# CHARLES UNIVERSITY

## FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies


# Uncertainty and House Prices: Empirical Evidence

Master's thesis


Author: Bc. Jiří Kos

Study program: Economics and Finance

Supervisor: prof. Roman Horváth, Ph.D.



Year of defense: 2023

## Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, January 3, 2023

Bc. Jiří Kos

# Abstract

This thesis studies the relationship between house prices, economic fundamentals and uncertainty using panel data from 10 OECD member countries and time series data from the United States. Traditional techniques, such as cointegration testing, are used to find a possible long-run link between house prices and their determinants. Employing both single-equation ARDL and multi-equation VEC models, we find evidence of a possible long-run relationship between house prices and fundamentals in the panel data. The results from the time series analysis are inconclusive, mostly leaning towards no presence of cointegration. A measure of interest rate is a vital determinant in most models., while income does not exhibit a long-run connection with house prices. Moreover, results indicate the importance of uncertainty in determining house price dynamics, exhibiting both negative and positive effects.



| | |
|---|---|
| **JEL Classification** | C22, D80, R20, R21, R28, R30, |
| **Keywords** | house prices, uncertainty, cointegration, economic fundamentals, interest rate |
| **Title** | Uncertainty and House Prices: Empirical Evidence |

## Abstrakt

Tato práce se zaobírá vztahem mezi cenami nemovitostí, ekonomickými fundamenty a nejistotou pomocí panelových dat z 10 členských zemí OECD a časových řad ze Spojených států amerických. Tradiční metody, jako je testování kointegrace, jsou použity k nalezení možného dlouhodobého vztahu mezi cenami nemovitostí a jejich determinanty. Využitím jak jednorovnicového ARDL, tak i vícerovnicových VEC modelů nacházíme v panelových datech důkazy o možném dlouhodobém vztahu mezi cenami nemovitostí a fundamenty. Výsledky analýzy časových řad USA jsou neprůkazné a spíše se přiklánějí k absenci kointegrace. Míra úrokové sazby je ve většině modelů zásadním determinantem, zatímco příjem domácností nevykazuje dlouhodobou souvislost s cenami nemovitostí. Výsledky navíc naznačují důležitost nejistoty při určování dynamiky cen nemovitostí, která má negativní i pozitivní dopady.



| | |
|---|---|
| **Klasifikace JEL** | C22, D80, R20, R21, R28, R30, |
| **Klíčová slova** | ceny nemovitostí, nejistota, kointegrace, ekonomické fundamenty, úrokové sazby |
| **Název práce** | Nejistota a ceny nemovitostí: Empirická evidence |

## Acknowledgments

I would like to thank my parents, friends and colleagues for their unwavering support and valuable advice. I also express my gratitude towards prof. Roman Horváth, Ph.D. for his positive approach and guidance throughout the process of writing this thesis.

Typeset in LaTeXusing the IES Thesis Template.

**Bibliographic Record**

Kos, Jiří: *Uncertainty and House Prices: Empirical Evidence.* Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2023, pages 116. Advisor: prof. Roman Horváth, Ph.D.

# Contents

# List of Tables

# List of Figures

# Acronyms

**ADF** Augmented Dickey-Fuller

**AIC** Akaike Information Criterion

**ARDL** Autoregressive Distributed Lag

**DF** Dickey-Fuller

**ECM** Error Correction Model

**GARCH** Generalised Autoregressive Conditional Heteroskedasticity

**HAC** Heteroskedasticity and Autocorrelation Consistent

**KPSS** Kwiatkowski-Phillips-Schmidt-Shin

**LM** Lagrange Multiplier

**LSDV** Least-square Dummy Variable

**OLS** Ordinary Least Squares

**PP** Phillips-Perron

Pacro:S&P**S&P** Pacro:S&PThe Standard and Poor's

**SIC** Schwarz Information Criterion

**VAR** Vector Autoregression

**VECM** Vector Error Correction Model

**VIX** CBOE Volatility Index - Chicago Board Options Exchange

# Master's Thesis Proposal

| | |
|---|---|
| **Author** | Bc. Jiří Kos |
| **Supervisor** | prof. Roman Horváth, Ph.D. |
| **Proposed topic** | Uncertainty and House Prices: Empirical Evidence |

**Motivation** After the crash of the housing bubble on the United States real estate market, which was one of the lead contributors to the Great Recession, we could see a renewed interest from the economists in the driving factors of the real estate prices. As a result of that, past decade witnessed a birth of many research studies regarding this topic. Past research mainly focused on the link between the economic fundamentals, such as interest rates and household income, and real estate prices in search for their main determinants. The results from the previous literature were rather mixed, when some of them rejected the link between the economic fundamentals and house prices in some countries, while others did not deny the possible connection. This thesis aspires to expand the previous research by adding a measure of financial market uncertainty among the possible factors, which influence the house prices, using time series data mainly from the United States.

Previous research linking uncertainty with real estate prices mostly conducted time series or panel data analysis using the vector error correction model (VECM). Bahmani-Oskooee & Ghodsi (2017) included a measure of economic policy uncertainty in their panel data analysis of the United States. They found out that uncertainty had short-run negative effects in 24 states, but only in 17 states in the long run. On the other hand, Kirikkaleli et al. (2020) used time series data and various causality tests to investigate house prices in Germany. For the studied period, there was a positive correlation between uncertainty and housing sector prices.

Another reason for further analysis of the housing market is a long-term climate of low interest rates, which prevailed in many developed countries during the past decade. Since the previous results (Kishor & Marfatia, 2017; Vizek & Posedel, 2011) were usually unclear in estimating the link between the economic fundamentals and real estate prices, additional research could shed some light on this issue. Finally, as world was struck by the latest pandemic crisis, house prices in many regions seemed unaffected by the spiking uncertainty, which contradicted the expectations and previous findings, for example Aye et al. (2019).

## Hypotheses

Hypothesis #1: Cointegration is present between economic fundamentals and house prices.

Hypothesis #2: Real estate prices tend to increase in low interest rate environment.

Hypothesis #3: House prices are negatively affected by increasing financial markets uncertainty.

**Methodology** As a first step, I will collect and adjust data of real estate prices. Here, I will rely mostly on the house prices database used in Jorda et al. (2019) and Knoll et al. (2017), which they used for examining the real rate of return of different assets and analyzing how the house prices evolved, respectively. This database is unique in its thoroughness and completeness, as it contains house and land prices of various countries for the last 140 years. This will ensure that our model will give reliable estimates using exhaustive long-span data. Next, I will focus the independent variables of the model. To test whether cointegration between house prices and economic fundamentals is present, I will use the mortgage interest rates and household income, as done by Case & Shiller (2003). Adding more variables could decrease the number of studied countries and degrees of freedom in the data panel analysis, since the needed data might not be available for them. This approach was also chosen by Bahmani-Oskooee & Ghodsi (2017). Data for these variables will be collected from the OECD Library (2021) and U.S. Census Bureau, (1984). Finally, financial market uncertainty will be included in the model. There are various methods which can be used to measure uncertainty. For example, AndrĂ© et al. (2017) applies news-based measure of economic policy uncertainty to his analysis of the house prices dynamics. In my research, I will focus on financial markets uncertainty and its relationship with house prices. I expect the estimate of mortgage interest rates to be negative and that of household income to be positive. However, the financial markets uncertainty estimate might be positive or negative. In times of financial distress, economic agents might spend less, which could also reduce the housing demand, therefore the estimate could be negative. On the other hand, the rise in the financial market uncertainty could lead the public to shift towards more safe assets, including housing, so the estimate would be positive (Bahmani-Oskooee & Ghodsi, 2017). I will use the uncertainty index constructed by Jurado et al., (2015), which is used to measure uncertainty for the broader macro economy and the financial sector.

To test the hypothesis and establish the relational model among economic variables in a non-structural way, I will estimate VAR models, more specifically the vector error correction model (VECM). Unlike simple VAR model, which is mostly used to analyze short-term link among variables through impulse responses, VECM is suitable for examining long-term relationships between the variables, since it considers cointegration factor among variables. This model or its variations is used for example by Bahmani-Oskooee & Ghodsi (2017) or Kishor & Marfatia (2017). Firstly, the concept of cointegration will be used on time series data from United States. The reason for using the USA for my research is the availability of data and the maturity of equity market. Next, based on VAR results, cointegrating regression will be estimated using fully modified ordinary least square and dynamic ordinary least square. These methods are used to estimate cointegration equation and were developed by Phillips and Hansen (1990) and Stock and Watson (1993), respectively. Finally, causality tests will be used to find the direction of causal relationship.

**Expected Contribution** In my thesis, I will estimate a vector error correction model to examine the cointegration between economic fundamentals and house prices on both panel and time series data. Unlike several previous studies, I plan to include a measure of financial markets uncertainty index among the explanatory variables. Connecting a comprehensive house prices dataset and versatile economic policy uncertainty index could shed more light on the relationship between economic fundamentals and housing sector, as well as answer the question whether financial uncertainty has negative or positive effect on real estate prices.

### Outline

1. Introduction and motivation: results from the past research on cointegration between housing sector and economic fundamentals were rather mixed. Adding a financial markets uncertainty index among explanatory variables in error correction model and using more comprehensive dataset could clarify the issue, as well as explain the relationship between house sector and uncertainty.

2. Previous literature: Here, I will describe the previous studies focusing on the cointegration of house sector and macroeconomic fundamentals, as well as the studies of uncertainty and its effect on investments.

3. Data: I will describe how I will collect and adjust the house prices data, as well as the data for the explanatory variables used. I will be collecting time series data in my studies.

4. Methods: I will explain causality tests and how to derive the vector error correction model from vector autoregression model and how the cointegration factor is explicitly included in the model. I will estimate cointegrating equation based on these results, as well as run causality tests (for example Granger causality).

5. Results: I will discuss the estimates of the VEC model and results for causality tests for time series data.

6. Concluding remarks: I will summarize my results and compare them with previous literature, explain how they differ and what it implies for future policy decision making.

## Core bibliography

André, C., Bonga-Bonga, L., Gupta, R., & Mwamba, J. W. M. (2017). Economic Policy Uncertainty, U.S. Real Housing Returns and Their Volatility: A Nonparametric Approach. The Journal of Real Estate Research, 39(4), 493-514.

Aye, G., Clance, M., & Gupta, R. (2019). The Effect of Economic Uncertainty on the Housing Market Cycle. Journal of Real Estate Portfolio Management, 25, 67-75. https://doi.org/10.1080/10835547.2019.12090024

Bahmani-Oskooee, M., & Ghodsi, S. H. (2017). Policy Uncertainty and House Prices in the United States. The Journal of Real Estate Portfolio Management, 23(1), 7-86.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty*. The Quarterly Journal of Economics, 131(4), 1593-1636. https://doi.org/10.1093/qje/qjw024

Case, K., & Shiller, R. (2003). Is There a Bubble in the Housing Market? Brookings Papers on Economic Activity, 34(2), 299-362.

Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M., & Taylor, A. M. (2019). The Rate of Return on Everything, 1870â€“2015*. The Quarterly Journal of Economics, 134(3), 1225-1298. https://doi.org/10.1093/qje/qjz012

Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring Uncertainty. American Economic Review, 105(3), 1177-1216.
https://doi.org/10.1257/aer.20131193

Kirikkaleli, D., Gokmenoglu, K., & Hesami, S. (2020). Economic policy uncertainty and house prices in Germany: Evidence from GSADF and wavelet coherence techniques. International Journal of Housing Markets and Analysis, ahead-of-print(ahead-of-print). https://doi.org/10.1108/IJHMA-07-2020-0084

Kishor, N. K., & Marfatia, H. A. (2017). The Dynamic Relationship Between Housing Prices and the Macroeconomy: Evidence from OECD Countries. The Journal of Real Estate Finance and Economics, 54(2), 237-268. https://doi.org/10.1007/s11146-015-9546-8

Knoll, K., Schularick, M., & Steger, T. (2017). No Price Like Home: Global House Prices, 1870-2012. The American Economic Review, 107(2), 331-353.

Neal, T. (2014). Panel Cointegration Analysis with Xtpedroni. The Stata Journal, 14(3), 684-692. https://doi.org/10.1177/1536867X1401400312

Pedroni, P. (1999). Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors. Oxford Bulletin of Economics and Statistics, 61(S1), 653-670. https://doi.org/10.1111/1468-0084.0610s1653

Pedroni, P. (2004). Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis (Department of Economics Working Paper No. 2004-15). Department of Economics, Williams College.
https://econpapers.repec.org/paper/wilwileco/2004-15.htm

Phillips, P., & Hansen, B. (1990). Statistical Inference in Instrumental Variables Regression with I(1) Processes. Review of Economic Studies, 57(1), 99-125.

Stock, J. H., & Watson, M. W. (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. Econometrica, 61(4), 783. https://doi.org/10.2307/2951763

Vizek, M., & Posedel, P. (2011). Are House Prices Characterized by Threshold Effects? Evidence from Developed and Post-Transition Countries. Czech Journal of Economics and Finance (Finance a Uver), 61, 584-600.

OECD (2021), Household disposable income (indicator), Long-term interest rates, Short-term interest rates. doi: 10.1787/dd50eddd-en (Accessed on 26 June 2021)

U.S. Census Bureau. (1984, January 1). Real Median Household Income in the United States. FRED, Federal Reserve Bank of St. Louis; FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/MEHOINUSA672N

# Chapter 1

# Introduction

> *'People aren't as impressed by homes anymore after they saw how they collapsed in price with the financial crisis.'*
>
> \- Robert J. Shiller, American Economist

House prices are the key indicator of the housing market's health, and their changes significantly impact the economy. While there has always been a discussion about the main determinants of house prices, it accelerated in the early 2000s after the dot-com bubble burst. This period was marked by the fast growth of house prices in the United States, which continued until 2006 when the burst of a bubble on the housing market and subsequent subprime mortgage crisis were one of the major contributing factors of the 2007-2008 global financial crisis. The scale at which the crisis in the housing market spilt to the other parts of the economy was the main reason why researchers became more interested in studying the house price dynamics. For this reason, researchers, investors and policymakers need to understand the behaviour of housing prices in specific market conditions and, most importantly, their driving factors. Understanding the main determinants of house prices is especially important for policymakers, as their policy measures might influence the housing market and, through it, the broader economy. The introduction of new tax rates or government expenditure measures, which directly affect the personal income of economic agents, could influence the economy indirectly through the housing market, which might not be the intention of policymakers.

The Great Recession also sparked interest in uncertainty and its possible economic influence. Until then, uncertainty was only viewed as an endogenous response to shocks in other macroeconomic fundamentals. Since then, many studies emerged trying to derive several uncertainty measures, each describing a different section of the economy. The subsequent studies proved that while uncertainty can be an endogenous response to business cycles, some types can also be a source of output fluctuations.

Previous literature usually included personal income and a measure of interest rate among the most essential determinants of house prices. Still, the results of a possible relationship were mixed or leaned towards the rejection of possible long-run links. This thesis expands the previous literature by utilizing both panel and time series data while including a measure of uncertainty among the possible factors affecting house prices. By including the uncertainty index among our explanatory variables, we hope to identify the potential factors driving house prices and examine the properties of this relationship.

The first part of this thesis provides an overview of the previous literature on the topic. In the beginning, several papers are presented, which examined the link between macroeconomic fundamentals and house prices. These papers' authors usually utilised the cointegration and error-correction model for finding a long-run relationship. The following sections provide an overview of the literature regarding uncertainty and its possible measures. Finally, these sections are connected, and several articles linking uncertainty, fundamentals and house prices are presented.

As a next step, we describe the properties of our data, as both panel and time series are utilized to analyze the determinants of house prices. Also, two measures of uncertainty are used throughout the thesis. The first measure, the Economic Policy Uncertainty index, is used in the panel analysis, as it is available for individual countries. The Financial Markets Uncertainty index is then utilized for time-series analysis.

The central part of this thesis is the empirical analysis. The panel section begins with introducing panel unit root tests, which are used to determine the order of integration in heterogeneous panels. This is followed by panel cointegration tests to reveal a possible cointegrating relationship between our variables. Based on the results of these tests, the panel autoregressive distributed lag (ARDL) model is estimated using the dynamic pooled-mean group (PMG) estimator, which allows heterogeneity of short-run coefficients while restricting the long-run homogeneity among panels.

The time-series analysis utilizes the Johansen cointegration testing procedure, a multi-equation system allowing more possible cointegrating relationships. Firstly, the order of integration is determined by the unit root tests. Next, the Johansen cointegration test is conducted to find cointegration among our series. Finally, the vector error correction model is estimated to analyze dynamic relationships between variables. The results are then interpreted using impulse response functions. To check the robustness of our results and explore possible links between the financial market uncertainty and house prices, the ARDL model is also estimated within the time-series analysis.

The last chapter briefly summarizes results from both panel and time-series estimations.

# Chapter 2

# Literature Review

There has been a lengthy discussion regarding housing price dynamics. The debate usually revolved around their driving factors, when researchers focused on the link with economic fundamentals, such as interest rates and income. The discussion further accelerated after the housing bubble crash in the United States real estate market, which was one of the main contributors to the Great Recession. As a result, we have witnessed the birth of many papers during the past decade, which usually tried to explain the relationship between economic fundamentals and the housing sector. With rising interest in determinants of real estate prices, the Great Recession also shifted the focus to redefine the way we look at systematic risk and evaluate uncertainty in the economy. Thus, many research studies emerged, which presented various methods to measure uncertainty and how these indexes can be used to estimate their effect on economic agents. This chapter first introduces previous literature focusing on the determinants of house prices. The second part explains how uncertainty is measured and how it affects various economic variables. Finally, these parts are connected in the final section, where several papers examining the relationship between property prices and uncertainty are introduced.

Past research mainly examined the relationship between economic fundamentals and the housing sector by establishing cointegration with time series or panel data models. Perhaps the most important determinants, which have been used in every research focusing on this relationship, are household income and interest rates. Those variables were the main ones used by Case & Shiller (2003) in their analysis of a housing bubble in the USA housing market. Using quarterly state-level data from 1985 to 2002, this essential paper found that income per capita alone explained changes in real estate prices in most states, while economic fundamentals were not correlated with the housing sector in only eight of them. These states, however, were characterised by significant volatility of housing prices, exhibited long-lasting inertia and thus cannot be easily explained by income.

The importance of including income among explanatory variables was also demonstrated in Quigley & Hwang (2006). Their research investigated the effects of regional and national economic factors on the outcomes of residential housing markets in US metropolitan regions. Using a two-stage least squares regression in an error components framework, they confirm the importance of changes in the economic fundamentals, including income and unemployment. Besides that, new supply on the market and other factors, such as regulation, variations in materials and labour costs, were strongly important in determining house prices. The model suggests that more regulation causes higher and more persistent real estate prices in response to endogenous shocks. McQuinn & O'Reilly (2006) took a slightly different approach to estimating the long-run relationship between fundamentals and real estate prices. Their theoretical model includes demand for housing driven by how much individuals can borrow for housing, which is determined by their disposable income and current mortgage rates. An empirical test of this model is then applied to the Irish housing market. Its results show the importance of a relationship between the borrowing ability of an individual and current house prices in the economy.

Gallin (2006), on the other hand, reached a different conclusion in his analysis of house price development before the Great Depression. Using 27 years of national-level data, he showed that fundamentals, including income, did not explain a rapid increase in house prices on data from the United States. Also, not even more detailed panel-data tests for metro areas data and the bootstrap approach did not reject the possibility of no cointegration between the fundamentals and the housing sector. This outcome raised serious questions about the suitability of error-correction models, which were the mainstream approach to analysing possible cointegration between fundamentals and house prices. Similar results were reported by Mikhed & Zemcik (2009), who also investigated a possible cointegration of various fundamentals, such as income, mortgage rates or stock market condition, with prices of the US house market. Using a univariate and even more suitable panel data unit root test, they found that the real estate sector prices did not align with fundamentals for any panel data subsamples before 2006. Finally, Tsatsaronis & Zhu (2004) also reached a similar outcome in their analysis of residential house price dynamics, when household income had little to no explanatory power over the dynamics of housing prices in all developed countries of interest.

So far, all mentioned studies have focused on the linear framework of house price dynamics, regardless of their conclusions. However, as the following research papers suggest, real estate prices have often shown nonlinear historical patterns. One of the first papers, which discovered this phenomenon, was Abelson *et al.* (2005), whose authors researched changes in Australian house prices from 1970 to 2003. Besides estimating a long-run equilibrium model to describe the long-term determinants of property prices, an asymmetric error correction model was developed to study the short-term house price changes. While fundamentals, including income and mortgage rate, appeared significant as long-run determinants of house prices, the authors also found that property prices react more strongly to changes in fundamentals during an economic boom, as opposed to economic depression periods. Kim & Bhattacharya (2009) examined the Smooth Transition Autoregressive model (STAR) to look for possible nonlinear properties of house price dynamics in the United States. Over the 1969-2004 period, property prices exhibited nonlinear behaviour in all analysed regions except one, which would mean that the nonlinear method should be used to study how the housing sector reacts to changes in various macro fundamental variables.

Outcomes of these papers were utilised by Zhou (2010), who also argued that past empirical research only examined linear cointegration between house prices and fundamentals, possibly ignoring the nonlinear cointegration. This could arguably lead to false conclusions that no cointegration is present. To test this hypothesis, a model is constructed to test for cointegration on data from ten cities across the United States by a two-step testing procedure. The results were mixed as in other mentioned studies. However, linear cointegration between the housing sector and macro fundamentals was found only in data from one city. In six different cities, there was evidence of a nonlinear relationship. Vizek & Posedel (2011) further expanded the nonlinear approach in their analysis of real estate price determinants and adjustment properties in developed and transition countries. The threshold cointegration method, unlike traditionally used linear cointegration, allows for threshold adjustment in the short run. At the same time, it maintains linearity in the long run, and their results were somewhat mixed. House prices were characterised by threshold effects in most transitory countries. Still, for developing countries, which experienced a dramatic increase in housing sector prices over the following period, there seemed to be no evidence of threshold cointegration. For some countries, however, the Granger causality test indicated that prices were not wholly disconnected from macroeconomic fundamentals, mainly disposable income and mortgage rates.

Regardless of exploring linear or nonlinear properties of house prices and their determinants, the results seem to be mixed using both methods, when some of the papers conclude that there is a connection between fundamentals and house prices, while others reject it. For that reason, many studies included various variables in their models to capture other drivers of housing prices.

Besides using already mentioned fundamental determinants, which were part of almost every research paper regarding this topic, Geng (2018) focused on including policy, institutional and structural factors in his model, such as tax incentives or rent controls. Based on cross-country data from twenty advanced economies belonging to OECD, they found that fundamentals have a long-run impact on equilibrium house prices. Still, the results and effects enormously vary across countries due to the different policy and structural factors. For example, income had a much higher impact on house price growth in the Netherlands, where housing finance has a huge tax relief, as opposed to Canada or New Zealand, which is not tax-favoured. Also, mortgage rates seem to have a much stronger relationship with house prices in countries with elastic supply markets, such as the USA, than in countries with less flexible markets.

Tsatsaronis & Zhu (2004) in their cross-country analysis provided evidence that inflation was a significant determinant of property price dynamics for the following period. Together with that, they incorporated many variables related to specific mortgage conditions, proving to be statistically significant in the model. For example, real estate prices were found more sensitive to short-term rates in those countries where floating rates prevailed as the most common mortgage interest payment setting.

Focusing on borrowing patterns and dynamics of the housing prices, Lamont & Stein (1999) found out that property prices in those countries, where the property owners have a high loan-to-value ratio, react more quickly to shocks in income per capita, as opposed to less leveraged markets. These findings again show how borrowing and leverage can influence the response of house prices to changes in fundamentals. Together with that, their research also showed the persistence of the house prices growth, meaning that past development of real estate prices was a strong predictor of future prices. The high growth persistence was also captured by Posedel & Vizek (2009) in their analysis of house price determinants in transition and EU countries. Besides examining the effect of fundamentals on house price development, Abelson *et al.* (2005) included housing supply and various equity price indexes, which proved to be statistically significant in the model and essential predictors of house price dynamics.

As with other economic phenomena, The Great Depression forced all researchers to redefine their approach to measuring systematic risk and uncertainty and study its effects on the real economy. This led to a sudden spike in literature, which tried to explain the properties of uncertainty and how it interacts with various economic assets, including housing. Before discussing this, however, we must first understand what uncertainty is and how it is measured.

When discussing uncertainty in this Thesis, we are referring to economic uncertainty, meaning uncertainty related to economic variables. The definition of the term is essential, because outside economics, it is easily confused with *risk*. However, the distinction is quite clear from an economic point of view. Knight (1921), in his landmark book *Risk, Uncertainty and Profit*, defined the terms as a result of his research of profit and its creation. According to him, the *risk* is present only if events in future happen with probability, which can be reasonably measured.

On the other hand, *uncertainty* is current if we cannot measure the possibility of future outcomes due to the lack of information at the moment. In other words, uncertainty applies to situations where the odds of future events are not quantifiable. Therefore, due to its broad definition, uncertainty does not have any objective measure, so analysing its behaviour and the nature of its relationship with economic activity is very challenging. Although this Thesis will mainly deal with the uncertainty index estimated for the financial sector, past literature used many different approximation techniques to assess the uncertainty levels. Generally, we can divide these approaches into four main categories: news-based approach, forecast disagreement and uncertainty based on data from financial markets.

The first method relies on the fact that information about uncertainty might be incorporated into the news. Here, the constructors of newspaper-based uncertainty indexes usually focus on using specific keywords and finding their frequencies in newspaper and online news web coverage data using various search algorithms. This, however, does not mean that newspapers and media distributors would cause uncertainty on purpose. Instead, it is based on the assumption that indicators of uncertainty might be present in the news and that search patterns of economic agents are different in times of financial distress or boom (Moore 2017). Alexopoulos & Cohen (2009) developed an uncertainty index based on articles from the New York Times, represented by several monthly reports that include keywords related to uncertainty and the economy. Next, the authors use VAR models to analyse if uncertainty shocks are the source of significant business cycle fluctuations. The authors then find out that macroeconomic aggregates such as output, productivity, employment or investment all decrease due to an unanticipated spike in uncertainty.

Perhaps the most popular policy-related uncertainty index based on newspaper coverage frequency was developed by Baker *et al.* (2013). Its popularity comes from its comprehensiveness, as the authors used three types of components while constructing the index. The first component, for the United States, is based on the frequency of articles published by ten major US media houses, including the Washington Post or the Wall Street Journal. These articles are then used to create a harmonised index, which reflects the volume of news discussing economic policy uncertainty. The second component of the index then utilises federal tax code provisions set to expire to account for uncertainty about the future progress of tax code levels imposed by policymakers. Finally, the third component uses dispersion between predictions of individual forecasters about the future status of the Consumer Price Index and federal and state public expenditures to estimate the uncertainty about economic-policy-related variables. The authors then used several methods for evaluation, including a human review of more than 5000 newspaper articles, to conclude that the index serves as a good proxy for changes in economic policy uncertainty. However, this index was later updated in Baker *et al.* (2016) and contains only the first news-based component, and it is available for other countries besides the United States.

The second approach, which can be used to measure uncertainty, is the measure of dispersion among individual economic forecasters. In times of high economic uncertainty, we should see higher dispersion among forecasts due to the wider potential distribution of outcomes (Moore 2017).

One of the first researchers to utilise this method was Bomberger (1996), who measured the uncertainty by the conditional variance of expected inflation about an individual forecast on Livingston data. For the period from 1946 to 1994, he found that there is a stable relationship between disagreement and uncertainty, indicated by a positive correlation between the conditional variance of inflation forecast errors and debate among individual forecasters at the time of the forecast. Sheng & Thevenot (2012) measured the dispersion of earnings forecast for several companies. Instead of building their model on the established proposition that uncertainty has idiosyncratic and standard components, they proposed a new method, which measures earnings forecast uncertainty as a sum of dispersion among forecasters and the variance of errors made estimated by generalised autoregressive conditional heteroskedasticity (GARCH) model. This approach is based on both private and public information available to analysts at the forecast time. Based on the earnings forecasts, they found that their new measure provides better estimates of forecast uncertainty than the older methods.

As Rich *et al.* (2012) argues, forecast dispersion might be an imperfect proxy for uncertainty since it could capture the disagreement among analysts instead of uncertainty. Because of these issues, Clark *et al.* (2020) derived a multiple-horizon stochastic volatility model from tracking time-varying uncertainty in forecast errors. In their case, uncertainty is estimated from survey data as point forecasts, gathered from pooling information embedded in different periods where the forecast errors happened. Their stochastic volatility model reported better survey forecast uncertainty measures accuracy than traditional variance approaches on data, which included forecasts made on several macroeconomic variables.

Lastly, measures of uncertainty based on financial market data usually utilise the volatility of stock markets. The importance of stock market volatility jumps was provided by Bloom (2009), who found a strong correlation between stock volatility and other uncertainty indicators, such as the spread of firm-level profit or GDP Livingston forecasts. Also, estimates from vector autoregression demonstrated that shocks in market volatility resulted in a short-run drop in industrial production. This outcome shows that jumps in volatility have an impact on the real economy. Baker & Bloom (2013) used stock market volatility and levels as proxies for business conditions in the economy in their analysis of the causal relationship between uncertainty and economic growth, with natural disasters, the threat of terrorism and various political shocks as instruments of the stock market proxies. Evidence from cross-country panel data showed that the first and second moment of business conditions proxies was highly statistically significant in explaining output growth.

The Chicago Board Options Exchange Market Volatility Index (the VIX index), which measures the volatility of S&P500 index options, was used as a proxy for uncertainty by Caggiano *et al.* (2014), who studied the impact of uncertainty shocks on dynamics of unemployment. Estimating non-linear vector autoregression using United States post-WWII data, they found shocks to uncertainty statistically significant in the model. Moreover, the magnitude of the response of unemployment levels to such shocks was much higher than using only standard linear VARs. As stated in Moore (2017), the biggest drawback of using stock volatility measures as a proxy for uncertainty is their indirect relationship to economic activities. Whereas the company's earnings are connected to the current financial state, most of the short-run stock variance is explained by other factors, as defined by Shiller (1981).

Attempting to overcome the inevitable shortcomings of these uncertainty measuring methods, other studies tried to implement combinations of these approaches. In the first version of their paper, Baker *et al.* (2013) integrated news-based and forecast disagreement uncertainty measures in constructing their Economic Policy Uncertainty Index for the United States. The uncertainty Index composed of all three measurement methods was used by Moore (2017) in his paper, where he estimated the effects of uncertainty on the Australian economy. Firstly computing the uncertainty measurement methods independently, he found that all of them behave similarly around significant events, meaning that some underlying process is present. His economic uncertainty index then tries to capture this process by smoothing noise brought in by any individual measure. The results show that jumps in uncertainty decrease investment and employment but increase household savings while simultaneously reducing consumption of durable goods.

In their highly influential paper, a different approach was taken by Jurado *et al.* (2015). They argued that the adequacy of using proxies or indicators, such as market volatility or dispersion of forecasts, to measure uncertainty relies too much on their correlation with latent stochastic process underlying uncertainty. However, the conditions under which those proxies would be tightly linked to the theoretical notion of uncertainty are unique. To illustrate, cross-sectional dispersion in company-level profit can vary over the business cycle due to the different cyclicalities of a specific industry. Also, besides uncertainty movement, stock market volatility can change due to adjustment of risk aversion or sentiment of the investors, as well as by swings in leverage. To address these issues, the authors propose new uncertainty measures directly related to macroeconomic activity, accessible from dependencies on observable economic indicators and the structure of specific theoretical models. Their indexes are based on the notion that financial decision-making depends more on the predictability of the development of the economy and less on the variability of various economic indicators. The authors used two post-war datasets to construct time-varying uncertainty indexes. The first, the macro dataset, uses data from hundreds of macroeconomic and financial indicators.

In contrast, the second firm-level dataset consists of profit growth observations from more than 150 companies. Their results show significant independent variance in their estimates compared to traditional measures of uncertainty using proxies. Most importantly, the results find far fewer important uncertainty episodes than the measures using proxies. Still, when they happen, those episodes are more significant, more persistent and correlated with actual macroeconomic fluctuations. These findings suggest that much of the proxies variation is not driven by uncertainty.

One of the paper's authors also further expanded this research in Ludvigson *et al.* (2021), which tried to explain whether uncertainty is a source of business cycles or an endogenous response to them. Besides using an extensive macroeconomic uncertainty index as suggested in Jurado *et al.* (2015), the authors followed the same approach to construct a broad-based financial uncertainty index, which is based solely on data from financial markets and has not been used in previous literature. Also, the economic policy uncertainty index from Baker *et al.* (2016) was included for comparison. Using novel structural VAR models, they have found that higher macroeconomic and policy uncertainty in recessions is usually an endogenous response to income fluctuations. This means that macro and policy uncertainty increases if the economy goes into recession. However, there is no evidence that positive shocks to either economic policy or macro uncertainty decrease economic activity, as several theoretical models suggest. On the other hand, financial market uncertainty was found to cause a sharp and consistent decline in economic activity, which implies that an increase in financial uncertainty exogenously impacts the economy and causes recessions. Also, no evidence suggests that adverse activity shocks would hurt financial uncertainty.

Our Thesis builds on the outcomes of this article mainly by incorporating their estimate of financial markets uncertainty among explanatory variables of the model explaining housing prices. Detailed estimation of financial market uncertainty is provided in the next chapter, using the method suggested by Jurado *et al.* (2015) and Ludvigson *et al.* (2021). The index is updated every half-year and is publicly available on a website created by one of the paper's authors.

The last part of this section presents an overview of research which studied the relationship between uncertainty and the housing sector. To the best of our knowledge, only a few studies tried to connect uncertainty with house prices, and most of them were written in the past decade. Bahmani-Oskooee & Ghodsi (2017) added a news-based measure of economic policy uncertainty from Baker *et al.* (2016) into their analysis of property prices in United States. Besides policy uncertainty, they added household income and mortgage rates among explanatory variables, following the approach of Case & Shiller (2003). According to their estimates, cointegration between uncertainty and house prices was found in 35 U.S. states. Uncertainty had a short-run negative effect on property prices, which lasted into the long run in some of them. Choudhry (2020) also included a measure of economic policy uncertainty in his analysis of real estate prices in England and Wales. Empirical analysis with autoregressive distributed lag bounds test for cointegration showed a stable long-run relationship of house prices with its determinants, including policy uncertainty. Also, uncertainty was found to have a strong negative impact on the housing sector.

The same measure of economic policy uncertainty was used by Kirikkaleli *et al.* (2021) in their investigation of a bubble in the German housing market. Through various causality tests, the authors reached a different conclusion, which contradicted outcomes from Bahmani-Oskooee & Ghodsi (2017) and Choudhry (2020). Uncertainty was found to have a strong positive correlation with the housing sector index throughout most of the studied period. Similar results were reported by Aye *et al.* (2019), who examined the spillover effect of uncertainty on the duration probability of the housing market cycle in 12 OECD countries. Using a discrete-time duration hazard model, their results show that higher economic uncertainty increases the likelihood of exiting busts in the housing market. However, uncertainty was found not to influence the possibility of leaving booms and normal times. Based on that, the authors suggest that housing might be a good protection against economic uncertainty. André *et al.* (2017) found evidence of structural breaks and non-linearity between housing returns and news-based measures of economic policy uncertainty, meaning uncertainty affected real housing return and their volatility. More importantly, these results stand even after controlling for other macroeconomic and financial determinants, which implies that uncertainty impacts housing returns directly and not only indirectly through its influence on the economy and financial markets, in particular. The results also implicate significant tail risk for real estate investors since large shocks in uncertainty generated disproportionate decreases in housing returns.

Lastly, Su *et al.* (2016) also investigated the causal link between economic policy uncertainty and housing returns in Germany. Testing the relationship's stability using the bootstrap rolling window causality test on the estimates of the vector autoregressive models, the authors find no stable link between the variables, meaning policy uncertainty had no effect on property prices in the studied period. On the other hand, feedback from several sub-periods indicates a possible causal link from housing sector dynamics to policy uncertainty, which varies over time. The causal connection of the opposite direction might be the stability of the German real estate market due to the social welfare nature and rational institutional and regulatory arrangements.

# Chapter 3

# Data

In our analysis of the relationships between the housing sector, economic fundamentals and uncertainty, we will utilize two different datasets and approaches. The first part of our paper uses a panel dataset of selected countries from the Jordà-Schularick-Taylor Macrohistory Jordà *et al.* (2017), and the Economic Policy Uncertainty index from Baker *et al.* (2016). The panel dataset will be subject to unit root and cointegration testing, after which we will estimate an appropriate model in a single-equation framework.

To examine possible multiple cointegrating relationships, we will also follow the Johansen (1988) approach of testing for cointegration. Based on the results from cointegration tests, the vector autoregression model will be specified for a multi-equation analysis of our variables. For this, we will use monthly data for the United States of America. Time series analysis will allow us to see the effect of uncertainty rising from financial markets, for which we will include the Financial Market Uncertainty index from Ludvigson *et al.* (2015) and recently Ludvigson *et al.* (2021).

Due to their complexity, this chapter first describes the methodology behind the uncertainty indices used throughout the analysis. The following sections present variables used in the panel dataset and data used for the time series analysis of the United States.

## 3.1 Uncertainty Indices

### 3.1.1 Financial Market Uncertainty

Throughout this section, we will describe the measure of financial markets uncertainty developed by Jurado *et al.* (2015) and Ludvigson *et al.* (2021).

Since there is not a single objective measure of uncertainty, past techniques usually utilized forms of proxies or indicators to estimate it, such as financial markets volatility, disagreement among forecasters or profit dispersions.

Although it is tempting to utilize one of these measures due to their commonness and relative simplicity, evidence suggests that the connection between them and the theoretical notion of uncertainty is questionable at best. In our case, using the volatility of a stock index as a proxy for uncertainty in financial markets could lead to serious bias in results since some changes in volatility might be unrelated to shifts in uncertainty. For example, asset volatility might be influenced simply by investor sentiment or interest rate structure changes, while these episodes do not have to be necessarily associated with increased uncertainty (Jurado *et al.* 2015).

This notion demonstrates the shortcomings of using such uncertainty measures, and the interpretation of results based on these tools would be flawed, at least. The issue is, however, how to solve these shortcomings and create a complex index.

To create ultimate uncertainty indexes, which are as free as possible from dependencies on various economic indicators and the structure of theoretical models, Jurado *et al.* (2015) and Ludvigson *et al.* (2021) introduced new uncertainty measure, which is tightly linked to the macroeconomic activity. The measure is based on the notion that economic decision-making is determined by changes in the predictability of the economy, not by individual variability of specific economic indicators. Within this framework, a large amount of estimated uncertainties compiled from panel data is aggregated. More formally, the authors define h-period ahead uncertainty in variable $y$ as:

$$U^y_{jt}(h) = \sqrt{E[(y_{jt+h} - E[y_{jt+h}|I_t])^2|I_t]}. \tag{3.1}$$

This equation defines $h$-period ahead uncertainty $U^y_{jt}$ in the variable $y_jt$ as a conditional volatility of the unforecastable components of the future value of the series. The expectation $(E(.|I_t)$ is made concerning the information $(I_t)$, which is available to economic agents in time $t$. From the equation, it is evident that the uncertainty in the variable rises as a result of a rise in the expectations today about the square error in forecasting $y_{jt+h}$. The conditional expectation of squared error is calculated from the stochastic volatility model. Based on this, Jurado *et al.* (2015) construct macroeconomic uncertainty by aggregating individual uncertainty at each date using weights $w$ as:

$$U_t^y(h) = plim_{N_y \to \infty} \sum_{j=1}^{N_y} w_j U_{jt}^y(h) = E_w[U_{jt}^y(h)] \tag{3.2}$$

As explained in Ludvigson *et al.* (2021), an augmented diffusion index forecast replaces the conditional expectation $E_w[U_{jt}^y(h)]$ in 3.2 to allow for nonlinearities. Assumed to have a factor structure, autoregression predictions are augmented with several common factors estimated from time series $x_{it}$. To account for nonlinearities, the elements are calculated from raw data and have polynomial terms included.

In this Thesis, we will use the financial markets uncertainty derived from a financial dataset, covering the sample 1987:01-2021:12 and consisting of 148 measures of economic indicators. The dataset is constructed solely from the financial market time series already used in Ludvigson & Ng (2007) and is updated to include more recent observations. It includes valuation ratios such as default and terms spreads, divided-price and earnings-price ratios, yields on corporate bonds or Treasuries, growth rates of dividends and prices, and a cross-section of various equity returns. The index is regularly updated by the authors of the Jurado *et al.* (2015) and Ludvigson *et al.* (2021) and it is available on the website of one of the authors[1].

Figure 3.1 is taken from Ludvigson *et al.* (2021) and shows the development of the Financial Market Uncertainty index from 1960 to 2015. Shaded areas in the plot represent recession dates, according to the National Bureau of Economic Research. In this period, uncertainty was highest during the Black Monday market crash in 1987 and at the start of the financial crisis in the 2008-2009 period. With this in mind, we see how the index reacts to various shocks in financial markets.

[1] https://www.sydneyludvigson.com/macro-and-financial-uncertainty-indexes

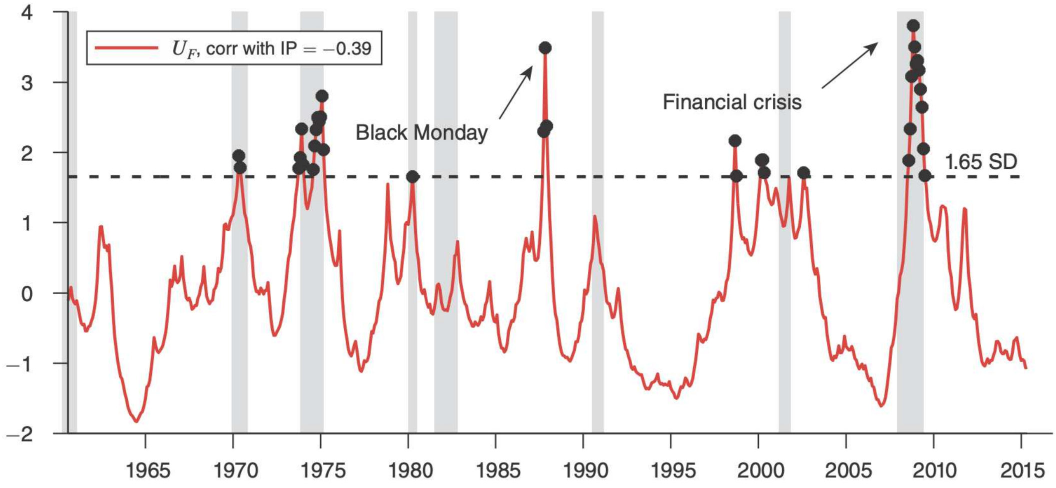


Figure 3.1: Financial Market Uncertainty index
*Source:* Ludvigson *et al.* (2021)

### 3.1.2 Economic Policy Uncertainty

Unfortunately, the Financial Market Uncertainty index from Jurado *et al.* (2015) and later Ludvigson *et al.* (2021) is not available on a country-specific level and cannot be used for panel analysis. Thus, we have decided to include Economic Policy Uncertainty Index from Baker *et al.* (2013)[2] and later Baker *et al.* (2016) or Davis (2016)[3], which is available on both global and country-level for several developed economies. Moreover, Horvath & Kapounek (2022) found that this index is also heavily correlated with the Financial Market Uncertainty index, making it a good alternative for our panel model.

Unlike the Financial Market Uncertainty index, derived from a macroeconomic model, this index is mainly based on newspaper coverage frequency. In the case of the United States, the index utilizes coverage of 10 leading newspapers[4]. This is followed by a textual search for keywords like "uncertain" or "uncertainty" connected with "economic", as well as some policy terms, for example, "congress" or "legislation". Quantifying newspaper coverage is the most flexible approach and is used for most of the country-specific indices constructed by the authors (Baker *et al.* 2016).

For the United States, however, Baker *et al.* (2016) report another two com-

[2]This is the original version of Baker *et al.* (2016) and contains additional policy uncertainty measures.

[3]Constructs the Global Economic Policy Uncertainty index

[4]The index relies on the following newspapers: USA Today, the Miami Herald, the Chicago Tribune, the Washington Post, the Los Angeles Times, the Boston Globe, the San Francisco Chronicle, the Dallas Morning News, the New York Times, and the Wall Street Journal.

ponents to create the index - the number of expiring federal provision tax codes and dispersion among forecasters. The first uses the present value of weighted tax code provisions set to expire in 10 years, reported by the Congressional Budget Office, which presents a notion about the future development of the federal tax code. The last component of the US policy uncertainty index utilizes dispersion among predictions of economic forecasters about future levels of federal, state and local expenditure and consumer price index. To increase the policy uncertainty measures over time and countries, however, data from these two components are used solely for the USA policy uncertainty index, while the newspaper coverage approach is used for the majority of country indices reported on the authors' web page[5].

In our panel analysis, we will be using country-specific policy uncertainty indices for the following countries: Australia, Canada, Germany, France, United Kingdom, Italy, Japan, Netherlands, Sweden and the United States, with data ranging from 1985 to 2020.

Figure 3.2 is taken from Baker *et al.* (2016) and presents the Economic Policy Uncertainty index specifically for the United States from 1985 to 2015. From the figure, we can see how the index reacts to various policy shocks and other events that might have triggered higher uncertainty about future policy paths. We see that the dispute about the level of the debt ceiling and the effects of government spending in 2011 caused the most significant spike in policy uncertainty during the period, followed by the events of 9/11 in 2001 and the beginning of the Second Gulf War in 2003. With this in mind, we see how the index reacts to various shocks in financial markets.

[5] All reported indices of Economic Policy Uncertainty are available at www.policyuncertainty.com

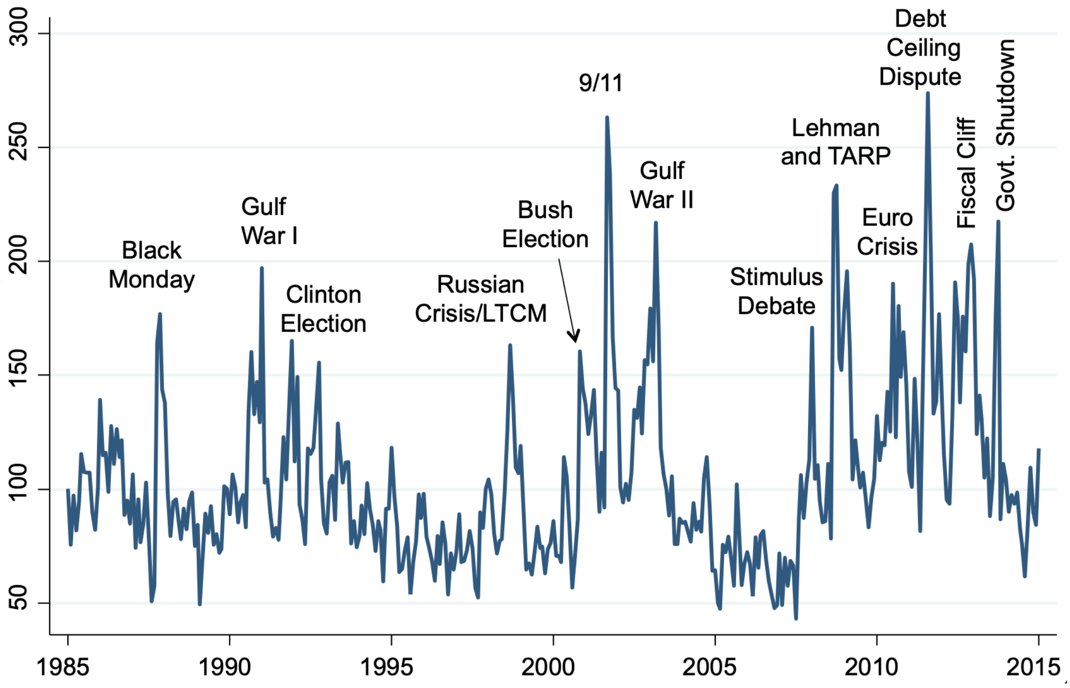


Figure 3.2: USA Economic Policy Uncertainty index
*Source:* Baker *et al.* (2016)

One of the paper's authors also created a Global Economic Policy Uncertainty index in Davis (2016), which is presented in Figure 3.3. Computed in current and PPP-adjusted prices, we can see that the global index behaves similarly to the USA-specific index, with minor exceptions. From both Figures 3.2 and 3.3, we also see how policy uncertainty reacts differently to some news than uncertainty arising from financial markets. For example, the debt ceiling dispute regarding government policies in 2011 caused a much more significant spike in the Economic Policy Uncertainty index. On the other hand, the Black Monday stock market crash in 1987 was a far more significant event for the Financial Market uncertainty index than for policy uncertainty.

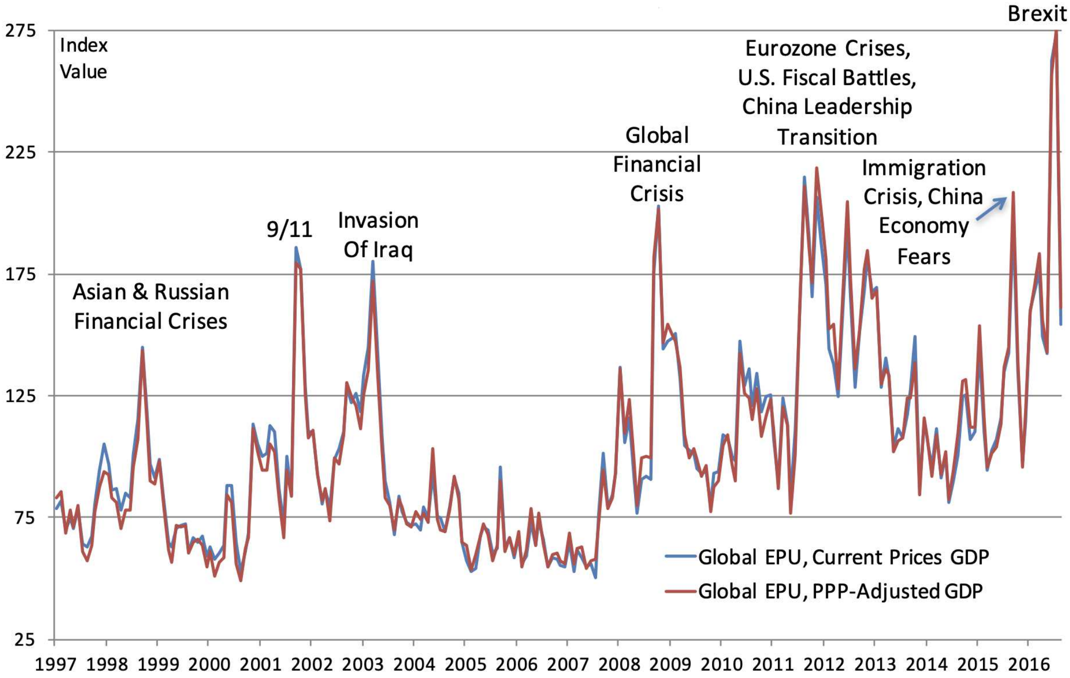


Figure 3.3: Global Economic Policy Uncertainty index
*Source:* Davis (2016)

## 3.2 Panel Data

In this Thesis, we will use two primary datasets to study the cointegration between fundamentals, house prices and uncertainty. The first part of our analysis builds on the exhaustive database of macroeconomic data from Jordà *et al.* (2017), which covers periods from 1870 to 2020. Jordà-Schularick-Taylor Macrohistory Database is unique in its thoroughness and completeness, covering 18 economies from 1870 annually. The compilation of such a long time frame was made possible by collecting data that were already available but scattered in various studies. This dataset will be used to perform panel cointegration on selected countries from OECD.

#### General Equation

Following Equation 3.3 in general, untransformed form denotes our basis for determining the house prices in our panel analysis:

$$hpnom_{it} = \alpha + wage_{it} + rconpc_{it} + unemp_{it} + ltrate_{it} + cpi_{it} + polunc_{it} + \epsilon_{it} \quad (3.3)$$

where[6]:

- *hpnom* is the index of nominal house prices with the base year 1990,
- *wage* is the index of nominal wages of all employees with the base year 1990,
- *rconpc* is the real consumption per capita index with the base year 2006,
- *unemp* is the average unemployment rate, denoted in percentage,
- *lrate* is the nominal long-term interest rate, denoted in percentage,
- *cpi* is the consumer price index with the base year 1990,
- *polunc* is the economic policy uncertainty index from Baker *et al.* (2016). As the index is reported using a monthly frequency, it was recalculated into yearly data by taking the average of the observations.

[6]More information about construction of the Jordà *et al.* (2017) dataset is available at the author's website: https://www.macrohistory.net/database/

Ideally, we would like to include as many relevant variables in the model as possible without overfitting it. However, to increase degrees of freedom for our model, keep a long time period and also include a lot of cross-sections, we have decided to include only the most commonly used in the relevant literature. In line with Case & Shiller (2003) and Bahmani-Oskooee & Ghodsi (2017), we included *wage* and *lret* variables, which will serve as a measure of household income and interest rate in our model. Also, we have decided to include average consumption per capita, as represented by *rconpc*, to better capture households' purchasing power over the studied period. We expand our model with measures of inflation *cpi* and unemployment *unemp*, which follows the methodology in Abelson *et al.* (2005), Quigley & Hwang (2006) and Tsatsaronis & Zhu (2004). Being aware of other determinants of house prices, which have been used in previous studies, we have decided to include only the most frequently used variables in our model to increase degrees of freedom and the number of case studies for such a long period.

The financial markets uncertainty index from Ludvigson *et al.* (2021) is not available for specific countries. Thus, we have decided to incorporate the economic policy uncertainty index developed by Baker *et al.* (2016), available on both global and country levels. Furthermore, this index is also heavily correlated with the Ludvigson *et al.* (2021) financial uncertainty index, which has been found in Horvath & Kapounek (2022). Using the policy uncertainty index should provide us with a clear understanding of the role of uncertainty in determining house prices.

Regarding the expected effect of exogenous variables, we expect wages to be positively correlated with real estate prices. Higher wages should result in higher housing demand, which should translate into higher prices. Also, we expect the CPI index to positively affect house prices, as they should be increasing with the overall price level. On the other hand, unemployment and long-term rates should negatively affect house prices. A higher unemployment rate dampens the general purchasing power, thus reducing the demand for housing. An increase in interest rates will make mortgage products more expensive, decreasing demand and prices.

The expected effects of consumption and uncertainty are not so easily predictable. If housing is deemed a consumption, increasing consumption should result in higher house prices. On the other hand, if housing is perceived as an investment, higher consumption could also decrease housing prices, as economic agents make a trade-off between consumption and investment. Also, the effect of uncertainty can be positive or negative. If higher uncertainty forces people to spend less, it could lead to lower housing demand and, thus, lower prices. On the contrary, if higher uncertainty causes people to adjust their investments towards safer assets, like housing, it could also lead to a price increase.

### Data Properties

Our primary panel data set consists of annual data ranging from 1870 to 2020 for 10 OECD member countries - Australia, Canada, Germany, France, United Kingdom, Italy, Japan, Netherlands, Sweden, and the United States of America. These countries were selected to analyze economies with long-established and developed free markets. Including other countries, such as Czechia or Slovakia, where the period of the free market economy is relatively recent, could seriously skew our results, as asset prices were determined by factors other than supply and demand during the communist era. Also, the country-specific uncertainty index is only available for selected countries.

Regarding the transformation of our data, we follow the example of Mikhed & Zemcik (2009) and Bahmani-Oskooee & Ghodsi (2017), who used the logarithmic form of variables for panel unit root and cointegration testing. Also, authors of our dataset Jordà *et al.* (2017) often utilize log transformations of several variables from their database, for example, in Knoll *et al.* (2017). Using logarithmic transformation will help to normalize our dataset, as well as with the interpretation of our results.

The only exception, where we will not take a logarithmic transformation, is the variable *ltrate*. The long-term rate series, which represents the interest rate level in our Equation 3.3, contains negative observations. Using a logarithm would deprive us of important information contained in the series. Furthermore, since *ltrate* is already expressed in percentage and the observations do not oscillate far from zero, using the non-transformed form of the variable should not significantly affect our results.

Summary statistics of all untransformed[7] variables at individual sample are displayed in Table 3.1:

[7]Descriptive statistics of transformed variables is available in Appendix A, Table 5.3

Table 3.1: Full-range Panel Summary Statistics

| | HPNOM | WAGE | CONSPC | UNEMP | LTRATE | CPI | POLUNC |
|---|---|---|---|---|---|---|---|
| Mean | 59.69 | 42.21 | 41.07 | 5.50 | 5.18 | 41.39 | 125.13 |
| Median | 8.04 | 3.98 | 25.67 | 4.87 | 4.56 | 11.12 | 107 |
| Maximum | 509.18 | 285.29 | 120.43 | 24.90 | 20.22 | 202.05 | 542.77 |
| Minimum | 0.00 | 0.00 | 4.07 | 0.20 | -0.51 | 0.00 | 37.60 |
| Std. Dev. | 96.17 | 66.62 | 32.22 | 3.64 | 2.70 | 55.22 | 70.39 |
| Skewness | 2.15 | 1.57 | 0.85 | 1.27 | 1.42 | 1.27 | 2.62 |
| Kurtosis | 7.49 | 4.35 | 2.27 | 5.48 | 6.05 | 3.20 | 11.98 |
| Observations | 1220 | 1507 | 1504 | 1182 | 1504 | 1510 | 300 |

As can be seen from Table 3.1, the number of observations available for each variable is different, with *polunc* having the lowest amount of observations and *rgdppc, cpi, lrrate* the highest. Since *polunc* variable has significantly lower observations than the rest of our dataset, we will have to perform more estimations to check our results' robustness. Thus, we will estimate Equation 3.3 on more variations of our dataset, using both balanced and unbalanced panel data.

By comparing estimates of several data structures, we will be able to analyze the cointegrating relationship between fundamentals and uncertainty and assess the importance of uncertainty in our model. Also, using both balanced and unbalanced panels will help us check our estimates' robustness.

## 3.3 Time Series Data

The second part of our study will focus on a time series analysis of the United States. Here, we will mostly follow the approach of Posedel & Vizek (2009) on USA monthly data, ranging from 1987 to 2021. In addition, to understand the effects of uncertainty on the housing market, we will include Financial Markets Uncertainty from Jurado *et al.* (2015) and later Ludvigson *et al.* (2021).

#### General Equation

While we had to limit the number of explanatory variables in our panel model to increase the number of cross-sections and dimensions, we did not face such restrictions in our time series analysis. The Equation 3.4 in the untransformed form denotes our basis for determining the housing market prices:

$$\begin{aligned} hpindex_t = \alpha + income_t + mortg_t + pop_t + rent_t + cpi_t \\ + costsw_t + unrate_t + fmunc_t + \epsilon \end{aligned} \tag{3.4}$$

where *hpindex* is the house price index, *income* is the personal income, *mortg* is the mortgage rate, *pop* is the population level, *rent* is the rent of a primary residence, *cpi* is the consumer price index, *unrate* is the unemployment rate and *fmunc* is the Financial Market Uncertainty index. To expand:

- **S&P/Case-Shiller U.S. National Home Price Index** (*hpindex*) represents housing prices and is the dependent variable of our model. It is the leading measure of U.S. residential real estate prices, tracking changes in the value of a residential real estate in the USA. The index is seasonally adjusted with a base year 2000[8] and it is being constructed by S&P Dow Jones Indices LLC. The data has been retrieved from FRED, Federal Reserve Bank of St. Louis.

- **Personal Income** (*income*) is received by individuals and includes provisions for labour, the land and capital used in current production, as well as net current transfer payments from businesses and the government. The data are seasonally adjusted and denoted in billions of USD. The source of the personal income is the U.S. Bureau of Economic Analysis and has been retrieved from FRED, Federal Reserve Bank of St. Louis.

- **30-Year Fixed Rate Mortgage Average in the United States** (*mortg*) represents the average interest rate on mortgages in our model. It is constructed by Freddie Mac, which surveyed lenders on the rates and points for their most popular 30-year fixed rate on their mortgage products. Since it is reported weekly, it has been recalculated into monthly frequency. Similarly to other series, it has been retrieved from FRED, Federal Reserve Bank of St. Louis.

- **Population** (*pop*) represents the population level in our model and includes the resident population and armed forces overseas. The observations are denoted in thousands and the series is reported by the U.S. Bureau of Economic Analysis. It has been retrieved from FRED, Federal Reserve Bank of St. Louis.

[8] Year 2000 = 100

- **Rent of Primary Residence** (*rent*) is an U.S. City average index of changes in rent expenditures, reported by U.S. Bureau of Labor Statistics. It is an index with base years 1982-1984 and has been retrieved from FRED, Federal Reserve Bank of St. Louis.

- **Consumer Price Index** (*cpi*) is the Consumer Price Index for All Urban Consumers: All Items in the U.S. City Average and represents the effect of inflation in our model. The index is constructed by the U.S. Bureau of Labor Statistics, with the years 1982-1984 serving as a base for our model. The data has been downloaded from FRED, Federal Reserve Bank of St. Louis.

- **Earnings in Construction** (*costsw*) is the Average Hourly Earnings of Production and Nonsupervisory Employees, Construction, which serves as an indication of construction costs in our Equation. It is denoted in dollars per hour and seasonally adjusted. The data source is the U.S. Bureau of Labor Statistics, retrieved from FRED, Federal Reserve Bank of St. Louis.

- **Unemployment Rate** (*unrate*) represents the number of unemployed as a percentage of the labour force, where the labour force data are restricted to people 16 years of age and older. It is a seasonally adjusted rate reported by the U.S. Bureau of Labor Statistics and has been retrieved from FRED, Federal Reserve Bank of St. Louis.

- **Financial Markets Uncertainty** (*fmunc*) represents uncertainty index arising from financial sector used in Jurado *et al.* (2015) and later Ludvigson *et al.* (2021), which has been further described in Section 3.1. The index is being updated regularly by the authors or the papers and has been retrieved from their website[9].

We expect personal income, population, CPI, earnings in construction and rent to be positively correlated with house prices. Increases in income and population level should result in higher demand for housing, which could further translate to house price increases. Increases in the earnings of construction workers mean higher building costs, which might be projected into house prices by the developers or the construction plans could be forfeited entirely, resulting

[9] Macro and Financial Uncertainty Indexes, available at https://www.sydneyludvigson.com/macro-and-financial-uncertainty-indexes.

in a supply decrease and higher prices. Also, we expect house prices to increase if the rents are rising since people will be more inclined to buy a house if rents go up.

On the contrary, mortgage and unemployment rates are expected to affect real estate prices negatively. An increase in interest rates on mortgage products will decrease the demand for housing and thus lower house prices. Lower housing demand and price decreases could also result from a higher unemployment rate.

However, the effects of financial market uncertainty are not so easily predictable. Higher uncertainty might stimulate people to decrease their spending, including mortgage payments expenditures, thus reducing housing demand and housing prices. On the other hand, people might also be motivated to adjust their portfolios towards real assets, housing included, which is viewed as a safer investment. The effect of uncertainty on real estate prices can therefore be positive or negative.

### Data Properties

In selecting our variables, we followed the same approach as Mikhed & Zemcik (2009) with two exceptions. Firstly, we used the S&P/Case-Shiller U.S. National Home Price Index instead of other alternatives since it is already reported monthly. Secondly, we did not include any stock market index as a proxy for the stock market in our model since financial uncertainty, which is part of our model, already captures essential market moves. As with the panel analysis, we will be using log transformation of the variables for unit root and cointegration testing and subsequent model specification.

Our time-series dataset consists of USA monthly data for the 1987:m1 to 2021:m12 period. Following Table 3.2 presents the descriptive statistics of our data in an untransformed form:

Table 3.2: USA Time-Series Summary Statistics

| | HPINDEX | INCOME | MORTG | POP | RENT | CPI | COSTSW | UNRATE | FMUNC |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 133.99 | 10732.87 | 6.34 | 291502.31 | 218.67 | 191.20 | 19.88 | 5.88 | 0.90 |
| Median | 138.98 | 10036.50 | 6.26 | 293226.50 | 211.15 | 189.00 | 19.21 | 5.50 | 0.87 |
| Maximum | 280.41 | 24371.90 | 11.28 | 332640.00 | 355.54 | 280.13 | 31.25 | 14.70 | 1.56 |
| Minimum | 63.96 | 3831.40 | 2.68 | 241857.00 | 121.20 | 111.40 | 11.99 | 3.50 | 0.63 |
| Std. Dev. | 51.07 | 4788.87 | 2.21 | 28004.37 | 65.97 | 44.71 | 5.33 | 1.65 | 0.19 |
| Skewness | 0.43 | 0.43 | 0.30 | -0.17 | 0.36 | -0.02 | 0.28 | 1.33 | 0.85 |
| Kurtosis | 2.32 | 2.20 | 2.10 | 1.77 | 2.00 | 1.81 | 1.89 | 5.53 | 3.38 |
| Observations | 420 | 420 | 420 | 420 | 420 | 420 | 420 | 420 | 420 |

Table B.1 in Appendix B provides summary statistics of variables after logarithmic transformation. Graphs of individual time series are presented in Appendix B, Figure B.1.

# Chapter 4

# Methodology

After acquiring and arranging data for panel and time series analysis, this chapter presents the methodology behind our approach to studying the cointegrating relationship between house prices, fundamentals and uncertainty. Firstly, we will lay out the "three-step procedure", which will be used to analyze possible cointegration in our panel dataset. Next, the methodology of the Johansen framework, as laid out in Johansen (1988) and Johansen (1991), is presented to utilize it for time series analysis of the USA.

## 4.1 Panel Analysis

Our panel analysis will consist of three steps in total. Firstly, we need to perform panel unit root tests on our variables, which will help us to determine the order of integration of our variables. If our variables have the same order of integration, we will continue with panel cointegration tests to find any cointegrating relationship between our variables. Finally, we will choose an appropriate single-equation model for our data.

### 4.1.1 Panel Unit Root Testing

As a first step, we will examine the order of integration and stationarity of our data. The order of integration denoted as *I(d)* is a statistis describing the unit root process in a time series. More specifically, it reports minimum differences needed to obtain a stationary series. A time series is said to be stationary if its properties, such as mean, variance or others, do not change over time. Formally, the marginal and all joint distributions of a stationary time series process are invariant across time (Wooldridge 2013).

As the general concept of cointegration states, which has been introduced in Granger (1981) and later formalized in Engle & Granger (1987), two-time series are said to be cointegrated if they have the same order of integration $d$ and their linear combination is integrated of the order less than $d$. Thus, before trying to find a cointegrating relationship between our variables, we must first analyze if our variables are integrated in the same order.

Fortunately, statistical tools can be utilized to examine the stationary of a series. For panel data, panel unit root tests are the most commonly used instrument. Though most of the panel unit root tests are tailored for a balanced dataset, some tests can also handle unbalanced data, which is essential in our case since our full range panel dataset from the Schularick library is unbalanced. The most common of these tests are the Im-Pesaran-Shin test and Fisher-type tests (Choi 2001).

The Im-Pesaran-Shin test is a t-test for unit roots in heterogeneous panels. It is based on the mean of individual Dickey-Fueller t-statistics and allows for individual effects, as well as time trends and time effects. The test was first introduced in Im *et al.* (2003) and is available for any commonly used software. Other tests which allow heterogeneous cross-sections for panel data are the Fisher-type tests developed by Maddala & Wu (1999) and later Choi (2001). Generally, they combine p-values of unit root tests from specific panels using four methods proposed by Choi (2001).

Both Fisher-Type tests and the Im-Pesaran-Shin test have the following hypothesis:

- $H_0$: All panels contain a unit root.
- $H_1$: One or more series in the panel are generated by a stationary process.

After conducting the panel unit root tests and determining whether our variables have the same order of integration, we will proceed with panel cointegration tests to find possible long-term relationships between variables.

### 4.1.2 Panel Cointegration Testing

As cointegration became a widely studied topic in econometrics, researchers quickly realized that finding a sufficiently long time series is not as easy as it would seem. Furthermore, the shortcomings of Dickey-Fuller unit roots tests for univariate analysis became increasingly apparent. Together with panel unit root testing, these issues led to the development of panel cointegration testing, which unlocked huge potential for multi-dimensional analysis (Örsal 2008).

Since then, many cointegration tests have been invented and can be divided into residual-based and maximum-likelihood-based categories. For our analysis, we will be using two-panel cointegration tests based on the residual method - Pedroni (1999) and Kao (1999).

#### Pedroni (1999)

Panel cointegration test developed by Pedroni (1999) is suitable for both univariate and multivariate models. It is based on four within-dimension and three between-dimension tests for possible cointegrating relationships. Firstly, the test builds on the computation of regression residuals of the hypothesized cointegrating regression, which can take the following form:

$$\begin{aligned} y_{i,t} = \alpha_i + \beta_{1i}x_{1i,t} + \beta_{2i}x_{2i,t} + ... + \beta_{Mi}x_{Mi,t} + \epsilon_{i,t} \\ t = 1, ...., T; i = 1, ..., N; m = 1, ..., M. \end{aligned} \tag{4.1}$$

where T is the number of observations over time, N refers to the number of panel members, and M is the number of regression variables. The Equation 4.1 is then estimated using the least squares method for every cross-section to obtain the cointegration test statistics.

The results from Equation 4.1 are then used to construct the panel cointegration tests. Seven statistics are created - four within-dimension and three between-dimension tests.

More specifically, the within-dimension tests are based on estimators, which pool the autoregressive coefficients across cross-sections for the unit root tests. In contrast, the between-dimension statistics are based on estimators computing the averages of individually estimated coefficients for each cross-section in the panel. The tested hypothesis is as follows:

- $H_0$: No cointegration (p = 0).
- $H_1$: All panels are cointegrated.

### Kao (1999)

Second-panel cointegration test, which we will use in our analysis, was invented by Kao (1999). His paper focused mainly on spurious regression in panel data and studied the asymptotic properties of the least-square dummy variable (LSDV) estimator. The outcomes of his analysis indicate that asymptotics from the LSDV estimator is different from those contained in spurious regression for a simple time series. Since this has substantial implications for residual-based cointegration tests, Kao (1999) introduced Dickey-Fuller and Augmented Dickey-Fuller types of cointegration tests. For the bivariate case, the following model can be considered:

$$y_{i,t} = \alpha_i + \beta x_{i,t} + \epsilon_{i,t}, \qquad t = 1, ...., T; i = 1, ..., N \tag{4.2}$$

where $\alpha$ is a fixed effect, which varies across observations in the panel, $\beta$ represents the slope parameter and $y_{i,t}$, $x_{i,t}$ follow a random walk process for all $i$. After defining a long-run covariance matrix, the Dickey-Fuller test is then applied to an estimated residual of the model. The tests are described under the following hypothesis:

- $H_0$: No cointegration (p = 0).
- $H_1$: All panels are cointegrated.

Once we perform the tests, we can determine whether there is a cointegrating relationship present between our variables.

### 4.1.3 Panel ARDL Model

In selecting the appropriate model for our panel data, previous studies on the topic of cointegration present us with two main approaches. If we have a low number of observations, and the results from panel unit root tests show a mixture of $I(0)$ and $I(1)$ series, the autoregressive distributed lag (ARDL) model should be the most suitable approach for our data. On the other hand, if we have a sufficient number of observations, variables are integrated of order $I(1)$ and we suspect more than one cointegrating relationship, a multi-equation vector error-correction model (VECM) would be a better option.

This section only briefly describes the properties of ARDL and the advantages and disadvantages of its use, while the VECM is described in the time-series section. The model selection and specification will be based on results from panel unit root and panel cointegration tests.

### Autoregressive Distributed Lag Model

The ARDL Model was first proposed in Pesaran *et al.* (1999) and later expanded in Pesaran *et al.* (2001) to explain both short and long-run effects between variables of interest. The main feature of this approach is the possibility of testing for a level relationship between an endogenous variable and its regressors, even for a series with mixed order of integration.

The motivation behind the invention of this model is that all previous approaches focused on variables integrated of order one, $I(1)$. Not only is this a very restrictive assumption, which could force us to exclude some important variables, but it requires us to perform additional pre-testing, bringing further uncertainty about the outcomes of our model. The ARDL model relaxes this assumption and allows the inclusion of both $I(0)$ and $I(1)$.

The generalized ARDL(p,q) model can take the following form:

$$Y_t = \gamma_{0i} + \sum_{i=1}^{p} \delta_i Y_{t-i} + \sum_{i=0}^{q} \beta_i' X_{t-i} + \epsilon_{it}, \qquad i = 1, ..., k \tag{4.3}$$

where $Y_t'$ is a vector and all series in $X_t'$ are allowed to be $I(0)$ and $I(1)$, $\gamma$ is a constant, $\beta$ $\delta$ are coefficients, $p$,$q$ represent optimal lag order and $\epsilon_{it}$ is a vector of error terms.

While providing consistent results even in small samples and allowing the inclusion of both $I(0)$ and $I(1)$, the panel ARDL model also has many shortcomings. Most importantly, it is a single-equation system, providing small thoroughness into our analysis, where more cointegrating relationships between the series could be present. Furthermore, as opposed to a reduced form of the VAR framework, the ARDL model requires that explanatory variables are at least weakly exogenous. Ignoring this requirement leads to many distorted conclusions if the endogeneity of the variables is not addressed.

## 4.2 Time Series Analysis

Our time series analysis of the USA will consist of two main parts. Firstly, we will test for a possible long-run cointegrating relationship between fundamentals and house prices using the Johansen testing procedure. In the second part, an appropriate model will be chosen based on the unit root and cointegration testing results.

### 4.2.1 Unit Root Testing

As with the panel analysis, we need to determine the order of integration of our time series before testing for cointegration. To do so, we will employ several types of unit root tests.

One of the first approaches to test for the presence of unit root was proposed by Dickey & Fuller (1979). Their Dickey-Fuller is based on a simple autoregressive model:

$$y_t = py_{t-1} + \mu \tag{4.4}$$

A unit root in a time series is present if $p$ equals zero. The test was further adjusted for larger and more complex time series, and this augmented version of the test in one of the most popular statistical tools used for unit root testing. While the testing procedure is the same as for the original version, the Augmented Dickey-Fuller (ADF) test is applied to the following model:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + ... + \delta_{p-1} \Delta y_{t-p+1} + \mu \tag{4.5}$$

where $\alpha$ is a constant, $T$ is a time trend and $p$ is the optimal lag order, as the model contains lagged values of the dependent variable $Y_t$. The test's null hypothesis states that a unit root is present in the time series, $\beta = 0$.

The ADF test has some drawbacks. Most importantly, it cannot distinguish between pure and unit-root-generating processes and becomes biased in rejecting the null hypothesis. This happens mainly for the moving average processes. Fortunately, there are other unit root tests which use a different methodology.

The second unit root test, which we will utilize, is the Kwiatkowski-Phillips-Schmidt-Shin (KPSS), developed by Kwiatkowski *et al.* (1992), and it is based on simple linear regression. Its most important feature, however, is that it tests the null hypothesis of stationarity against the alternative that it follows a unit root process. This is in contrast with most alternatives, which test the null hypothesis of a unit root.

Lastly, the Phillip-Perron unit root test, proposed by Phillips & Perron (1988), will be performed for robustness checks. The test builds on the ADF by addressing the issue that the $y$ generating process might have a higher order of autocorrelation than the original equation explains. However, unlike the ADF, which solves this issue by adding more lags of the dependent variable into the equation, the Phillip-Perron test performs a non-parametric adjustment to the t-statistic of the test.

### 4.2.2 Cointegration Testing

#### Engle-Granger Procedure

There are two main methodologies to examine cointegrating relationships between multiple series. First is the Engle-Granger procedure, developed in Engle & Granger (1987) and uses a single-equation model to test for cointegration, considering only one cointegration vector. The relationship is tested with the following procedure:

1. After the series are tested for stationarity using unit root tests, and their order of integration is identified to be $I(1)$, we estimate a linear regression model using OLS, with one of the time series as the dependent variable and the other time series as the independent variables:

$$y_t = \beta x_t + \mu_t \tag{4.6}$$

2. The estimated residuals from the model are then tested for stationarity using ADF or KPSS model.

3. If the results from unit root tests indicate that the residuals are stationary, meaning $H_0$ is rejected by the ADF test, the cointegration is present in the model. On the contrary, cointegration is not present if the residual series contains a unit root.

The main drawback of the single-equation Engle-Granger approach is that if the model has more than two variables, there can be more than one cointegrating relationship. For this reason, we will use the multi-equation Johansen procedure, which tests the cointegration between more non-stationary variables.

#### Johansen Procedure

The Johansen cointegration testing is used to determine the possible cointegrating relationship among two or more $I(1)$ series. It is based on a concept of multivariate cointegration, which was developed by Søren Johansen in a series of papers, mainly Johansen (1988), Johansen & Juselius (1990) and Johansen (1991). The approach is built on a generalization of a simple univariate cointegration test used to study cointegration between two-time series.

The starting point of the Johansen methodology is the general form of a vector autoregressive (VAR) model of order $p$, without drift:

$$y_t = \mu + A_1 y_{t-1} + ... + A_p y_{t-p} + w_t \tag{4.7}$$

where $\mu$ is the vector-valued mean, $A_i$ are the matrices of coefficients for each lag, and $w_t$ is the zero-mean multivariate Gaussian noise term. This can be further differenced and rewritten into the form of the Vector Error Correction Model (VECM):

$$\begin{gathered} \Delta y_t = \mu + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + w_t; \\ \Pi = \sum_{i=1}^{p} A_i - I; \qquad \Gamma_i = -\sum_{j=i+1}^{p} A_j \end{gathered} \tag{4.8}$$

Johansen estimates the VAR/VECM model by a maximum likelihood (ML) method to determine the number of cointegrating vectors $r$ in the series. Johansen proposes two different likelihood tests for this purpose: The Trace test and the Maximum Eigenvalue test (Hjalmarsson & Österholm 2007).

The Trace test evaluates several linear combinations in series by testing the null hypothesis of a $r$ cointegrated vectors against the alternative of $n$ vectors:

$$\gamma_{trace} = -T \sum_{i=r+1}^{n} ln(1 - \hat{\lambda}_i) \tag{4.9}$$

where $\lambda_i$ is the estimated eigenvalue, defined as a nonzero vector that changes at most by a scalar factor from the matrix $\Pi$. If the null hypothesis is rejected, a cointegrating relationship in the sample is confirmed.

The Maximum Eigenvalue test is similar to the Trace test but has a different hypothesis. Here, the null hypothesis of $r$ cointegrated vectors is tested against the alternative of $r$+1.

$$\gamma_{max} = -T ln(1 - \hat{\lambda}_{r+1}) \tag{4.10}$$

None of these test statistics follows a chi-square distribution but is tested on critical values provided in Johansen & Juselius (1990). If the null hypothesis is rejected, cointegration is present in our series (Hjalmarsson & Österholm 2007).

The main disadvantage of the Johansen test is its sensitivity to lag selection. Thus, we will use one of the information criteria to choose the optimal lag length.

### 4.2.3 VAR Framework and VECM

Vector Autoregression (VAR) is a statistical approach to modelling dynamic relationships between multiple series. The concept was first introduced by Sims (1980) and extended the univariate autoregressive models (AR).

While the univariate AR models are used to analyze the link between a dependent variable and its lagged values, VAR generalizes the AR model by allowing for multivariate series. Each variable in VAR is then modelled as a linear combination of its lagged values and lagged values of other variables. This allows us to study dynamic relationships between variables and analyze the relative importance of each series in determining the other variables in the model.

Treating all variables in the model as endogenous is the model's main advantage. The variables depend on many features than just its lagged values, as opposed to univariate AR models, which means that they might be suited to fit the data more comprehensively. This is important for the subsequent interpretation of results or forecasting. On the other hand, VAR is a system of estimations, which makes interpreting coefficients rather tricky. Also, there is no straightforward approach to lag selection for VAR. Instead, we have to rely on other methods, like information criteria.

A VAR model of order $p$ with $q$ exogenous variables can be defined as:

$$y_t = \alpha + A_1 y_{t-1} + ... + A_p y_{t-p} + B_1 x_t + ... + B_q x_{t-q} + \epsilon_t \tag{4.11}$$

where $x_t$ is the vector of independent variables.

All series are assumed to be stationary for the VAR system to be stable. Including non-stationary variables could lead to spurious results, and the interpretation of estimated coefficients would not be appropriate. Solving these issues often requires us to differentiate our series, meaning we lose crucial long-run information about the relationship between our series. In this case, VAR would only capture the short-run effects among variables.

### Vector Error-Correction Model

In case our series are $I(1)$, and we suspect one or more cointegrating relationships, it is more suitable to use the Vector Error-Correction Model (VECM), which is a restricted form of VAR with cointegration restrictions in its specification. It is an expansion of the error-correction model, which is used to study both short-run and long-run effects between variables. The idea behind the error-correction model is that short-run dynamics are also influenced by the last period's deviation, error, from long-run equilibrium. The speed of adjustment to the long-run equilibrium is then directly estimated and is an essential factor driving the dependent variable's short-run movements. The VECM expands the single-equation error-correction model into a multi-equation system, where more cointegrating relationships and, thus, long-run equilibriums are possible.

If cointegration is found in series, the error-correction model, as proposed in Engle & Granger (1987), can be defined as:

$$\Delta y_t = \alpha + \beta_1 \Delta x_t + \beta_2 (y_{t-1} - \beta_0 - \gamma x_{t-1}) + \epsilon_t \tag{4.12}$$

where the term in brackets $y_{t-1} - \beta_0 - \gamma x_{t-1}$ is the error correction term and $\beta_2$ measures the speed of adjustment to long-run equilibrium.

In the presence of cointegration between $x_t$ and $y_t$, the resulting error-correction term is $I(0)$ even though $x_t$ and $y_t$ are $I(1)$. Since the error-correction model analyzes only one cointegrating relationship, we can define VECM for the VAR(p) process as:

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + ... + \Gamma_{p-1} \Delta y_{t-p+1} + \epsilon_t \tag{4.13}$$

where $\Pi$ is rank (number) of cointegrating relationships and $\Pi y_{t-1}$ is the error-correction term. The VECM is also used within the Johansen cointegration testing procedure, as noted in Equation 4.8. Since VECM(p-1) is an equal representation of the VAR(p) process, we have to consider this while determining the appropriate number of lags using information criteria.

### Impulse Responses

As VAR/VECM is a system of equations, interpreting estimated coefficients is rather complicated. Instead, we will interpret our results using impulse responses. If a shock occurs in one of the series, it could affect the other series through the lagged variables system in VAR. The impulse response functions (IRF) show the impact of this shock on other variables. The IRFs are calculated by applying a series of possible shocks to specific variables and observing the reaction of other variables in the model. The functions can then be plotted into a graph to better interpret the responses to shocks. Although the impulse responses for non-cointegrated stationary systems describe only the transitory effects of the shocks, we can also observe permanent shocks if our variables have the same order of integration and are cointegrated (Pesaran & Smith 1998).

# Chapter 5

# Results

## 5.1 Panel Data

### 5.1.1 Unit Root Tests

Though most unit root tests for panel data are tailored for a balanced dataset, few tests can also handle unbalanced data. The most common of these tests are the Im-Pesaran-Shin test (Im *et al.* 2003) and Fisher-type tests (Choi 2001), whose null hypothesis states that all panels contain a unit root. Table 5.1 summarizes the results of these tests with individual effect, individual effect and linear trend, and without individual effect and trend for the *lhpnom* variable at full sample size:

Table 5.1: Results of selected panel unit root tests for *lhpnom*

| Individual Intercept | | | | | | |
|---|---|---|---|---|---|---|
| | Level | | | First difference | | |
| Test | Stat. | Prob. | Obs. | Stat. | Prob. | Obs. |
| Im, Pesaran, Shin | 5.92 | 1.00 | 1130 | -9.99 | 0.00 | 1122 |
| Fisher - ADF | 16.53 | 0.68 | 1130 | 169.60 | 0.00 | 1122 |
| Fisher - PP | 13.81 | 0.79 | 1205 | 404.56 | 0.00 | 1190 |
| Individual Intercept & Trend | | | | | | |
| | Level | | | First Difference | | |
| | Stat. | Prob. | Obs. | Stat. | Prob. | Obs. |
| Im, Pesaran, Shin | 0.37 | 0.65 | 1128 | -11.00 | 0.00 | 1124 |
| Fisher - ADF | 16.08 | 0.71 | 1128 | 177.59 | 0.00 | 1124 |
| Fisher - PP | 9.86 | 0.97 | 1205 | 383.18 | 0.00 | 1190 |
| None | | | | | | |
| | Level | | | First Difference | | |
| | Stat. | Prob. | Obs. | Stat. | Prob. | Obs. |
| Im, Pesaran, Shin | - | - | - | - | - | - |
| Fisher - ADF | 7.99 | 0.99 | 1132 | 133.61 | 0.00 | 1106 |
| Fisher - PP | 8.89 | 0.98 | 1205 | 724.67 | 0.00 | 1190 |

** Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.
* Automatic selection of maximum lags was used based on the Akaike Information Criterion.

The results of unit root tests for *lhpnom* at level confirm what is mostly observable from plots of the panel series. At level, all tests cannot reject the null hypothesis about the presence of a unit root in the panels at a 10% or lower significance level. Opposite results are present for the first difference of the series. Here, every test rejects the null hypothesis of a unit root in the series at a 1 % level of significance.

As none of the tests could reject the null hypothesis at a level and every test could reject it at a first difference, we can conclude that the series is integrated of order one or $I(1)$.

Following Table 5.2 presents unit root tests results for the *lwage* variable:

Table 5.2: Results of selected panel unit root tests for *lwage*

| Individual Intercept | | | | | | |
|---|---|---|---|---|---|---|
| | Level | | | First difference | | |
| Test | Stat. | Prob. | Obs. | Stat. | Prob. | Obs. |
| Im, Pesaran, Shin | 5.63 | 1.00 | 1437 | -8.59 | 0.00 | 1434 |
| Fisher - ADF | 3.43 | 1.00 | 1437 | 167.45 | 0.00 | 1434 |
| Fisher - PP | 0.97 | 1.00 | 1495 | 290.59 | 0.00 | 1483 |
| Individual Intercept Trend | | | | | | |
| | Level | | | First Difference | | |
| | Stat. | Prob. | Obs. | Stat. | Prob. | Obs. |
| Im, Pesaran, Shin | -0.13 | 0.45 | 1427 | -7.35 | 0.00 | 1435 |
| Fisher - ADF | 16.4 | 0.69 | 1427 | 144.16 | 0.00 | 1435 |
| Fisher - PP | 12.2 | 0.90 | 1465 | 256.52 | 0.00 | 1483 |
| None | | | | | | |
| | Level | | | First Difference | | |
| | Stat. | Prob. | Obs. | Stat. | Prob. | Obs. |
| Im, Pesaran, Shin | - | - | - | - | - | - |
| Fisher - ADF | 10.56 | 0.96 | 1428 | 66.76 | 0.00 | 1415 |
| Fisher - PP | 10.48 | 0.96 | 1495 | 297.09 | 0.00 | 1483 |

** Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.
* Automatic selection of maximum lags was used based on the Akaike Information Criterion.

For the *lwage* variable, the results of selected unit root tests are also quite clear. At level, all tests cannot reject the null hypothesis about the presence of unit root in series at a 1 % significance level. On the other hand, tests performed on differenced data show the opposite patterns. The null hypothesis is rejected for all unit root tests at a 1% significance level. Thus, we can conclude that *lwage* series is integrated of order one $I(1)$.

Unit root test results for other exogenous variables used in the analysis of the unbalanced panel dataset are presented in Appendix A, Tables A.2, A.3, A.4 and A.5. Generally, results for all variables except *lunemp* show a similar pattern as *lhpnom* and *lwage*. The series displays non-stationary properties at a level while differencing the data removes the unit root from the series, and the null hypothesis can be rejected in most cases. Due to the construction of the index, we expected the uncertainty index *polunc* to be stationary at level. The tests might falsely identify the series as $I(0)$ due to structural breaks in the data. For this reason, we also tried to perform the Hadri (2000) panel unit root tests, whose hypotheses are opposite to other panel unit root tests and the null hypothesis indicates stationarity. The resulting test statistics rejected the null hypothesis of stationarity in the data, meaning that we cannot conclude that the series is $I(0)$ and will treat it as $I(0)$ from now onwards.

As for the *lunemp* variable, most tests could not reject the null hypothesis of a unit root in the series. As a result of that, we assume stationarity of the series at both level and first difference, which means we have to treat it as integrated of order 0, $I(0)$[1].

Even though we could try to exploit the ARDL model approach without testing for cointegration, which would allow us to include both $I(0)$ and $I(1)$ series, we believe that the correct decision is to continue with panel cointegration tests without the *lunemp* variable. Firstly, all other series seem to be $I(1)$. Many studies, like Bahmani-Oskooee & Ghodsi (2017) or Mikhed & Zemcik (2009), did not include unemployment rate among explanatory variables and instead focused on other factors, sometimes using only measures of income and interest rate. Secondly, the possibility of using a vector error-correction model if our variables are cointegrated could help us study the cointegrating relationships in multiple directions since it is a multi-equation system and accounts for endogeneity.

### 5.1.2 Panel Cointegration Tests

This section provides results from panel cointegration tests proposed by Pedroni (1999) and Kao (1999). From the panel unit root tests, we know that all of our variables, except unemployment, are integrated of order one. We may therefore perform cointegration tests on our variables only if we exclude unemployment from our analysis. Cointegration tests will be conducted with and without *polunc* variable to see the effects of uncertainty. Both Pedroni (1999) and Kao (1999) panel cointegration tests have a null hypothesis of no cointegration. We used automatic lag length selection based on the Schwarz Information Criterion (SIC) for all our tests.

#### Panel Cointegration Tests without Uncertainty

Table 5.3 presents statistics of Pedroni (1999) panel cointegration test performed on a full-sample of our series, without *polunc*:

[1] We also tried to perform panel unit root tests on the non-log form of the variable, but the tests also indicated stationarity at both level and first difference

Table 5.3: Statistics of Pedroni (1999) Panel Cointegration Test without Uncertainty

| Alternative hypothesis: common AR coefs. (within-dimension) | | | | |
|---|---|---|---|---|
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 1.66 | 0.05 | 1.61 | 0.05 |
| Panel rho-Statistic | -0.43 | 0.33 | -0.87 | 0.19 |
| Panel PP-Statistic | -0.15 | 0.44 | -0.28 | 0.39 |
| Panel ADF-Statistic | 0.56 | 0.71 | -0.62 | 0.26 |
| Alternative hypothesis: individual AR coefs. (between-dimension) | | | | |
| | Statistic | Prob. | | |
| Group rho-Statistic | -0.64 | 0.26 | | |
| Group PP-Statistic | -0.29 | 0.38 | | |
| Group ADF-Statistic | -0.85 | 0.19 | | |

From Table 5.3, we see that only two statistics out of eleven reject the null hypothesis of no cointegration at a 5 % level of significance while including the individual intercept. A summary of additional conducted tests with the inclusion of individual intercepts and trends is available in Appendix A, Table A.6. In total, only eight test statistics out of thirty-three rejected the null hypothesis at 10 % or lower level of significance for our entire sample without uncertainty.

Next, Kao (1999) panel cointegration test was conducted, and the results of its ADF-type test are provided in Table 5.4 below:

Table 5.4: Statistics of Kao (1999) Panel Cointegration Test without Uncertainty

| | t-Stat. | Prob. |
|---|---|---|
| Augmented Dickey-Fuller Test | -2.62 | 0.00 |
| Residual variance | 0.01 | |
| HAC variance | 0.01 | |

A more detailed summary of the test with the ADF Test equation is provided in Appendix A, Table A.7. For this test, the null hypothesis of no cointegration in panels is rejected at a 1 % significance level.

The results of panel cointegration tests for our full-sample dataset without *polunc* variable are contradictory at best. The majority of test statistics from Pedroni (1999) did not reject the null hypothesis of no cointegration, whereas Kao (1999) cointegration test confidently rejected the null hypothesis. Thus, we have decided to perform cointegration tests on our balanced subsample with data ranging from 1997 to 2020. Results of these tests are provided in Appendix A, Tables A.8 and A.9. For the balanced dataset, the null hypothesis of the Pedroni (1999) cointegration test could not be rejected based on every test statistic. On the other hand, Kao (1999) rejected the null hypothesis of no cointegration at a 1 % level of significance. Lastly, we have tried to remove most of our explanatory variables and use only *wage* and *ltrate*. Still, most test statistics in Tables A.10 and A.11 pointed towards no cointegration.

### Panel Cointegration Tests with Uncertainty

As the panel cointegration test without uncertainty provided unclear results, to say the least, we can now proceed with tests containing the uncertainty measure.

Table 5.5 presents test statistics of Pedroni (1999) panel cointegration test performed on full-sample data with the *polunc* variable:

Table 5.5: Statistics of Pedroni (1999) Panel Cointegration Test with Uncertainty

| Alternative hypothesis: common AR coefs. (within-dimension) | | | | |
|---|---|---|---|---|
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 1.51 | 0.07 | 1.18 | 0.11 |
| Panel rho-Statistic | 1.72 | 0.96 | 1.67 | 0.95 |
| Panel PP-Statistic | 2.02 | 0.98 | 1.57 | 0.94 |
| Panel ADF-Statistic | 0.32 | 0.63 | -0.44 | 0.33 |
| Alternative hypothesis: individual AR coefs. (between-dimension) | | | | |
| | Statistic | Prob. | | |
| Group rho-Statistic | 3.36 | 0.99 | | |
| Group PP-Statistic | 2.82 | 0.99 | | |
| Group ADF-Statistic | -0.38 | 0.35 | | |

From Table 5.5, we see that only one statistic of Pedroni (1999) test with individual intercept rejects the null hypothesis of no cointegration at 10 % level of significance.

The results of Kao (1999) cointegration test for full sample data with *polunc* are displayed in Table 5.6:

Table 5.6: Statistics of Kao (1999) Panel Cointegration Test with Uncertainty

| | t-Stat. | Prob. |
|---|---|---|
| ADF | -3.74 | 0.00 |
| | | |
| Residual variance | 0.00 | |
| HAC variance | 0.00 | |

Similarly to testing without *polunc* variable, we can reject the null hypothesis of no cointegration at 1% significance level based on Kao (1999) cointegration test.

#### Implications of Panel Cointegration Tests

Results from cointegration tests gave us contradictory implications about the presence of cointegration in the panel series. While Pedroni (1999) test mostly pointed towards no cointegration, Kao (1999) test rejected the null hypothesis of no cointegration in all cases. The inclusion of uncertainty did not significantly affect the results of the tests, with the Pedroni (1999) test statistics rejecting the null hypothesis in more cases. Based on the results, we cannot conclude with certainty if cointegration is present in our series.

Given the contrasting results, we believe that adopting the VAR framework would not be the correct approach. As the vector error-correction model requires a clear idea about the presence of cointegrating relationships between the panels, estimating and interpreting results from it would be at least doubtful. On the other hand, if we would differentiate our variables to have $I(0)$ series for standard vector autoregression, we would lose all information about long-run effects.

Thus, we believe that the correct approach is to utilize the panel autoregressive distributed lag model proposed by Pesaran *et al.* (2001). Although it is not a multi-equation system like VAR or VECM, it allows us to keep both $I(0)$ and $I(1)$ series for estimation and does not require any cointegration pre-testing.

### 5.1.3 Panel ARDL Model

Since the ARDL model can contain both $I(0)$ and $I(1)$ series, we can include unemployment in our model, which was removed for the cointegration testing.

Our log-transformed Equation 3.3 for determining house prices can be rewritten for estimation as:

$$\begin{aligned} lhpnom_{it} = \beta_0 + \beta_1 lwage_{it} + \beta_2 lrconpc_{it} + \beta_3 lunemp_{it} + \\ \beta_4 ltrate_{it} + \beta_5 lcpi_{it} + \beta_6 lpolunc_{it} + \mu \end{aligned} \tag{5.1}$$

where the estimated parameters $\beta$ are elasticities for log-transformed variables[2]. Based on Pesaran *et al.* (1999), we can then define the panel ARDL specification of our model as:

$$\begin{aligned} \Delta lhpnom_{it} = \beta_0 + \sum_{i=1}^{p} \beta_1 \Delta lhpnom_{i,t-1} + \sum_{i=1}^{p} \beta_2 \Delta lwage_{i,t-1} \\ + \sum_{i=1}^{p} \beta_3 \Delta lrconpc_{i,t-1} + \sum_{i=1}^{p} \beta_4 \Delta lunemp_{i,t-1} + \sum_{i=1}^{p} \beta_5 \Delta ltrate_{i,t-1} \\ + \sum_{i=1}^{p} \beta_6 \Delta lcpi_{i,t-1} + \sum_{i=1}^{p} \beta_7 \Delta lpolunc_{i,t-1} \\ + \gamma(lhpnom_{i,t-1} - \alpha_0 - \alpha_1 lwage_{i,t} + \alpha_2 lrconpc_{i,t} + \alpha_3 lunemp_{i,t} + \\ \alpha_4 ltrate_{i,t} + \alpha_5 lcpi_{i,t} + \alpha_6 lpolunc_{i,t}) + \epsilon_{it} \end{aligned} \tag{5.2}$$

where $p$ is the optimally selected lag based on the information criterion chosen, $\beta_i'$ are vectors of coefficients of the estimated short-run parameters. The second part of the equation represents the error-correction term, where $\gamma$ is the speed of adjustment and $\alpha_i'$ are vectors of coefficients of the estimated long-run parameters.

If cointegration is present in our model, we expect $\gamma$ to be statistically significant and negative if variables return to a long-run equilibrium after a shock. On the other hand, if the coefficient is zero and not statistically significant, we do not have evidence of cointegrating relationships between variables.

### Pooled Mean Group Estimation

For the estimation of the ARDL model, we will use the Pooled Mean Group (PMG) estimation of heterogeneous panels, proposed by Pesaran *et al.* (1999). Using averages of individual countries, the PMG estimator provides consistent estimates of the mean short-run coefficients across all cross-sections[3]. Further-

[2] The long-term rate *ltrate* is the only untransformed explanatory variable since it contains negative observations.

[3] Assuming that a large number of cross-sections is available.

more, it constrains equality among long-run coefficients of the cross-sections but allows a cross-dimensional difference of short-run coefficients, as well as of the intercept and the speed of adjustment coefficient.

Unfortunately, the PMG estimator also does not cope well with highly unbalanced datasets, which is the case of our full-sample dataset. Thus, we have decided to perform several estimates on a balanced subsample, with data ranging from 1997 to 2020. Another option would be to remove some of the cross-section, but we would risk the consistency of the short-run estimates due to the cross-dimensional nature of the estimator.

Similarly to panel cointegration testing, we will estimate the model with and without the *polunc* variable to see the uncertainty's effect on the error-correction term. In total, four alterations of Equation 5.2 will be estimated:

- Full-scale model with all variables, except uncertainty,
- Full-scale model with all variables, including uncertainty,
- Reduced form of the model with only *wage* and *ltrate* among explanatory variables,
- Reduced form of the model with only *wage*, *ltrate* and *polunc* among explanatory variables.

Estimating more variations of our original model will give us an idea about the effects of uncertainty on the housing market and provide us with robustness checks for our results. Moreover, we will estimate a reduced form of our model to follow the approach of Case & Shiller (2003) and Bahmani-Oskooee & Ghodsi (2017), who used only income, interest rate and uncertainty proxies in their analysis. We will use automatic lag selection based on Schwarz Information Criterion (SIC) for all estimations.

### Panel Estimation Results

Firstly, the ARDL model is estimated with all variables from Equation 5.2, with and without the uncertainty index. The results of the estimation are summarised in Table 5.7:

Table 5.7: PMG Estimation Results - Full-scale ARDL (1,1,1,1,1,1,1)

| | Without Uncertainty | | With Uncertainty | |
|---|---|---|---|---|
| Variable | Coefficient | Prob.* | Coefficient | Prob.* |
| Long Run Equation | | | | |
| LWAGE | -6.07 | 0.00 | -8.51 | 0.00 |
| LCONSP | 6.62 | 0.00 | 8.00 | 0.00 |
| LTRATE | -0.28 | 0.00 | -0.32 | 0.00 |
| LCPI | 2.90 | 0.06 | 5.17 | 0.04 |
| LUNEMP | -0.46 | 0.00 | -0.60 | 0.00 |
| LPOLUNC | | | -0.03 | 0.65 |
| Short Run Equation | | | | |
| COINTEQ01 | -0.09 | 0.00 | -0.07 | 0.00 |
| D(LWAGE) | 0.84 | 0.00 | 0.89 | 0.00 |
| D(LCONSP) | -0.21 | 0.32 | 0.00 | 0.98 |
| D(LTRATE) | 0.02 | 0.01 | 0.20 | 0.02 |
| D(LCPI) | -0.13 | 0.75 | -0.46 | 0.33 |
| D(LUNEMP) | 0.02 | 0.54 | 0.04 | 0.34 |
| D(LPOLUNC) | | | 0.01 | 0.37 |
| C | -0.57 | 0.00 | -0.80 | 0.00 |

Looking first at the model without uncertainty index and its estimated long-run coefficients, we see that most coefficients are significant on the 1 % level, with the consumer price index at just a 10 % level. All coefficients have mostly expected signs, except *lwage*. As wages increase, which might translate into increased demand for housing, we would expect house prices to grow. Other long-run coefficients seem to have a sign that we would expect. House prices seem to increase due to increases in inflation and consumption.

On the other hand, hikes in interest rates impair the ability to finance mortgages, which might lead to a decrease in demand for housing. Finally, an increase in unemployment in our model leads to housing prices decrease. This might be the result of a decreasing purchasing power of the population during periods of high unemployment.

Compared to the long-run, the short-run estimates give us conflicting results. Only wage and interest rates are statistically significant at the 1 % level, but they have opposing signs. The wage coefficient is now positive, as we expected, but an increase in interest rates now indicates a rise in house prices. The positive reaction of house prices to hikes in interest rates might result from a supply decrease because of less favourable financing conditions for developers.

The inclusion of the uncertainty index had only a minor effect on the model. Moreover, the index was not statistically significant for both short-run and long-run equations. Other coefficients did not change signs or statistical significance except CPI, which is now statistically significant at the 5 % level.

Finally, looking at the error-correction term of our model, we see statistically significant at one % and negative in either case, meaning that the null hypothesis of no cointegration is rejected. This contrasts with conducted panel cointegration tests, whose results were inconclusive but mostly leaned towards not rejecting the null hypothesis of no cointegration. To interpret the coefficient, we can say that the speed of adjustment to the long-run equilibrium of the relationship ranges from 7 % (with uncertainty index) to 9 % on average per year in our model. In other words, it takes approximately 10-11 years to restore the long-run equilibrium after an initial shock.

Next, reduced form of the model with only *lwage*, *ltrate* and *polunc* was estimated, whose result is displayed in Table 5.8:

Table 5.8: PMG Estimation Results - Reduced-form ARDL (2,1,1)

| | Without uncertainty | | With Uncertainty | |
|---|---|---|---|---|
| Variable | Coefficient | Prob.* | Coefficient | Prob.* |
| Long Run Equation | | | | |
| LWAGE | 0.04 | 0.92 | -0.20 | 0.65 |
| LTRATE | -0.16 | 0.00 | -0.16 | 0.00 |
| LPOLUNC | | | 0.19 | 0.01 |
| Short Run Equation | | | | |
| COINTEQ01 | -0.08 | 0.00 | -0.08 | 0.00 |
| D(LHPNOM(-1)) | 0.55 | 0.00 | 0.56 | 0.00 |
| D(LWAGE) | 0.30 | 0.01 | 0.31 | 0.04 |
| D(LTRATE) | 0.01 | 0.03 | 0.02 | 0.02 |
| D(LPOLUNC) | | | -0.01 | 0.37 |
| C | 0.49 | 0.00 | 0.48 | 0.00 |

Regarding the long-run estimates of the reduced model, we see that the wage is not statistically significant in the estimates with and without uncertainty, in contrast to the full-scale model. The long-term rate has been found significant again at one % level with a negative sign, suggesting that rate hikes dampen the house price.

The short-run estimates follow the same pattern as the full-scale model, with both wage and long-term rates found to be statistically significant and with a positive sign. However, based on the selected information criterion, the lag of the dependent variable was included in the short-run equation and is statistically significant, meaning that part of an increase in house prices is explained by its previous increases.

While the uncertainty index remains insignificant in the short run, as in the previous full-scale model, we see a change in its relevance in the long run. Here, uncertainty was found to be statistically significant in the long run with a positive sign, meaning that increases in policy uncertainty lead to increases in housing market prices. Although we might have expected the coefficient to be negative since higher uncertainty could force people to spend less, a positive sign does not have to be some abnormality. During periods of higher uncertainty, the public could adjust their investments towards real assets, which they consider safer, including housing. This might be why the relationship is positive in the long-run equation of our model with fewer variables.

### Cross-sectional Estimation Results

Since the PMG estimator allows short-run coefficients, intercept and the speed of adjustment coefficient to differ across cross-sections, we can also analyse the individual short-run coefficients for each country. The individual short-run estimations are provided in Appendix A, Table A.12. We will interpret cross-sectional short-run estimates only for the full-scale model with all variables.

For all countries, the error-correction term was statistically significant and negative at a 1 % level, with coefficients ranging from -0,18 to -0,02. The fastest speed of adjustment to the long-run equilibrium was found in the United States, while the United Kingdom and Japan had the slowest rate. Surprisingly, the long-term rate positively correlated with house prices in all countries except Japan. Policy uncertainty was found statistically significant in every country, but the coefficient was very low for most cases. Finally, the wage was found statistically significant at a 10 % level or less only in half of the countries.

## 5.2 Time Series Data of USA

### 5.2.1 Unit Root Tests

To determine the order of integration of our variables, we performed the Phillip-Perron, Augmented Dickey-Fuller and KPSS tests for unit root, whose results are available in Appendix B, tables B.2 and B.3. The automatic lag selection was used based on the Schwarz information criterion.

Following Table 5.9 summarises the results of unit root tests for each variable:

Table 5.9: Summary of Unit Root Tests

| | At level | At first difference |
|---|---|---|
| lhpindex | Unit root | Stationarity |
| lincome | Unit root | Stationarity |
| lmortg | Unit root | Stationarity |
| lpop | Unit root | Unit root |
| lrent | Unit root | Stationarity |
| lcpi | Unit root | Stationarity |
| lcostsw | Unit root | Stationarity |
| lunrate | Unit root | Stationarity |
| lfmunc | Stationarity | Stationarity |

From Table 5.9, we see that majority of the series are integrated of order one, except *lfmunc* and *lpop*. The Financial Uncertainty index is developed to have the same mean over time, making the series stationary. The results from the tests thus confirm what is evident from the theory behind the index or Figure 3.1.

The results for the *lpop* variable were inconclusive but mostly leaned towards non-stationarity at both level and first difference. Thus, we tried to take a second difference of the variable, whose unit root test results pointed towards $I(2)$.

As we know from the cointegration definition of Engle & Granger (1987), series must have the same order of integration to be cointegrated. For this reason, financial market uncertainty and population level cannot be cointegrated with house prices, meaning Johansen cointegration testing will be performed without them.

### 5.2.2 Johansen Cointegration Test

#### Lag Selection

Before performing the Johansen cointegration test, we need to determine the appropriate lag length. For this, we will estimate a simple VAR model with all variables in levels and choose a lag length which minimises the selected information criterion.

Table 5.10 presents the results for individual information criterion:

Table 5.10: Lag Length Selection by Individual Information Criterion

| Lag | LogL | LR | FPE | AIC | SC | HQ |
|---|---|---|---|---|---|---|
| 0 | 3623.15 | NA | 5.61e-17 | -17.55412 | -17.48580 | -17.52710 |
| 1 | 10216.48 | 12930.62 | 8.96e-31 | -49.32274 | -48.77619 | -49.10655 |
| 2 | 10740.18 | 1009.265 | 8.94e-32* | -51.62710* | -50.60232* | -51.22175* |
| 3 | 10786.03 | 86.80759 | 9.08e-32 | -51.61182 | -50.10882 | -51.01731 |
| 4 | 10815.01 | 53.87263 | 1.00e-31 | -51.51462 | -49.53338 | -50.73094 |
| 5 | 10857.83 | 78.16239* | 1.03e-31 | -51.48463 | -49.02517 | -50.51179 |
| 6 | 10894.04 | 64.84864 | 1.10e-31 | -51.42251 | -48.48482 | -50.26051 |
| 7 | 10928.30 | 60.21056 | 1.19e-31 | -51.35097 | -47.93505 | -49.99981 |
| 8 | 10953.23 | 42.96436 | 1.34e-31 | -51.23413 | -47.33999 | -49.69380 |

* indicates lag order selected by the criterion

We see that most selected information criterion is minimised using two lags, while the likelihood ratio uses five lags. Thus, we will use one lag[4] for the Johansen cointegration test, as it minimises both Akaike and Schwarz information criteria.

#### Johansen Test

The following Table 5.11 summarises outcomes of the Johansen Cointegration test for various data trends. The tests statistics of both Trace tests and Maximum Eigenvalue tests were compared with critical values from Mackinnon *et al.* (1999), which were found to be more precise than the original values calculated in the Johansen and Engle-Granger papers.

[4] As VAR(p) has equal representation of VECM(p-1).

Table 5.11: Summary of the Johansen Cointegration Test

| Data Trend: | None<br>No Intercept<br>No Trend | None<br>Intercept<br>No Trend | Linear<br>Intercept<br>No Trend | Linear<br>Intercept<br>Trend | Quadratic<br>Intercept<br>Trend |
|---|---|---|---|---|---|
| Trace | 2 | 3 | 3 | 3 | 3 |
| Max-Eig | 2 | 3 | 3 | 3 | 3 |

*Critical values based on Mackinnon *et al.* (1999)

We see that the Trace and Maximum Eigenvalue tests do not give contradictory results, which is a good sign. Also, for most data trends, the results indicate three cointegrating equations at a five % significance level. Two cointegrating equations were found only if we assume no intercept or trend present, which is highly unlikely for our series. Thus, we will go with the majority in this case and conclude that we have three cointegrating equations in our model.

The Trace and Maximum Eigenvalue test statistics for linear trend and intercept are displayed in Table 5.12. For both tests, we begin at $r = 0$ (null hypothesis of cointegrating relationships) and move upwards to higher ranks, until the null hypothesis is rejected. For both tests, the null hypothesis is rejected at rank $r = 3$, from which we can conclude that there are three cointegrating equations.

Estimated cointegrating equations are displayed in Appendix B, Tables B.4 and B.5.

Table 5.12: Trace and Maximum Eigenvalue Test Statistics - Linear Trend

Unrestricted Cointegration Rank Test (Trace)

| Hypothesized No. of CE(s) | Eigenvalue | Trace Statistic | 0.05 Critical Value | Prob.** |
|---|---|---|---|---|
| None * | 0.28 | 284.29 | 125.62 | 0.00 |
| At most 1 * | 0.15 | 144.42 | 95.75 | 0.00 |
| At most 2 * | 0.11 | 76.99 | 69.82 | 0.01 |
| At most 3 | 0.05 | 30.56 | 47.86 | 0.69 |
| At most 4 | 0.01 | 9.67 | 29.80 | 0.98 |
| At most 5 | 0.01 | 3.77 | 15.49 | 0.92 |
| At most 6 | 0.00 | 0.43 | 3.84 | 0.51 |

Trace test indicates 3 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

| Hypothesized No. of CE(s) | Eigenvalue | Max-Eigen Statistic | 0.05 Critical Value | Prob.** |
|---|---|---|---|---|
| None * | 0.28 | 139.88 | 46.23 | 0.00 |
| At most 1 * | 0.15 | 67.43 | 40.08 | 0.00 |
| At most 2 * | 0.11 | 46.42 | 33.88 | 0.00 |
| At most 3 | 0.05 | 20.89 | 27.58 | 0.28 |
| At most 4 | 0.01 | 5.90 | 21.13 | 0.98 |
| At most 5 | 0.01 | 3.34 | 14.26 | 0.92 |
| At most 6 | 0.00 | 0.43 | 3.84 | 0.51 |

Max-eigenvalue test indicates 3 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

### 5.2.3 Vector Error-Correction Model

Once we have identified the number of cointegrating equations, we can construct the vector error-correction model to see the speed of adjustment to long-run equilibrium and generate impulse responses to innovations in our series.

As VAR(p) model has an equal representation as a VEC(p-1) model, we include only one lag in our model. Cointegration rank was determined by the Johansen test, which indicated three cointegrating equations. Thus, VECM(1,3) is estimated, and Table 5.13 presents the long-run coefficients with standard errors in parentheses and t-statistics in square brackets.

Table 5.13: Cointegrating Equations for VECM(1,3)

| Cointegrating Eq: | CointEq1 | CointEq2 | CointEq3 |
|---|---|---|---|
| LHPINDEX(-1) | 1.00 | 0.00 | 0.00 |
| LINCOME(-1) | 0.00 | 1.00 | 0.00 |
| LMORTG(-1) | 0.00 | 0.00 | 1.00 |
| LRENT(-1) | -7.06 | -0.31 | -0.29 |
| | (2.02) | (0.21) | (1.12) |
| | [-3.49985] | [-1.50952] | [-0.26213] |
| LCPI(-1) | -1.55 | -1.45 | -0.01 |
| | (1.09) | (0.11) | (0.61) |
| | [-1.42057] | [-12.9992] | [-0.02204] |
| LCOSTSW(-1) | 7.79 | -0.10 | 1.56 |
| | (2.16) | (0.22) | (1.20) |
| | [ 3.59995] | [-0.45066] | [ 1.29726] |
| LUNRATE(-1) | -1.59 | -0.05 | -0.29 |
| | (0.16) | (0.02) | (0.09) |
| | [-9.88642] | [-3.21253] | [-3.20216] |
| C | 20.77 | 0.43 | -4.26 |

The long-run cointegrating equations with estimated coefficients can be rewritten as:

$$
\begin{aligned}
lhpindex_{t-1} &= 20.77 - 7.06lrent_{t-1} - 1.55lcpi_{t-1} \\
&\quad +7.79lcostsw_{t-1} - 1.59lunrate_{t-1}
\end{aligned} \tag{5.3}
$$

$$
\begin{aligned}
lincome_{t-1} &= 0.43 - 0.31lrent_{t-1} - 1.45lcpi_{t-1} \\
&\quad -0.10lcostsw_{t-1} - 0.05lunrate_{t-1}
\end{aligned} \tag{5.4}
$$

$$
\begin{aligned}
lmortg_{t-1} &= -4.26 - 0.29lrent_{t-1} - 0.01lcpi_{t-1} \\
&\quad +1.56lcostsw_{t-1} - 0.29lunrate_{t-1}
\end{aligned} \tag{5.5}
$$

Finally, the VECM short-run parameters are estimated:

$$
\begin{bmatrix} \Delta lhpindex_t \\ \Delta lincome_t \\ \Delta lmortg_t \\ \Delta lrent_t \\ \Delta lcpi_t \\ \Delta lcostsw_t \\ \Delta lunrate_t \end{bmatrix}
=
\begin{bmatrix} 0.0006 \\ 0.0082 \\ -0.0002 \\ 0.0025 \\ 0.0020 \\ 0.0019 \\ 0.0230 \end{bmatrix}
+
\begin{bmatrix}
0.8756 & 0.1178 & -0.8573 & -0.0331 & -0.0570 & -0.0882 & -2.0550 \\
-0.0090 & -0.5149 & -0.0273 & -0.0018 & -0.0140 & -0.0070 & -0.2676 \\
0.0036 & 0.0250 & 0.2993 & -0.0002 & 0.0064 & 0.0030 & 0.0300 \\
-0.0964 & -0.7918 & -0.5663 & 0.0638 & -0.2170 & 0.5182 & -3.1127 \\
0.0428 & -0.0035 & 1.6268 & 0.0046 & 0.4652 & 0.0911 & -2.9864 \\
0.0042 & -0.1463 & -0.2290 & -0.0113 & -0.0085 & -0.3374 & -0.5315 \\
-0.0021 & 0.0335 & 0.0357 & 0.0014 & -0.0004 & -0.0057 & 0.1711
\end{bmatrix}
*
\begin{bmatrix} \Delta lhpindex_{t-1} \\ \Delta lincome_{t-1} \\ \Delta lmortg_{t-1} \\ \Delta lrent_{t-1} \\ \Delta lcpi_{t-1} \\ \Delta lcostsw_{t-1} \\ \Delta lunrate_{t-1} \end{bmatrix} \tag{5.6}
$$

Tables B.6 and B.7 in Appendix B also summarize the error-correction and short-run coefficients with standard errors and t-statistics of our estimation. For our series to converge towards some long-run equilibrium, the coefficients of the error-correction term need to be negative and statistically significant. For house prices, we see that the error correction terms from the first two cointegrating equations do not satisfy these conditions, with one being statistically insignificant and the other positive. The third error-correction coefficient is statistically significant and negative, meaning that after the initial short-run shock, the series should slowly converge towards long-run equilibrium.

However, interpreting these results as the existence of a long-run equilibrium of house prices would not be correct. The second error-correction term is positive, statistically significant and more extensive than the negative term. This means that overall, the error-correction terms are increasing the disequilibrium of the series. Thus, we will check the diagnostics of our model.

### Model Diagnostics

First of all, we will check the stability of our model. This can be done by checking the inverse roots of the AR the characteristic polynomial in Figure B.2. For the model to be stable, all the inverse roots should lie inside the circle, while the roots representing the cointegration vectors should lie on the circle. The figure shows that these assumptions are satisfied, and the model is, therefore, stable.

Next, the normality of residuals, homoskedasticity and possible autocorrelation is checked using selected tests, whose probabilities are summarized in Table 5.14.

Table 5.14: P-values of Selected Diagnostics Tests

| Test Type | Prob. |
|---|---|
| Jarque-Bera Test | 0.00 |
| Portmanteau Autocorrelation Test | 0.41 |
| White Heteroskedasticity Test | 0.00 |

Based on the p-value of the Jarque Bera test, we reject the null hypothesis of multivariate normality of residuals. Although this violates one of the model's assumptions, the normality of residuals is often violated in VAR/VEC models due to their inability to distinguish between regular and extraordinary shocks. As Juselius (2006) argues, the residuals do not need to be normally distributed if their non-normality is caused primarily by excess kurtosis, which is our case. Thus, we will continue the interpretation while keeping this assumption's possible violation in mind.

Next, the Portmanteau Autocorrelation test was conducted and based on its p-value, the null hypothesis of no autocorrelation was not rejected, meaning that we do not have autocorrelation in our model. Finally, the White Heteroskedasticity test was performed, and statistics rejected the null hypothesis of constant variance of the residuals. This implies that the ARCH effects are present in our model. This means that our model can be consistent and unbiased but not efficient.

### Impulse Responses

Our results cannot be interpreted on a level-to-level basis, as this would only be beneficial with one cointegrating equation in our model. More cointegrating equations might describe several equilibriums, which makes it almost impossible to interpret coefficients this way.

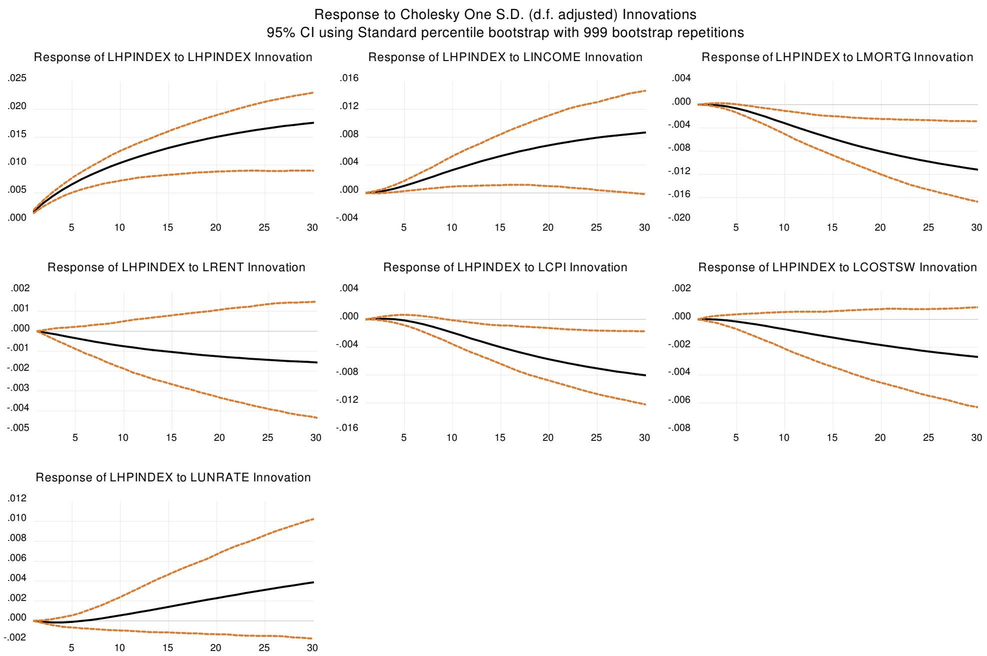


Figure 5.1: Impulse Responses of House Prices

Besides giving us information about relations between variables, the VECM represents changes that are also permanent and not only transitory, as opposed to the unrestricted VAR model, where the effects of shocks disappear over time. A more suitable approach to interpretation is to look at impulse response functions, as done in unrestricted VAR models (Pesaran & Smith 1998). The impulse responses of house prices to shocks in other variables of our models are presented in Figure 5.1.

First of all, the majority of the shocks from our explanatory variables have to be generated by coefficients of our long-run Equations 5.3, 5.4 and 5.5, as the short-run coefficients were found to be low and insignificant in the short-run equation for house prices in B.7. From Figure 5.1, we see that shock in personal income causes a permanent spike in house prices, which is the expected outcome and is in line with economic theory.

The mortgage rates effects are also consistent with the economic theory, with shocks causing a decrease in house prices. Together with the financial consistency, the mortgage rates and house prices provide evidence of a possible cointegrating relationship. Moreover, the error-correction term of the third cointegrating equation 5.5, which included mortgage rates, was the only term with both a negative sign and statistical significance.

We see a gradual decrease in house prices for the rent variable after a shock in average rents. Even though rent increases could decrease house prices, as people would instead rent as a substitute for taking a mortgage, we see that the confidence interval is too large. Its upper bound is increasing above zero with additional time. This makes the interpretation of the rent effect rather tricky.

Regarding the last three variables, their effects on house prices are not in line with economic theory. Although innovation in inflation slightly increases house prices in the first months, the response is then negative and gradually decreases with time. This could also mean that housing will decrease shortly after house prices increase due to higher inflation. For this to be true, however, we would have to assume that the income of consumers did not increase with higher price levels or that it was increasing at a slower rate than the house prices. The same applies to the shocks to the costs series. We expect increased costs to limit the housing supply, translating into increasing prices. Lastly, innovation in unemployment slightly decreases housing prices initially, followed by increased house prices. It must be noted that the confidence interval is relatively large, and its lower bound gradually drops below zero.

Overall, the interpretation of the estimated model is tricky at best. As short-run coefficients were found to be not statistically significant, we have to rely on the effects caused by long-run coefficients. From the impulse responses, we see that income, mortgage rates and rent exhibit some economic consistency, while inflation, costs and unemployment do not behave as expected. Regarding possible cointegrating relationships, the only error-correction term, which had a negative sign and statistical significance, was associated with the third cointegrating equation, including mortgage rates. However, the sign from the second equation was positive and more prominent than the negative term, which means that overall, the error-correction terms are increasing the disequilibrium of the model, indicating a possible absence of cointegration.

Aware of possible specification issues of our model, we tried to repeat the analysis using only income and mortgage rates as house price determinants, following the approach of Case & Shiller (2003) or Bahmani-Oskooee & Ghodsi (2017). After selecting the proper lag length, the Johansen cointegration test indicated one cointegrating equation. The estimated error-correction term for the limited VECM model is displayed in Appendix B, Table B.8. From the table, we see that the error-correction term is also positive, implying that there is no cointegration between fundamentals and house prices, as the error-correction term pushes the model towards disequilibrium.

### 5.2.4 ARDL model with Uncertainty

As the uncertainty index was not integrated in the same order as other variables in the model, we could not include it in the Johansen testing procedure and VECM. Although it might be tempting to differentiate our series and utilize a VAR system with all variables stationary, it would mean losing all the long-run information from our series. Thus, we will use a single-equation ARDL model, which was already utilized for panel data, as it allows us to include both $I(0)$ and $I(1)$.

For our time series, we can specify the ARDL model as:

$$
\begin{aligned}
\Delta lhpindex_t = \beta_0 &+ \sum_{i=1}^{p} \Delta\beta_1 lhpindex_{t-1} + \sum_{i=0}^{p} \Delta\beta_2 lincome_{t-1} + \sum_{i=0}^{p} \Delta\beta_3 lmortg_{t-1} \\
&+ \sum_{i=0}^{p} \Delta\beta_4 lrent_{t-1} + \sum_{i=0}^{p} \Delta\beta_5 lcpi_{t-1} + \sum_{i=0}^{p} \Delta\beta_6 lcostsw_{t-1} \\
&+ \sum_{i=0}^{p} \Delta\beta_7 luntrate_{t-1} + \sum_{i=0}^{p} \Delta\beta_8 lfmunc_{t-1} \\
&+ \gamma(lhpindex_{t-1} - \alpha_0 - \alpha_1 lincome_t + \alpha_2 lmortg_t + \alpha_3 lrent_t \\
&+ \alpha_4 lcpi_t + \alpha_5 lcostsw_t + \alpha_6 lunrate_t + \alpha_7 lfmunc_t) + \epsilon_t
\end{aligned}
\tag{5.7}
$$

where the $\beta$'s are the short-run coefficents, $\gamma$ is the error-correction term and $\alpha$'s are the long-run coefficients.

The estimation procedure for time series is different to the panel data. First, the Bound cointegration test has to be conducted within the ARDL framework. If cointegration is found in the model, an error-correction form of the model is defined to capture long-run effects.

### ARDL Bound Test

The ARDL Bound test is used to test the null hypothesis that all long-run parameters are equal to zero, which means no cointegration:

- $H_0$: $\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = 0$

The obtained statistics from the test are compared to critical values introduced by Pesaran *et al.* (2001). If the resulting statistic is lower than the $I(0)$ lower bound of critical values, the null hypothesis of no cointegration cannot be rejected. Statistics higher than the $I(1)$ upper bound will result in the rejection of the null hypothesis, which implies the presence of cointegration in our model. Finally, if the resulting statistics lie between those bounds, our model has a possible specification problem.

Results of the ARDL Bound test are provided in Appendix B, B.9. We see that the F-statistics lie strictly between lower and upper bounds for most significance levels, which means that the cointegration test is inconclusive, and there might be some specification issues. The t-statistic of the test is lower than the $($I)(0), which points towards not rejecting the null hypothesis about no cointegration. The F-statistic is the most critical determinant for the bound test, so we assume we have a specification problem.

Thus, we will again try to reduce our model to only income and mortgage rates as the main determinants of house prices, and the uncertainty index. Table B.10 in Appendix B presents results for the ARDL Bound test on the reduced model. Both F-statistic and t-statistics are lower than the respective $I(0)$ bounds, indicating that we cannot reject the null hypothesis of no cointegration of our model.

In line with the results from the VEC model, we did not find evidence of a cointegrating relationship for our full-scale and reduced model. To see the effect of uncertainty on house prices, we can perform model diagnostics and try to interpret the results from the reduced short-run ARDL model.

### Model Diagnostics

We will perform tests for autocorrelation, normality and possible heteroskedasticity to check our model's validity. The results from these tests are summarized in Table 5.15:

Table 5.15: P-values of Selected Diagnostics Tests

| Test Type | Prob. |
|---|---|
| Jarque-Bera Test | 0.00 |
| Serial Correlation LM Test | 0.41 |
| Breusch-Pagan-Godfrey Test | 0.01 |

The Jarque-Bera test rejects the null hypothesis of normally distributed residuals. The violation of this assumption is again caused mainly by excess kurtosis. We will therefore continue with the interpretation while bearing in mind the possible violation. Possible autocorrelation was checked by the Breusch-Godfrey LM test, which failed to reject the null hypothesis of no serial correlation. We will assume that there is no autocorrelation in our model. Lastly, Breusch-Pagan-Godfrey Test was conducted to test for a possible presence of heteroskedasticity. The null hypothesis of homoskedasticity was rejected by this test, meaning that ARCH effects are present in our model. We will try to limit the effects by reestimating the model with HAC standard errors and a covariance matrix.

### ARDL(1,0,0,2)

We used automatic lag selection based on the Schwarz information criterion. The resulting model is ARDL(1,0,0,2) with one lag of the dependent variable and two lags of the uncertainty index. The estimation results are provided in Table 5.16:

Table 5.16: ARDL(1,0,0,2) Model with Uncertainty

| Variable | Coefficient | Std. Error | t-Statistic | Prob.* |
|---|---|---|---|---|
| DLHPINDEX(-1) | 0.9408 | 0.0213 | 44.1916 | 0.0000 |
| DLINCOME | -0.0007 | 0.0036 | -0.1809 | 0.8566 |
| DLMORTG | 0.0061 | 0.0026 | 2.3347 | 0.0200 |
| LFMUNC | 0.0018 | 0.0022 | 0.8117 | 0.4174 |
| LFMUNC(-1) | -0.0120 | 0.0045 | -2.7019 | 0.0072 |
| LFMUNC(-2) | 0.0109 | 0.0030 | 3.5901 | 0.0004 |
| C | 0.0003 | 0.0001 | 2.2992 | 0.0220 |
| R-squared | 0.8984 | Mean dependent var | | 0.0035 |
| Adjusted R-squared | 0.8969 | S.D. dependent var | | 0.0052 |
| S.E. of regression | 0.0017 | Akaike info criterion | | -9.9188 |
| Sum squared resid | 0.0012 | Schwarz criterion | | -9.8512 |
| Log likelihood | 2080.0312 | Hannan-Quinn criter. | | -9.8921 |
| F-statistic | 605.8174 | Durbin-Watson stat | | 2.0453 |
| Prob(F-statistic) | 0.0000 | | | |

The ARDL model estimates show that most of the short-run movement in house prices is explained by their previous values. Moreover, personal income was found insignificant as a short-run determinant of house prices.

The mortgage rate was found statistically significant in the model but was positively correlated with house prices, which is not in line with our expectations. We would expect higher rates to lower demand for housing, which would then translate into prices, as was shown by our VECM. On the other hand, the VECM impulse responses explained the long-run effects, which means that our models do not necessarily contradict themselves.

Finally, we can examine the link between the Financial Markets Uncertainty index and house prices. Although the level variable was deemed insignificant in the model, its lags are significant. The first lag of uncertainty seems to have a negative effect on house prices. However, this effect's magnitude is almost entirely erased by the second lag of uncertainty. Interpreting the overall effect of uncertainty on house prices is rather tricky since the presence of heteroskedasticity could influence the standard errors. It seems that house prices tend to decrease firstly, as uncertainty could stimulate people to spend less. However, higher uncertainty could also force people to adjust their portfolios towards safer assets, like real estate, which would increase demand for housing and thus increase prices.

## 5.3 Discussion of Results

#### Panel Analysis Summary

In our panel analysis, we first studied the order of integration in our variables, followed by panel cointegration testing. Using the Pedroni (1999) and Kao Kao (1999) panel cointegration tests, we could not conclude that there was a cointegrating relationship. Most test statistics from the Pedroni cointegration test pointed towards no cointegration among series. The Kao test decisively rejected the null hypothesis of no cointegration for balanced and unbalanced datasets.

As the conclusion about cointegration could not be made, we estimated the dynamic panel ARDL model using the pooled mean method for heterogeneous panels. Not only this allowed us to study cointegration without additional pre-testing, but we could also include $I(0)$ variables into our model since ARDL allows both $I(0)$ and $I(1)$ series. For all variations of the ARDL model, the estimated error-correction term was negative and statistically significant. Based on this, we concluded that there is cointegration present between the series.

Next, we examined the determinants of house prices and their expected effects. For the ARDL model with all exogenous variables, income was found statistically significant and negative in the long run, which is not in line with economic theory and signals possible miss-specification issues. The probable cause of this could be the inclusion of consumption among explanatory variables since its coefficient entirely erases the effect of income. Thus, a reduced form of the model was estimated with only income, mortgage rates and uncertainty. Here, income was not statistically significant in the long run. This is in line with the findings of Gallin (2006) or Tsatsaronis & Zhu (2004), who did not find income to be a determining factor of house prices in the long run. In the short run, however, it was found significant in both models and with positive coefficients, which we would expect. Other determinants were mainly exhibiting the expected patterns. Although the mortgage rate was positive and significant in the short run, the long-run estimates demonstrated a negative relationship with house prices.

Finally, uncertainty was included in the model to study its effects on house prices. In the ARDL model with all variables, uncertainty was not statistically significant in both the short and long run. On the contrary, the reduced form of the model indicated a positive long-run relationship between house prices and uncertainty while still being insignificant in the short run. This implies that due to higher policy uncertainty, people will deem housing a safer asset and adjust their portfolios accordingly rather than reducing their spending on real estate. This contrasts with the findings of Bahmani-Oskooee & Ghodsi (2017), who found policy uncertainty to have short-run and primarily adverse effects on house prices in the USA.

We see the main shortcomings of our panel analysis in the reduced period in ARDL estimation. The PMG estimator did not cope well with our highly unbalanced dataset, with a near-singular matrix estimated in this case, which meant we had to reduce the number of observations significantly. On the other hand, as stated in Pesaran *et al.* (2001), the PMG estimator provides consistent estimates even in small samples if the amount of cross-sections is sufficiently high, which is our case, as we used data from ten countries with more than two hundred observations in total.

Another issue of our analysis could be the choice of the PMG estimator. Being considered the best compromise among dynamic estimators for heterogenous panels, the PMG estimator allows heterogeneity among the short-run coefficients while forcing homogeneity in the long run. This assumption is the best description of what we would expect, as short-run effects are subject to the country-specific factors while the determinants driving long-run relationships should be the same for all cross-sections. While this assumption is the most probable for developed economies in our panels, there is also a possibility that the long-run factors are heterogeneous. Thus, we suggest estimating the ARDL model using other dynamic estimators, such as mean-group or dynamic fixed effects estimators, and comparing the results as the potential expansion of our study.

### Time Series Analysis Summary

We first examined a possible link between macroeconomic fundamentals and house prices in the time-series analysis. While the Johansen cointegration test suggested three possible cointegrating equations, the estimated VEC model revealed only one possible long-run equilibrium. However, since the error-correction term of the other cointegrating equation was more significant and positive, the coefficient adjustments pushed the whole model into disequilibrium. As a result, we could not confirm the presence of cointegration in the series. This was also the case for a reduced form of the model, where only personal income and mortgage rates were taken into account. Thus, our results align with the analysis of Mikhed & Zemcik (2009), who also did not find evidence of a cointegrating relationship in the USA data.

Since the results of the Johansen cointegration test and estimated VEC models conflicted, a single-equation ARDL model was estimated. The ARDL model allows the inclusion of both $I(0)$ and $I(1)$ series and does not require any additional pre-testing. This allowed us to include the uncertainty index in our model and further test for cointegration using the ARDL Bounds test. Firstly, a model including all variables was estimated, but the ARDL Bounds test pointed towards possible miss-specification. Thus, a reduced form of the model with only income and mortgage rates was estimated. Here, the Bound test could not reject the null hypothesis of no cointegration. Based on the results from VECM and ARDL models, we concluded that there is no cointegration between our series. This somehow contrasts with the results from our panel analysis, where we found evidence of a cointegrating relationship in the USA cross-section. We see two possible causes for this outcome. Firstly, we used a different period and data frequency for our time-series analysis, where some shocks, which were not present in the panel analysis, could influence the results to a different outcome. Secondly, different measures of house prices, income and mortgage rates were used in both models, which could also be one of the possible factors for different results.

Regarding the driving factors of house prices, VECM indicated that most short-run coefficients were not statistically significant, and most effects came from the cointegrating equations. The analysis of impulse response functions revealed that income, mortgage rates and rent had the expected impact on house prices. On the other hand, the house prices did not have expected responses to shocks in building costs, unemployment rate and possibly CPI. While shocks to income and unemployment rates mainly caused the increases in house prices, the decreases were associated with shocks to mortgage rates, rents, building costs and inflation.

The relationship between uncertainty, income, mortgage rates and house prices was studied within the ARDL framework. The estimated short-run coefficient of income was found to be statistically insignificant in the model, which corresponds to VECM model findings. On the other hand, mortgage rates were found to be significant and positive. Although we would expect house prices to decrease with increasing mortgage rates, these results somehow correspond to our panel ARDL analysis findings in Tables 5.7 and 5.8. The negative relationship between house prices and mortgage rates was in the long-run equation, while the short-run effect was positive. Lastly, the coefficient of the financial market uncertainty index was analyzed. The uncertainty coefficient was found insignificant at the level, but its two lags were included in the model and found significant. Although financial uncertainty seemed to have a negative relationship with house prices, it is difficult to interpret its overall impact as our model contained ARCH effects.

We see the primary shortage of our analysis in the possible effects of structural breaks in the series, which could force some unit root tests to identify the order of integration of our series falsely. Even though the dynamic specification of our models should be robust enough for outliers, our results could be influenced by large shocks generated by the financial crisis in 2008 and the COVID-19 pandemic in 2020, which caused sudden spikes in some series, mainly unemployment and personal income. For future analysis, we suggest the usage of unit root testing, which is robust enough for a possible structural break in the data, and possibly utilizing the time-varying cointegration method proposed by Bierens & Martins (2010), where the cointegration relationship varies smoothly over time, as opposed to traditional VECM, where the cointegrating vector is assumed to not change over time. Through this method, we might be able to identify the structural breaks and take them into account while determining the possible long-run equilibrium.

# Chapter 6

# Conclusion

There are many reasons why it is desirable to understand the main determinants of house prices. The potential buyers or sellers of real estate can use this knowledge to decide whether to pursue the potential deal. If the potential buyers understand how current market conditions and their possible changes affect house prices, it could help them know if the asking price for the asset is reasonable. Understanding the main drivers of housing prices is also important for policy makers, as changes in the housing market significantly impact the overall economy. Before the policy makers introduce new measures, such as changes in tax rate or government expenditures, it is desirable to understand the full-scale effect of their actions. For example, suppose personal income is positively linked with housing prices. In that case, introducing additional taxes on income could also affect the housing market and indirectly other sectors of the economy, such as construction, investments or retail, which might not be the intention of policymakers.

The previous empirical findings tried to identify a long-run relationship between house prices and their determinants through the concept of cointegration. While some studies established the relationship, the results were mainly mixed or leaning towards no long-run equilibrium between house prices and fundamentals. This thesis tried to expand the existing literature by examining panel and time series datasets of advanced economies and by including measures of economic uncertainty among possible explanatory variables.

The results from the time-series analysis mostly corresponded to previous findings of Mikhed & Zemcik (2009) or Gallin (2006). While the cointegration test indicated three possible cointegrating equations, estimated error-correction terms from the vector error-correction model pushed the model towards disequilibrium, which meant no cointegration could be identified. This was later confirmed by the ARDL Bound test, which also pointed towards no cointegration in the series. Different outcomes were presented in the panel analysis. While the results of panel cointegration tests were mainly inconclusive, the subsequent estimation of the ARDL using a dynamic pooled-mean group estimator indicated the long-run equilibrium.

Regarding the relationship between house prices and the most important fundamentals in our model, the income was found not statistically significant in the long runlong run by most models in both panel and time series analysis. On the other hand, the mortgage rate was found to have a a negative long-run relationship with house prices, which corresponds to previous literature and economic theory findings.

The overall effect of uncertainty index on house prices was not unequivocal. Panel analysis provided some evidence of a possible positive long-run relationship between economic policy uncertainty and house prices. This would imply that economic agents do not reduce spending on real estate during high periods of uncertainty but consider it a safe asset while adjusting their portfolios. On the other hand, the time series analysis did not find evidence of a long-run relationship between housing and financial market uncertainty. Instead, financial uncertainty had a short-run and mostly negative effects on house prices. Although both indexes are constructed differently, it is clear that additional research is needed to understand their overall behaviour.

# Bibliography

ABELSON, P., R. JOYEUX, G. MILUNOVICH, & D. CHUNG (2005): "Explaining House Prices in Australia: 1970-2003*." *Economic Record* **81(s1)**: pp. S96–S103. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-4932.2005.00243.x.

ALEXOPOULOS, M. & J. COHEN (2009): "Uncertain Times, uncertain measures." *Technical Report tecipa-352*, University of Toronto, Department of Economics. Publication Title: Working Papers.

ANDRÉ, C., L. BONGA-BONGA, R. GUPTA, & J. W. M. MWAMBA (2017): "Economic Policy Uncertainty, U.S. Real Housing Returns and Their Volatility: A Nonparametric Approach." *The Journal of Real Estate Research* **39(4)**: pp. 493–514. Publisher: American Real Estate Society.

AYE, G., M. CLANCE, & R. GUPTA (2019): "The Effect of Economic Uncertainty on the Housing Market Cycle." *Journal of Real Estate Portfolio Management* **25**: pp. 67–75.

BAHMANI-OSKOOEE, M. & S. H. GHODSI (2017): "Policy Uncertainty and House Prices in the United States." *The Journal of Real Estate Portfolio Management* **23(1)**: pp. 73–86. Publisher: American Real Estate Society.

BAKER, S. R. & N. BLOOM (2013): "Does Uncertainty Reduce Growth? Using Disasters as Natural Experiments." *Working Paper 19475*, National Bureau of Economic Research. Series: Working Paper Series.

BAKER, S. R., N. BLOOM, & S. J. DAVIS (2013): "Measuring Economic Policy Uncertainty." *SSRN Scholarly Paper ID 2198490*, Social Science Research Network, Rochester, NY.

BAKER, S. R., N. BLOOM, & S. J. DAVIS (2016): "Measuring Economic Policy Uncertainty*." *The Quarterly Journal of Economics* **131(4)**: pp. 1593–1636.

Bierens, H. J. & L. F. Martins (2010): "Time-varying cointegration." *Econometric Theory* **26(5)**: pp. 1453–1490.

Bloom, N. (2009): "The Impact of Uncertainty Shocks." *Econometrica* **77(3)**: pp. 623–685. Publisher: Econometric Society.

Bomberger, W. A. (1996): "Disagreement as a Measure of Uncertainty." *Journal of Money, Credit and Banking* **28(3)**: pp. 381–392. Publisher: [Wiley, Ohio State University Press].

Caggiano, G., E. Castelnuovo, & N. Groshenny (2014): "Uncertainty shocks and unemployment dynamics in U.S. recessions." *Journal of Monetary Economics* **67(C)**: pp. 78–92. Publisher: Elsevier.

Case, K. & R. Shiller (2003): "Is There a Bubble in the Housing Market?" *Brookings Papers on Economic Activity* **34(2)**: pp. 299–362. Publisher: Economic Studies Program, The Brookings Institution.

Choi, I. (2001): "Unit root tests for panel data." *Journal of International Money and Finance* **20(2)**: pp. 249–272.

Choudhry, T. (2020): "Economic Policy Uncertainty and House Prices: Evidence from Geographical Regions of England and Wales." *Real Estate Economics* **48(2)**: pp. 504–529. Publisher: American Real Estate and Urban Economics Association.

Clark, T. E., M. W. McCracken, & E. Mertens (2020): "Modeling Time-Varying Uncertainty of Multiple-Horizon Forecast Errors." *The Review of Economics and Statistics* **102(1)**: pp. 17–33.

Davis, S. J. (2016): "An index of global economic policy uncertainty." *Working Paper 22740*, National Bureau of Economic Research.

Dickey, D. A. & W. A. Fuller (1979): "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American Statistical Association* **74(366)**: pp. 427–431.

Engle, R. F. & C. W. J. Granger (1987): "Co-integration and error correction: Representation, estimation, and testing." *Econometrica* **55(2)**: pp. 251–276.

Gallin, J. (2006): "The Long-Run Relationship between House Prices and Income: Evidence from Local Housing Markets." *Real Estate Economics* **34(3)**: pp. 417–438. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6229.2006.00172.x.

Geng, N. (2018): "Fundamental Drivers of House Prices in Advanced Economies." *IMF Working Papers* **18**: p. 1.

Granger, C. (1981): "Some properties of time series data and their use in econometric model specification." *Journal of Econometrics* **16(1)**: pp. 121–130.

Hadri, K. (2000): "Testing for stationarity in heterogeneous panel data." *The Econometrics Journal* **3(2)**: pp. 148–161.

Hjalmarsson, E. & P. Österholm (2007): "Testing for Cointegration Using the Johansen Methodology When Variables are Near-Integrated."

Horvath, R. & S. Kapounek (2022): "Financial market uncertainty in the US: Measurement, trends, and effects." *Journal of Banking & Finance* **forthcoming**.

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

Johansen, S. (1988): "Statistical analysis of cointegration vectors." *Journal of Economic Dynamics and Control* **12(2)**: pp. 231–254.

Johansen, S. (1991): "Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models." *Econometrica* **59(6)**: pp. 1551–1580.

Johansen, S. & K. Juselius (1990): "Maximum likelihood estimation and inference on cointegration â€" with applications to the demand for money." *Oxford Bulletin of Economics and Statistics* **52(2)**: pp. 169–210.

Jordà, O., M. Schularick, & A. M. Taylor (2017): "Macrofinancial history and the new business cycle facts." *NBER Macroeconomics Annual* **31**: pp. 213–263, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press. Retrieved on 15.11.2022.

Jurado, K., S. C. Ludvigson, & S. Ng (2015): "Measuring Uncertainty." *American Economic Review* **105(3)**: pp. 1177–1216.

Juselius, K. (2006): *The Cointegrated VAR Model: Methodology and Applications*. Advanced Texts in Econometrics. OUP Oxford.

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

Kim, S.-W. & R. Bhattacharya (2009): "Regional Housing Prices in the USA: An Empirical Investigation of Nonlinearity." *The Journal of Real Estate Finance and Economics* **38(4)**: pp. 443–460. Publisher: Springer.

Kirikkaleli, D., K. Gokmenoglu, & S. Hesami (2021): "Economic policy uncertainty and house prices in Germany: evidence from GSADF and wavelet coherence techniques." *International Journal of Housing Markets and Analysis* **Vol. 14, No. 5**: pp. 842–859.

Knight, F. H. (1921): "Risk, Uncertainty and Profit." *SSRN Scholarly Paper ID 1496192*, Social Science Research Network, Rochester, NY.

Knoll, K., M. Schularick, & T. Steger (2017): "No price like home: Global house prices, 1870-2012." *American Economic Review* **107(2)**: pp. 331–53.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, & Y. Shin (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of Econometrics* **54(1)**: pp. 159–178.

Lamont, O. & J. Stein (1999): "Leverage and House-Price Dynamics in U.S. Cities." *RAND Journal of Economics* **30(3)**: pp. 498–514.

Ludvigson, S. & S. Ng (2007): "The empirical risk-return relation: A factor analysis approach." *Journal of Financial Economics* **83(1)**: pp. 171–222. Publisher: Elsevier.

Ludvigson, S. C., S. Ma, & S. Ng (2015): "Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?" *Working Paper 21803*, National Bureau of Economic Research. Series: Working Paper Series.

Ludvigson, S. C., S. Ma, & S. Ng (2021): "Uncertainty and business cycles: Exogenous impulse or endogenous response?" *American Economic Journal: Macroeconomics* **13(4)**: pp. 369–410.

MacKinnon, J. G., A. A. Haug, & L. Michelis (1999): "Numerical distribution functions of likelihood ratio tests for cointegration." *Journal of Applied Econometrics* **14(5)**: pp. 563–577.

Maddala, G. S. & S. Wu (1999): "A comparative study of unit root tests with panel data and a new simple test." *Oxford Bulletin of Economics and Statistics* **61(S1)**: pp. 631–652.

McQuinn, K. & G. O'Reilly (2006): "Assessing the Role of Income and Interest Rates in Determining House Prices." *Technical Report 15/RT/06*, Central Bank of Ireland. Publication Title: Research Technical Papers.

Mikhed, V. & P. Zemcik (2009): "Do house prices reflect fundamentals? Aggregate and panel data evidence." *Journal of Housing Economics* **18(2)**: pp. 140–149. Publisher: Elsevier.

Moore, A. (2017): "Measuring Economic Uncertainty and Its Effects." *Economic Record* **93(303)**: pp. 550–575. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-4932.12356.

Örsal, D. D. K. (2008): "Comparison of Panel Cointegration Tests." *Economics Bulletin* **3(6)**: pp. 1–20.

Pedroni, P. (1999): "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors." *Oxford Bulletin of Economics and Statistics* **61(S1)**: pp. 653–670. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0084.0610s1653.

Pesaran, M. H., Y. Shin, & R. J. Smith (2001): "Bounds testing approaches to the analysis of level relationships." *Journal of Applied Econometrics* **16(3)**: pp. 289–326.

Pesaran, M. H., Y. Shin, & R. P. Smith (1999): "Pooled mean group estimation of dynamic heterogeneous panels." *Journal of the American Statistical Association* **94(446)**: pp. 621–634.

Pesaran, M. H. & R. P. Smith (1998): "Structural analysis of cointegrating vars." *Journal of Economic Surveys* **12(5)**: pp. 471–505.

Phillips, P. C. B. & P. Perron (1988): "Testing for a unit root in time series regression." *Biometrika* **75(2)**: pp. 335–346.

Posedel, P. & M. Vizek (2009): "House price determinants in transition and EU-15 countries." *Post-Communist Economies* **21(3)**: pp. 327–343. Publisher: Taylor & Francis Journals.

Quigley, J. & M. Hwang (2006): "Economic Fundamentals in Local Housing Markets: Evidence from U.S. Metropolitan Regions." *Journal of Regional Science* **46**: pp. 425–453.

Rich, R., J. Song, & J. Tracy (2012): "The measurement and behavior of uncertainty: evidence from the ECB Survey of Professional Forecasters." *Staff Reports 588*, Federal Reserve Bank of New York.

Sheng, X. & M. Thevenot (2012): "A new measure of earnings forecast uncertainty." *Journal of Accounting and Economics* **53(1)**: pp. 21–33. Publisher: Elsevier.

Shiller, R. (1981): "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review* **71**: pp. 421–36.

Sims, C. A. (1980): "Macroeconomics and reality." *Econometrica* **48(1)**: pp. 1–48.

Su, D., X. Li, O.-R. Lobonţ, & Y. Zhao (2016): "Economic policy uncertainty and housing returns in Germany: Evidence from a bootstrap rolling window." *Zbornik Radova Ekonomskog Fakulteta u Rijeci : Časopis za Ekonomsku Teoriju i Praksu* **34**: pp. 43–61.

Tsatsaronis, K. & H. Zhu (2004): "What Drives Housing Price Dynamics: Cross-Country Evidence." *SSRN Scholarly Paper ID 1968425*, Social Science Research Network, Rochester, NY.

Vizek, M. & P. Posedel (2011): "Are House Prices Characterized by Threshold Effects? Evidence from Developed and Post-Transition Countries." *Czech Journal of Economics and Finance (Finance a uver)* **61**: pp. 584–600.

Wooldridge, J. M. (2013): *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning. Google-Books-ID: 4TZnpwAACAAJ.

ZHOU, J. (2010): "Testing for Cointegration between House Prices and Economic Fundamentals." *Real Estate Economics* **38(4)**: pp. 599–632. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6229.2010.00273.x.

# Appendix A

# Panel Analysis

Table A.1: Transformed full-range panel summary statistics

| | LHPNOM | LWAGE | LCONSP | LUNEMP | LLTRATE | LCPI | LPOLUNC |
|---|---|---|---|---|---|---|---|
| Mean | 1.02 | 0.29 | 3.40 | 1.47 | 1.52 | 1.21 | 4.72 |
| Median | 2.08 | 1.38 | 3.25 | 1.58 | 1.52 | 2.41 | 4.67 |
| Maximum | 6.23 | 5.65 | 4.79 | 3.21 | 3.01 | 5.31 | 6.30 |
| Minimum | -26.53 | -27.81 | 1.40 | -1.61 | -3.06 | -25.19 | 3.63 |
| Std. Dev. | 6.27 | 6.12 | 0.81 | 0.74 | 0.57 | 5.56 | 0.44 |
| Skewness | -3.31 | -3.12 | 0.14 | -0.58 | -1.86 | -3.59 | 0.77 |
| Kurtosis | 14.40 | 14.33 | 1.78 | 3.25 | 14.19 | 17.05 | 4.16 |
| Observations | 1220 | 1507 | 1504 | 1182 | 1495 | 1510 | 300 |

Table A.2: Results of Panel Unit Root Tests at Level - Unbalanced Dataset

| | Im, Pesaran and Shin | | ADF-Fisher type | | PP - Fisher type | |
|---|---|---|---|---|---|---|
| | Statistic | Prob. | Statistic | Prob. | Statistic | Prob. |
| lrconpc | 4.70 | 1.00 | 1.82 | 1.00 | 1.88 | 1.00 |
| lunemp | -3.76 | 0.00*** | 49.25 | 0.00*** | 67.14 | 0.00*** |
| ltrate | 1.60 | 0.95 | 6.84 | 0.99 | 10.68 | 0.95 |
| lcpi | 5.69 | 1.00 | 1.86 | 1.00 | 1.39 | 1.00 |
| lpolunc | -0.78 | 0.22 | 24.04 | 0.24 | 22.25 | 0.32 |

*** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level

Table A.3: Results of Panel Unit Root Tests at First Difference - Unbalanced Dataset

| | Im, Pesaran and Shin | | ADF-Fisher type | | PP - Fisher type | |
|---|---|---|---|---|---|---|
| | Statistic | Prob. | Statistic | Prob. | Statistic | Prob. |
| lrconpc | -25.15 | 0.00*** | 510.96 | 0.00*** | 674.30 | 0.00*** |
| lunemp | -30.53 | 0.00*** | 548.75 | 0.00*** | 552.78 | 0.00*** |
| ltrate | -22.34 | 0.00*** | 461.97 | 0.00*** | 675.43 | 0.00*** |
| lcpi | -16.25 | 0.00*** | 290.16 | 0.00*** | 313.34 | 0.00*** |
| lpolunc | -14.28 | 0.00*** | 187.79 | 0.00*** | 196.79 | 0.00*** |

*** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level

Table A.4: Results of Panel Unit Root Tests at Level - Balanced Dataset

| | Im, Pesaran and Shin | | ADF-Fisher type | | PP - Fisher type | |
|---|---|---|---|---|---|---|
| | Statistic | Prob. | Statistic | Prob. | Statistic | Prob. |
| lhpnom | -1.06 | 0.14 | 33.73 | 0.02** | 22.37 | 0.32 |
| lwage | 3.66 | 1.00 | 8.80 | 0.99 | 4.10 | 0.99 |
| lrconpc | 3.71 | 0.99 | 7.48 | 0.99 | 7.45 | 0.99 |
| lunemp | -4.30 | 0.00*** | 57.06 | 0.00*** | 25.88 | 0.16 |
| ltrate | 1.46 | 0.93 | 8.82 | 0.98 | 11.80 | 0.92 |
| lcpi | 0.69 | 0.75 | 15.68 | 0.74 | 15.19 | 0.76 |
| lpolunc | -0.57 | 0.28 | 22.87 | 0.30 | 21.31 | 0.38 |

*** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level

Table A.5: Results of Panel Unit Root Tests at First Difference - Balanced Dataset

| | Im, Pesaran and Shin | | ADF-Fisher type | | PP - Fisher type | |
|---|---|---|---|---|---|---|
| | Statistic | Prob. | Statistic | Prob. | Statistic | Prob. |
| lhpnom | -2.38 | 0.00*** | 36.20 | 0.01*** | 33.82 | 0.03** |
| lwage | -3.50 | 0.00*** | 51.64 | 0.00*** | 50.80 | 0.00*** |
| lrconpc | -1.52 | 0.06* | 53.15 | 0.00*** | 52.46 | 0.00*** |
| lunemp | -5.59 | 0.00*** | 72.42 | 0.00*** | 60.98 | 0.00*** |
| ltrate | -10.45 | 0.00*** | 129.96 | 0.00*** | 159.84 | 0.00*** |
| lcpi | -6.26 | 0.00*** | 74.74 | 0.00*** | 69.17 | 0.00*** |
| lpolunc | -12.58 | 0.00*** | 156.87 | 0.00*** | 163.05 | 0.00*** |

*** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level

Table A.6: Statistics of Pedroni (1999) Panel Cointegation Test for Full Range Sample, without Uncertainty

| | | | | |
|---|---|---|---|---|
| | Individual Intercept | | | |
| | Alternative hypothesis: common AR coefs. (within-dimension) | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 1.66 | 0.05 | 1.61 | 0.05 |
| Panel rho-Statistic | -0.43 | 0.33 | -0.87 | 0.19 |
| Panel PP-Statistic | -0.15 | 0.44 | -0.28 | 0.39 |
| Panel ADF-Statistic | 0.56 | 0.71 | -0.62 | 0.26 |
| | Alternative hypothesis: individual AR coefs. (between-dimension) | | | |
| | Statistic | Prob. | | |
| Group rho-Statistic | -0.64 | 0.26 | | |
| Group PP-Statistic | -0.29 | 0.38 | | |
| Group ADF-Statistic | -0.85 | 0.19 | | |
| | Individual Intercept & Trend | | | |
| | Alternative hypothesis: common AR coefs. (within-dimension) | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 0.74 | 0.23 | 0.90 | 0.18 |
| Panel rho-Statistic | -0.86 | 0.19 | -2.15 | 0.02 |
| Panel PP-Statistic | -1.09 | 0.14 | -2.24 | 0.01 |
| Panel ADF-Statistic | -1.53 | 0.06 | -2.77 | 0.00 |
| | Alternative hypothesis: individual AR coefs. (between-dimension) | | | |
| | Stat. | Prob. | | |
| Group rho-Statistic | -1.04 | 0.14 | | |
| Group PP-Statistic | -1.55 | 0.06 | | |
| Group ADF-Statistic | -2.81 | 0.00 | | |
| | None | | | |
| | Alternative hypothesis: common AR coefs. (within-dimension) | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | -0.54 | 0.70 | 0.39 | 0.35 |
| Panel rho-Statistic | 0.04 | 0.52 | -0.30 | 0.38 |
| Panel PP-Statistic | 0.10 | 0.54 | 0.05 | 0.52 |
| Panel ADF-Statistic | -0.37 | 0.36 | 0.06 | 0.52 |
| | Alternative hypothesis: individual AR coefs. (between-dimension) | | | |
| | Stat. | Prob. | | |
| Group rho-Statistic | 0.51 | 0.69 | | |
| Group PP-Statistic | 0.27 | 0.60 | | |
| Group ADF-Statistic | 0.29 | 0.61 | | |

Table A.7: Results of Kao (1999) Panel Cointegration Test for Unbalanced Dataset, without Uncertainty

| | | | | |
|---|---|---|---|---|
| | | | t-Statistic | Prob. |
| ADF | | | -3.74 | 0.00 |
| | | | | |
| Residual variance | | | 0.00 | |
| HAC variance | | | 0.00 | |
| | | | | |
| Augmented Dickey-Fuller Test Equation | | | | |
| Dependent Variable: D(RESID) | | | | |
| Method: Least Squares | | | | |
| Date: 12/13/22 Time: 13:34 | | | | |
| Sample (adjusted): 1988 2020 | | | | |
| Included observations: 269 after adjustments | | | | |
| | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| | | | | |
| RESID(-1) | -0.17 | 0.03 | -5.27 | 0.00 |
| D(RESID(-1)) | 0.23 | 0.06 | 3.66 | 0.00 |
| D(RESID(-2)) | 0.14 | 0.06 | 2.26 | 0.02 |
| | | | | |
| R-squared | 0.13 | Mean dependent var | | 0.00 |
| Adjusted R-squared | 0.12 | S.D. dependent var | | 0.06 |
| S.E. of regression | 0.05 | Akaike info criterion | | -3.00 |
| Sum squared resid | 0.76 | Schwarz criterion | | -2.96 |
| Log likelihood | 406.96 | Hannan-Quinn criter. | | -2.99 |
| Durbin-Watson stat | 1.88 | | | |

Table A.8: Statistics of Pedroni (1999) Panel Cointegation Test for Balanced Dataset

| | | | | |
|---|---|---|---|---|
| Individual Intercept | | | | |
| Alternative hypothesis: common AR coefs. (within-dimension) | | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 1.58 | 0.05 | 1.12 | 0.13 |
| Panel rho-Statistic | 1.73 | 0.95 | 1.70 | 0.95 |
| Panel PP-Statistic | 1.67 | 0.95 | 1.15 | 0.87 |
| Panel ADF-Statistic | 0.89 | 0.81 | 0.27 | 0.61 |
| Alternative hypothesis: individual AR coefs. (between-dimension) | | | | |
| | Statistic | Prob. | | |
| Group rho-Statistic | 3.15 | 0.99 | | |
| Group PP-Statistic | 1.98 | 0.97 | | |
| Group ADF-Statistic | 0.64 | 0.74 | | |
| Individual Intercept Trend | | | | |
| Alternative hypothesis: common AR coefs. (within-dimension) | | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 1.13 | 0.13 | 0.55 | 0.29 |
| Panel rho-Statistic | 3.08 | 0.99 | 2.87 | 0.99 |
| Panel PP-Statistic | 3.79 | 0.99 | 2.69 | 0.99 |
| Panel ADF-Statistic | 2.75 | 0.99 | 1.68 | 0.95 |
| Alternative hypothesis: individual AR coefs. (between-dimension) | | | | |
| | Stat. | Prob. | | |
| Group rho-Statistic | 4.05 | 1.00 | | |
| Group PP-Statistic | 3.23 | 0.99 | | |
| Group ADF-Statistic | 0.99 | 0.84 | | |
| None | | | | |
| Alternative hypothesis: common AR coefs. (within-dimension) | | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | -0.06 | 0.53 | -0.03 | 0.51 |
| Panel rho-Statistic | 1.06 | 0.86 | 0.98 | 0.83 |
| Panel PP-Statistic | 0.21 | 0.58 | -0.01 | 0.49 |
| Panel ADF-Statistic | -0.32 | 0.37 | -0.73 | 0.23 |
| Alternative hypothesis: individual AR coefs. (between-dimension) | | | | |
| | Stat. | Prob. | | |
| Group rho-Statistic | 2.28 | 0.98 | | |
| Group PP-Statistic | 0.55 | 0.71 | | |
| Group ADF-Statistic | -1.21 | 0.12 | | |

Table A.9: Results of Kao (1999) Panel Cointegration Test for Balanced Dataset

| | | | | |
|---|---|---|---|---|
| | | | t-Statistic | Prob. |
| ADF | | | -4.195181 | 0.0000 |
| Residual variance | | | 0.002027 | |
| HAC variance | | | 0.002878 | |
| Augmented Dickey-Fuller Test Equation | | | | |
| Dependent Variable: D(RESID) | | | | |
| Method: Least Squares | | | | |
| Date: 12/31/22 Time: 15:23 | | | | |
| Sample (adjusted): 1999 2020 | | | | |
| Included observations: 219 after adjustments | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| RESID(-1) | -0.202355 | 0.035894 | -5.637614 | 0.0000 |
| D(RESID(-1)) | 0.454727 | 0.069456 | 6.547017 | 0.0000 |
| R-squared | 0.222276 | Mean dependent var | | 0.003086 |
| Adjusted R-squared | 0.218692 | S.D. dependent var | | 0.051873 |
| S.E. of regression | 0.045851 | Akaike info criterion | | -3.317738 |
| Sum squared resid | 0.456207 | Schwarz criterion | | -3.286788 |
| Log likelihood | 365.2923 | Hannan-Quinn criter. | | -3.305238 |
| Durbin-Watson stat | 1.697445 | | | |

Table A.10: Statistics of Pedroni (1999) Panel Cointegration Test with Income and Long Term Rate as Determinants

| | | | | |
|---|---|---|---|---|
| Individual Intercept | | | | |
| Alternative hypothesis: common AR coefs. (within-dimension) | | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 1.51 | 0.12 | 1.53 | 0.06 |
| Panel rho-Statistic | -0.01 | 0.50 | 0.03 | 0.51 |
| Panel PP-Statistic | 0.28 | 0.61 | 0.54 | 0.70 |
| Panel ADF-Statistic | -1.19 | 0.12 | -0.56 | 0.28 |
| Alternative hypothesis: individual AR coefs. (between-dimension) | | | | |
| | Statistic | Prob. | | |
| Group rho-Statistic | 1.13 | 0.87 | | |
| Group PP-Statistic | 0.92 | 0.82 | | |
| Group ADF-Statistic | -1.08 | 0.13 | | |
| Individual Intercept Trend | | | | |
| Alternative hypothesis: common AR coefs. (within-dimension) | | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 0.38 | 0.35 | 0.47 | 0.32 |
| Panel rho-Statistic | 0.15 | 0.56 | -1.06 | 0.14 |
| Panel PP-Statistic | 0.29 | 0.61 | -0.35 | 0.36 |
| Panel ADF-Statistic | -0.17 | 0.43 | -1.66 | 0.04 |
| Alternative hypothesis: individual AR coefs. (between-dimension) | | | | |
| | Stat. | Prob. | | |
| Group rho-Statistic | 0.49 | 0.68 | | |
| Group PP-Statistic | 0.41 | 0.66 | | |
| Group ADF-Statistic | -1.28 | 0.10 | | |
| None | | | | |
| Alternative hypothesis: common AR coefs. (within-dimension) | | | | |
| | Stat. | Prob. | Weig. Stat. | Prob. |
| Panel v-Statistic | 0.39 | 0.34 | -0.31 | 0.62 |
| Panel rho-Statistic | 0.96 | 0.83 | 1.44 | 0.92 |
| Panel PP-Statistic | 1.12 | 0.87 | 1.89 | 0.97 |
| Panel ADF-Statistic | 1.27 | 0.89 | 1.78 | 0.96 |
| Alternative hypothesis: individual AR coefs. (between-dimension) | | | | |
| | Stat. | Prob. | | |
| Group rho-Statistic | 1.34 | 0.91 | | |
| Group PP-Statistic | 1.34 | 0.91 | | |
| Group ADF-Statistic | -0.18 | 0.42 | | |

Table A.11: Statistics of Kao (1999) Panel Cointegration Test with Income and Long Term Rate as Determinants

| | | | | |
|---|---|---|---|---|
| | | | t-Statistic | Prob. |
| ADF | | | -2.048672 | 0.0202 |
| | | | | |
| Residual variance | | | 0.009426 | |
| HAC variance | | | 0.011537 | |
| | | | | |
| Augmented Dickey-Fuller Test Equation | | | | |
| Dependent Variable: D(RESID) | | | | |
| Method: Least Squares | | | | |
| Date: 12/31/22 Time: 15:47 | | | | |
| Sample (adjusted): 1872 2020 | | | | |
| Included observations: 1189 after adjustments | | | | |
| | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| | | | | |
| RESID(-1) | -0.027508 | 0.006576 | -4.183178 | 0.0000 |
| D(RESID(-1)) | 0.174650 | 0.028342 | 6.162209 | 0.0000 |
| | | | | |
| R-squared | 0.039950 | Mean dependent var | | -0.002209 |
| Adjusted R-squared | 0.039141 | S.D. dependent var | | 0.107332 |
| S.E. of regression | 0.105211 | Akaike info criterion | | -1.664020 |
| Sum squared resid | 13.13928 | Schwarz criterion | | -1.655474 |
| Log likelihood | 991.2601 | Hannan-Quinn criter. | | -1.660799 |
| Durbin-Watson stat | 1.971322 | | | |

Table A.12: PMG Estimates for Individual Countries

**Australia**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.05 | 0.00 |
| D(LWAGE) | 0.07 | 0.94 |
| D(LCONSP) | 1.47 | 0.35 |
| D(LTRATE) | 0.07 | 0.00 |
| D(LCPI) | -2.19 | 0.40 |
| D(LUNEMP) | 0.28 | 0.00 |
| D(LPOLUNC) | 0.06 | 0.00 |
| C | -0.36 | 0.04 |

**Canada**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.11 | 0.00 |
| D(LWAGE) | 0.96 | 0.04 |
| D(LCONSP) | 0.99 | 0.10 |
| D(LTRATE) | -0.00 | 0.00 |
| D(LCPI) | 1.47 | 0.36 |
| D(LUNEMP) | 0.25 | 0.00 |
| D(LPOLUNC) | 0.00 | 0.01 |
| C | -1.24 | 0.01 |

**Germany**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.05 | 0.00 |
| D(LWAGE) | 0.37 | 0.48 |
| D(LCONSP) | -0.13 | 0.44 |
| D(LTRATE) | 0.01 | 0.00 |
| D(LCPI) | -0.32 | 0.43 |
| D(LUNEMP) | 0.02 | 0.04 |
| D(LPOLUNC) | 0.02 | 0.00 |
| C | -0.56 | 0.00 |

**France**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.12 | 0.00 |
| D(LWAGE) | 3.16 | 0.17 |
| D(LCONSP) | -1.82 | 0.03 |
| D(LTRATE) | 0.02 | 0.00 |
| D(LCPI) | 1.27 | 0.52 |
| D(LUNEMP) | -0.08 | 0.01 |
| D(LPOLUNC) | -0.04 | 0.00 |
| C | -1.42 | 0.02 |

**UK**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.02 | 0.00 |
| D(LWAGE) | 1.02 | 0.14 |
| D(LCONSP) | 0.31 | 0.15 |
| D(LTRATE) | 0.03 | 0.00 |
| D(LCPI) | -2.17 | 0.24 |
| D(LUNEMP) | -0.06 | 0.01 |
| D(LPOLUNC) | -0.01 | 0.00 |
| C | -0.17 | 0.23 |

**Italy**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.06 | 0.00 |
| D(LWAGE) | 0.74 | 0.00 |
| D(LCONSP) | -0.60 | 0.00 |
| D(LTRATE) | 0.01 | 0.00 |
| D(LCPI) | 0.69 | 0.09 |
| D(LUNEMP) | -0.10 | 0.00 |
| D(LPOLUNC) | -0.01 | 0.00 |
| C | -0.62 | 0.00 |

**Japan**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.02 | 0.00 |
| D(LWAGE) | 0.42 | 0.05 |
| D(LCONSP) | 0.14 | 0.55 |
| D(LTRATE) | -0.02 | 0.00 |
| D(LCPI) | 0.71 | 0.10 |
| D(LUNEMP) | 0.06 | 0.00 |
| D(LPOLUNC) | -0.01 | 0.00 |

| | | |
|---|---|---|
| C | -0.33 | 0.00 |

**Netherlands**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.11 | 0.00 |
| D(LWAGE) | 0.54 | 0.00 |
| D(LCONSP) | -0.71 | 0.00 |
| D(LTRATE) | 0.03 | 0.00 |
| D(LCPI) | -2.12 | 0.00 |
| D(LUNEMP) | -0.08 | 0.00 |
| D(LPOLUNC) | 0.02 | 0.00 |
| C | -1.20 | 0.01 |

**Sweden**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.03 | 0.00 |
| D(LWAGE) | 0.46 | 0.48 |
| D(LCONSP) | 0.60 | 0.04 |
| D(LTRATE) | 0.01 | 0.00 |
| D(LCPI) | -1.73 | 0.04 |
| D(LUNEMP) | 0.02 | 0.01 |
| D(LPOLUNC) | 0.05 | 0.00 |
| C | -0.18 | 0.00 |

**USA**

| Variable | Coefficient | Prob. * |
|---|---|---|
| COINTEQ01 | -0.18 | 0.00 |
| D(LWAGE) | 1.15 | 0.03 |
| D(LCONSP) | -0.18 | 0.80 |
| D(LTRATE) | 0.04 | 0.00 |
| D(LCPI) | -0.22 | 0.56 |

| | | |
|---|---|---|
| D(LUNEMP) | 0.11 | 0.00 |
| D(LPOLUNC) | -0.01 | 0.00 |
| C | -1.96 | 0.05 |

# Appendix B

# Time-Series Analysis

Table B.1: USA Time-Series Summary Statistics - Transformed

| | LHPINDEX | LINCOME | LMORTG | LPOP | LRENT | LCPI | LCOSTSW | LUNRATE | LFMUNC |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 4.82 | 9.17 | 1.78 | 12.58 | 5.34 | 5.22 | 2.95 | 1.74 | -0.12 |
| Median | 4.93 | 9.21 | 1.83 | 12.59 | 5.35 | 5.24 | 2.96 | 1.70 | -0.14 |
| Maximum | 5.64 | 10.10 | 2.42 | 12.71 | 5.87 | 5.64 | 3.44 | 2.69 | 0.44 |
| Minimum | 4.16 | 8.25 | 0.99 | 12.40 | 4.80 | 4.71 | 2.48 | 1.25 | -0.45 |
| Std. Dev. | 0.39 | 0.47 | 0.36 | 0.10 | 0.31 | 0.24 | 0.27 | 0.26 | 0.20 |
| Skewness | -0.05 | -0.20 | -0.22 | -0.28 | -0.02 | -0.32 | -0.03 | 0.55 | 0.45 |
| Kurtosis | 1.67 | 1.91 | 2.00 | 1.83 | 1.82 | 1.98 | 1.75 | 3.05 | 2.42 |
| Observations | 420 | 420 | 420 | 420 | 420 | 420 | 420 | 420 | 420 |

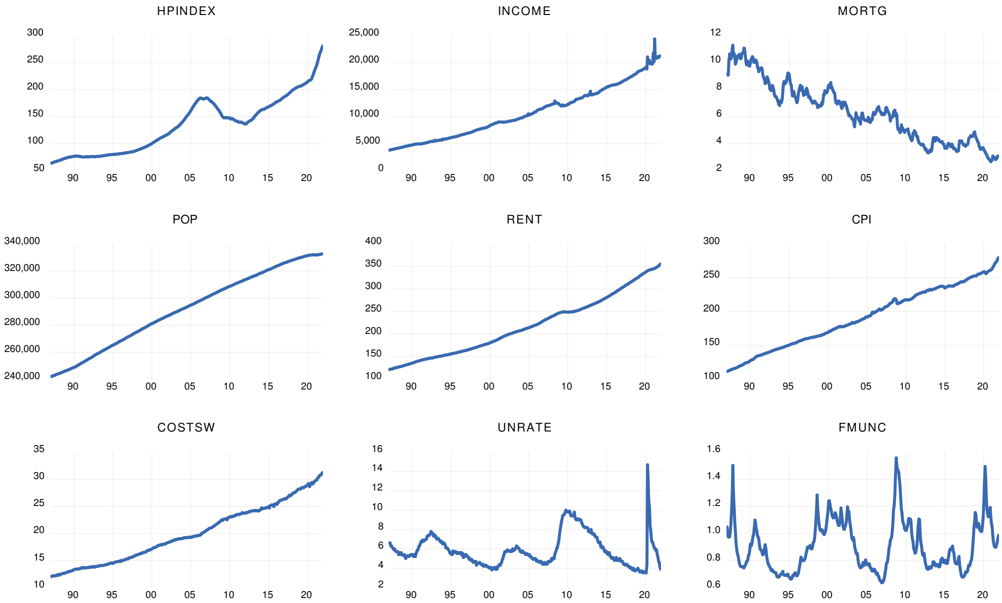


Figure B.1: Graphs of Variables in Time Series Analysis

## Table B.2: Phillips-Perron and Augmented Dickey-Fuller Unit Root Test Results

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Phillips-Perron test** | | | | | | | | | | |
| | At Level | | | | | | | | | |
| | | LHPINDEX | LINCOME | LMORTG | LPOP | LRENT | LCPI | LCOSTSW | LUNRATE | LFMUNC |
| With Constant | t-Statistic | 0.25 | -1.88 | -0.60 | -7.15 | -0.48 | -2.78 | 0.51 | -2.77 | -2.92 |
| | Prob. | 0.98 | 0.34 | 0.87 | 0.00 | 0.89 | 0.06 | 0.99 | 0.06 | 0.04 |
| | | | | | *** | | * | | * | ** |
| With Constant & Trend | t-Statistic | -1.29 | -3.02 | -3.72 | 4.46 | -2.08 | -3.07 | -1.96 | -2.78 | -2.98 |
| | Prob. | 0.89 | 0.13 | 0.02 | 1.00 | 0.56 | 0.11 | 0.62 | 0.21 | 0.14 |
| | | | | ** | | | | | | |
| Without Constant & Trend | t-Statistic | 3.87 | 10.44 | -1.52 | 16.80 | 18.64 | 10.95 | 14.70 | -0.75 | -2.46 |
| | Prob. | 1.00 | 1.00 | 0.12 | 1.00 | 1.00 | 1.00 | 1.00 | 0.39 | 0.01 |
| | | | | | | | | | | ** |
| | At First Difference | | | | | | | | | |
| | | d(LHPINDEX) | d(LINCOME) | d(LMORTG) | d(LPOP) | d(LRENT) | d(LCPI) | d(LCOSTSW) | d(LUNRATE) | d(LFMUNC) |
| With Constant | t-Statistic | -3.10 | -40.59 | -15.09 | -1.30 | -19.00 | -12.02 | -27.76 | -17.92 | -9.65 |
| | Prob. | 0.03 | 0.00 | 0.00 | 0.63 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | | ** | *** | *** | | *** | *** | *** | *** | *** |
| With Constant & Trend | t-Statistic | -3.17 | -42.15 | -15.08 | -3.69 | -19.00 | -12.09 | -27.75 | -17.90 | -9.63 |
| | Prob. | 0.09 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | | * | *** | *** | ** | *** | *** | *** | *** | *** |
| Without Constant & Trend | t-Statistic | -2.34 | -29.32 | -15.11 | -0.78 | -3.80 | -8.76 | -26.11 | -17.94 | -9.66 |
| | Prob. | 0.02 | 0.00 | 0.00 | 0.38 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | | ** | *** | *** | | *** | *** | *** | *** | *** |
| **Augmented Dickey-Fuller test** | | | | | | | | | | |
| | At Level | | | | | | | | | |
| | | LHPINDEX | LINCOME | LMORTG | LPOP | LRENT | LCPI | LCOSTSW | LUNRATE | LFMUNC |
| With Constant | t-Statistic | 0.27 | -1.66 | -0.60 | -3.64 | 0.48 | -2.24 | 0.53 | -2.94 | -3.62 |
| | Prob. | 0.98 | 0.45 | 0.87 | 0.01 | 0.99 | 0.19 | 0.99 | 0.04 | 0.01 |
| | | | | | *** | | | | ** | *** |
| With Constant & Trend | t-Statistic | -2.55 | -2.15 | -3.88 | 1.36 | -2.42 | -3.12 | -3.20 | -2.94 | -3.67 |
| | Prob. | 0.30 | 0.52 | 0.01 | 1.00 | 0.37 | 0.10 | 0.09 | 0.15 | 0.03 |
| | | | | ** | | | | * | | ** |
| Without Constant & Trend | t-Statistic | 1.70 | 3.55 | -1.50 | 0.60 | 3.87 | 3.17 | 2.74 | -0.76 | -2.87 |
| | Prob. | 0.98 | 1.00 | 0.12 | 0.85 | 1.00 | 1.00 | 1.00 | 0.39 | 0.00 |
| | | | | | | | | | | *** |
| | At First Difference | | | | | | | | | |
| | | d(LHPINDEX) | d(LINCOME) | d(LMORTG) | d(LPOP) | d(LRENT) | d(LCPI) | d(LCOSTSW) | d(LUNRATE) | d(LFMUNC) |
| With Constant | t-Statistic | -2.41 | -4.30 | -13.89 | -0.85 | -4.00 | -3.54 | -2.72 | -18.06 | -10.05 |
| | Prob. | 0.14 | 0.00 | 0.00 | 0.80 | 0.00 | 0.01 | 0.07 | 0.00 | 0.00 |
| | | | *** | *** | | *** | *** | * | *** | *** |
| With Constant & Trend | t-Statistic | -2.47 | -4.54 | -13.89 | -3.74 | -3.93 | -5.25 | -2.72 | -18.04 | -10.04 |
| | Prob. | 0.34 | 0.00 | 0.00 | 0.02 | 0.01 | 0.00 | 0.23 | 0.00 | 0.00 |
| | | | *** | *** | ** | ** | *** | | *** | *** |
| Without Constant & Trend | t-Statistic | -1.75 | -1.99 | -13.80 | -0.83 | -1.03 | -0.99 | -0.58 | -18.08 | -10.06 |
| | Prob. | 0.08 | 0.04 | 0.00 | 0.36 | 0.27 | 0.29 | 0.47 | 0.00 | 0.00 |
| | | * | ** | *** | | | | | *** | *** |

Notes: (*)Significant at the 10%; (**)Significant at the 5%; (***) Significant at the 1%. and (no) Not Significant

## Table B.3: KPSS Unit Root Test Results

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **KPSS Test** | | | | | | | | | | |
| | At Level | | | | | | | | | |
| | | LHPINDEX | LINCOME | LMORTG | LPOP | LRENT | LCPI | LCOSTSW | LUNRATE | LFMUNC |
| With Constant | t-Statistic | 2.38 | 2.55 | 2.42 | 2.55 | 2.57 | 2.55 | 2.58 | 0.14 | 0.10 |
| | Prob. | *** | *** | *** | *** | *** | *** | *** | | |
| With Constant & Trend | t-Statistic | 0.20 | 0.53 | 0.09 | 0.61 | 0.15 | 0.50 | 0.16 | 0.14 | 0.07 |
| | Prob. | ** | *** | | *** | ** | *** | ** | * | |
| | At First Difference | | | | | | | | | |
| | | d(LHPINDEX) | d(LINCOME) | d(LMORTG) | d(LPOP) | d(LRENT) | d(LCPI) | d(LCOSTSW) | d(LUNRATE) | d(LFMUNC) |
| With Constant | t-Statistic | 0.15 | 0.31 | 0.04 | 1.90 | 0.10 | 0.66 | 0.11 | 0.05 | 0.05 |
| | Prob. | | | | *** | | ** | | | |
| With Constant & Trend | t-Statistic | 0.14 | 0.08 | 0.02 | 0.28 | 0.08 | 0.16 | 0.09 | 0.05 | 0.03 |
| | Prob. | * | | | *** | | ** | | | |

Notes:
(*)Significant at the 10%; (**)Significant at the 5%; (***) Significant at the 1% and (no) Not Significant

Table B.4: Cointegration Equation 1 for Johansen test

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 Cointegrating Equation(s): | | Log likelihood | 10815.84 | | | |
| | | | | | | |
| Normalized cointegrating coefficients (standard error in parentheses) | | | | | | |
| LHPINDEX | LINCOME | LMORTG | LRENT | LCPI | LCOSTSW | LUNRATE |
| 1.00 | 0.44 | -1.25 | -6.83 | -2.17 | 5.79 | -1.25 |
| | (0.84) | (0.22) | (1.17) | (1.14) | (1.23) | (0.11) |
| | | | | | | |
| Adjustment coefficients (standard error in parentheses) | | | | | | |
| D(LHPINDEX) | 0.00 | | | | | |
| | (0.00) | | | | | |
| D(LINCOME) | 0.01 | | | | | |
| | (0.00) | | | | | |
| D(LMORTG) | 0.02 | | | | | |
| | (0.01) | | | | | |
| D(LRENT) | 0.00 | | | | | |
| | (0.00) | | | | | |
| D(LCPI) | 0.00 | | | | | |
| | (0.00) | | | | | |
| D(LCOSTSW) | 0.00 | | | | | |
| | (0.00) | | | | | |
| D(LUNRATE) | 0.06 | | | | | |
| | (0.01) | | | | | |

Table B.5: Cointegrating Equations 2 & 3 for a Full-scale Model

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 Cointegrating Equation(s): | | Log likelihood | 10849.55 | | | |
| | | | | | | |
| Normalized cointegrating coefficients (standard error in parentheses) | | | | | | |
| LHPINDEX | LINCOME | LMORTG | LRENT | LCPI | LCOSTSW | LUNRATE |
| 1.00 | 0.00 | -1.21 | -6.70 | -1.54 | 5.90 | -1.24 |
| | | (0.21) | (1.11) | (0.59) | (1.16) | (0.10) |
| 0.00 | 1.00 | -0.10 | -0.28 | -1.45 | -0.26 | -0.02 |
| | | (0.03) | (0.17) | (0.09) | (0.18) | (0.02) |
| | | | | | | |
| Adjustment coefficients (standard error in parentheses) | | | | | | |
| D(LHPINDEX) | -0.00 | 0.02 | | | | |
| | (0.00) | (0.00) | | | | |
| D(LINCOME) | 0.02 | -0.11 | | | | |
| | (0.00) | (0.03) | | | | |
| D(LMORTG) | 0.02 | -0.03 | | | | |
| | (0.01) | (0.07) | | | | |
| D(LRENT) | 0.00 | 0.00 | | | | |
| | (0.00) | (0.00) | | | | |
| D(LCPI) | -0.00 | 0.02 | | | | |
| | (0.00) | (0.01) | | | | |
| D(LCOSTSW) | -0.00 | 0.04 | | | | |
| | (0.00) | (0.01) | | | | |
| D(LUNRATE) | 0.09 | -0.43 | | | | |
| | (0.02) | (0.15) | | | | |
| | | | | | | |
| 3 Cointegrating Equation(s): | | Log likelihood | 10872.76 | | | |
| | | | | | | |
| Normalized cointegrating coefficients (standard error in parentheses) | | | | | | |
| LHPINDEX | LINCOME | LMORTG | LRENT | LCPI | LCOSTSW | LUNRATE |
| 1.00 | 0.00 | 0.00 | -7.06 | -1.55 | 7.79 | -1.59 |
| | | | (2.01) | (1.09) | (2.16) | (0.16) |
| 0.00 | 1.00 | 0.00 | -0.31 | -1.45 | -0.10 | -0.05 |
| | | | (0.20) | (0.11) | (0.22) | (0.02) |
| 0.00 | 0.00 | 1.00 | -0.29 | -0.01 | 1.56 | -0.29 |
| | | | (1.12) | (0.61) | (1.20) | (0.09) |
| | | | | | | |
| Adjustment coefficients (standard error in parentheses) | | | | | | |
| D(LHPINDEX) | -0.00 | 0.02 | -0.00 | | | |
| | (0.00) | (0.00) | (0.00) | | | |
| D(LINCOME) | 0.02 | -0.12 | -0.02 | | | |
| | (0.00) | (0.03) | (0.01) | | | |
| D(LMORTG) | 0.05 | -0.09 | -0.12 | | | |
| | (0.01) | (0.07) | (0.02) | | | |
| D(LRENT) | 0.00 | 0.00 | -0.00 | | | |
| | (0.00) | (0.00) | (0.00) | | | |
| D(LCPI) | -0.00 | 0.02 | -0.00 | | | |
| | (0.00) | (0.01) | (0.00) | | | |
| D(LCOSTSW) | -0.00 | 0.04 | 0.00 | | | |
| | (0.00) | (0.01) | (0.00) | | | |
| D(LUNRATE) | 0.09 | -0.43 | -0.07 | | | |
| | (0.02) | (0.15) | (0.04) | | | |

## Table B.6: Error-correction Terms for the VECM (1,3)

| Error Correction: | D(LHPINDEX) | D(LINCOME) | D(LMORTG) | D(LRENT) | D(LCPI) | D(LCOSTSW) | D(LUNRATE) |
|---|---|---|---|---|---|---|---|
| CointEq1 | -0.0001 | 0.0186 | 0.0488 | 0.0018 | -0.0001 | -0.0035 | 0.0904 |
| | (0.0005) | (0.0040) | (0.0092) | (0.0003) | (0.0007) | (0.0012) | (0.0206) |
| | [-0.19862] | [ 4.61206] | [ 5.31170] | [ 6.21739] | [-0.14501] | [-3.00958] | [ 4.37887] |
| CointEq2 | 0.0161 | -0.1213 | -0.0887 | 0.0025 | 0.0224 | 0.0398 | -0.4321 |
| | (0.0039) | (0.0298) | (0.0678) | (0.0021) | (0.0053) | (0.0085) | (0.1522) |
| | [ 4.14775] | [-4.07000] | [-1.30859] | [ 1.17184] | [ 4.22667] | [ 4.65753] | [-2.83897] |
| CointEq3 | -0.0038 | -0.0185 | -0.1222 | -0.0019 | -0.0030 | 0.0032 | -0.0726 |
| | (0.0011) | (0.0085) | (0.0194) | (0.0006) | (0.0015) | (0.0024) | (0.0436) |
| | [-3.37032] | [-2.16850] | [-6.28562] | [-3.02445] | [-1.99717] | [ 1.32557] | [-1.66318] |

## Table B.7: Short-run Coefficients for VECM (1,3)

| Error Correction: | D(LHPINDEX) | D(LINCOME) | D(LMORTG) | D(LRENT) | D(LCPI) | D(LCOSTSW) | D(LUNRATE) |
|---|---|---|---|---|---|---|---|
| D(LHPINDEX(-1)) | 0.8753 | 0.1178 | -0.8573 | -0.0331 | -0.0570 | -0.0882 | -2.0550 |
| | (0.0221) | (0.1692) | (0.3848) | (0.0122) | (0.0301) | (0.0485) | (0.8643) |
| | [ 39.5870] | [ 0.69652] | [-2.22761] | [-2.71987] | [-1.89620] | [-1.81833] | [-2.37761] |
| D(LINCOME(-1)) | -0.0090 | -0.5149 | -0.0273 | -0.0018 | -0.0133 | -0.0070 | -0.2676 |
| | (0.0060) | (0.0456) | (0.1037) | (0.0033) | (0.0081) | (0.0131) | (0.2329) |
| | [-1.50356] | [-11.2977] | [-0.26321] | [-0.53786] | [-1.64106] | [-0.53702] | [-1.14941] |
| D(LMORTG(-1)) | 0.0036 | 0.0249 | 0.2993 | -0.0002 | 0.0064 | 0.0030 | 0.0300 |
| | (0.0028) | (0.0211) | (0.0479) | (0.0015) | (0.0037) | (0.0060) | (0.1076) |
| | [ 1.30459] | [ 1.18196] | [ 6.24949] | [-0.13856] | [ 1.70135] | [ 0.49286] | [ 0.27877] |
| D(LRENT(-1)) | -0.0964 | -0.7918 | -0.5663 | 0.0638 | -0.2170 | 0.5182 | -3.1127 |
| | (0.0888) | (0.6794) | (1.5455) | (0.0488) | (0.1208) | (0.1947) | (3.4709) |
| | [-1.08610] | [-1.16534] | [-0.36641] | [ 1.30631] | [-1.79673] | [ 2.66079] | [-0.89680] |
| D(LCPI(-1)) | 0.0428 | -0.0035 | 1.6268 | 0.0046 | 0.4652 | 0.0911 | -2.9864 |
| | (0.0331) | (0.2535) | (0.5765) | (0.0182) | (0.0451) | (0.0726) | (1.2948) |
| | [ 1.29209] | [-0.01370] | [ 2.82160] | [ 0.25274] | [ 10.3230] | [ 1.25339] | [-2.30645] |
| D(LCOSTSW(-1)) | 0.0042 | -0.1463 | -0.2290 | -0.0113 | -0.0085 | -0.3374 | -0.5315 |
| | (0.0217) | (0.1662) | (0.3781) | (0.0119) | (0.0296) | (0.0476) | (0.8491) |
| | [ 0.19442] | [-0.88015] | [-0.60561] | [-0.94595] | [-0.28723] | [-7.08210] | [-0.62588] |
| D(LUNRATE(-1)) | -0.0021 | 0.0335 | 0.0357 | 0.0014 | -0.0004 | -0.0057 | 0.1711 |
| | (0.0014) | (0.0105) | (0.0238) | (0.0008) | (0.0019) | (0.0030) | (0.0535) |
| | [-1.51558] | [ 3.19389] | [ 1.49735] | [ 1.88156] | [-0.21249] | [-1.88875] | [ 3.19694] |
| C | 0.0006 | 0.0082 | -0.0002 | 0.0025 | 0.0020 | 0.0019 | 0.0230 |
| | (0.0003) | (0.0022) | (0.0050) | (0.0002) | (0.0004) | (0.0006) | (0.0111) |
| | [ 2.24531] | [ 3.78088] | [-0.04656] | [ 16.2320] | [ 5.22019] | [ 2.98693] | [ 2.06778] |
| R-squared | 0.9015 | 0.3396 | 0.1802 | 0.3214 | 0.2669 | 0.1828 | 0.1074 |
| Adj. R-squared | 0.8991 | 0.3234 | 0.1600 | 0.3048 | 0.2489 | 0.1627 | 0.0855 |
| Sum sq. resids | 0.0011 | 0.0662 | 0.3424 | 0.0003 | 0.0021 | 0.0054 | 1.7271 |
| S.E. equation | 0.0017 | 0.0128 | 0.0290 | 0.0009 | 0.0023 | 0.0037 | 0.0651 |
| F-statistic | 372.4282 | 20.9305 | 8.9457 | 19.2792 | 14.8188 | 9.1031 | 4.8985 |
| Log likelihood | 2086.4397 | 1235.8170 | 892.2851 | 2336.5946 | 1957.7477 | 1758.1271 | 554.0889 |
| Akaike AIC | -9.9303 | -5.8604 | -4.2167 | -11.1272 | -9.3146 | -8.3595 | -2.5985 |
| Schwarz SC | -9.8241 | -5.7542 | -4.1105 | -11.0211 | -9.2084 | -8.2533 | -2.4923 |
| Mean dependent | 0.0035 | 0.0041 | -0.0026 | 0.0026 | 0.0022 | 0.0023 | -0.0013 |
| S.D. dependent | 0.0052 | 0.0155 | 0.0316 | 0.0011 | 0.0026 | 0.0040 | 0.0681 |
| Determinant resid covariance (dof adj.) | | 0.0000 | | | | | |
| Determinant resid covariance | | 0.0000 | | | | | |
| Log likelihood | | 10872.7618 | | | | | |
| Akaike information criterion | | -51.5539 | | | | | |
| Schwarz criterion | | -50.6078 | | | | | |
| Number of coefficients | | 98.0000 | | | | | |

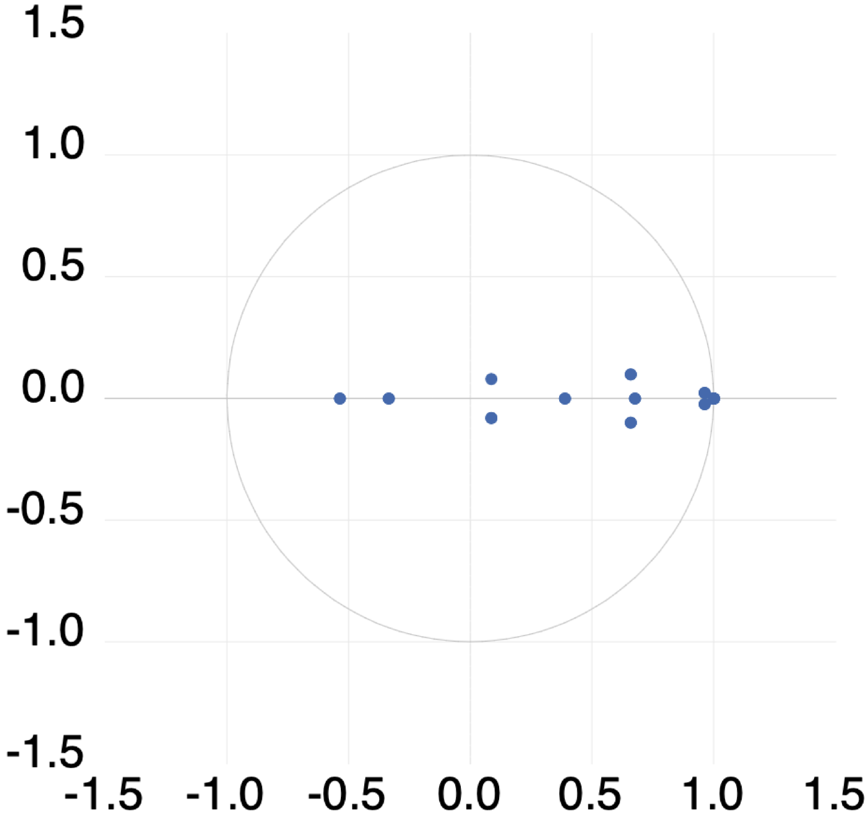


Figure B.2: Inverse Roots of AR Characteristic Polynomial

Table B.8: Error-correction Term for VECM(1,1)

| Error Correction: | D(LHPINDEX) | D(LINCOME) | D(LMORTG) |
|---|---|---|---|
| CointEq1 | 0.0007 | 0.0018 | 0.0181 |
| | (0.0002) | (0.0019) | (0.0043) |
| | [ 3.03229] | [ 0.93606] | [ 4.20908] |

Table B.9: ARDL Bounds Test - Full Model

| F-Bounds Test | | Null Hypothesis: No levels relationship | | |
|---|---|---|---|---|
| Test Statistic | Value | Signif. | I(0) | I(1) |
| | | | Asymptotic: n=1000 | |
| F-statistic | 3.027557 | 10% | 2.03 | 3.13 |
| k | 7 | 5% | 2.32 | 3.5 |
| | | 2.5% | 2.6 | 3.84 |
| | | 1% | 2.96 | 4.26 |
| Actual Sample Size | 418 | | Finite Sample: n=80 | |
| | | 10% | 2.129 | 3.289 |
| | | 5% | 2.476 | 3.746 |
| | | 1% | 3.233 | 4.76 |
| t-Bounds Test | | Null Hypothesis: No levels relationship | | |
| Test Statistic | Value | Signif. | I(0) | I(1) |
| t-statistic | -1.533291 | 10% | -2.57 | -4.23 |
| | | 5% | -2.86 | -4.57 |
| | | 2.5% | -3.13 | -4.85 |
| | | 1% | -3.43 | -5.19 |

Table B.10: ARDL Bounds Test - Reduced Model

| | | | | |
|---|---|---|---|---|
| F-Bounds Test | | Null Hypothesis: No levels relationship | | |
| Test Statistic | Value | Signif. | I(0) | I(1) |
| | | | Asymptotic: n=1000 | |
| F-statistic | 2.730430 | 10% | 2.72 | 3.77 |
| k | 3 | 5% | 3.23 | 4.35 |
| | | 2.5% | 3.69 | 4.89 |
| | | 1% | 4.29 | 5.61 |
| Actual Sample Size | 418 | | Finite Sample: n=80 | |
| | | 10% | 2.823 | 3.885 |
| | | 5% | 3.363 | 4.515 |
| | | 1% | 4.568 | 5.96 |
| t-Bounds Test | | Null Hypothesis: No levels relationship | | |
| Test Statistic | Value | Signif. | I(0) | I(1) |
| t-statistic | -0.975210 | 10% | -2.57 | -3.46 |
| | | 5% | -2.86 | -3.78 |
| | | 2.5% | -3.13 | -4.05 |
| | | 1% | -3.43 | -4.37 |

# Appendix C

# Data Sources

Baker, S. R., N. Bloom, & S. J. Davis (2016): 'Measuring Economic Policy Uncertainty*.' The Quarterly Journal of Economics 131(4): pp. 1593-1636.

Freddie Mac, 30-Year Fixed Rate Mortgage Average in the United States [MORTGAGE30US], retrieved from FRED, Federal Reserve Bank of St. Louis;
https://fred.stlouisfed.org/series/MORTGAGE30US, November 30, 2022.

Òscar Jordà, Moritz Schularick, and Alan M. Taylor. 2017. 'Macrofinancial History and the New Business Cycle Facts.' in NBER Macroeconomics Annual 2016, volume 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press. Retrieved on November 30, 2022.

Uncertainty Data: Macro and Financial Indexes. (n.d.). Sydney C. Ludvigson, Professor of Economics at New York University and NBER. Retrieved November 30, 2022, from
www.sydneyludvigson.com/macro-and-financial-uncertainty-indexes

S&P Dow Jones Indices LLC, S&P/Case-Shiller U.S. National Home Price Index [CSUSHPISA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CSUSHPISA, November 29, 2022.

U.S. Bureau of Labor Statistics, Average Hourly Earnings of Production and Nonsupervisory Employees, Construction [CES2000000008], re-

trieved from FRED, Federal Reserve Bank of St. Louis;
https://fred.stlouisfed.org/series/CES2000000008, November 30, 2022.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis;
https://fred.stlouisfed.org/series/CPIAUCSL, November 29, 2022.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Rent of Primary Residence in U.S. City Average [CUSR0000SEHA], retrieved from FRED, Federal Reserve Bank of St. Louis;
https://fred.stlouisfed.org/series/CUSR0000SEHA, November 30, 2022.

U.S. Bureau of Economic Analysis, Personal Income [PI], retrieved from FRED, Federal Reserve Bank of St. Louis;
https://fred.stlouisfed.org/series/PI, November 30, 2022.

U.S. Bureau of Economic Analysis, Population [POPTHM], retrieved from FRED, Federal Reserve Bank of St. Louis;
https://fred.stlouisfed.org/series/POPTHM, November 30, 2022.

U.S. Bureau of Labor Statistics, Unemployment Rate [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis;
https://fred.stlouisfed.org/series/UNRATE, November 30, 2022.