\hat{\beta}_{\text{FGLS}}$ is the FGLS estimate of the parameters β. The estimation results are shown in Table 7.

Variable	FGLS		
	Coefficient	*p*-value	Wald chi^2
ln *R&D*	0.527*** (0.111)	0.000	940.190
ln *GREEN*	−0.779 (0.498)	0.118	
ln *EE*	−0.400*** (0.134)	0.003	
ln *REPG*	−0.461*** (0.018)	0.000	
ln *GDP*	3.795* (2.257)	0.093	
ln *GDP*^{2}	−0.175 (0.109)	0.109	
ln *URB*	2.090*** (0.409)	0.000	
Number of observations	360		
Number of groups	**30**		

Table 7: Estimation results using FGLS in China (Azimi and Bian, 2023)

Note: R&D: Research and development investment per unit of GDP, $GREEN$: Green-covered area in cities, EE: Industrial output value per unit of energy consumption, $REPG$: Percentage of power generation from renewable sources, GDP: GDP per capita, GDP^2: Square of GDP per capita, URB: Percentage of urban population from the total population

The FGLS results imply that for each 1% increase in green space development, energy efficiency, and renewable energy power generation, there corresponds to a respective decrease in per capita CO_2 emissions by 0.779%, 0.400%, and 0.461%. Conversely, investment in research and development per unit of GDP, as well as a rise in the urban population percentage, are linked to a significant increase in per capita CO_2 emissions. The positive coefficient on GDP and the negative coefficient on GDP squared suggest the presence of an Environmental Kuznets Curve. This indicates a turning point of GDP after which increases in GDP start to reduce CO_2 emissions, and a reduction effect exists on CO_2 as GDP continues to grow.

In general, the variables representing carbon neutrality policy significantly contribute to reducing carbon emissions, with the exceptions being ln R&D, GDP, and URB. The results reveal that an improvement in energy efficiency and renewable energy power generation decreases the per capita CO_2 emissions from the source control perspective. From the sink increase perspective, only green space development affects CO_2 emissions reduction; the development of carbon capture, utilization, and storage (R&D) appears not to have a significant effect.

4.2 Generalized Method of Moments

The Generalized Method of Moments (GMM) is used for estimating statistical model parameters. It was developed by Lars Peter Hansen and Robert Hodrick (Hansen, 1982) as an extension of the Method of Moments.

Assume that the population distribution has an unknown mean μ and variance equal to 1. In this case, the moment condition states that $E(x_i) = \mu$. If $\{x_i : i = 1, 2, \ldots, n\}$ is a distribution of independent and identically distributed samples, the sample average $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample analogue to the population mean $E(x_i)$. By using this analogy principle, the method of moments (MM) estimator for $E(x_i) = \mu$ is simply given by $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \hat{\mu}_n$.

Basically, it is to calculate the first moment, then replace it with the sample analogue and solve the equation for the unknown parameter. The case we consider is simply that the number of moment conditions q is equal to the number of unknown parameters p. Assuming functionally independent moment equations, the resulting system of equations provided by the moment conditions can be solved to obtain the MM estimator.

If $q < p$, there is insufficient information, and the model is not identified. In the case of $p < q$, the model is overidentified, and in most cases, it is not possible to solve the system of equations. However, estimation can still proceed in this situation with GMM (Hansen, 1982).

Suppose there is an observed sample $\{x_i : i = 1, 2, \ldots, T\}$ from which we want to estimate an unknown parameter vector $\theta \in \Theta \subseteq \mathbb{R}^p$ with true value θ_0. Let $f(x_i, \theta)$ be a continuous and continuously differentiable $\mathbb{R}^p \to \mathbb{R}^q$ function of θ, and let $m(\theta_0) \equiv E(f(x_i, \theta_0)) = 0$ be a set of q moment conditions. The basic idea behind GMM is to replace the theoretical expected value with its empirical analog, the sample average: $\hat{m}(\theta) \equiv \frac{1}{T}\sum_{t=1}^{T} f(x_i, \theta_0)$, and then minimize the norm of this expression with respect to θ:

$$\|\hat{m}(\theta)\|_W^2 = \hat{m}(\theta)^\top \hat{W}(\theta)\hat{m}(\theta),$$

where W is a positive-definite weighting matrix. In practice, the weighting matrix W is calculated from the available data and will be denoted as $\hat{W}$. Then, the GMM estimator is given by:

$$\hat{\theta} = \arg\min_{\theta\in\Theta} \left(\frac{1}{T}\sum_{t=1}^{T} f(x_i, \theta_0)\right)^\top \hat{W} \left(\frac{1}{T}\sum_{t=1}^{T} f(x_i, \theta_0)\right)$$

The GMM estimator exploits information from the general form of population moment conditions. When the number of moment conditions equals the number of unknown parameters, then GMM and MM are equivalent. When $q < p$, the GMM estimator is the value of θ closest to solving the sample moment conditions.

The GMM approach is particularly useful when dealing with models that have more parameters than identifying equations, or when the assumptions of classical estimation techniques such as ordinary least squares are violated. It provides a flexible framework for estimating parameters by matching sample moments with their population counterparts.

4.2.1 Application in Selected Papers

To obtain the long-run coefficients, some investigated studies (Wang *et al.*, 2022; Lau *et al.*, 2023; Khan *et al.*, 2021) use the system-generalized method of moments (SYS-GMM) type system dynamic panel data estimation proposed by Blundell and Bond (1998). It is an extension of GMM technique that captures significant autoregressive parameters and a relatively large ratio of the variance of the panel-level effects to the variance of idiosyncratic error terms.

Through the system-generalized method of moments method, Wang *et al.* (2022) empirically examined the potential carbon-reduction effect of China's high-quality energy development (HED). To ensure the accuracy of the regression, several control variables are incorporated into the model, including energy efficiency (EE), economic development (GDP), urbanization (Urb), and trade openness (Tra). Natural logarithm transformations are applied to the data to mitigate the effects of data fluctuation and heteroscedasticity. Additionally, due to the potential lagged impact of CO_2 emissions, a lagged term for CO_2 emissions is incorporated into the empirical model. Thus, the empirical model is represented as follows:

$$\begin{aligned}\ln(\mathrm{CO}_{2it}) = \alpha_0 + \alpha_1 \ln(\mathrm{CO}_{2(i,t-1)}) + \alpha_2 \ln(\mathrm{HED}_{it}) + \alpha_3 \ln(\mathrm{EE}_{it}) + \alpha_4 \ln(\mathrm{Urb}_{it}) \\ + \alpha_5 \ln(\mathrm{Tra}_{it}) + \alpha_6 \ln(\mathrm{GDP}_{it}) + \varepsilon_{it} \qquad (4)\end{aligned}$$

where $\alpha_1 - \alpha_6$ are the estimated parameters. To improve the dynamic explanatory ability of the estimation model, this study introduces the lag term of the explained variable as the explanatory variable of the regression model. However, including a lag term in the model leads to an endogeneity issue, because the lagged dependent variable is also correlated with the error term.

To address the endogeneity problem, Arellano and Bond (1991) developed the "differential GMM method (DIF-GMM)", which employs instrumental variables and the corresponding moment conditions. Although the DIF-GMM method uses a lag term to reduce the impact of endogeneity, this method may encounter a serious problem of weak instrumental variables under limited sample conditions. Based on DIF-GMM, SYS-GMM combines the estimation of the original level model with the difference conversion model simultaneously, effectively mitigating the problem of weak instrumental variables. This robustness makes SYS-GMM particularly suitable for estimating Eq((4)). This study also lists the results of OLS and FGLS for comparative analysis to highlight the accuracy and effectiveness of the SYS-GMM technique. The results of the three estimation methods are shown in Table 8.

The significant coefficient of the lagged dependent variable $\ln\left(\mathrm{CO}_{2(i,t-1)}\right)$ in the model suggests that past levels of CO_2 emissions have a significant impact on the current level of emissions, validating the dynamic nature of the model. Unlike OLS and FGLS, which are static panel data estimators and do not directly address the lagged nature of variables, SYS-GMM dynamically accommodates such characteristics, enhancing the model's reliability. The Sargan test, which checks the validity of overidentifying restrictions in the SYS-GMM, shows a high p-value (0.9443), indicating that this model passes the test and the instruments used are valid.

Dependent variable: *LnCO₂*			
Variable	Estimation		
	OLS	FGLS	SYS-GMM
$lnCO_{2i,\ t\text{-}1}$			0.463***
			(15.42)
lnHED	−0.120***	−0.075***	−0.032***
	(−3.52)	(−4.74)	(−2.68)
lnEE	−1.153***	−1.138***	−0.743***
	(−40.63)	(−100.58)	(−26.03)
lnGDP	1.078***	1.070***	0.675***
	(69.59)	(171.95)	(20.54)
lnTra	0.030***	0.022***	−0.129***
	(−2.62)	(4.96)	(−6.99)
lnUrb	−0.052	−0.087***	−0.299***
	(−0.69)	(−3.94)	(−4.54)
_Cons	−4.57***	−4.288***	−2.186***
	(−11.56)	(−33.13)	(−9.13)
R-squared	0.9430		
AR (1)			0.0122
AR (2)			0.6469
Sargan test			0.9443

Table 8: Results of the impact of HED on CO_2 in China (Wang *et al.*, 2022)

Note: *HED*: High-Quality Energy Development, *EE*: Energy Efficiency, *GDP*: Economic Development, *Urb*: Urbanization, *Tra*:Trade Openness. The values of t- and z-statistics are included in parentheses of static and dynamic panel estimation, respectively.

The results of the three estimation methods show that the impacts of the variables on CO_2 emissions are basically in line with the researcher's expectations, with an exception being trade openness. In this regard, the negative coefficient of trade openness under the SYS-GMM estimation is in line with the researcher's expectations, thereby further illustrating the efficacy of the SYS-GMM.

The key explanatory variable, High-Quality Energy Development, is negative at the 1% significance level. In particular, every 1% increase in the HED can reduce CO_2 emissions by 0.032%, illustrating the potential of the HED strategy to promote effective carbon emission reduction in China. In addition, energy efficiency, trade openness, and urbanization also display a negative relationship with CO_2, indicating their potential to reduce CO_2 emissions.

4.3 Method of Moment Quantile-Regression

The Method Moments Quantile Regression (MMQR) extends the Quantile Regression Model by considering the conditional moments of the dependent variable. This method provides a more detailed picture of the distributional effects of the explanatory variables.

The MMQR approach particularly makes it possible to capture the conditional heterogeneous covariance effects of the influences of CO_2 emissions by considering the individual effects, which affect the whole distribution, rather than just shifting means as in the other panel quantile regression approaches (Ike *et al.*, 2020). In other words, this method estimates conditional quantile effects through known location and scale functions, both of these estimates identified by conditional expectations of appropriately defined variables (Machado and Santos Silva, 2019).

The location-scale model is a class of statistical models that postulates the relationship between the cumulative distribution function of the dependent variable and the independent variables in terms of location (i.e., mean or median) and scale (i.e., variance or standard deviation). The estimation of the conditional quantiles $Q_Y(\tau|X)$ for a model of the location-scale variant takes the following form:

$$Y_{it} = \alpha_i + X'_{it}\beta + \sigma(\delta_i + Z'_{it}\gamma)U_{it} \tag{5}$$

where $\alpha_i + X'_{it}\beta$ represents the location while $\sigma(\delta_i + Z'_{it}\gamma)$ represents the scale. The probability, $P\{\sigma(\delta_i + Z'_{it}\gamma) > 0\} = 1$, $(\alpha, \beta', \delta, \gamma')$ are parameters to be estimated. (α_i, δ_i), $i = 1, \ldots, n$, capture the individual i fixed effects, and Z_{it} is a k-vector of identified components of X_{it} which are differentiable transformations with element l given by: $Z_l = Z_l(X)$, $l = 1, \ldots, k$.

X_{it} is independently and identically distributed for any fixed i and is independent across time t. U_{it} is an unobserved random variable, which is independently and identically distributed across individuals i and through time t. U_{it} is also orthogonal to X_{it} and normalized to satisfy the moment conditions:

$$E(U_{it}) = 0, \quad E(|U_{it}|) = 1. \tag{6}$$

$E(U_{it}) = 0$ ensures that U_{it} does not introduce bias to the overall model predictions. $E(|U_{it}|) = 1$ is related to the normalization of U_{it}, scaling its magnitude to a standardized unit across the entire dataset. In the case where $\sigma(\cdot)$ is the identity function and $Z = X$, Eq. (5) implies that:

$$Q_Y(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + X'_{it}\beta + Z'_{it}\gamma q(\tau)$$

where $q(\tau) = F_U^{-1}(\tau)$, and thus $P(U < q(\tau)) = \tau$. $Q_Y(\tau|X_{it})$ indicates the quantile distribution of the dependent variable Y_{it} (e.g., natural logarithm of CO_2 emissions) which is conditional on the location of independent variable X_{it}. The scalar coefficient $\alpha_i(\tau) \equiv \alpha_i + \delta_i q(\tau)$ is indicative of the quantile-τ fixed effect for individual i, or the distributional effect at τ. The individual effect does not denote an intercept shift, unlike the usual least-squares fixed effects. They are time-invariant parameters whose heterogeneous impacts are allowed to differ across the quantiles of the conditional distribution of the endogenous variable Y.

Consider now the MMQR estimator implied by moment conditions:

$$\begin{aligned}
E[RX] &= 0,\\
E[R] &= 0,\\
E[(|R| - \sigma(\delta + Z'\gamma))D_\gamma^\sigma] &= 0,\\
E[(|R| - \sigma(\delta + Z'\gamma))D_\delta^\sigma] &= 0,\\
E[I(R \leq q(\tau)\sigma(\delta + Z'\gamma)) - \tau] &= 0,
\end{aligned}$$

where $R = Y - (\alpha + X'\beta) = \sigma(\delta + Z'\gamma)U$, $D_\gamma^\sigma = \frac{\partial\sigma(\delta+Z'\gamma)}{\partial\gamma}$, and $D_\delta^\sigma = \frac{\partial\sigma(\delta+Z'\delta)}{\partial\delta}$. For this model, the moment conditions have a convenient triangular structure with respect to the model parameters that allows the GMM estimator to be calculated sequentially (Machado and Silva, 2019):

1. Regress $(Y_{it} - \frac{1}{T}\sum_t Y_{it})$ on $(X_{it} - \frac{1}{T}\sum_t X_{it})$ by least squares to obtain $\hat{\beta}$;
2. Estimate $\hat{\alpha}_i = \frac{1}{T}\sum_t(Y_{it} - X_{it}'\hat{\beta})$ and obtain the residuals $\hat{R}_{it} = Y_{it} - \hat{\alpha}_i - X_{it}'\hat{\beta}$;
3. Regress $(|\hat{R}_{it}| - \frac{1}{T}\sum_t |\hat{R}_{it}|)$ on $(Z_{it} - \frac{1}{T}\sum_t Z_{it})$ by least squares to obtain $\hat{\gamma}$;
4. Estimate $\hat{\delta}_i = \frac{1}{T}\sum_t(|\hat{R}_{it}| - Z_{it}'\hat{\gamma})$;
5. Estimate $q(\tau)$ by $\hat{q}$, the solution to $\min_q \sum_i \sum_t \rho_\tau(\hat{R}_{it} - (\hat{\delta}_i + Z_{it}'\hat{\gamma})q)$, where $\rho_\tau(A) = (\tau - 1)I\{A \leq 0\} + \tau I\{A > 0\}$ is the check-function.

To summarize, the provided equations and steps outline the estimation process for the MMQR estimator. Overall, the MMQR estimator offers a framework for investigating the quantile-τ fixed effect and distributional effects of the dependent variable, given the independent variables' locations. By considering the moment conditions and implementing the sequential estimation steps, the model provides insights into the heterogeneous impacts across quantiles of the conditional distribution of the endogenous variable Y.

4.3.1 Application in Selected Papers

Jeon (2022) examined the linkages between energy-related CO_2 emissions, economic growth, and renewable energy consumption for the 48 U.S. states over the period 1997–2017 by employing the MMQR with fixed effects. The following dynamic panel model is suggested:

$$\begin{aligned}\ln \mathrm{CO}_{2it} = {} & \gamma_0 + \gamma_1 \ln \mathrm{CO}_{2i,t-1} + \gamma_2 \ln \mathrm{GDP}_{it} + \gamma_3 \ln \mathrm{GDP}_{it}^2 + \gamma_4 \mathrm{S\_RE}_{it} \\ & + \gamma_5 \ln \mathrm{HDD}_{it} + \gamma_6 \ln \mathrm{CDD}_{it} + \gamma_7 \ln \mathrm{E\_PRICE}_{it} \\ & + \rho_i + \varepsilon_{it}\end{aligned}$$

where the subscripts i and t denote the state and time periods, respectively. CO_{2it} is per capita energy-related CO_2 emissions, $\mathrm{CO}_{2i,t-1}$ is the lagged dependent variable; GDP denotes per capita real gross domestic product by state; S_RE denotes share of renewable energy in total energy consumption; HDD denotes heating degree days; CDD denotes cooling degree days; E_PRICE denotes electricity price; F_PRICE denotes primary energy price; ρ_i denotes time-invariant state-specific fixed effect; ε_{it} denotes error term.

The MMQR approach particularly makes it possible to capture the conditional heterogeneous covariance effects of the influences of CO_2 emissions by considering the individual effects, which affect the whole distribution. The estimation of the conditional quantiles of CO_2 emissions $Q_{\mathrm{CO}_2}(\tau|X)$, for a location-scale model takes the following form:

$$\mathrm{CO}_{2it} = \alpha_i + X_{it}'\rho + (\delta_i + Z_{it}'\gamma)U_{it}$$

where the probability $P\{(\delta_i + Z_{it}'\gamma) > 0\} = 1$. The parameters (α_i, δ_i), $i = 1, \ldots, 48$, capture the individual state fixed effects. X_{it} is independently and identically distributed for any fixed state i and independent across time t. ρ represents a vector of estimated parameters, which vary on different quantile τ of CO_2. U_{it} is an unobserved random variable and independently and identically distributed across the individual state i at time t. And U_{it} is statistically independent of X_{it} and normalized to meet the moment conditions stated in (6).

$$\begin{aligned}Q_{\ln \mathrm{CO}_{2it}}(\tau|X_{it}) = {} & (\alpha_i + \delta_i q(\tau)) + \rho_1 \ln \mathrm{GDP}_{it} + \rho_2 \ln \mathrm{GDP}_{it}^2 + \rho_3 S_{RE_{it}} \\ & + \rho_4 \ln \mathrm{HDD}_{it} + \rho_5 \ln \mathrm{CDD}_{it} + \rho_6 \ln \mathrm{E\_PRICE}_{it} + \rho_7 \ln \mathrm{F\_PRICE}_{it} + \varepsilon_{it}\end{aligned}$$

where $Q_{\ln \mathrm{CO}_{2it}}(\tau|X_{it})$ represents the quantile distribution of the dependent variable CO_2 (natural logarithm of energy-related CO_2 emissions per capita) which is conditional on the location of independent variable X_{it}. Scalar coefficient $\alpha_i(\tau) \equiv \alpha_i + \delta_i q(\tau)$

is the quantile-τ fixed effect for state i, or the distributional effect at τ. Unlike the usual fixed effects, the distributional effect indicates the heterogeneous impacts of time-invariant characteristics, which allows capturing different effects on different states of the conditional distribution of CO_2. $q(\tau)$ indicates the τ-th sample quantile which satisfies the condition $\int_0^1 q(\tau)d\tau = 0$, which implies that α_i can be interpreted as the average effect for state i. Fig 3 shows the estimation results of MMQR with fixed effects.

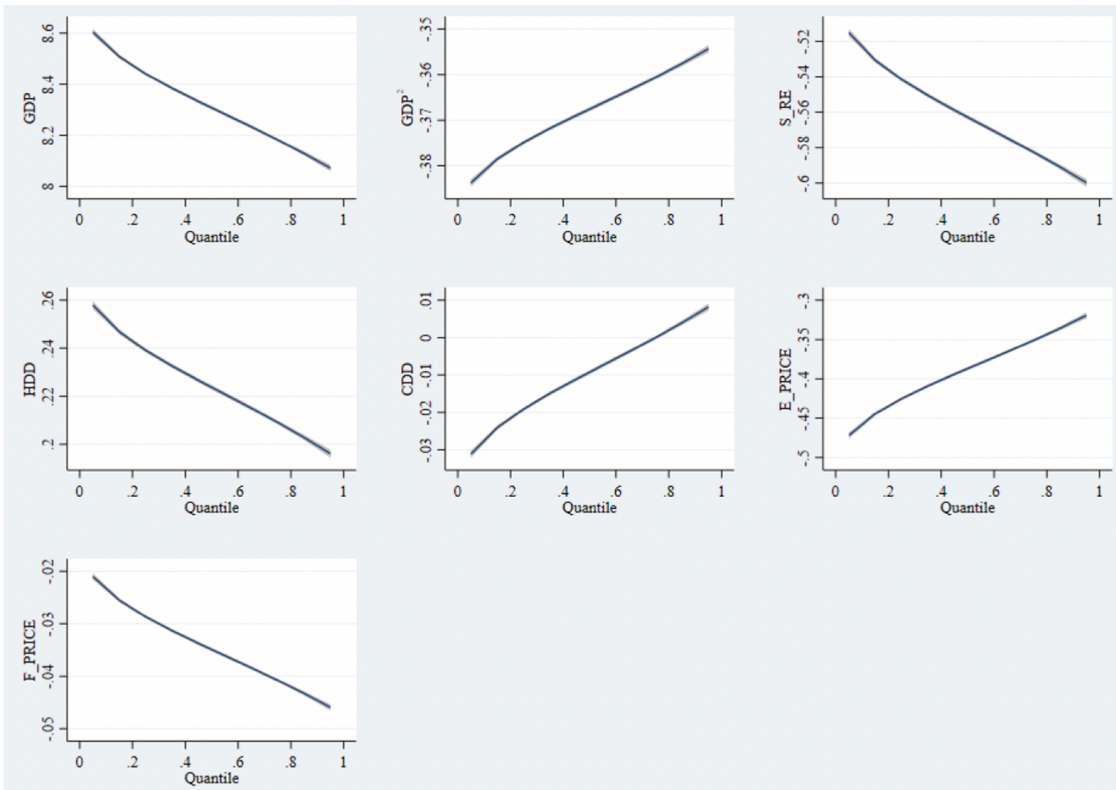

Figure 3: Change in the fixed-effect MMQR coefficients across 48 U.S. states (Jeon, 2022)

Note: GDP: Per capita real gross domestic product by state, S_RE: Share of renewable energy in total energy consumption (%), HDD: A unit of measure to relate the day's temperature to the energy demand of heating at a residence or place of business, CDD: A unit of measure to relate the day's temperature to the energy demand of cooling at a residence or place of business, E_PRICE: Average electricity price, F_PRICE: Average primary energy price. Y-axes explains the coefficients of explanatory variables (GDP, GDP^2, S_RE, HDD, CDD, E_PRICE, F_PRICE) and X-axes portrays the quantiles of the dependent variable (CO_2). The solid line in each panel shows the point MMQR estimation results at different CO_2 quantiles; the shaded bands depict the corresponding 95% confidence intervals.

Figure 3 presents the parameter estimates from the MMQR model, illustrating that the impacts of GDP, renewable energy consumption, and all other explanatory variables on per capita energy-related CO_2 emissions are heterogeneous at different quantiles. Specifically, the effect of per capita GDP is positively significant and heterogeneous for per capita energy-related CO_2 emissions across all quantiles. This result supports that economic growth, which is often accompanied by increased energy consumption, and thus generates CO_2 emissions. The results indicate that a 1% rise in per capita GDP stimulates per capita energy-related CO_2 emissions by 8.064%–8.604%. The positive effects of increases in GDP on CO_2 emissions gradually decreased from lower to higher quantiles.

Another finding is that renewable energy consumption (S_RE) has a negative effect on CO_2 emissions throughout the conditional distribution, a 1% increase in the share of renewable energy leads to a reduction in emissions by 0.515%–0.601%. This implies that investments in renewable energy stimulate technological innovation and facilitate access to clean technologies which contribute to emission reduction. This mitigation effect is stronger at the higher quantiles, potentially due to the fact that states with higher energy-related CO_2 are more dependent on fossil fuels than states with lower CO_2. As a result, the high-CO_2 states experience a more pronounced marginal effect from renewable energy consumption on emission reduction. This pattern highlights the critical role of renewable energy in reducing CO_2 emissions, especially in the case of high CO_2.

Moreover, the results illustrate that an increase in heating degree days (HDD) tends to generate more CO_2 emissions at all quantiles, while the effect of electricity price (E_PRICE) on CO_2 is negative across all quantiles. These effects could be attributed to the positive correlation between residential energy consumption and hotter days, and negative correlation between residential energy demand and electricity price. Notably, results reveal that both heating degree days and electricity price exert a more pronounced impact at the lower quantiles, implying that residential energy consumption may account for a larger share of total energy consumption in states with lower CO_2 emissions.

Furthermore, the primary energy prices (F_PRICE) negatively influence on CO_2 emissions. The mitigating effects of energy prices are stronger at the higher quantiles. This pattern could arise from the varying sectoral energy demands across states. In particular, the industrial or electric power sectors may be the main drivers of energy consumption in high-emission states, rather than the residential sector.

In general, employing the MMQR estimation technique allows researchers to produce relatively robust results by considering the effects of outliers and reducing the potential covariates and heterogeneity.

5 Discussion

Due to the diversity of modelling approaches and the heterogeneity of the analyzed data and results, a meta-analysis could not be performed. Therefore, this thesis consists of the statistical methods and their application in the selected studies. Each method has its strengths, but the selection should be in accordance with the properties of the data and the specific research subject.

The first thing to note is that time series data have specific characteristics that can affect the quality of the subsequent analysis. Specifically, non-stationary data often exhibit trends over time, which could lead to spurious regression where there appears to be a relationship between the variables that actually do not have a real causal association. Therefore, it is essential to conduct a series of tests, such as unit root and cointegration tests, to ensure the data are stationary and meet the preconditions for a regression analysis. In addition, given the sequential nature of time series data, where observations are interrelated over time, lagged terms can effectively capture temporal dynamics and dependencies. As the evolution of CO_2 and its determinants over time is crucial, lagged values of these variables are included in some models to capture dynamic changes.

To explore the influencing factors on CO_2 emissions, the selected studies used mainly Two-Way Fixed-Effect model, time-varying Difference-in-differences model, and Panel Quantile Regression model with fixed effects to perform the empirical analysis. The common feature of these three models is that they all provide better control for unobservable heterogeneity and allow for more precise inferences about the relationships between dependent and explanatory variables by exploiting the temporal and cross-sectional dimensions of panel data.

The 2FE model can eliminate the effects of cross-sectional and time-series heterogeneity by simultaneously controlling for individual and temporal fixed effects for each observation. After adding the control variables that may have influences, the 2FE model can further guarantee the accuracy and consistency of the independent variable coefficients in the panel data. However, Imai and kim (2020) show that the 2FE's ability to adjust for the two types of unobserved confounders simultaneously hinges upon the assumption of linear additive effects, this means that it cannot adjust for unit-specific and time-specific unobserved confounders nonparametrically unless certain functional-form assumptions are made.

Additionally, the 2FE estimator assumes that the unobserved confounding factors are uncorrelated with the treatment variable. If this assumption is violated, the 2FE estimator may produce biased estimates of the treatment effect. For instance, in a study by Jing *et al.* (2022), if the relationship between AI development and carbon emissions is non-linear or interactive with time-varying unobserved confounding factors, the 2FE estimator may not fully capture these effects.

Compared to 2FE, the time-varying DID method provides a more flexible way to control for time-invariant unobserved factors. It can better explain treatment effects and is well-suited for policy evaluation and treatment effects estimation. Unlike common influencing factors, the implementation of public policies is nonrandom. Thus, direct comparison of changes in the mean value of outcome variables before and after the implementation of policies will lead to estimation errors.

By introducing the DID method, researchers are able to better address the issue of treatment timing variation across multiple time periods and units, and can achieve more accurate estimates under the assumption of conditional parallel trends. The assumption of parallel trends requires that CO_2 emissions in the treatment and control groups follow similar trends before the policy intervention. While the DID method excellently handles time-invariant unobserved factors and treatment timing variation, the presence of heteroskedasticity and endogeneity can still bias the results.

FGLS estimator can be used in the presence of heteroskedasticity or autocorrelation (Greene, 2003) and provides efficient parameter estimates under certain assumptions about error structures. FGLS estimates the structure of the error term in advance, and then conducts GLS estimation based on this predicted structure. Therefore, in circumstances where the panel data presents both fixed effects and a complex error structure, the 2FE model and FGLS can be jointly applied. The 2FE model can first be used to eliminate unobservable entity and time effects that might affect CO_2 emissions, after which the FGLS can be employed for efficient parameter estimation.

However, FGLS technique also has some challenges. First, the estimation process requires iterations and complex calculations. Second, it requires a consistent estimate of the error covariance matrix of the model. Under FGLS, we assume a specific structure of the error term covariance matrix, and use that structure to transform our model in order to apply the OLS method, when fixed effects are present, these assumptions might be violated due to the individual-specific effect on the error term. This means that the FGLS may give misleading results if we have wrong assumptions about the structure of the errors. Considering this issue, GMM, with their ability to handle complex error structures and endogeneity problems, may be preferable.

The basic idea of GMM is to estimate model parameters by minimizing the deviation of the moment conditions of the model. GMM, unlike OLS, FE and RE estimation, do not require distributional assumptions, like normality, and can allow for heteroscedasticity of unknown form (Greene, 2003). SYS-GMM estimation is an extension of GMM technique for solving specific problems in panel data analysis. It utilizes a two-step estimator that can correct standard errors, enabling more robust inference in the presence of weak instruments (Blundell and Bond, 1998). In SYS-GMM, fixed effects are differenced out so that the possible effects of invariant individual characteristics (such as the geographical location or culture of a particular country) on the dependent variable can be eliminated. The GMM dynamic models can be one-step system GMM models and two-step GMM models with one lag or two lags. The two-step system GMM models employ residuals of the first-step estimation to estimate the variance–covariance matrix when there is no assumption for independency and homoscedasticity of error terms (Salari and Javid, 2016). Thus, SYS-GMM uses the lagged levels of the variables as instruments for their first differences, which helps to tackle the endogeneity arising from the inclusion of lagged terms (Roodman, 2009).

SYS-GMM also addresses the possible weak instrumental variables issue. An instrumental variable is considered weak when it lacks sufficient correlation with the endogenous regressors it is supposed to serve as an instrument for. Consider a study investigating the impact of GDP per capita on CO_2 emission, an instrumental variable in this case could be trade openness (the sum of exports and imports as a share of GDP). The idea is that trade openness might affect GDP per capita, as nations that trade more may be wealthier due to higher levels of economic activity. The assumption here is that trade openness would not directly affect CO_2 emissions, except through its impact on GDP per capita. However, if the relationship between trade openness and GDP per capita is weak, maybe because other factors like technological advancements or government policies have a much stronger impact on GDP, then trade openness would be a weak instrument in this context. In such cases, it would be difficult to effectively estimate the causal impact of GDP per capita on CO_2 emissions using just a standard instrumental variable approach. Then we have a situation where the instrumental variable is weak. Techniques such as the SYS-GMM, which is more robust to weak instruments, might then be more appropriate for estimating this relationship.

System GMM has an advantage in situations where there is a strong autocorrelation in the variables or when the time series is relatively short and the cross-section is large, because SYS-GMM utilizes both lagged differences and lagged levels of the variables as instruments, thereby increasing the number of available instruments. Thereby, SYS-GMM allows for more instruments and can dramatically improve efficiency (Roodman 2009). However, SYS-GMM ignores cross-sectional dependence and assumes that the panel members have homogenous slope coefficients. For instance, if each studied country has specific characteristics not captured in the model that affect CO_2 emissions, such as differences in policies or economic development, these differences can introduce bias into the estimations. This limitation is especially pertinent given the diverse developmental stages of the regions within the panel. To address this limitation, a panel quantile regression technique was employed to examine the distributional and heterogeneous effect across quantiles (Sarkodie and Strezov, 2019).

Panel Quantile Regression (PQR) approaches are typically employed when influencing factors exhibit different effects based on the conditional distribution of CO_2 emissions. PQR techniques can provide reliable estimates in the presence of outliers, and is the most suitable technique in cases where there is little or no relationship between two variables' conditional means (Binder and Coad, 2011). Moreover, the Method of Moments Quantile Regression (MMQR) estimates the quantile regression model by moments, this involves specifying moment conditions that capture the relationship between the exogenous variables (like policies or socio-economic factors) and the quantiles of interest. By considering these different moments, MMQR captures more details about the error structure, including asymmetry and heavy tails, and can handle more complex error structures because it does not assume a specific distribution of errors and can flexibly model various distributions by matching the moments.

The advantage of MMQR is that it allows the use of methods that are only valid in the estimation of conditional means, while still providing information on how the regressors affect the entire conditional distribution (Machado and Santos Silva, 2019). By estimating the model with fixed effects, the MMQR allows for the identification of the latent effects of the exogenous variables across the conditional distribution of CO_2 emissions. The application of MMQR makes it possible to identify the conditional heterogeneous covariance effects of the determinants of CO_2 emissions by allowing the individual effects to affect the entire distribution. However, when the number of individuals greatly exceeds the length of the study period, the accuracy of the estimates derived from MMQR can decrease dramatically (Machado and Santos Silva, 2019).

The array of statistical methods investigated in this thesis provide a diverse perspectives on the complex interplay between various influencing factors and CO_2 emissions. The selection of methodods should be guided by the quality and quantity of available data, and the plausibility of the underlying assumptions in the specific research context. It is also crucial to carefully consider the limitations of each method, and robustness checks should be conducted to ensure the validity of the results.

Considering the results, the magnitude and direction of the relationships between the influencing factors and CO_2 emissions appear to vary depending on the regions, the period studied, and the specific method employed. This requires a careful and context-specific interpretation of the results.

6 Summary and Prospects

This thesis has reviewed 30 selected articles that applied diverse statistical models to unravel the complex influencing factors of CO_2 emissions. These studies spanned across different countries and regions, each presenting its unique perspective based on the selected variables, methods, and periods analyzed.

Through the systematic review and content analysis, the research focus predominantly lies in exploring socio-economic, energy-related, and policy-related factors as critical determinants of CO_2 emissions. Socio-economic factors, such as urbanization, population growth, per capita GDP, and financial development, can either exacerbate or mitigate CO2 emissions depending on the specific context. Energy consumption and the source of energy significantly influence CO2 emissions. Despite variations across individual studies, the investigated studies generally identified that the use of renewable energy and the transition from fossil fuels can substantially reduce these emissions. Innovations such as AI and industrial robots can also help mitigate emissions by optimizing the industrial structure. Moreover, the impact of these factors can also be influenced by the region's specific circumstances and policies. Implementations of specific policies like low-carbon city pilot policy and smart city pilot policy can effectively reduce emissions by enhancing environmental regulations and promoting green technology innovation.

Moreover, a number of statistical methods were identified in investigating the determinants of CO_2 emissions. In the main part of this thesis, the theoretical foundations of the statistical methods used in the reviewed literature are illustrated, explaining how these methods were applied in the respective studies.

First, a series of time series tests were conducted to understand the characteristics of the panel data. Following this, various regression models are constructed, enabling the use of diverse estimation methods for parameter estimation. The varied modelling approaches encountered in this thesis indicate the flexibility and multiplicity of tools available to researchers in this domain. Then, the results derived from these methods are presented. Through this process, a comprehensive understanding of the factors that influence CO_2 emissions was carried out. The determination of the nexus among parameters quantifying the factors above is crucial to provide guidelines for policymakers to formulate effective policies concerning, e.g., the reduction of greenhouse gases emissions and the energy sources diversification to diminish the dependence on non-renewable sources (Balsamo *et al.*, 2023).

In terms of future developments, the growth of data science stimulated by intelligent systems, enhances the capability to solve complex problems and enables accurate predictions. As such, these advancements aid in the automation of analytical model construction, the identification of patterns, and data-driven decision making with minimal human intervention. These approaches are best suited to cases in which there are nonlinear relationships between the target and response variables. This underscores the potential of these methods in not only understanding the intricate, nonlinear relationships between CO_2 emissions and their key drivers like GDP and energy consumption, but also in outperforming conventional approaches in accuracy.

Several Artificial Intelligence and Machine learning techniques such as artificial neural networks, genetic algorithms and decision trees were identified as being instrumental in predicting and optimizing CO_2 emissions in various aspects of construction. Mardani *et al.* (2020) demonstrated the superiority of artificial neural network (ANN) in a study carried out for G20 countries' CO_2 emissions. The combined method of neuro-fuzzy inference system and ANN techniques showed highly accurate estimation compared to multiple-linear regression for estimating CO_2 emissions based on energy consumption and economic growth, with average absolute error values of 0.065 and 0.522, respectively.

With the innovation of approaches and the expansion of computational capabilities, there is great potential for the development and application of more advanced and precise models in the future. This will contribute to providing robust, data-driven insights to address the complex factors influencing CO_2 emissions. The incorporation of these analytical tools into fields such as the devise effective strategies for reducing CO_2 emissions and tackling climate change, holds significant potential for notable advancements in both academic research and practical applications.

A References

Acheampong, A.O., Amponsah, M. and Boateng, E. (2020) ‘Does financial development mitigate carbon emissions? Evidence from heterogeneous financial economies’, Energy Economics, 88, p. 104768.

Aitken, A. (1936). IV.—On Least Squares and Linear Combination of Observations. Proceedings of the Royal Society of Edinburgh, 55, 42-48.

Ali, A. (2017) ‘Panel Data Models’, p.4.

Anser, M.K.*et al.*(2020) ‘Impact of urbanization, economic growth, and population size on residential carbon emissions in the SAARC countries’, Clean Technologies and Environmental Policy, 22(4), pp. 923–936.

Arellano, M. and 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), pp. 277–297.

Attílio, L.A., Faria, J.R. and Rodrigues, M. (2023) ‘Does monetary policy impact CO_2 emissions? A GVAR analysis’, Energy Economics, 119, p. 106559.

Awan, A. *et al.*(2022) ‘The impact of renewable energy, internet use and foreign direct investment on carbon dioxide emissions: A method of moments quantile analysis’, Renewable Energy, 189, pp. 454–466.

Azam, A. *et al.*(2021) ‘An empirical analysis of the non-linear effects of natural gas, nuclear energy, renewable energy and ICT-Trade in leading CO_2 emitter countries: Policy towards CO_2 mitigation and economic sustainability’, Journal of Environmental Management, 286, p. 112232.

Azimi, M. and Bian, Z. (2023) ‘Provincial and regional analysis of carbon neutrality policy and the environmental Kuznets curve: examining their effect on CO_2 emissions in China’, Environmental Science and Pollution Research, 30(16), pp. 46234–46247.

Balsamo, M., Montagnaro, F. and Anthony, E.J. (2023) ‘Socio-economic parameters affect CO2 emissions and energy consumption – An analysis over the United Nations Countries’, Current Opinion in Green and Sustainable Chemistry, 40, p. 100740.

Baltagi, B.H. (2008) Econometrics. Berlin, Heidelberg: Springer Berlin Heidelberg, p. 174, 222

Barbieri, L. (2005) ‘Panel Unit Root Tests: A Review’, Università Cattolica del Sacro Cuore, Dipartimento di Scienze Economiche e Sociali - Piacenza, p. 9.

Blundell, R. and Bond, S. (1998) 'Initial conditions and moment restrictions in dynamic panel data models', Journal of Econometrics .

Canay, I.A. (2011) 'A simple approach to quantile regression for panel data', The Econometrics Journal, 14(3), pp. 368–386.

Chen, P.*et al.*(2022) 'Do Artificial Intelligence Applications Affect Carbon Emission Performance?—Evidence from Panel Data Analysis of Chinese Cities', Energies, 15(15), p. 5730.

Chen, Y. and Lee, C.-C. (2020) 'Does technological innovation reduce CO_2 emissions?Cross-country evidence', Journal of Cleaner Production, 263, p.121550.

Cheng, C.*et al.* (2021) 'How does technological innovation mitigate CO_2 emissions in OECD countries? Heterogeneous analysis using panel quantile regression', Journal of Environmental Management, 280, p. 111818.

Debone, D., Leite, V.P. and Miraglia, S.G.E.K. (2021) 'Modelling approach for carbon emissions, energy consumption and economic growth: A systematic review', Urban Climate, 37, p. 100849.

Dickey, D.A. and Fuller, W.A. (1979) 'Distribution of the Estimators for Autoregressive Time Series With a Unit Root'.

Dietz, T. and Wa, P. (1994) 'Rethinking the Environmental Impacts of Population, Affluence and Technology'.

Dumitrescu, E.-I. and Hurlin, C. (2012) 'Testing for Granger non-causality in heterogeneous panels', Economic Modelling, 29(4), p. 1453.

Essandoh, O.K., Islam, M. and Kakinaka, M. (2020) 'Linking international trade and foreign direct investment to CO_2 emissions: Any differences between developed and developing countries?', Science of The Total Environment, 712, p. 136437.

Fahrmeir, L. *et al.* (2013) 'Extensions of the Classical Linear Model', in Regression: Models, Methods and Applications. Berlin, Heidelberg: Springer Berlin Heidelberg, p. 180.

Ghouse, G. and Khan, S.A. (2018) 'ARDL model as a remedy for spurious regression: problems, performance and prospectus', p. 14.

Greene, W.H. (2003) Econometric analysis. 5th ed. Upper Saddle River, N.J: Prentice Hall.

Greene and William, H. (2012) Econometric Analysis. Upper Saddle River, NJ: Pearson

Prentice Hall.

Grossman, G.M. and Krueger, A.B. (1995) 'Economic Growth and the Environment', The Quarterly Journal of Economics.

Hansen, C.B. (2007) 'Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects', Journal of Econometrics, 140(2), pp. 670–694.

Hansen, L.P. (1982) 'Large Sample Properties of Generalized Method of Moments Estimators', Econometrica, 50(4), p. 1029.

Harris, R.I.D. and Sollis, R. (2003) Applied time series modelling and forecasting. Chichester, West Sussex, England; Hoboken, NJ: J. Wiley.

Hausman, J.A. (1978) 'Specification Tests in Econometrics', Econometrica, 46(6), pp. 1251–1271.

He, Z. and Maekawa, K. (2001) 'On Spurious Granger Causality', Economics Letters, 73, pp. 307–313.

Ike, G.N., Usman, O. and Sarkodie, S.A. (2020) 'Testing the role of oil production in the environmental Kuznets curve of oil producing countries: New insights from Method of Moments Quantile Regression', Science of The Total Environment, 711, p. 135208.

Im, K.S., Pesaran, M.H. and Shin, Y. (2003) 'Testing for unit roots in heterogeneous panels', Journal of Econometrics, 115(1), pp. 53–74.

Imam, A. (2016) 'On Consistency of Tests for Stationarity in Autoregressive and Moving Average Models of Different Orders', American Journal of Theoretical and Applied Statistics, 5, p. 146.

Jareño, F.*et al.* (2020) 'Bitcoin and gold price returns: A quantile regression and NARDL analysis', Resources Policy, 67, p. 101666.

Jeon, H. (2022) 'CO_2 emissions, renewable energy and economic growth in the US', The Electricity Journal, 35(7), p. 107170.

Jing, Q.-L. *et al.*(2022) 'The Impact of Public Transportation on Carbon Emissions—From the Perspective of Energy Consumption', Sustainability, 14(10), p. 6248.

Johnston, J. (1972) Econometric Methods. New York: McGraw-Hill.

Kao, C. (1999) 'Spurious regression and residual-based tests for cointegration in panel data', Journal of Econometrics, 90(1), pp. 1–44.

Kao, C. and Chiang, M.-H. (2001) 'ON THE ESTIMATION AND INFERENCE OF A COINTEGRATED REGRESSION IN PANEL DATA', pp. 179-222.

Keho, Y. (2020) 'An Empirical Study of Some Driving Factors of CO_2 Emissions: Evidence from Quantile Regression', Bulletin of Applied Economics, pp. 87–95.

Khan, I., Hou, F. and Le, H.P. (2021) 'The impact of natural resources, energy consumption, and population growth on environmental quality: Fresh evidence from the United States of America', Science of The Total Environment, 754, p. 142222.

Koenker, R. and Machado, J.A.F. (1999) 'Goodness of Fit and Related Inference Processes for Quantile Regression', Journal of the American Statistical Association, 94(448), pp. 1296–1310.

Lau, C.K. *et al.* (2023) 'Introducing a new measure of energy transition: Green quality of energy mix and its impact on CO_2 emissions', Energy Economics, 122, p. 106702.

Lechner, M. (2010) 'The Estimation of Causal Effects by Difference-in-Difference Methods Estimation of Spatial Panels', Foundations and Trends® in Econometrics, 4(3), pp. 165–224.

Lin, B. and Xu, B. (2018) 'Factors affecting CO_2 emissions in China's agriculture sector: A quantile regression', Renewable and Sustainable Energy Reviews, 94, pp. 15–27.

Liu, J. *et al.*(2022) 'The effect of artificial intelligence on carbon intensity: Evidence from China's industrial sector', Socio-Economic Planning Sciences, 83, p. 101002.

Liu, X. *et al.*(2022) 'Evaluation of low carbon city pilot policy effect on carbon abatement in China: An empirical evidence based on time-varying DID model', Cities, 123, p. 103582.

Machado, J.A.F. and Santos Silva, J.M.C. (2019) 'Quantiles via moments', Journal of Econometrics, 213(1), pp. 145–173.

Mardani, A. *et al.*(2020) 'A multi-stage method to predict carbon dioxide emissions using dimensionality reduction, clustering, and machine learning techniques', Journal of Cleaner Production, 275, p. 122942.

McNown, R., Sam, C.Y. and Goh, S.K. (2018) 'Bootstrapping the autoregressive distributed lag test for cointegration', Applied Economics, 50(13), pp. 1509–1521.

Pedroni, P. (2004) 'PANEL COINTEGRATION: ASYMPTOTIC AND FINITE SAMPLE PROPERTIES OF POOLED TIME SERIES TESTS WITH AN APPLICATION TO THE PPP HYPOTHESIS', Econometric Theory, 20(03).

Pesaran, M.H. and Smith, R. (1995) 'Estimating long-run relationships from dynamic heterogeneous panels', Journal of Econometrics, 68(1), pp. 79–113.

Rahman, M.M., Alam, K. and Velayutham, E. (2022) 'Reduction of CO_2 emissions: The role of renewable energy, technological innovation and export quality', Energy Reports, 8, pp. 2793–2805.

Roth, J. *et al.* (2022) 'What's trending in difference-in-differences? A synthesis of the recent econometrics literature', Journal of Econometrics, 235(2), pp. 2218–2244.

Salari, M. and Javid, R.J. (2016) 'Residential energy demand in the United States: Analysis using static and dynamic approaches', Energy Policy, 98, p. 643.

Shahbaz, M. *et al.*(2020) 'Public-private partnerships investment in energy as new determinant of CO_2 emissions: The role of technological innovations in China', Energy Economics, 86, p. 104664.

Su, C.-W. *et al.*(2020) 'Trade and technological innovation: The catalysts for climate change and way forward for COP21', Journal of Environmental Management, 269, p. 110774.

Su, C.W.*et al.*(2023) 'Does technology innovation help to achieve carbon neutrality?', Economic Analysis and Policy, 78, pp. 1–14.

Su, K. and Lee, C.-M. (2020) 'When will China achieve its carbon emission peak? A scenario analysis based on optimal control and the STIRPAT model', Ecological Indicators, 112, p. 106138.

Wang, B. *et al.* (2022) 'High-quality energy development in China: Comprehensive assessment and its impact on CO_2 emissions', Energy Economics, 110, p. 106027.

Wang, J.*et al.*(2023) 'Can industrial robots reduce carbon emissions? Based on the perspective of energy rebound effect and labor factor flow in China', Technology in Society, 72, p. 102208.

Wang, Z. and Zhu, Y. (2020) 'Do energy technology innovations contribute to CO_2 emissions abatement? A spatial perspective', Science of The Total Environment, 726, p. 138574.

Wen, S. and Liu, H. (2022) 'Research on energy conservation and carbon emission reduction effects and mechanism: Quasi-experimental evidence from China', Energy Policy, 169, p. 113180.

Wohlrabe, K. (2023) 'Univariate Time Series Analysis', Institute for Statistics, Ludwig-

Maximilians-Universität München, (Course), p.script_part2,5.

Wu, R. *et al.*(2021) 'The drivers of declining CO_2 emissions trends in developed nations using an extended STIRPAT model: A historical and prospective analysis', Renewable and Sustainable Energy Reviews, 149, p. 111328.

Xu, A., Wang, W. and Zhu, Y. (2023) 'Does smart city pilot policy reduce CO_2 emissions from industrial firms? Insights from China', Journal of Innovation & Knowledge, 8(3), p. 100367.

Xu, D. *et al.*(2022) 'Financial development, renewable energy and CO_2 emission in G7 countries: New evidence from non-linear and asymmetric analysis', Energy Economics, 109, p. 105994.

Yirong, Q. (2022) 'Does environmental policy stringency reduce CO_2 emissions? Evidence from high-polluted economies', Journal of Cleaner Production, 341, p. 130648.

Declaration of authorship

I hereby declare that this thesis submitted is my own unaided work. All direct or indirect sources used are acknowledged as references. I am aware that the Thesis in digital form can be examined for the use of unauthorized aid and in order to determine whether the report as a whole or parts incorporated in it may be deemed as plagiarism. For the comparison of my work with existing sources I agree that it shall be entered in a database where it shall also remain after examination, to enable comparison with future Theses submitted.

Munich, July 17$^{\text{th}}$, 2023

Weiyun Hong