|  |  |  |
| --- | --- | --- |
| صورة تحتوي على نص, دائرة, الخط, رسم بياني  تم إنشاء الوصف تلقائياً | **Cairo University**  **Faculty of Graduate Studies for Statistical Research** | صورة تحتوي على نص, رمز, قصاصة فنية, الخط  تم إنشاء الوصف تلقائياً |

**Study On** **Panel Data Methodologies**

**With**

**Application for** **Macroeconometrics**

**(****Inflation Forecasting)**

**By:** [**Mohamed Sayed Abd AL-mgyd**](https://www.linkedin.com/in/mohamed-abd-al-mgyd-9886151b2/)

Statistician and Data Scientist

**Supervised By**

**Dr.** **[Amal Mohamed Abd-Elfatah](https://www.linkedin.com/in/amal-mohamed-abdelfattah-phd-3a6bb6a4/)**

Assistant Professor of Statistics

Department of Applied Statistics and Econometrics

Faculty of Graduate Studies for Statistical Research (FSSR)

# Abstract:

Panel data methodologies have become essential tools in applied Macroeconometrics, enabling researchers to exploit both cross-sectional and time-series variations. This study applies panel data econometric techniques to the problem of inflation forecasting, using a comprehensive macroeconomic dataset spanning 70 countries from 2000 to 2024. The primary target variable is PCPIPCH (Inflation, average consumer prices), with explanatory variables including government fiscal indicators, trade volumes, investment ratios, and labor market conditions.

All variables, except *TRWMA* (from the World Bank), are sourced from the International Monetary Fund (IMF) World Economic Outlook (WEO) database. The dataset, titled WEO\_Data\_Countries, includes macroeconomic indicators such as NGDP\_RPCH (GDP growth), GGR\_NGDP (government revenue), GGXWDG\_NGDP (government debt), LUR (unemployment rate), NID\_NGDP (investment), TM\_RPCH (imports), TX\_RPCH (exports), and several PPP-related measures, all harmonized across countries and years.

The empirical analysis compares multiple panel data models, including Pooled OLS, Fixed Effects (FE), Random Effects (RE), and Dynamic Panel Models such as the Arellano-Bond GMM estimator, to assess their performance in inflation forecasting. Specification tests—such as the Hausman test, Wald test, and tests for serial correlation, stationarity, and heteroskedasticity—are conducted to guide appropriate model selection and to validate assumptions.

The forecasting performance of the models is evaluated using out-of-sample prediction accuracy, measured by Root Mean Square Error (RMSE). Results indicate that dynamic panel models significantly outperform static models, particularly in handling endogeneity and country-specific shocks. The Arellano-Bond GMM approach proves most effective when lagged inflation and macroeconomic fundamentals are used as predictors.

Overall, the study demonstrates that careful model selection, informed by both theory and diagnostic testing, enhances the reliability of inflation forecasts in Macroeconometrics contexts. These insights are valuable for policymakers and researchers aiming to understand inflation dynamics and guide economic policy using panel data econometrics.

This study emphasizes the importance of methodological rigor and appropriate panel estimation techniques in macroeconomic forecasting, offering evidence-based guidance for policymakers and economists concerned with inflation dynamics in a global context.

**Keywords:** Panel Data, Macroeconometrics, Python, Statistical Models, Inflation Forecasting.

**Table of Contents**

[Abstract: 2](#_Toc164015606)

[Introduction: 2](#_Toc164015607)

[Acknowledgments: Thanks, and appreciation to my supervisor Prof. Salah Mahdi. 3](#_Toc164015608)

[Content 3](#_Toc164015609)

[Historical perspective: 3](#_Toc164015610)

[Definition and conditions: 3](#_Toc164015611)

[Distribution and Regression Model: 4](#_Toc164015612)

[Switching from one model to another and vice versa: 5](#_Toc164015613)

[Summary: 6](#_Toc164015614)

[References: 6](#_Toc164015615)

[Selecting a negative binomial Or Poisson regression model by applying it to COVID-19 data: 7](#_Toc164015616)

# Introduction:

Inflation forecasting is a central concern in macroeconomics, as accurate projections of price movements are essential for effective monetary policy, fiscal planning, and investment decision-making. In an increasingly interconnected global economy, inflation dynamics are influenced by a multitude of macroeconomic factors that vary across countries and over time. Traditional time-series methods often fall short in capturing this multidimensional complexity, especially when trying to account for heterogeneity between countries or exploit the full depth of available economic data. This has led to a growing reliance on panel data methodologies, which integrate both cross-sectional and time-series dimensions, allowing for richer and more robust econometric analysis.

Panel data techniques have become indispensable in modern Macroeconometrics research due to their capacity to control unobserved heterogeneity, increase degrees of freedom, and improve the efficiency of estimators. These advantages are particularly relevant when modeling inflation, a variable that is inherently affected by both domestic policies and international economic conditions. This study applies panel data econometric models to the task of forecasting inflation, using a rich and harmonized dataset that spans 70 countries over the period 2000–2024. The focal variable is PCPIPCH (Inflation, average consumer prices), sourced primarily from the International Monetary Fund (IMF) World Economic Outlook (WEO) database, along with a wide array of macroeconomic indicators related to output, government finance, trade, labor markets, and investment.

The study employs and compares a range of panel data estimation methods—Pooled Ordinary Least Squares (Pooled OLS), Fixed Effects (FE), Random Effects (RE), and Dynamic Panel Estimators such as the Arellano-Bond Generalized Method of Moments (GMM). In addition to model estimation, the analysis includes rigorous specification testing, including Hausman tests for model selection, as well as diagnostic checks for stationarity, serial correlation, and heteroskedasticity. Out-of-sample forecast accuracy is evaluated using the Root Mean Square Error (RMSE) as a performance metric.

By identifying the most suitable modeling framework for inflation forecasting within a macro-panel context, this study contributes to the growing literature on empirical macroeconomics. The findings emphasize the effectiveness of dynamic panel models in addressing issues of endogeneity and serial correlation, common challenges in macroeconomic data. Ultimately, the insights gained are intended to support evidence-based policymaking and to advance the methodological foundations of Macroeconometrics forecasting.

# Acknowledgments: Thanks, and appreciation to my supervisor Prof. Amal Mohamed Abd-Elfatah.

overview of the historical development of Poisson regression models and negative binomial regression models: