Josephine Dufitinema

Financial risk forecasting

A case study of the Finnish housing market

ACTA WASAENSIA 463

Vaasan yliopisto
UNIVERSITY OF VAASA

ACADEMIC DISSERTATION

To be presented, with the permission of the Board of the School of Technology and Innovation of the University of Vaasa, for public examination on the 10^{th} *of September, 2021, at noon.*

Reviewers

Professor Markku Juhani Lanne
University of Helsinki
Faculty of Social Sciences
Economics
P.O. Box 17 (Arkadiankatu 7)
00014 University of Helsinki
FINLAND

Professor Mikael Felix Linden
University of Eastern Finland
Faculty of Social Sciences and Business Studies
Department of Health and Social Management
BOX 1627, 70211 Kuopio
FINLAND

Julkaisija	Julkaisupäivämäärä
Vaasan yliopisto	Elokuu 2021

Tekijä(t) Josephine Dufitinema	**Julkaisun tyyppi** Artikkeliväitöskirja	
ORCID tunniste https://orcid.org/0000-0001-6716-1229	**Julkaisusarjan nimi, osan numero** Acta Wasaensia, 463	
Yhteystiedot Vaasan yliopisto Tekniikan ja innovaatiojohtamisen yksikkö Tilastotiede PL 700 FI-65101 VAASA	**ISBN** 978-952-476-962-4 (painettu) 978-952-476-963-1 (verkkoaineisto) http://urn.fi/URN:ISBN:978-952-476-963-1 **ISSN** 0355-2667 (Acta Wasaensia 463, painettu) 2323-9123 (Acta Wasaensia 463, verkkoaineisto)	
	Sivumäärä 187	**Kieli** English

Julkaisun nimike

Rahoitusriskien ennustaminen: Sovelluksena Suomen asuntomarkkinat

Tiivistelmä

Väitöskirjassa tutkitaan asuntojen hintojen volatiilisuuden aikariippuvuutta. Aineisto koostuu viideltätoista eri asuntoalueelta Suomesta. Sen mukaan, osoittautuuko hintavaihtelu aikariippuvaksi tai vakioksi, mallinnetaan hintasarjojen generointiprosessi, jota voidaan hyödyntää parhaiden ennusteiden tuottamisen tukena. Tutkimuksen tarkoituksena on myös tarjota markkinaosapuolille näkemys Suomen asuntomarkkinoiden tulevaisuudenkuvasta. Alueilla, joilla volatiilisuus osoittautuu vakioksi, tutkimuksessa verrataan autoregressiivisen liukuvan keskiarvon (ARMA) mallien ja fraktionaalisten integroituneiden ARMA-mallien (ARFIMA) ennusteominaisuuksia. Alueilla, joissa volatiilisuus osoittautuu ajasta riippuvaiseksi, verrataan kahta eri volatiliteetin mallintamistapaa: yleistettyä autoregressiivistä ehdollista heteroskedastisuusmallia (Genetralized AutoRegressive Conditional Heteroskedasticity GARCH) ja stokastisen volatiliteetin (SV) mallia. Tutkimuksen tulokset osoittavat vahvasti, että hintavaihtelu on aikariippuvaista useimmilla tutkituista alueista. Alueilla, joilla hintavaihtelu (hintojen tuottosarja) ei ole aikariippuvaista, ARFIMA- ja ARMA-mallit tuottavat kahtalaisia tuloksia asunnon koon ja ennusteiden mukaan. Ennusteissa pitkän aikavälin riippuvuuksia huomioiva ARFIMA osoittautuu ARMA-mallia paremmaksi. Alueilla, joissa hintavaihtelu on aikariippuvaista, niin sanottu leverage-efekti osoittautuu tärkeäksi komponentiksi volatilisuuden mallintamisessa niin deterministissä kuin stokastisessakin tapauksessa. Näillä alueilla deterministisen prosessin GARCH-mallit tuottavat kuitenkin parempia volatiliteettiennusteita kuin SV-mallit. Suomen asuntomarkkinoiden näkymien suhteen tutkimusten tulokset antavat olettaa, että useimmilla alueilla markkinoiden kasvu jatkuisi periodilla 2019–2021. Hintojen nousun ja sitä kautta tuottojen kasvun voidaan odottaa kuitenkin laantuvan joillakin alueilla ja samalla hintojen vaihtelun lisääntyvän.

Asiasanat

Suomen alueet; Asuntojen hinnat; Mallinnus; Ennustaminen; Tuotot; Volatiliteetti

Publisher Vaasan yliopisto	Date of publication August 2021	
Author(s) Josephine Dufitinema	**Type of publication** Doctoral thesis by publication	
ORCID identifier https://orcid.org/0000-0001-6716-1229	**Name and number of series** Acta Wasaensia, 463	
Contact information University of Vaasa School of Technology and Innovation Statistics P.O. Box 700 FI-65101 Vaasa Finland	**ISBN** 978-952-476-962-4 (print) 978-952-476-963-1 (online) http://urn.fi/URN:ISBN:978-952-476-963-1 **ISSN** 0355-2667 (Acta Wasaensia 463, print) 2323-9123 (Acta Wasaensia 463, online)	
	Number of pages 187	**Language** English

Title of publication
Financial risk forecasting: A case study of the Finnish housing market

Abstract

The housing market sector is an essential component of the economy of most developed countries. Forecasting house price movements is crucial for investment decision making, designing housing policies, asset allocation, and risk management. This dissertation aims to examine whether the Finnish house prices of fifteen main regions display constant or time-varying variances. Depending on whether the variance is constant or time-varying, a time-series generating process will be established that provides superior forecast. Finally, the thesis aims to offer to the market players an outlook of the Finnish housing markets. The methodology used compares the Autoregressive Moving Average (ARMA) models and AR Fractionally Integrated MA (ARFIMA) models for regions with constant variances. For regions with time-varying variances, two classes of time-series volatility models are compared: Generalised AR Conditional Heteroscedasticity (GARCH)-type and Stochastic Volatility (SV) models.

The study outcomes reveal strong evidence of clustering effects in the house price returns of most of the studied regions. The two models for modelling house price returns for areas with homoscedastic errors yield mixed results in the in-sample performance. In out-of-sample (forecasting), ARFIMA models tend to outclass ARMA models in return predictions. In areas with time varying volatility, models accounting for leverage yield the best in-sample fits for both deterministic and stochastic volatility models. However, in forecasting (out-of-sample) price changes and volatilities of these regions, the GARCH-types models outperform their SV counterparts. Regarding the Finnish housing market outlook, it is predicted that most regions will experience continuous growth in the 2019 – 2021 period. However, some areas are expected to experience a decline in house price returns and a high price fluctuation.

Keywords
Finland regions, House prices, Modelling, Forecasting, Returns, Volatility

Dedication

To our mother, Phoibe Nyiramucyo, who sacrificed everything for our education. She always says to us "The only heritage I can give you is education". To our father, Haguminema Joseph, from him, we inherited the love of science.

ACKNOWLEDGEMENTS

The completion of this doctoral thesis would not have been possible without the help and support of several people I would gratefully like to thank. First and foremost, I would like to express my gratitude to my supervisor Professor Seppo Pynnönen, who continually offered support, advice and guidance throughout my doctoral studies. From our weekly meetings to the Zoom calls, he was always there to provide helpful comments and research advice. He led me to research in the applied statistics and econometric fields and taught me an intuitive way to look at data and pay careful attention to details. I am sincerely indebted to him to ensure that I have the necessary resources and time to do my research properly. It has been an honour for me to work under his supervision and benefit from his scientific knowledge. I would also like to thank Professor Tommi Sottinen, whose help has been extremely beneficial and essential during all phases of my doctoral studies. His ideas and suggestions have significantly contributed to the improvement of this study.

During my studies, I had the privilege of getting to know and working with bright individuals from the Mathematics and Statistics Unit, led by Professor Seppo Hassi. Many special thanks to you all for showing great trust in me and giving me the opportunity to do postgraduate studies and improve my expertise in an inspiring working environment.

I would also like to express my warmest thanks to Professor Markku Juhani Lanne (University of Helsinki) and Professor Mikael Felix Linden (University of Eastern Finland) for acting as the pre-examiners of my dissertation. Their constructive suggestions and comments helped improve the thesis. Additionally, I would like to thank Professor Juho Kanniainen (University of Tampere) for accepting to act as the opponent for the public examination.

I gratefully acknowledge the University of Vaasa, the Foundation for Economic Education (Liikesivistyrahasto), and the Finnish Cultural Foundation for providing financial support and sources that made my PhD studies possible.

I want to thank my father, Joseph Haguminema, and my mother, Phoibe Nyiramucyo. They guided my first steps and made sure that I have the right and proper education. Even when times were very hard, they sacrificed everything for the education of my sisters and me and always reminded us of the importance of hard work. I wish to thank my sisters Josepha Ufitinema, Joseline Uwinema, and the youngest member of our family, Inema Teta Jenny Louange, for their unconditional support and love. I appreciate your encouragements throughout my studies since day one.

I extend my gratitude to my in-laws family, the Belinga family, for their support. Last but not least, special recognition goes to my husband, Dr Mvola Belinga Eric, who has been the front and centre of this exciting journey. He has been a true guide and a source of motivation, inspiration and strong and reliable support. Our daily discussions helped me navigate the challenges of PhD life and understand the tricks of academic writing. He worked tirelessly to make this achievement possible, and in the last stages of my studies, when I felt overwhelmed by the writing process, he told me, “Every long journey starts with a single step, so one step at the time.”

I thank all my friends who showed me generous support, love and kindness.

Most importantly, I thank God for this opportunity and all the graces.

Dufitinema Josephine
August 2021
Vaasa, Finland

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

ACF	AutoCorrelation Function
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criteria
APARCH	Asymmetric Power ARCH
ARCH	Autoregressive Conditional Heteroscedasticity
ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BIC	Bayesian Information Criteria
BM	Brownian Motion
CAPM	Capital Asset Pricing Model
CGARCH-M	Component Generalised ARCH-in-mean
COVID-19	Corona Virus Disease-19
DIC	Deviance Information Criterion
DMA	Dynamic Model Averaging
DMS	Dynamic Model Selection
ECM	Error-Correction Model
EGARCH-M	Exponential GARCH-in-mean
EWMA	Exponential Weighted Moving Average
FBM	Fractional Brownian Motion
FGN	Fractional Gaussian Noise
FIGARCH	Fractional Integrated GARCH
GAR	Generalised AR
GARCH	Generalised Autoregressive Conditional Heteroscedasticity
GARCH-M	GARCH-in-mean
GDP	Gross Domestic Product

GJR-GARCH	Glosten, Jagannthan, and Runkle GARCH
GMP	Gross Metropolitan Product
KTI	Kiinteistötieto Oy
LRD	Long-Range Dependence
MAE	Mean Absolute Error
MCMC	Markov Chain Monte Carlo
MFBM	Mixed Fractional Brownian Motion
MLE	Maximum Likelihood Estimation
MSA	Metropolitan Statistical Area
MSM	Markov-Switching Multifractal
OECD	Organisation of Economic Co-operation and Development
P-P	Phillips-Perron
PCA	Principal Component Analysis
PLS	Partial Least Squares
RMSE	Root Mean Squared Error
SPLS	Sparse Partial Least Squares
SV	Stochastic Volatility
SWARCH	Switching ARCH
TGARCH	Threshold GARCH
UK	United Kingdom
US	United States
VAR	Vector Autoregressions
VEC	Vector Error-Correction
ZIP	Zone Improvement Plan

The dissertation is based on the following four refereed articles:

(I) Dufitinema, J. and Pynnönen, S. (2020). Long-range dependence in the returns and volatility of the Finnish housing market, *Journal of European Real Estate Research*, Volume 13, Issue 1, pp. 29-50.

(II) Dufitinema, J. (2020). Volatility clustering, risk-return relationship and asymmetric adjustment in the Finnish housing market, *International Journal of Housing Markets and Analysis*, Volume 13, Issue 4, pp. 661-688.

(III) Dufitinema, J. (2021). Stochastic volatility forecasting of the Finnish housing market, *Applied Economics*, Volume 53, Issue 1, pp. 98-114.

(IV) Dufitinema, J. (2021). Forecasting the Finnish house price returns and volatility: A comparison of Time Series Models, *International Journal of Housing Markets and Analysis*, DOI: https://doi.org/10.1108/IJHMA-12-2020-0145.

AUTHOR'S CONTRIBUTION

Publication I: "Long-range dependence in the returns and volatility of the Finnish housing market"

The author is the principal investigator in this article. She conducted the data collection, analysis and reporting. The other author contributed by commenting and reviewing the work to improve its quality.

Publication II: "Volatility clustering, risk-return relationship and asymmetric adjustment in the Finnish housing market"

This article is the independent work of the author. Constructive discussions, comments, suggestions and advice from Professor Seppo Pynnönen are acknowledged.

Publication III: "Stochastic volatility forecasting of the Finnish housing market"

This article is the independent work of the author. Constructive discussions, comments, suggestions and advice from Professor Seppo Pynnönen are acknowledged.

Publication IV: "Forecasting the Finnish house price returns and volatility: A comparison of Time Series Models"

This article is the independent work of the author. Constructive discussions, comments, suggestions and advice from Professor Seppo Pynnönen are acknowledged.

1 INTRODUCTION

Modelling, forecasting, and monitoring risk are at the heart of economic, financial theories and practices. Maximising an investment return while limiting risk is the general target for investors and asset managers. Thus, financial risk management has become a vast field, and one of its evolving components is risk measurement. The quantification of the asset prices volatility – the widely used measure of risk - is a good foundation for assessing investment risk. Accordingly, the field of financial econometrics dedicates substantial attention to asset volatility and the tools for its measuring, modelling, and forecasting. As a result, studies focusing on the volatility of different asset classes, such as equities, bonds, commodities, and currencies, have been gaining importance. One asset class, however, the housing assets are notably different from others in terms of return, risk, and liquidity. Housing market risk modelling and forecasting, particularly, the Finnish housing market is the subject of this work.

The introductory section of this dissertation presents the research framework. That is, the motivation, goals and objectives of the study, the methodology and the study's contribution to the housing market analysis.

1.1 Background

Purchasing a dwelling is by far the single significant transaction in the lifetime of most people. In addition to being a dwelling, a house also composes a significant portion in many households' wealth. For instance, Campbell and Cocco (2003) have shown that assets of more than half of middle-class American families are in the form of housing. This wealth effect of housing on consumption is substantial, and it is even larger than the wealth effects of financial assets (Oikarinen, 2007, and references therein). In Finland, over half of the households' total wealth (50.3 per cent) is in the form of housing (Statistics Finland, 2016). In the United States (US), housing is the largest component of household wealth; it represented, respectively, 28.3 and 24.6 per cent of the total households' net worth and households' asset (Financial Accounts Data, 2018). The United Kingdom (UK) housing stock total value is estimated to £7.29 trillion (Savills, 2019). Thus, housing, its market, in particular, is a vital component of the country's economy.

Housing assets, in addition to being consumed, they are also an investment. This dual role of investment and consumption makes them unique among various categories of consumption and investment. The durability makes the housing consumption aspect special and stands out from other forms of consumption. As an investment, housing, unlike other assets, has a material presence and provides a flow of housing services to the owner. Given the critical role of the housing sector in the economy and the fact that a substantial proportion of total economic wealth and individual assets are maintained through housing, Shiller (1998) asserted that housing market risk is one of the most considerable personal financial risks encountered by individuals.

1.2 Motivation

The housing market, as it is a tremendous component of the country economy, developments in house prices is essential for economic activities. Several factors contribute to the importance of house prices and the reasons why investors, policymakers, and consumers should monitor house price movements. First, house prices drive the housing construction sector activities. An increase in house prices in relation to building costs will boost the profit of the new housing construction projects, and thereby stimulate the housing investment. In Finland, the housing construction sector has also been a major booster of the substantial investment in residential properties, with the most active construction in apartment buildings. In 2020, up to 5000 flats were completed in the Helsinki metropolitan region, and up to 3000 in other largest cities (KTI, 2020). It is an increase compared to the previous year, which saw approximately 4000 completed apartments in the metropolitan area and 2000 in other regions. These flats are 100 per cent targeted for the investment market.

Second, housing as a significant element of households' wealth, a decline in house prices will reduce this household wealth, and possibly give rise to the mortgage-secured loans, curbing the housing consumption and the level of activity in the country's economy (Jacobsen & Naug, 2005). Third, the banking sector is perhaps the major party exposed to housing and mortgage activities, as housing plays a crucial role as collateral in the mortgage-security loans. Thus, a fall in house prices, meaning a collateral value decline will induce the banks' loan losses, which can increase even further if the borrowers are unable to service their debt. Subsequently, mortgage lenders may become more reserved in providing loans, putting a strain to the country's financial system, and leading to an economic downturn (Bori & Lowe, 2002; International Monetary Fund, April, 2003). Last, other parties are also heavily involved in housing and related securities, and thereby they are key components of the country overall economy. Those are mortgage market, mortgage insurance, and backed securities (Miller & Peng, 2006).

Having noted the importance of the housing market, analysing, and understanding individual house prices dynamics would be beneficial and play a vital role in investment decision making, designing housing policies, asset allocation, portfolio, and risk management. Particularly, the examination of house prices of not only the national level but also the region, city and sub-area levels can facilitate investors in their portfolio diversification. This aspect is due to the factor that within a country, housing markets are distinct, and it is not recommended to analyse a country's house prices as if it formed one consistent market. Moreover, Oikarinen and Asposalo (2004) emphasised that within a country, variances in regional's social and economic structures contribute to the local housing diversification accessibility. Furthermore, regarding risk in the financial market, the results associated with other assets such as stocks might not be directly applicable in the case of housing. This investment and consequently, risk difference is due to various housing special features, namely, heterogeneity, lack of public market place, durability, asset maintenance costs (taxes), and the housing's dual role (consumption and investment). Therefore, the investigation of the risk in the housing market, in particular, the risk-return relationship is an essential task and have profound implications on investment and policy decision making. An extensive discussion on risk in the housing market is provided in section 2.3.

1.3 Research questions

The research questions are:

i. **Is there long-range dependence behaviour in both returns and volatility of the Finnish housing market?**
This question assesses the presence of long memory behaviour in the returns and volatility of the studied type of dwellings. The presence of long-range dependence behaviour in the house price returns offers evidence of a high degree of predictability of the asset based on historical information. The evidence of long memory in the house price volatility plays a crucial role in the development of appropriate time series volatility forecasting models of the studied market.

ii. **Is there volatility clustering in the Finnish housing market? If yes, what is the nature of the time-varying volatility? What is the relation between house price returns and house price risk? Are there asymmetric effects in the house price volatility?**
These questions propose to investigate whether the studied types of dwellings (apartments) manifest the Autoregressive Conditional Heteroscedasticity (ARCH) effects. Further, they explore the volatility properties of the Finnish housing market.

iii. **Does the Finnish house price volatility follow a stochastic evolution?**
This question identifies whether the volatility of the studied type of dwellings follows a stochastic evolution. That is, it evaluates the volatility using a

stochastic equation and assesses whether Stochastic Volatility (SV) models can capture and forecast the volatility of the studied type of dwellings.

iv. **Which time series models perform better in-sample and out-of-sample forecasting for the Finnish housing market?**
This question aims to compare different time series models' performances in modelling and forecasting the Finnish house price returns and volatility; in the viewpoint of developing suitable time-series forecasting models of this housing market.

1.4 Objectives

The objectives of the work are four-fold. The first is to examine, in fifteen regions of Finland, whether the variance of their house price returns is constant or time-varying. The second is to investigate in both cases, constant or time-varying variance, whether time-series data generating processes can be used to model the considered return series. The third is to establish which time series forecasting model yielded superior out-of-sample forecasts for house price returns as well as house price volatilities. The fourth is to provide an outlook of the Finnish housing market in terms of three-years forecasts for both house price returns and volatilities. Each of these goals is expanded on below.

First, the examination of constant or time-varying variance in the house price returns is crucial in order to detect volatility clustering effects in the studied series. These effects also called ARCH effects refer to the presence of periods of higher volatility swing followed by calm volatility periods. If a series is found to exhibit these ARCH structures, it implies that the variance of the corresponding regional housing market is time-dependent. In this case, a volatility model is required to capture the series data generating mechanism. On the other hand, if a series manifest a constant structure, an Autoregressive Moving Average (ARMA) model is used to mimic the price returns behaviour. Second, within the two groups, regions with linear forms and regions with ARCH structures, the abilities of the short and long memory time-series models are compared, to model the considered house price returns and volatilities. This approach is due to the higher degree of long-range dependence found in the house price returns as well as volatilities for a greater number of the Finnish regions. Third, the out-of-sample forecasting performance of the competing models is assessed, to provide for every region, the accurate model for house price returns and volatility forecasting. Last, prospects of the Finnish housing market are given in terms of three-years forecasts, to offer to the market players the likelihood of market development.

1.5 Research methods and research development

The research approaches used are literature review and empirical analysis. The data are collected, followed by data wrangling and analysis to answer the above research questions and meet the study's goals.

The following part of the dissertation presents the overall structure of the research context. The current state of the art in the housing market is described with an emphasis on the Finnish housing market. After that, a discussion on risks and the risk-return relationship in the studied market is given. Next, a review on modelling and forecasting the returns behaviour of the considered market is provided, followed by an introduction of categories of methods used in financial modelling and forecasting. Finally, the study presents and examines the empirical analysis results. The results of these analyses, done in conjunction with the four publications presented in the second part of the dissertation, are analysed, and conclusions are drawn. Table 1 gives an overview of the dissertation structure and content. It consists of two parts, where Part I, the main part of the work, is based on the results of the four publications presented in Part II of the study.

Table 1. Overview of the dissertation.

Part I		
Chapter	**Aim of chapter**	**Record**
Introduction	Outline of the area of interest	Background, research motivation, questions, objectives, approaches, and an indication of the research gap.
State of the art	Literature review	A view of the area of study, and its latest level of development.
Methods	Research methodology	The setting of the empirical analysis, data acquisition, wrangling, and analysis.
Results and Discussions	Outcomes from all the published articles and their discussions	Presentation of the empirical findings and an analysis of the study's findings with regards to the research questions.

Conclusions	A synthesis of the work with the main key points	Conclusions of the study, noting the implications and providing perspectives for further research.
Part II: Published Articles		

1.6 Contribution to the housing market analysis

This study is a contribution to the empirical assessment of the housing markets, with a focus on the Finnish housing market. There is a wide range of empirical studies on financial markets, however, only a restricted number on housing. The importance of the housing market and house prices in relation to various economic and financial sectors makes the review timely. Moreover, in the current state of lower interest rates, higher and turbulent stock prices, investors are seeking new targets of which housing markets, and real estate in general, have been of paramount interests lately. Hence, insights into house price dynamics are the fundamental input in asset allocation and investment decision making.

Furthermore, among the housing markets analysis studies, only a few focused on the region's level. However, the housing market is subjected to local social and economic structures. Therefore, instead of analysing the housing market at the national level, the study examines house prices at the city and sub-area level for a cross-assessment and comparison of housing investment on the city and sub-market levels. Similarly, previous studies used the family-home property type data sets; this research, however, employs apartments (also referred to as, block of flats) type data. The number of rooms categorises the studied dwellings: one-room, two-rooms, and larger apartments (over three rooms) types. The reasons for using flats property type data are their fast-growing popularity as a place to live in Finland and their increased attractiveness in both consumers and investors. At the end of 2018, Statistics Finland Overview reported that apartments counted for nearly half of all occupied dwellings, they represented 46 per cent. Detached and semi-detached was the second favourable house type, with 39 per cent, followed by terraced with 14 per cent. Regarding the investment aspect, apartments continue to strengthen their position in the Finnish residential property market with foreign, domestic as well as individual investors continue to increase their portfolios across the country (KTI, 2020). Thus, a thorough knowledge and understanding of the house price dynamics of these types of dwellings favoured by investors is primary for risk and portfolio management.

Additionally, to quantify the risk in the housing market, the used approaches allow modelling the time-dependent mean and variance of the studied series simultaneously, instead of assuming a constant variance. The techniques are commonly used in finance research; however, their application in the housing market is still quite limited. The use of these methods is relevant as the time-varying volatility is also observed in the housing market; meaning that during the turbulent periods, the probability of large losses is higher than what the standard mean-variance portfolio analysis would indicate. Therefore, the employed methods enable to recognise this time-varying characteristic of the house price volatility, as failing to do so would underestimate the actual house price risk.

Moreover, the long-range dependence aspect is at the core of the used methodology. Financial assets such as equities have been found to exhibit a high persistence in both their returns and volatility. The presence of the long-range dependence property, also known as long memory, in returns implies a high level of predictability of the asset returns at long horizons, questioning the validity of weak form efficiency. The evidence of long memory behaviour in asset volatility sustains the development of suitable time series models that describe and forecast asset volatility. This high level of autocorrelation was also found in real estate returns and volatility, and individual housing markets. Notably, in this study, a higher degree of persistence was found in both house price returns and volatilities of the studied type of dwellings. Therefore, incorporating this crucial property in modelling and forecasting of the house prices dynamics enables to capture and describe the magnitude and the pattern of house price returns and volatilities, and consequently assess the price risk of investing in residential markets. Again, a failure to account for this persistence nature can cause an underestimation of the probability of large losses or prices drops.

2 STATE OF THE ART

In recent years, there has been an increasing interest in understanding the housing markets dynamics of different countries and regions. However, compared to other financial assets, the empirical analysis of housing assets is considerably limited. The theoretical motives, employed approaches and empirical outcomes vary depending on the aspect under study. In the viewpoint of this study's goals, this section reviews the current state of the housing markets in general, and the Finnish housing market, in particular. Discussions on the risk, risk-return relationship and long-range dependence in the housing markets follow. Next, an overview of the literature on modelling and forecasting house prices is provided. Finally, an outline on house price volatility forecasting is presented.

2.1 The housing markets developments

House prices in most industrialised countries have experienced significant growth since the 1990s, with a notable decline in awake of the 2008 economic and financial crisis. In the Organisation of Economic Co-operation and Development (OECD) countries, between 1970 and 2003, house prices underwent four periods of expansion, followed by three periods of contraction (Lecat & Mésonnier, 2005). With the use of a statistical method that unveils boom and bust periods, Lecat and Mésonnier demonstrated that house prices of half of the sample OECD countries experienced an exceptionally substantial period of expansion between 1995 and 2003. These countries include Australia, France, Spain, the Netherlands, Denmark, UK, Canada, and the US. In the mid-2004, house price indices of some countries started to flag; these include the UK, the Netherland, and Australia (Reserve Bank of Australia, 2004). Meanwhile, house prices in France were continuing to face a sharp increase at a rate considered unsustainable (Moëc, 2004).

By 2006, problems in the run-up to the 2008 financial crisis were started to loom, house prices in some countries first surged to the unprecedented heights, then undertake a sinking trend from 2007 onwards. As Baugnet et al. (2011) reported, during the 1996-2007 upward phase, on average, house prices climbed up to 44 per cent with a substantial increase in countries such as UK, France, Spain, and Ireland. An acceleration phase was also recorded in the US, Belgium, the Netherlands, and Finland; however, to a lower extent than the other four countries. The exception was for Germany and Japan; in the past decades, the two countries did not report housing booms, their house price movements have drifted from this overall pattern (Agnello & Schuknecht, 2010; Engsted & Pedersen, 2014). Between 2007 and 2008, housing markets entered the downturn phase, with the exception of the US housing market, which in 2005 has already begun to experience a decline in house sales. The severity of the downward phase varied from one county to another. For instance, Ireland recorded a 35 per cent fall in cumulative house price, whereas the figures were between 15 and 20 per cent in Spain, the UK, and the US. Finland, same as the Netherland, and France experienced a price reduction ranging between 5 and 10 per cent (Baugnet et al., 2011).

In the late 2009 and beginning of 2010, several countries, including Finland, were in the recovery stage. In the last few years, the after crisis stabilisation of the various housing markets has been noted. However, with the ongoing health crisis – the COVID-19 pandemic – the increased uncertainty contributes to the loss of market players confidence and potentially impact the development of the housing markets (Cecchetti et al., 2020). On the other hand, there is some optimism among researchers that the current situation would considerably differ from the 2000s financial crisis due to measures such as the central banks' stimulus packages (KTI, 2020). Though, with forecasts subjected to significant uncertainty and the pandemic effects on various parts of the economy remain unknown; the extent to which the crisis will impact the housing market is still obscure.

2.1.1 Housing market characteristics

Housing assets hold special features which distinguish them from other financial assets such as equities and bonds. These differences are significant as they have substantial implications in asset allocation, portfolio diversification and investment decision making. First, housing is heterogeneous. Unlike other homogenous financial assets commonly found on the trading market; a dwelling is an individual unit defined by its size, age, and unique location. These features participate in determining the dwelling's market value. Thus, with no immediate information available, market participants may have to use other means such as previous sales in the nearby area to evaluate various dwellings' market values. Second, housing has no public market place. This aspect contributes even more to the lack of house price information. For instance, unlike for stock markets, where information on prices is available on their public market places; collections of details on a particular dwelling such as location, neighbourhood, and residential environments are often tasks for market participants. Moreover, even with more efforts put in the information-gathering process, there is the imperfect knowledge or asymmetric information. That is, the seller is usually more informed about the actual dwelling's characteristics than the buyer.

Third, housing markets are characterised by higher transaction costs. Initially, direct investment in housing requires a significant amount of capital. On top of that, other costs such as searches costs, moving expenses, and taxation contribute to these housing's high transactions. Finally, housing is a long-term or durable good. The process of planning and building new dwellings takes time, which means that a significant proportion of the housing supply consists of already existed dwellings, making the housing supply slow in adjusting to changes in market conditions. Similarly, housing devaluation is slow, meaning that house prices may exhibit a significant autocorrelation in short to medium run.

2.2 The Finnish housing market

The Finnish housing market is no exception with regards to the above-discussed housing developments. Since the 1970s, it has experienced significant upswings and downswings. In the beginning of the 1970s with the baby-boom generation, the housing demand increased, house prices surged, and housing construction climbed to its height (70,000 dwellings per year). Following the 1973 oil crisis, the Finnish house prices took a lengthly downturn (Kivistö, 2012). The early 1980s, prices started to gather pace; in the second half of 1980s, the Finnish economy entered an overheating period, and house prices skyrocketed in 1987-1989. During the period of a little over two years, house prices rose by 60 per cent, followed by a steep downturn that lasted four years (1989-1993). This striking development of the Finnish housing market ranks as the most well-known event of the entire period (Laakso, 2000). Moreover, the housing sector was recognised to play an essential part in the 1980s overheating period and 1900s depression. One of the reasons for this housing boom was the structural changes in the housing markets. That is the 1986 financial deregulation by Bank of Finland, which improved access to the mortgages and relaxed the down payment ratios. As a result, there was a rapid increase in bank lending and large capital inflows accompanied by the financial system's inadequate supervision, eventually leading to the housing market boom (Honkapohja, 2009). Additionally, Kosonen (1997) noted that house prices' fast fall was further accelerated by a decrease in household real income and mass unemployment.

It was only in 1996 that house prices turned to a permanent trend. This period coincided with Finland's entree into the European Union and an inflation decline. A small downward trend in house prices was short-lived in 2001 due to the ”dot-com” bubble, after which a fast growth was renewed and continued until the 2008 financial crisis. The sub-prime crisis forced prices into a downturn; however, the downward trend lasted only one year. Thereafter, house prices have risen, and by 2011 they were above the pre-crisis levels. Compared to other Nordic countries, Finland presents a more moderate house price developments with a 27 per cent increase in real house prices between 2000 and 2019. Denmark follows with a 47 per cent increase in the same period. Whereas, Norway and Sweden have experienced remarkable house prices growth with 109 and 147 per cent rates, respectively (Anundsen, 2020).

Laakso (2000) pointed out that demographic factors such as population growth and its age structure are among the elements that influence house prices developments. In addition, the size and form of households also play a crucial role in determining house prices. Currently, in Finland, urbanisation is the main booster of regional house price fluctuations. In 1990, urban area counted for 60 per cent of the Finnish population; by 2017, the share has climbed up to 70 per cent (Kaleva, 2019). Moreover, Mankiw and Weil (1989) and Kuismanen et al. (1999) have stressed the importance of the age groups with regards to the housing consumption demand, highlighting that the per capita housing consumption rapidly increases within the 20 to 29 years old group age. Therefore, as the majority of regional migrants in Finland are working-age populations or young adults, there has been an increased housing demand in urban areas such as the Helsinki metropolitan area (the main economic centre in Finland). This strong demand led to a rise in house prices, and these price changes diffused in the surrounding regions (Oikarinen, 2005). Furthermore, the Finnish household's average size decreased continuously; it stood to 2.01 persons in 2017. In the largest cities such as Helsinki, 48 per cent of the households are single-person (Kaleva, 2019). This pattern puts pressure on house prices, especially for small and well-located flats, and boost housing construction of apartment buildings. In 2018, studios and one-bedrooms flats occupied 75 per cent of the newly built dwellings (Statistics Finland, 2019).

2.2.1 Structure and features of the Finnish housing market

Over the past years, the Finnish residential properties have manifested a fast growth and set up their position as the property investment market's largest sector. Having amounted up to €20 billion in 2018; it rose to €25 billion by the end of 2019 and represented 32 per cent of the total invested universe. Kiinteistötieto Oy (KTI) indicated that substantial increase was due to brisk capital growth and increased volume through new developments. The residential market development has, however, showed differentiation between the various parts of the country. Largest cities such as the Helsinki metropolitan area continue to stand out as a result of its accelerate growth in demand and higher transaction volumes. Other regions that indicate positive development are Turku, Tampere and Oulu. Also, Kuopio, Jyväskylä, and Lahti demonstrate a stable outlook development. In this Finnish residential market increasing polarisation, regions beyond these main cities' limits have found themselves outside the largest investors' radar.

In Finland, there are two forms of housing owner-occupied and rental. The former is the widespread form. In 2018, around 63 per cent of Finnish households lived in owner-occupied homes; the figure was down by one per cent compared to the previous year (Kaleva, 2020). The latter form has also increased its attractiveness; Statistics Finland Overview reported that by the end of 2019, more than one-quarter of the population were living in rented dwellings. Renting has become a common status, especially in major cities. In the viewpoint of the thesis's purpose of providing insights into housing investment decision making and asset allocation; the rental housing form is thoroughly discussed below.

At the end of 2019, there were approximately 3 million dwellings in Finland; 33 per cent of the total stock were rental dwellings. Rented housing has become more profound in large cities such as Helsinki, where currently, the share of rental exceeds owner-occupied; 49 per cent of all dwellings are rented. The increase of the rental housing form appreciation is associated with the decline of the Finnish household's average size. Among those living in rental dwellings, 87 per cent were a household of one or two persons. Age also plays a crucial role in preferring a rental status; it attracts the young generation in particular. Statistics Finland Overview (2019) recorded that, among the rental dwellings habitants, 79 per cent were a household with the oldest person aged under 30, 38 per cent were aged between 30 to 44, and 22 per cent between 45 to 74.

The rental market comprises of two sectors, subsidised and the non-subsidised. The former refers to the rental dwellings that have been provided with particular public subsidies such as a state-guaranteed loan or an interest subsidy. The Finnish municipalities own the majority of subsidised housing stock. In this sector, selling is controlled and bonded to various regulations. The latter includes all investor groups, professional investors, property companies and funds, institutions, foreign investors, and private investors. Finnish households, as well as small companies, have also entered this market segment recently. By 2018, the share of rental dwellings between the two sectors was as follows: among a total of 900,000 dwellings, 42 per cent were subsidised and 58 per cent non-subsidised [36 per cent owned by households/private investors, while professional investors owned 22 per cent] (Kaleva, 2020). The non-subsidised stock continues to increase, mainly due to the withdrawal of the subsidised dwellings' restrictions. It is also the primary driver of the housing construction, especially in apartment buildings, to address the high demand caused by urbanisation increase and household size decrease. At the end of 2019, nearly 13,000 apartment buildings were in completion; all these flats were targeted 100 per cent for investment/rental market.

2.3 Risk in housing markets

The topic of risk is deeply considered in the financial area. Having its roots in Markowitz (1952)'s mean-variance analysis framework; it developed into the Capital Asset Pricing Model (CAPM) (Fama, 1968; French, 2003; Sharpe, 1964). In the case of housing assets with their investment and consumption properties, a fundamental scientific question that can be asked is "If housing assets are risky; risky

for whom?". In other words, it is crucial to describe whose perspectives the risk assessment applies to in housing markets analysis. To answer this question, one has first to specify housing risk and its measure. Berg et al. (2012) define housing market risk as the expected distance between the observed percentage return and the expected percentage return. It is measured by the standard deviation computed from a set of observations as follows:
Suppose an observed time series of house prices, $p_0, p_1, ...p_N$, where p_t represent an index value or a single property value observed at time t. The time $t+1$ return approximated by natural logarithmic is defined as:

$$R_{t+1} = \ln(p_{t+1}) - \ln(p_t). \tag{2.1}$$

With a sample of N returns $R_1, R_2, .., R_N$, the expected return is calculated as a sample arithmetic average:

$$\bar{R} = \frac{1}{N}\sum_{n=1}^{N} R_N.$$

The variance of housing returns is estimated as :

$$\sigma^2 = \frac{1}{N-1}\sum_{n=1}^{N}(R_N - \bar{R})^2.$$

The estimated standard deviation of housing returns is, therefore, the square root of the variance.
Another term used in risk analysis is volatility. It is widely used in private-sector and academic research worlds, finance, baking, and beyond. It is the synonym of the above-defined standard deviation distance measure. Both refer to how spread out percentage returns are relative to the average return. Hence, when the term "house price volatility" is used in the text, it refers to the housing risk.

Keeping the housing risk definition and its measure in mind, and in the standpoint of the study's goals of offering perceptions in housing investment and portfolio allocation, the above-asked question is answered by providing interpretations of "housing investment" and "housing investor". The former term refers to the rented out dwellings, meaning those dwellings whose purpose is not to fulfil the proprietor's housing consumption demands, rather generate the capital in the form of rental cash flows. The latter term refers to an individual or other entity (institutions, property companies and funds) with a small or large housing portfolio. Therefore, this study's housing risk analysis perspectives apply to the residential investors investing in properties at which they do not reside, not to the homeowner (owner-occupant) who enjoys the housing services. Wilhelmsson and Zhao (2018) provides a discussion on housing risks from the homeowner's perspectives.

Several forms of risks are associated with housing investment feature. First, the rental risk; it is linked to the uncertainty of finding the appropriate tenant, paying the agreed rent, and damaging the property that yields to its value reduction. Second, the industry risk; unlike other financial assets that require a small proportion of capital for their investment, proper housing investment and portfolio diversification call for a considerable amount of money. This risk is due to the large unit size and indivisibility of housing investments in the real estate market. Consequently, individual investors/households possess at most one or two dwellings, whereas property companies, funds and institutions own large housing portfolios. Third, the planning risk; it is associated with the real estate, including housing market sensibility to the changes in social-economic characteristics such as interest rates. Fourth, the expense risk; it is related to the unforeseen maintenance and repair costs, and physical structure depreciation. Fifth, the tax risk; it is affiliated to the tax changes or new taxes consideration of particular investments. Sixth, the legislative and juridical risks; they are associated with rental agreement juridical issues and legislation changes. Finally, liquidity risk; it is linked to high transactions and low liquidity criteria of housing assets. Huffman (2003) has analysed and divided the real estate risks into physical, financial, and regulatory risks.

2.4 Risk-return relationship for housing

The risk-return relation is the core of asset valuation. With a considerable literature focusing on assets class such as stocks (see, for instance, Guo & Nelly, 2008); no agreement has been attained regarding the sign of the risk-return interaction in the finance mainstream literature. On the one hand, some findings support Merton (1973) concept of risk-averse investors' requirement of a high return as compensation for the increased risk; thereby, observing a positive risk-return relationship. On the other hand, some outcomes take the Glosten et al. (1993) stand on the investors' acceptance of a lower return during volatile periods as they sense a riskier future. In other words, investors may be willing to accept a lower return if they feel that it is a better hedging strategy against a higher riskier future. This effect could lead to more significant savings and lower return; thereby, a negative risk-return relationship would be documented. Additionally, Whitelaw (1994) proposed the time-varying behaviour of the asset's risk-return interaction. Therefore, no consensus has been reached to the claim that this relation would be negative or positive. Several studies such as Guo and Whitelaw (2006) and Scruggs (1998) have targeted to reconcile these conflicting findings; though neither focused on housing.

In the housing literature, the risk-return relationship has been extensively studied in the US housing market. Using the asset pricing framework, empirical findings in favour of a positive risk-return association for housing have been provided by Meyer and Wieand (1996), Crone and Voith (1999), Cannon et al. (2006) and Case et al. (2011). However, it is the use of Generalised Autoregressive Conditional Heteroscedasticity (GARCH)-based models that dominate the literature. Concrete examples include Dolde and Tirtiroglu (1997) who documented, a positive and negative risk-return tradeoff for respectively San Francisco and Connecticut areas. Miller and Peng (2006) took a step further in their risk-return investigation; the authors compared Vector Autoregressions (VAR) with GARCH models on the Metropolitan Statistical Area (MSA) level. Their findings also yield mixed results. However, Milles (2008a) criticised the aforementioned studies by pointing out the lack of the ARCH effects evidence in the Dolde and Tirtiroglu's considered municipality areas, as the authors did not first test these clustering effects. Moreover, Milles highlighted that risk could be at a more extensive region for real estate investors than the metropolitan area as covered by Miller and Peng. Specifically, the author investigated the risk-return relationship at the state level. Plus precisely, on twenty-eight states which were found to exhibit clustering effects; and mixed results of negative and positive risk-returns, were also noted in eight states. Han (2013) addressed the housing risk-return issue on the MSA level by examining cross-market differences. The author further explained the identified negative relationship through the lenses of three local housing factors, namely housing supply constraints, household's hedging incentives, and urban market growth.

Various international studies also reported similar evidence in the housing markets of different countries such as the UK where Milles (2011b) found a significant positive risk-return relation in Wales and a negative one in East Midlands. Morley and Thomas (2011; 2016) analysed England regions and Wales; and a positive risk-return link was evident in most of the studied areas except for the South West where a negative link was found. Cook and Watson (2017) studied the London submarket, emphasising on demonstrating how empirical design decisions impact the housing risk-return inferences. The authors' empirical design components that could influence the risk-return examination outcome included sample selection, variable descriptions, modelling techniques, optimisation methods, regional disaggregation, and dynamic specification. The frequently discussed housing's negative risk-return relation did rise in their findings when particular options were selected from the studied components. On the other hand, relatively balanced mixed results were obtained when thorough analysis involving optimisation was conducted. Therefore, the authors recommended that empiricists carefully consider the sometimes overlooked and implicit empirical design assumptions and decisions in their risk-return analysis, as they can impact inferences.

In the Canadian housing market, Lin and Fuerst (2014) examined the risk-return relationship across provinces. Their findings yield to a positive one in Ontario and Quebec and a negative one in British Columbus. In contrast to these mixed results, homogenous outcomes were obtained by C. L. Lee (2017) in the case of the Australian housing market, in both national and capital city level. The author assigned these conclusive results across the studied regions to use the enhanced model, the Component-GARCH-in-mean (CGARCH-M) model. In the Finnish housing market, this dissertation's second article investigates the matter by employing the GARCH-in-mean (GARCH-M) model. Similar to the studies above, the risk-return relation sign's variation was observed in all three apartment types and across considered cities and sub-areas. Therefore, it can be observed that even in the growing housing market analysis literature, various countries' empirical results provide inconclusive evidence with regards to the risk-return relationship. These differences can be noted across regions; that is why it is recommended not to analyse the country's housing market as one coherent market, rather as distinct regional housing markets. In this thesis, the risk-return relation is investigated on the city and submarket level for cross-analysis and comparison for Finnish cities and sub-areas housing investments.

2.5 Long-range dependence - An overview

The empirical evidence of long-range dependence (LRD) or long memory in time series data has emerged in various financial and economic studies. The concept refers to a phenomenon that describes strong correlations in a time series. In other words, a long-range dependent time series process is characterised by a very slow decay of its autocorrelation function (ACF) as the number of lags increases. The presence of long-range dependence in the assets returns calls into question the efficient market hypothesis. That is, the suggestion that assets returns are unpredictable. Under this hypothesis, the asset prices require to follow a martingale process such as a random walk, in which the asset's price history does not affect its price change. In case that the asset return series are long-term dependent, then there is a positive autocorrelation between distant observations. In this context, future returns are predicted by past price returns; a complete departure from the efficient market hypothesis. Therefore, the investigation of financial market efficiency is directly linked to the presence of long memory in the asset returns.

The importance of long memory processes came to light in the 1960s with a series of papers of the leading figure in developing these processes, the late Benoit B. Mandelbrot. Motivated by a fascinating study in hydrology by Hurst (1951) whose research purpose was to control the Nile River flows by developing a system with a series of reservoirs and dams. The early publications of Mandelbrot and his co-authors, namely Mandelbrot (1963; 1965), Mandelbrot and Van Ness (1968), Mandelbrot and Wallis (1968), and Mandelbrot (1983) paved the way in this field. These articles started the scientific community's debates with a detailed study of fractional Brownian motion (FBM) and long memory processes. The Mandelbrot's work was further discretised independently by Granger and Joyeux (1980) and Hosking (1981) to introduce the Autoregressive Fractionally Integrated Moving Average (ARFIMA) process, a simple extension of the classical Box and Jenkins (1970)'s Autoregressive Integrated Moving Average (ARIMA) models. Since then, long-range dependence has also been found to be of great importance not only in econometrics (Robinson, 2003) and finance (Lo, 2001) but also in other fields such as internet modelling (Dileep & Gupta, 2020), linguistics (Alvarez-Lacalle et al., 2006), DNA sequencing (Karmeshu & Krishnamachari, 2004), and climate studies (Varotsos & Kirk-Davidoff, 2006). A brief history of long memory processes is given in Graves et al. (2017), and an excellent review is outlined in Samorodnitsky (2006) and Beran et al. (2013).

Evidence of long term persistence has been investigated in various assets classes, either in their returns and/or their volatility. Christodoulou-Volos and Siokis (2006) and Ólan (2002) noted long-range dependence in stock returns. Baillie et al. (2007) identified evidence of long memory properties in the volatility of daily commodity futures. In energy futures such as propane, gasoline, oil, and heating oil, Cunado et al. (2010) found that their volatility exhibited a long term persistence. In housing markets, a high volatility persistence was documented by Tsai et al. (2010) in the UK housing market. Barros et al. (2013) assessed the degree of persistence of house prices of 69 Chinese cities. A high degree of dependence was mainly found in house prices of Sanya, Shanghai, Haikou, and Shenzhen. Other cities manifested a mean reversion behaviour. In the US housing market, out of 62 MSA studied by Milles (2011a), over half of them, especially the MSA on the West Coast displayed long memory in their house price volatility. Moreover, a higher degree of long-term behaviour was found by Elder and Villupuram (2012) in the house price returns as well as the volatility of 14 and 10 US city and composite indices, respectively. Barros et al. (2015) investigated the matter on the metropolitan and state levels; their empirical analysis also yielded similar evidence in house price volatility of the considered sample.

In the Finnish housing market, this dissertation's first article explores the evidence of long-range dependence in the returns and volatilities of the studied regions' house prices. Analysing the degree of persistence in house price returns and volatilities is of paramount importance for investment, portfolio and risk management. One motive is that the presence of long memory behaviour in the house price returns indicates a higher level of the asset predictability; again, a complete deviation from the efficient market hypothesis. The other motive is that the evidence of long-term dependence in the house price volatilities supports developing adequate time series forecasting models for the studied housing market volatilities. Across all three apartment types, a high level of persistence was found in price returns. Moreover, house price volatility of over half of the studied sample exhibited a long memory behaviour. These results were incorporated in the modelling as well as forecasting procedures of the house price returns and volatilities dynamics of the considered dwellings types.

2.6 Modelling and forecasting house prices (returns)

Understanding house prices dynamics through modelling and forecasting is of utmost importance to a significant number of sectors: real estate investors monitor the trend of the current and future house prices as drivers of their investment decision-making. Construction companies and housing policymakers outline their policies and plans based on the recommendations taken from house price dynamics insights. Also, consumers allocate their current and future consumptions based on accurate predictions of house prices' future movement. Therefore, academic and professional research has recently intensified to shed light on various countries' house prices dynamics.

Regarding modelling house prices, as discussed above, housing assets have special characteristics, including heterogeneity associated with each dwelling's unique hedonic features, such as age, size and location. Moreover, as house prices are observed at a low frequency – infrequently sales – is another aspect of housing markets. In light of these features, researchers have employed two approaches to model individual house prices. Those are hedonic models and repeat sales approach. The former method based upon those hedonic characteristics of dwellings, it allows a comparison of house prices (Gupta & Miller, 2012). The latter approach based upon multiple sales of the same homes, it attempts to include all relevant information on the dwellings' quality, such as depreciation and renovations (Nagaraja et al., 2011). Both approaches are mainly used to construct the house price indices that reflect the housing market changes over time.

Regression-based models have also been used to explain the house prices dynamics with respect to a set of explanatory variables. In this regard, the US and UK housing markets have received much attention. Mankiw and Weil (1989) started by investigating the effects of demographic changes on the US housing market. The authors used multiple regression and observed that the primary driver of increasing house prices in the 1970s was the Baby Boom generation's entry into the house-buying years. Malpezzi (1996) modelled the US house prices relative to various features. The author concluded that incomes and population changes were the significant determinants of these house prices. Cho (1996) offered a survey outlining theoretical as well as empirical concerns around house price dynamics. On the theoretical side, Cho emphasised that real estate markets were not efficient. On the practical side, methods available at the time for estimating excess returns and house price indices were at the centre of the debate. Quigley (1999) examined the US housing markets on the metropolitan areas level using logarithmic and percentage changes specification models. Whilst Case and Shiller (2003) investigated on the state level the impact of fundamentals such as income growth on the observed increasing pattern in the house prices. In the UK housing market, Nellis and Longbottom (1981) employed an error-correction model (ECM) to study the influence that building societies had on house prices' fast growth. Drake (1993) and Munro and Tu (1996) studied the dynamics of the UK national and regional house prices using the Johansen cointegration technique. The ECM model was also used by Barot and Yang (2002) to investigate the investment supply and housing demand in the UK in comparison to Sweden. On an international level, authors such as Goodhart and Hofmann (2008) evaluated the interlinkages between 17 industrialised countries' house prices with factors such as credits, money, and economic activity. House price changes and fluctuations in the OECD countries were analysed by Englund and Ioannides (1997) and Hirata et al. (2013).

The literature above displays the dominance of the use of regression-based modelling techniques. Jadevicius and Huston (2015) pointed out the gap in modelling housing markets - the use of the ARIMA modelling procedure -. Although in many economics and finance areas, the ARIMA models have been a primary major of modelling and forecasting; their application to the housing markets is quite limited. Specifically, Jadevicius and Huston examined its performance in modelling the Lithuanian house prices. In the same viewpoint and in order to model the Finnish house price returns, the fourth article of the dissertation employs the ARMA modelling framework for each studied region with no substantial clustering effects. The paper takes a step further and compares the ARMA model's performances with its long-memory peer the ARFIMA model, to investigate between short and long memory features, which is vital for modelling the Finnish house price returns.

Regarding forecasting house prices, a particular focus on the US housing market is also noted. DiPasquale and Wheaton (1994) studied the 1980s US house prices dynamics and their future trends using various macroeconomic features. The authors concluded that the employed variables enhanced the house price forecasts accuracy. Case and Shiller (1990) examined the house prices and excess returns forecastability of San Francisco, Chicago, Atlanta, and Dallas cities. The authors utilised multiple regression and found that forecasting variables such as population growth, income, and construction costs were positively associated with excess returns or house price changes. Zhou (1997) developed a vector autoregressive (VAR) model with error correction and tested its accuracy in forecasting 1991-1994 US single-family home sales and prices. Crawford and Fratantoni (2003) considered three univariate time series models, namely GARCH, ARIMA, and regime-switching, and compared their abilities to forecast house prices of Florida, Texas California, Ohio, and Massachusetts states. Their findings revealed the regime-switching model's better in-sample performance and the ARIMA model's superior out-of-sample forecasting performance.

Rapach and Strauss (2009) explored the forecastability differences of the house prices of 20 major states ranked by their population number. They compared autoregressive models with the models incorporating information from various economic features and found that the former delivered fairly accurate interior states house price forecasts. Rapach and Strauss's work was extended by Bork and Møller (2015) in their 50 states house prices forecasting exercise. The authors included the 2007-2008 US housing collapse period that Rapach and Strauss did not consider. Moreover, they employed the Dynamic Model Averaging (DMA) and Dynamic Model Selection (DMS); the approaches that allow the forecasting models and parameters to change over time. Gupta et al. (2011) compared different classical time series and Bayesian models in forecasting the US house price index. In their subsequent article, Gupta and Miller (2012b) investigated the time-series association between house prices in Phoenix, Las Vegas, and Los Angeles. The authors further evaluated, in each market, the out-of-sample forecasts employing vector autoregressive (VAR) and Vector error-correction (VEC) models, together with Bayesian-based models. Barari et al. (2014) first identified potential structural breaks in the 1995-2010 US house prices. Then, the authors conducted a forecasting procedure by comparing linear and non-linear time series models with their structural breaks peers. More recent, Bork and Møller (2018) used the partial least squares (PLS), principal component analysis (PCA), and sparse PLS (SPLS) in the assessment of the forecastability of the US house prices.

In a comparative analysis in the UK housing market, Brown et al. (1997) observed that the time-varying parameter models outperformed several regression-based models, in forecasting 1968 to 1992 UK house prices. The considered constant parameters models were an autoregressive (AR), an error correction, and a VAR model. Studies regarding forecasting house prices of other countries include, for instance, Hadavandi et al. (2011) who employed a fixed-effects model to analyse the house price changes in 20 regions of Iran. Wei and Cao (2017) compared standard time series models and a DMA method to predict house price growth in 30 major cities in China. Hepşen and Vatansever (2011) used ARIMA approach to predict Dubai housing market future trends. Boitan (2016) also used the ARIMA model to explore future courses of the chosen European Union countries' residential property prices. On an international level, Kishor and Marfatia (2018) forecasted future house price trends of 16 OECD countries based on domestic and global macroeconomic measures. An excellent review of house price forecasting emphasising the shortcomings and issues of the standard forecasting models is given by Ghysels et al. (2013). In the Finnish housing market, the fourth article of the dissertation, in addition to modelling house price returns employing the ARMA and ARFIMA models. The paper also assesses the two models capabilities in forecasting the house price returns of the studied regions with no substantial clustering effects.

2.7 Modelling and forecasting house prices volatility

Volatility is widely recognised as a practical risk measure, and consequently, its measuring, modelling, and forecasting have gained enormous importance. In housing literature, the latest research highlighted the significance of examining and comprehending house price volatility for policy decision making, housing risk dynamics understanding, housing investment, and suitable portfolio allocation and management (C. L. Lee & Reed, 2014). In various countries, the housing market volatility research focuses on investigating whether the studied market holds the volatility's stylised facts, namely volatility clustering and leverage effects. The literature on house price volatility modelling and/or forecasting covers remarkable studies on the US and UK housing markets, along with some countries such as Canada and Australia.

Dolde and Tirtiroglu (1997) were among the first to study the US housing market volatility. The authors noted that house prices of San Francisco and Connecticut areas exhibited time-varying volatility. In their following article, Dolde and Tirtiroglu (2002) investigated volatility changes in four US regions: Midwest, West, South, and Northeast. They determined 36 events (12 and 24 volatility increases and decreases, respectively), and reported that these shifts resulted from regional events and economic conditions. Miller and Peng (2006) examined the evidence of time variation of single-family home price volatility. Out of 277 studied MSAs, 34 (17 per cent) were found to exhibit conditional heteroscedasticity. Moreover, the authors analysed the interaction between economic features and house price volatility. Employing GARCH and VAR models; home value appreciation and per capita gross metropolitan product (GMP) factors were found to impact house price volatility considerably, while this volatility substantially Granger-caused the rate of increase of per capita personal income. Milles (2008a) also documented clustering effects, but this time, on the state level. Out of 50 states, ARCH effects were present in over half of them, 28 states, to be precise. Miao et al. (2011) explored the spatial dependencies within 16 metropolitan areas. Based on return transmission, the authors placed San Francisco, New York, and Miami among the most influential markets. Moreover, their results indicated substantial volatility persistence and dependence across each area. Spatial linkages in house price returns and volatility were also investigated by Zhu et al. (2013) in 19 regional markets. Both studies – Miao et al. and Zhu et al. – used the Case & Shiller home price indices.

Karoglou et al. (2013) identified structural instability in several major US cities in their 1987 to 2009 data range, notably for the early 1990s and throughout the post-2007 financial crisis. Large house price changes (jumps) were also reported by Webb et al. (2016) in various US cities both during non-financial and financial crisis periods. These jump intensities were associated with the national, state, and city-level fundamentals. More recent, Apergis and Payne (2020) extended this vast literature by modelling the house price volatility of five major US metropolitan condominium markets. Those are Boston, New York, San Francisco, Chicago, and Los Angeles. On the US housing market's volatility forecasting aspect, Guirguis et al. (2005) used six estimation approaches. The authors argued that among these techniques, their house price forecasting empirical results favoured models that allow time-varying parameters, in this case, the rolling GARCH and the Kalman Filter with an Autoregressive Presentation. Milles (2008b) compared the Generalised AR (GAR), ARMA and GARCH models to forecast five states' house prices. The author's results indicated that GAR model outclassed the other two in an out-of-sample forecasting procedure. Li (2012) conducted an in-sample and out-of-sample assessment of three models' performances on the US real estate data pre- and post-financial crisis. Those models were the RiskMetrics, GARCH, and Asymmetric Power ARCH (APARCH). The RiskMetrics model was found to perform sufficiently for the in-sample estimation, whereas all models reached unsatisfactory out-of-sample forecasts during the post-crisis period. More recent, Segnon et al. (2020) predicted ten US major cities' house price volatility using their developed Markov-Switching Multifractal (MSM) model and compared its capabilities to the GARCH-based models. Their outcomes implied an improved prediction accuracy of the Fractional Integrated GARCH (FIGARCH) and MSM approaches.

Another country with an immense house price volatility literature is the UK. Willcocks (2009) examined which univariate data generating process would explain the UK house price index dynamics. The author identified that the exponential GARCH in mean (EGARCH-M) model sufficiently captured the underlying index. In his following work, Willcocks (2010) investigated whether clustering effects were present in 13 UK regions house prices. Out of those 13 regions, seven exhibited ARCH effects, with the EGARCH model results providing little evidence of volatility asymmetry in six of them. Morley and Thomas (2011) also employed the EGARCH-M model to analyse asymmetric effects in 10 UK regions. A significant positive asymmetric volatility was found in four regions. The asymmetric volatility effects were also assessed by Tsai and Chen (2009) on the nation-wide level. Using the Glosten, Jagannthan, and Runkle GARCH (GJR-GARCH) model, their outcomes revealed that asymmetric volatility effects were significant in both new house and all house price data. In their subsequent article, Tsai et al. (2010) employed the switching ARCH (SWARCH) models to evaluate the switching states in the UK older and new housing market. Their estimation suggested relatively stable volatility states.

Milles (2010) explored whether the house price volatility of one region transmitted to another. The author concluded that indeed a high condition covariance is observed between adjoining areas and that as the distance between regions increases, this covariance declines. Using data on the regional and national level, Tsai (2014) also studied these spillover effects and noted that housing prices between the UK's northern and southern regions would increase whenever a new financial crisis occurs. Milles (2008a) examined conditional heteroscedasticity effects in 12 UK regions and found that over half of them displayed these volatility clustering effects. By extending his work, Milles (2015) conducted structural breaks tests in 13 UK regions and national home prices, on a quest to detect the presence of bubbles. The author's findings suggested that more areas experienced structural changes between the late 1980s and early 1990s. More recent, Begiazi and Katsiampa (2019) took a step further by examining structural breaks in the UK regional house prices by property types (semi-detached, detached, flats, and terraced). The structural changes were found in 7 out of 13 considered regions and in 3 out of 4 types of property.

Other countries with a growing literature on house price volatility include Australia, where C. L. Lee (2009) examined the 1987- 2007 house price volatility of eight capital cities. C. L. Lee and Reed (2014) discussed the house price volatility pattern again for these eight Australian capital cities by decomposing it into the permanent and transitory components. Canada follows with the works of the authors such as Hossain and Latif (2009) and Lin and Fuerst (2014). Hossain and Latif identified house price volatility determinants and further analysed the interaction between house price volatility and various macroeconomic variables. Lin and Fuerst tested volatility clustering and studied the asymmetric effects in Canadian regional house prices. Other countries include Scotland (Katsiampa & Begiazi, 2019), Spain (Guirguis et al., 2007), Czech Republic (Sunega et al., 2014), Hong Kong (Zheng, 2015), Turkey (Coskun & Ertugrul, 2016), and China (Tian & Gallagher, 2015). Not only industrialised countries have been targeted when it comes to house prices volatility investigations, developing countries such as Malaysia (Reen & Razali, 2016) and small one such as Cyprus island (Savva & Michail, 2017) have also drawn the attention of some researchers. On an international level, scholars such as Engsted and Pedersen (2014) conducted a detailed house price volatility investigation of 18 OECD countries. In the Finnish housing market, the second article of the dissertation uses the EGARCH model to investigate the asymmetric volatility in the studied regions. Moreover, the fourth article compares the short memory GARCH models, in this case, the EGARCH, to their long memory counterparts, the CGARCH and FIGARCH models in a volatility forecasting exercise of the considered regions.

3 DATA AND METHODS

The research methods used in this dissertation are illustrated in the flowchart in Figure 1 below. Those are qualitative and descriptive research approaches. Data were collected and analysed to answer the specific analytic questions. Four scientific publications respond to the considered research questions to fulfil the study's goals.

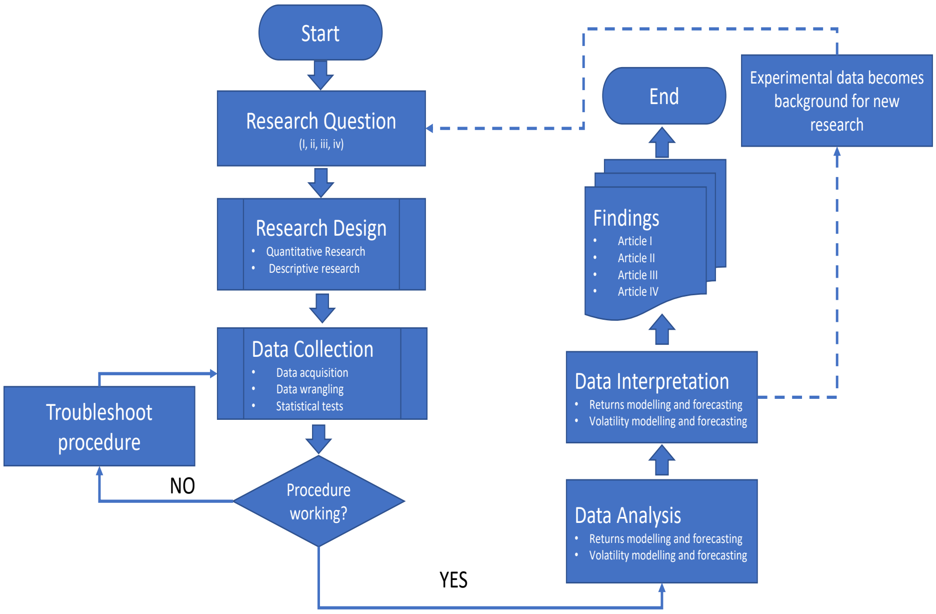


Figure 1. Research flowchart.

3.1 Data and its acquisition and wrangling

The data used in this dissertation are quarterly house price indices obtained from PxWeb databases of Statistics Finland. The studied period covers the first quarter of 1988 to the fourth quarter of 2018, with 124 observations. The type of dwellings considered is apartments, also referred to as blocks of flats, categorised by the number of rooms. That is one-room, two-rooms and larger (more than three rooms) apartments types. The studied regions are fifteen main regions in Finland grouped into three tiers in relation to their number of inhabitants. The first tier comprises four regions with a population exceeding 250,000; those are Helsinki, Turku, Tampere, and Oulu. The three first form a significant and growing area called the growth triangle in Southern Finland. Currently, the area accounts for, respectively, 49 and 55.5 per cent of the Finnish population and total Gross Domestic Product (GDP). The Oulu region, called the Northern Finland growth centre, is also among the well-performing region with substantial economic development and population growth.

The second tier contains seven regions with the number of residents above 100,000; those are Lahti, Joensuu, Kuopio, Jyväskylä, Seinäjoki, Vaasa, and Pori. The third tier consists of four regions with the number of inhabitants within 80,000 to 90,000; those are Hämeenlina, Kouvola, Lappeenranta, and Kotka. The regions in these two tiers also demonstrate significant expansion and economic growth performance. Moreover, with respect to their postcode or Zone Improvement Plan (ZIP) code numbers, these regions are distributed into forty-five cities and sub-areas. A thorough geographical division of each city and sub-area with corresponding postcode number is given in the dissertation's second article (Table 8 in Appendix).

From the standpoint of the thesis purpose of providing investors and risk managers the information related to diversify a housing portfolio investment geographically across Finland, the regional classification through postcode numbers offers a more localised delineation of housing submarkets. For instance, Helsinki-city, a citywide index data, differs from Helsinki-area1 a sub-area market. This way, one can examine various issues across submarkets. Other studies have also used ZIP code-level house prices in their housing submarket analysis. Authors such as Goetzmann and Spiegel (1997) found that postcodes as a spatial unit provide a well characteristic of the submarket and a spatial delineation correlated with essential variables impacting property values. Moreover, the author emphasised that the postcode numbers are the most straightforward submarket indicator to use as everyone knows their postcode numbers.

The data wrangling phase starts by transforming the house price indices of each city and sub-area in each apartment category into house price returns. Plus, precisely, logarithmic returns or continuous compound returns calculated as in Equation (2.1). Measuring house price returns using price changes is a matter that deserves some discussion. In their work, Bayer et al. (2010) considered house price returns as a combination of the price change and rental income. Although this approach has evident ground, studies that measure house price returns using only price changes, excluding the rental income dominate the literature. Among the works that opt for this approach include Dolde and Tirtiroglu (2002), Miller and Peng (2006), Milles (2008a; 2011a; 2011b), Morley and Thomas (2011; 2016), Lin and Fuerst (2014), C. L. Lee (2017), and Cook and Watson (2017). The aim of opting this method put forward by author such as C. L. Lee is to quantify the total return to holding housing as an asset. Moreover, the grounds proposed are that housing assets are not frequently traded as stock indices; hence house price indices are constructed differently. Following the above studies and in light of the prevalence, this procedure of measuring house price returns using price changes is adopted in this thesis. Statistics Finland uses the hedonic method to construct the house price indices used in this study.

3.2 Stationary tests

Next, unit root tests are conducted. That is, testing whether the house price returns R_t follow a random walk with drift,

$$R_t = \phi_0 + \phi_1 R_{t-1} + e_t,$$

where e_t denotes the error term, and assess the null hypothesis of unit root $H_0 : \phi = 1$ versus the alternative hypothesis $H_1 : \phi < 1$. Two standard procedures are employed, the Augmented Dickey-Fuller (ADF) (Dickey & Fuller, 1979) and Phillips-Perron (P-P) (Phillips & Perron, 1988) tests. The null hypothesis of unit root is rejected when the resulting test statistic is more negative than the critical value at a reasonable significance level; the studied series is therefore concluded to be stationary.

3.3 ARMA modelling

Once the stationarity is ensured, each return series in all three apartment categories is modelled by an ARMA *(p,q)* process of the following form:

$$R_t = \varphi_0 + \sum_{i=1}^{p} \varphi_i R_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i},$$

where a_t is a white noise series and p and q are non-negative integers. According to Akaike information criteria (AIC) and Bayesian information criteria (BIC), each model's appropriate lag is selected. Following the assumption of Miller and Peng (2006), Willcocks (2010), and Katsiampa and Begiazi (2019) that expectations in distinct areas and for each property types are heterogeneous; different lag orders are allowed across all studied cities and sub-areas in all three flats categories. Moreover, before testing for ARCH effects, the Ljung-Box test is performed to test the residuals' whiteness, that is, checking in the residuals whether there are any remaining autocorrelations.

3.4 ARCH tests

The assumption of the constant variance of the error term (homoscedasticity) is conventional in econometric models. However, various economic time series have demonstrated periods of substantial volatility followed by relative calm volatility for other periods. In such cases, the homoscedasticity assumption is inappropriate. The reason why it is crucial to test for the autoregressive conditional heteroscedasticity (ARCH) effects prior to model and/or forecast time series returns. The ARCH tests are performed on the squared residuals series from the estimated ARMA models. If the null hypothesis of homoscedasticity is rejected, a GARCH-type model or an SV model is required for the residuals. Two statistical tests candidates are available for ARCH effects testing and are presented below.

Portmanteau test for squared residual autocorrelation Q(h). It tests the null hypothesis that there are no remaining residual autocorrelations in the squared residuals up to lag h against the alternative hypothesis that at least one of the autocorrelations is nonzero. In other words, it is the standard Ljung-Box on squared residuals that checks whether the squared residuals are a sequence of white noise. If a small value of the p-value is obtained with respect to the specific level of significance, the null hypothesis is rejected favouring the alternative; indicating a presence of ARCH effects in the squared residuals.

ARCH–LM test - Lagrange Multiplier test of Engle (1982). It is based on fitting a linear regression of the form:

$$\hat{e}_t^2 = \beta_0 + \beta_1 \hat{e}_{t-1}^2 + ... + \beta_q \hat{e}_{t-q}^2 + u_t,$$

where $\hat{e}_t^2$'s are the squared residuals from the estimated ARMA models and u_t is the random error.
Then, testing the following hypotheses:

$$H_0 : \beta_1 = \beta_2 = ... = \beta_q = 0 \quad \text{(Homoscedasticity)}$$

versus

$$H_1 : \beta_q \neq 0, \quad \text{for at least one } q \quad \text{(Heteroscedasticity)}.$$

In other words, the test is the usual F-statistic test. Again, a small p-value for the specific critical level suggests a rejection of the null hypothesis, implying favourable alternative hypothesis.

3.5 Long-range dependence estimation

The thesis's following task is to investigate house price returns and volatility's long memory behaviour. The issue is crucial for developing suitable time-series forecasting models for the studied cities and sub-areas housing markets. Long-memory processes refer to processes with persistent autocorrelation. Two methodologies, namely Hurst exponent H of fractal analysis and the fractional-differencing parameter d of the ARFIMA process are beneficial in estimating long-range dependence and capturing these serial correlations. These methodologies and their key theoretical results are outlined below.

Hurst exponent H ***approach*** has its roots in the Hurst (1951) work. In his investigations on the river flows, Hurst introduced the rescaled range or the R/S statistic. The statistic is defined as follows:
Given k observations $X_1, X_2, ..., X_k$, define the n^{th} partial sum sequence $Y_n = Y_1 + Y_2 + ... + Y_n$ for $n = 0, 1, ...$ (with $Y_0 = 0$).

$$R/S(n) = \frac{\max_{1\leq k\leq n}\{Y_k - \frac{k}{n}Y_n\} - \min_{1\leq k\leq n}\{Y_k - \frac{k}{n}Y_n\}}{\sqrt{\frac{1}{n}\sum_{k=1}^{n}(X_k - \frac{1}{n}Y_n)^2}}. \tag{3.1}$$

The statistic's attraction is that for a specific period of, for instance, n years, $R/S(n)$ represents the observations' range. That is, it is a proxy for the ideal dam and reservoir height over that time. By examining various time series covering several geophysical phenomena, Hurst found that the empirical rate growth of the R/S statistic behaved as $R/S(n) \propto n^k$ for some k. The estimated k of 0.72 obtained using linear regression, was found to hold for some global significance. The phenomena which become known as the "Hurst phenomena". That is, the fact that the $R/S(n)$ increased faster than $n^{1/2}$. This Hurst finding's significant potential implication was that reservoirs and dams intended for long horizons could be insufficient, with floods as one possible effect.

In the early 1960s, there was a lot of debate around the Hurst' discoveries; however, no satisfactory explanation was found. Until the father of the long memory processes, Benoît Mandelbrot recognised the Hurst phenomena and laid the grounds of these processes. Mandelbrot proposed that instead of Gaussinality assumption, the Hurst phenomena could be explained by the heavy-tailed processes. Together with his colleagues [Mandelbrot and Van Ness (1968) and Mandelbrot and Wallis (1969)], they introduced the self-similarity concept, presented as follow: Let $X(t)$ be a continuous-time stochastic process. $X(t)$ is said to be self-similar, if for all positive a, $X(at) \stackrel{d}{=} a^H X(t)$ with H being a self-similarity parameter. The concept implies a process that possesses similar statistical features at various scales. Using this concept, they introduced two models which can be anti-persistent or persistent; the fractional Gaussian noise (FGN) and fractional Brownian motion (FBM). The FBM is a continuous-time stochastic process and generalises the standard Brownian motion (BM) with a supplementary parameter H, ranging between zero and one. The case $H = 0.5$ corresponds to the standard BM. H takes the values in the range [0, 0.5) for anti-persistent processes and (0.5,1] for persistent processes. As the FBM is nonstationary, its modified version is necessary for practical applications to use standard time series analysis tools. Hence, the introduction of the FGN; a discrete-time approximation of the FBM. Mandelbrot and Van Ness showed that the FGN is stationary, and primarily that it demonstrates the Hurst phenomena: for some $a > 0, R/S(n) \sim an^H$.

Moreover, Mandelbrot and Wallis developed the approach for estimating the FGN parameter H; the well-known rescaled range (R/S) analysis method. Following Equation (3.1), for the FGN, the following relationship holds:

$$\mathbb{E}[R/S(n)] \propto n^H \text{ as } n \to \infty.$$

Therefore, the H estimate is obtained as the slope of the averaged R/S as a function of n. They applied the R/S method to various data sets, including the Hurst's data types, and their outcomes yield the same evidence in favour of the long-range dependence hypothesis. Since then, the R/S has been recognised as a measure of long-range dependence. Building from it, other researchers have used various methods in the estimation of the FBM parameter H. Those include maximum likelihood estimation (MLE) approach (Sun et al., 2018). Furthermore, a combination of the BM and the independent FBM, namely the mixed fractional Brownian motion (MFBM) has recently been the core component in constructing financial models that capture financial assets' fluctuation such as stock prices. The key interest of the MFBM is that for $H \in (3/4, 1)$, it has been proved to be equivalent to the model driven by the BM process; hence it is arbitrage-free (Cheridito, 2001). The MLE approach has also been used in the estimation of the MFBM parameters. More recently, Dufitinema et al. (2020) used the hybrid method, MLE combined with a numerical approach, to simultaneously estimate all MFBM parameters, with the H parameter obtained numerically.

The ARFIMA approach. The ARFIMA processes which are the extension of the ARIMA processes to respond to the long memory behaviour of time series were introduced independently by the distinguished econometricians Granger and Joyeux (1980) and the hydrologist Hosking (1981). The ARFIMA process has become the standard replacement of the FGN in econometrics and Statistics literature. The process is defined as follows:

$$\Phi(L)(1-L)^d Y_t = \Theta(L)\epsilon_t, \quad t = 1, 2, ... \tag{3.2}$$

The back-shift or lag operator L is defined by $LY_t = Y_{t-1}$. The polynomials $\Phi(L)$ and $\Theta(L)$ describe the AR and MA terms respectively. ϵ_t is a white noise with $\mathbb{E}(\epsilon_t) = 0$, and variance σ_ϵ^2, and d is the fractional differencing parameter.

The Hosking's objection on the Mandelbrot's FGN was that the FGN was obtained by fractionally differentiating, followed by discretising. In other words, FGN was a discrete approximation to a continuous process. Thus, Hosking proposed to reverse this operations order, starting with discretisation followed by fractionally differencing. This method's interest is that the Brownian motion discretisation has an insightful explanation: it is the simple random walk or ARIMA(0,1,0) model. This model can then be fractionally differentiated using the fractional-differencing parameter d to get the ARFIMA (0,d,0) process which for $d \in (0, 0.5)$ manifests a long memory. A point that highlights how the ARFIMA processes are simply the extension of the classical Box-Jenkins framework. Another advantage of the ARFIMA (p, d, q) process is that it could model short-memory properties through its p, q parameters, and long-memory properties. If $d = 0$, the process reduces to the ARIMA, and the process is stated to demonstrate short memory. If $d \in (-0.5; 0)$, it is categorised as anti-persistence or long-range negative dependence.

The approaches to estimating the d parameter in Equation (3.2) can be categorised into two classes: semiparametric and parametric. In the former category, the long memory parameter d is first estimated separately, followed by the estimation of other parameters. The most widely used estimator in this group is the GPH estimator, named after its developers Geweke and Porter-Hudak (1983). The estimate of d is obtained by applying the ordinary least squares to the log periodogram. In the latter category, all three parameters, namely the autoregressive, moving average, and differencing, are simultaneously estimated. The most commonly used estimator in this category is the Whittle estimator proposed by Whittle (1953) and altered by Fox and Taqqu (1986). The MLE approach is used for the estimation. In practical estimations, the semiparametric approaches are preferred over their parametric peers. One reason is that parametric methods require to choose the adequate ARMA specification, which is the challenge of identifying and estimating explicitly p and q parameters. Whereas, in the semiparametric approach, the data generating process's full specification is not required to estimate the d parameter. The other reason is the heavy computations that parametric methods demand while the semiparametric are implemented smoothly (Reisen et al., 2001). Consequently, different studies investigating the long-range dependence through the ARFIMA approach have considered either the parametric or the semiparametric technique or a combination of the two methods. In this study, both approaches are used to analyse the house price returns' long-term memory behaviour, whilst the semiparametric technique is used to explore long-range dependence in house price volatility.

3.6 Time Series volatility forecasting models

An essential measure in finance is the risk of holding an asset, and the asset volatility is preferably the most frequently used measure of this risk. Thus, volatility, especially volatility forecasts, play a crucial role in various economic and financial situations, ranging from options and asset pricing, portfolio allocation to risk management. In analysing and forecasting the asset volatility, some of its characteristics drive these procedures. These volatility's properties often called stylised facts have been identified in most, if not all, financial returns. The first fact is the *volatility clustering*, a phenomenon which describes that large swings in financial returns tend to be followed by substantial fluctuations, and small fluctuations follow small changes. These events can lead to volatility clusters over time, and several studies have argued that these phenomena are likely caused by economic news. The second fact is *fat tails*; it stresses the fact that financial returns are not normally distributed. That is, they frequently have large negative or positive returns. The third fact is the *leverage effect*; it describes different impacts that positive shocks (a significant price increase) and negative shocks (a considerable price drop) have on asset volatility. That is the fact that volatility seems to respond differently to the two events with the latter having a more significant impact. Together with the volatility's mean reversion feature and its comovements across assets classes, these three properties have played a significant role and continue to do so in the development of volatility models. Several models have been suggested particularly to correct the existing ones' weaknesses for their incapability to capture these volatility characteristics.

In everyday language, volatility describes the swings or fluctuations noted in several phenomena over time. It is usually defined as the standard deviation of the asset returns; that is the square root of the variance of returns σ^2; a measure of returns dispersion around their mean. Its unbiased estimator is computed as a sample variance as follows:

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{t=1}^{N}(R_t - \bar{R})^2,$$

where $\bar{R}$ is the mean return.
However, this definition serves as the unconditional volatility over an entire period. One must distinguish it from conditional volatility, usually referred to as conditional standard deviation, denoted by σ_t. It is defined as the volatility in a specified period and is conditional on past information. This usual definition of volatility is one of the three types of volatility measures, which are:

- Volatility as the conditional standard deviation of asset returns – The focus of the volatility models discussed and employed in this thesis.
- Implied volatility: This volatility is deduced from prices from options markets using a pricing formula such as the Black and Scholes (1973) formula. It is used in the pricing of derivative securities. One of the implied volatilities' attractions is that they are based on the current market prices instead of using historical data. Thus, they are often referred to as the volatility's forward-looking estimators.
- Realised volatility: This volatility is computed as the sum of the squared returns. It is beneficial as volatility during a given period can be estimated precisely as the returns frequency increases. For instance, the realised volatility can be the sum of squared intraday returns for a particular day. Moreover, the realised volatility is also beneficial as it offers an adequately accurate measure of volatility useful for various purposes such as forecast evaluation and

volatility forecasting.

In modelling and forecasting asset volatility, one has to deal with the fact that asset volatility is not directly observable; it is estimated from the observed asset prices. That is, the volatility is inferred by looking at the asset price movement.
Let R_t be the log-return of an asset at time t. The conditional mean and variance of R_t given I_{t-1} are defined respectively as:

$$\mu_t = \mathbb{E}(R_t|I_{t-1}),$$
$$\sigma_t^2 = \text{Var}(R_t|I_{t-1}) = \mathbb{E}\left[(R_t - \mu_t)^2|I_{t-1}\right],$$

where I_{t-1} indicates the available information at time $t-1$ and consists of all past returns' linear functions.
The R_t is assumed to follow an ARMA(p,q) model so that $R_t = \mu_t + a_t$, where μ_t is given by

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i R_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j}, \tag{3.3}$$

and,

$$\sigma_t^2 = \text{Var}(R_t|I_{t-1}) = \text{Var}(a_t|I_{t-1}). \tag{3.4}$$

Equation (3.3) is referred to as the mean equation, while Equation (3.4) is the volatility equation, a_t denotes the innovation or shocks of an asset return. Equation 3.3) is the first step in building a volatility model; it removes the serial dependence in the asset returns. Next step is to test for ARCH effects in the mean equation's squared residuals. If these ARCH effects are found to be statistically significant, a volatility model constructed next is concerned with the time evolution and prediction of the σ_t^2. Below, time-series volatility modelling and forecasting models are reviewed, and a summary is given in Figure 2.

3.6.1 Historical volatility models

In this group of models, volatility σ_t is obtained from a sample of previous observations returns and is created from realised volatility calculated over a specific time window. The simplest historical price volatility model is the **random walk**, where tomorrow's optimum volatility forecast is today's volatility. Although this model accommodates the persistence to a near-term extent, it suffers not incorporating the mean-reverting feature of volatility. Next model is the **historical average method**. As the name suggests, the forecasts constructed using this approach are based on the whole volatility history, and all observations are given equal weights. This aspect helps this method to capture the volatility's mean-reverting feature in the long run; however, it ignores the short-term persistence and the fluctuations nature of volatility. Next on the list is the **moving average model** and its improved version, the **exponential weighted moving average (EWMA) model**. The moving average method discards older observations. Hence, more weights are put on recent observations, thereby capturing more recent volatility fluctuations and making it very sensitive to the estimation window length choice; for instance, when asset returns demonstrate volatility clustering. The improved version of the moving average model, the EWMA model, places greater weights on the most recent returns. This aspect improves its ability in volatility forecasting by capturing the tendency of volatility clustering. Finally, it is the **simple regression method**; where within this method, the simplest one is the AR model that expresses volatility as a function of its lagged values plus an error term. By including the past volatility errors, one obtains the ARMA model for volatility. Poon and Granger (2003) stressed, in their excellent forecasting volatility review, that these historically based prediction models' out-of-sample successful applications lie in optimal lag length searching or weighting scheme. Hence, the authors argued that a more refined forecasting approach would require parameter estimates' constant updating when new information becomes available and absorbed into the assessment through the estimation period. Thus, the introduction of the next group of models.

3.6.2 ARCH class conditional volatility models

In this group of models, the conditional variance σ_t^2 is a function of full available information until time $t-1$; hence it is called time-varying conditional variance. The first of this kind, the **ARCH model** was introduced by Engle (1982) in his groundbreaking article, and it has ever since been used extensively in the field of financial economics. ARCH model is tailored to capture the feature such as volatility clustering that time-series returns exhibit. It approximates volatility dynamics in an autoregressive manner, hence its name. The ARCH model of order q is defined as a linear representation of σ_t^2 and lagged values of the error term:

$$\begin{aligned} &R_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,\, \sigma_t^2) \\ &\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \ldots + \alpha_q \epsilon_{t-q}^2, \end{aligned} \tag{3.5}$$

where R_t is the log-return of an asset, μ_t is the conditional mean, $\omega > 0$ is the intercept, and ϵ_t is the error term.

Although the ARCH model could produce volatility clusters, it suffers a shortcoming of requiring a high lag order to capture the historical returns' full impact on the current volatility. To respond to this weakness, the generalisations of the ARCH model, the **GARCH model** was introduced by Bollerslev (1986). In tracking the time evolution of σ_t^2, the GARCH model incorporates the error term changes and volatility persistence. That is, the GARCH (p, q) expands the ARCH (q) model with a direct conditional variance p lag:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2. \tag{3.6}$$

To investigate the impact that asset volatility may have on asset returns, Engle et al. (1987) introduced the **GARCH-M model**; "M" stands for *in mean*. A parameter δ is integrated into the conditional mean model to quantify the form of the relationship between asset volatility and return as follows:

$$R_t = \mu_t + \delta\sigma_t^2 + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,\, \sigma_t^2)$$
$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2.$$

A significant positive δ denotes that the asset return is positively associated with its past volatility. Otherwise stated that high-risk results into a greater mean return. This study [Article 2] employs the GARCH-M model to examine the liaison between house price returns and volatilities in the Finnish housing market.

The GARCH model's limitation is its incapability to respond to the distinct effects of positive or negative shocks on the conditional variance. This phenomenon termed to asymmetric volatility has two explanations: the volatility feedback and leverage effect. The former refers to the positive correlations that are observed between asset returns and volatility. The latter refers to an observation that the asset volatility increases more for a negative return than it does for a positive return. To account for these drawbacks, Nelson (1991) introduced the **Exponential GARCH (EGARCH) model**. The model defines the conditional variance in a logarithmic form as follows:

$$R_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,\, \sigma_t^2)$$
$$\log(\sigma_t^2) = \omega + \sum_{i=1}^{q} \left[\alpha_i \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| + \tau_i \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right] + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2.$$

This aspect allows volatility to be always positive, and hence no further restrictions are necessarily imposed in the estimation. The estimated $\tau \neq 0$ indicates asymmetric effect, with leverage effect exciting if $\tau < 0$. Other models that take into account the asymmetric impacts of volatility include the **Threshold GARCH (TGARCH) model** of Rabemananjara and Zakoian (1993) and Zakoian (1994). The **GJR-GARCH model** proposed by Glosten et al. (1993), and the **Asymmetric Power ARCH (APARCH) model** by Ding et al. (1993). The TGARCH and GJR-GARCH model are often used interchangeably depending on whether one deals with the conditional standard deviations in the former or conditional variance in the latter. Both EGARCH and GJR-GARCH models were used in this dissertation [Article 2] to investigate the Finnish housing markets' asymmetric volatility.

As already mentioned, volatility persistence is a characteristic that several time series models are intended to accommodate. A GARCH model captivates an exponential decay in the autocorrelation of conditional variances. However, the evidence of a slow decline in correlations of squared and absolute financial assets returns has also been noted. These correlations that are slow to decay imply that the asset volatility appears to have a long memory, affecting future volatility movements over a long horizon. This aspect gave rise to the **Integrated GARCH (IGARCH) model** of Engle and Bollerslev (1986). However, The IGARCH model was found to impose an infinite persistence and shocks persist forever; hence, Baillie et al. (1996) introduced the **Fractional IGARCH (FIGARCH) model**. The FIGARCH model adds the fractional differences in the standard GARCH and replaces the first difference operator of the IGARCH by the fractional differencing operator $0 < d < 1$ of Equation (3.2). This facet allows the FIGARCH model to explain and capture a high degree of autocorrelation found in asset volatility.

In this study, motivated by the persistence behaviour found in the studied cities/subareas' house price volatility, the FIGARCH model as well as the Component GARCH (CGARCH) model are used to model and forecast their house price volatility. The CGARCH model developed by G. J. Lee and Engle (1999) also accommodates volatility's long memory component. On top of that, it also decomposes the conditional variance into permanent and transitory components, thereby investigating the long- and short-run volatility movements. In this group of ARCH class models, remark that the conditional variance is deterministic, meaning that it is a function of conditioning information. However, volatility may be subject to innovations sources that may or may not be associated with those that drive returns. Hence, the introduction of this third group of models.

3.6.3 Stochastic volatility models

The stochastic volatility (SV) framework was introduced by Taylor (1982; 1986). Under this approach, the time-varying volatility is presented as an unobserved component that follows a stochastic process. The stochastic nature of volatility may arise from various factors other than the lagged squared residuals and conditional variances. These factors can include investors' behaviours, economic and political information. Therefore, volatility is considered to comprise certain random elements. The SV model specifically incorporates an unobserved random shock into the volatility dynamics' characterisation and allows it to follow a latent stochastic process. The standard model of this group is the **SV(1) model** where the latent volatility σ_t^2 follows an AR(1) process; it is expressed as follows:

$$\begin{aligned} Y_t &= \sigma_t \epsilon_t, \quad t = 1, ..., n \\ \epsilon_t &\sim \mathcal{N}(0,\, 1) \\ \sigma_t^2 &= \exp(\lambda_0 + \lambda_1 \log \sigma_{t-1}^2 + \lambda_2 \eta_t), \end{aligned}$$

where Y_t represents the demeaned return process. The model has two innovations ϵ_t and η_t. The former is for the return itself while that latter is for the log-variance process or the latent stochastic process, and both are uncorrelated. λ_1 parameter characterises the persistence in the latent stochastic process, while λ_2 is its standard deviation or volatility component. As it can be noted, the SV(1) model's errors are normally distributed. However, for the asset returns analysis, the non-normal conditional residual distributions are appropriate due to the volatility characteristics such as fat tails. Thus, the **SV model with Student's *t* errors (SVt) model** has been proposed. Furthermore, the same as in the GARCH-type models, to accommodate the asymmetric effects feature of volatility, the **SV model with leverage effects (SVl) model** has also been suggested. The three models as well as the SV model whose latent volatility process mimics a stationary AR(2) process, the **SV–2 model**, were used in this thesis [Article 3] in house price volatility modelling and forecasting exercise, and their performances were compared.

Although the SV models have various advantages over GARCH models with their theoretical attractiveness and empirical evidence in their support over their peers (Nakajima & Omori, 2012); their practical application has drawn little attention. The constraints faced are mainly their highly non-linear estimations, leading to the absence of these approaches' software packages implementations. Researchers such as Chan and Grant (2016a; 2016b) and Hosszejni and Kastner (2020) have worked to overcome these challenges and provided the means of the SV models' estimation and forecasting procedures. This dissertation uses both the SV models [Article 3] and GARCH-based models [Article 4] to model and forecasts the studied house price volatility for city/sub-area with considerable ARCH effects. The best performing model for city/sub-area is selected for in-sample fit and out-of-sample forecasting based on the used metrics. Moreover, the out-of-sample results are employed to provide an outlook of the Finnish housing market in terms of three-year forecasts for house price returns and volatilities. That is, once for each city and sub-area, the out-of-sample best performing model has been established, the model is re-estimated on the whole sample data 1988:$Q1$ to 2018:$Q4$. Next, the three-year house price returns and volatility forecasts are built, 2019:$Q1$ to 2021:$Q4$.

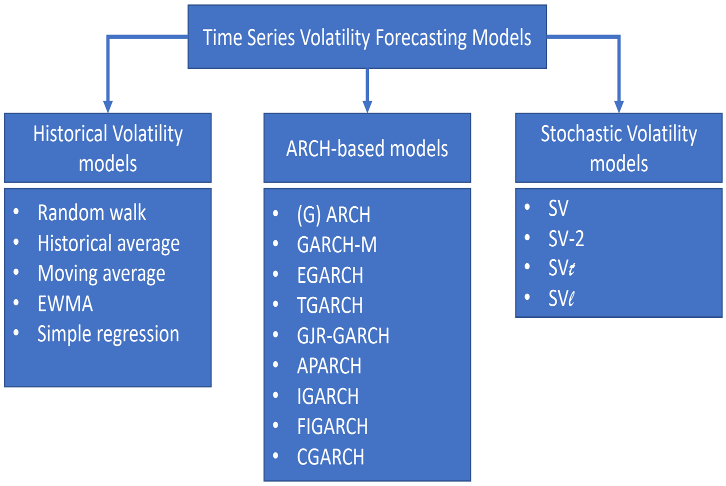


Figure 2. Volatility forecasting methods.

3.7 Model estimation

As mentioned earlier, the models in group one in their estimation make use of the sample standard deviation. However, the models in group two construct the asset returns' conditional variance via the maximum likelihood estimation procedure. That is, given the Equation (3.5) simplify written as of order one:

$$R_t = \mu_t + \epsilon_t, \quad \text{where} \quad \epsilon_t | I_{t-1} \sim \mathcal{N}(0,\, \sigma_t^2) \quad \text{and} \quad \sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2.$$

Thus, $R_t | I_{t-1} \sim \mathcal{N}(\mu,\, \sigma_t^2), \quad \text{with} \quad \sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2$.
Assuming that $\epsilon_0 = 0$, the likelihood function is given by

$$\mathcal{L}(\mu, \omega, \alpha | \epsilon_0) = \prod_{t=1}^{T} \mathcal{L}(R_t | I_{t-1}, \epsilon_0),$$

and the log-likelihood is:

$$l(\mu, \omega, \alpha | \epsilon_0) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \ln \sigma_t^2 - \sum_{t=1}^{T} \frac{(R_t - \mu)^2}{2\sigma_t^2}.$$

The first order conditions are computed with respect to the three parameters μ, ω, α.

The individual GARCH models estimations are carried out using the *rugarch* package (Ghalanos, 2020) in R (R Core Team, 2020). It follows the following three steps:

1. Model specification performed by the *ugarchspec()* function,
2. Model estimation using the *ugarchit()* function,
3. Model adequacy checking through tests of standardised residuals.

Again, it is worth mentioning that although in practice, the error term ϵ_t is often assumed to follow the normal distribution, various researchers pointed out that this approach is not plausible. One reason being, the fat tails features that financial returns exhibit; in fact, it is one of their stylised facts discussed above. Therefore, appropriate distributions suited to house this characteristic and others, such as skewness, should be used. Among these distributions include the Student t ("Std"), Generalized Error ("GED"), and their skew variants ("sStd", "sGED"). During the estimation, the appropriate distribution for each city and sub-area in every flat category is chosen based on AIC and BIC metrics.

Regarding the models in group three, the SV models, they are very challenging to estimate as discussed above. The estimation challenge lies in the fact that the likelihood function does not have a closed-form; hence the MLE approach can not be employed. Several methods have been put forward in this quest. Among them, including the quasi-maximum likelihood estimation procedure (Harvey et al., 1994), the general method of moments approach (Duffie & Singleton, 1993), and the Bayesian estimation technique through Markov Chain Monte Carlo (MCMC) methods (Kim et al., 1998). The last approach, the Bayesian estimation, opted in this study, uses the Bayes' theorem, also called Bayes' rule. The theorem works so that given the prior knowledge about the parameters' values, these beliefs are updated as new data are observed. The prior refers to the parameters' probability distribution before the data are observed, and it is called prior distribution. Next, combining these parameters' prior knowledge with the data's information, the conditional distribution of parameters is computed. This aspect of determining the parameters' conditional distribution given the data and the parameters' prior beliefs is the crucial step in Bayesian inference. This distribution conditional on observed data is called posterior distribution. In several practical problems, the posterior analytic computation is not possible; thus, numerical approaches are used instead. Among those Bayesian numerical techniques, the highly effective is the MCMC method; it uses the Markov Chain simulation to obtain the posterior distribution. That is creating a Markov process whose stationary transition distribution is the posterior distribution. The most popular MCMC algorithms include Gibbs sampling and Metropolis-Hastings.

3.8 Model selection and model comparison

3.8.1 In-sample fit

A crucial practical problem is estimating each model's quality given a collection of statistical models, thereby comparing two or more models, and choosing the one that might be appropriately representing the studied data generating mechanism. Given a vector of parameters θ, the value that maximises the log-likelihood, $\log(\mathcal{L}(\hat{\theta}))$, can be utilised to compare the fits of two or more models or to assess to what extent a model fits the data. However, the use of $\log(\mathcal{L}(\hat{\theta}))$ can result in over-fitting problem. This problem refers to adding parameters to the models that can increase its log-likelihood, even though the additional parameter does not necessarily mean that the model is a better data's description. It increases model complexity that may be fitting random noise in the data. Therefore, a parsimonious model is found by dealing with the tradeoff between the model's goodness of fit (maximising fit) and its simplicity (minimising model complexity). Two metrics are used to achieve this task; those the Akaike's information criterion (AIC) and Bayesian information criterion (BIC). They are defined as follows:

$$\text{AIC} = -2\log(\mathcal{L}(\hat{\theta})) + 2k,$$
$$\text{BIC} = -2\log(\mathcal{L}(\hat{\theta})) + \log(n)k,$$

where n is the sample size and k is the number of the estimated parameters in the model. For both criteria, the favourable model is the one with the minimum value. The minus twice log-likelihood is termed to *deviance* and measures the model's goodness of fit. The complexity penalties terms, $2k$ and $\log(n)k$, discourage over-fitting and penalise larger models. As the BIC depends on the sample size and $\log(n) > 2$ given that $n > 8$, the BIC penalises more model complexity and pursues a simpler model than the AIC. In this study, both metrics are used to select and compare the in-sample fit performance of the univariate time series models, namely ARMA and ARFIMA models for house price returns and GARCH-type models for house price volatility. Both criteria commonly choose the same or almost the same model for each city and sub-area.

For the SV models whose estimations are carried out through the Bayesian framework, the models' in-sample fit performances are assessed and compared using the Deviance information criterion (DIC); a Bayesian analogue of the AIC. The DIC is also a tradeoff between the model's complexity and its goodness of fit. The fit is assessed by the deviance, defined as:

$$D(\theta) = -2\log\mathcal{L}(data \mid \theta).$$

However, the complexity is measured by a Bayesian analogue of k number of parameters. That is an estimate of the *effective number of parameters* p_D, defined as the change between the posterior mean deviance and the deviance evaluated at the posterior mean of parameters:

$$p_{\mathrm{D}} = \overline{D} - D(\bar{\theta}).$$

Hence, the DIC is then defined analogously to AIC as:

$$\begin{aligned} \mathrm{DIC} &= D(\bar{\theta}) + 2p_{\mathrm{D}} \\ &= \overline{D} + p_{\mathrm{D}}. \end{aligned}$$

That is the sum between the average deviance from complete sets of iterations or the effective number of parameters and the Monte Carlo estimated posterior mean deviance. Just like the AIC and BIC, the smaller the DIC, the favourable the model.

3.8.2 Out-of-sample forecasts

To assess and weigh up the forecasting capabilities of the considered models; the data is split into training and test sets. The former set is used to build the models, whereas the latter is used to evaluate the models' predictive accuracy. Once each model is estimated using the training data set, volatility forecasts are constructed as the one-step-ahead (quarter), and last, they are compared to the volatility proxy. As discussed above, the asset volatility is unobserved, which is the first problem that one encounters when evaluating volatility forecasts. Several studies have suggested suitable volatility proxies, including squared returns (Sadorsky, 2006). However, other researchers such as Patton (2011) have pointed out that squared returns were rather noisy proxies of the conditional variance, and turned their attention toward the realised volatility as a volatility evaluation tool. In this study, the volatility is also proxied by realised volatility built as a rolling sample. That is, the volatility is measured as a quarter to quarter rolling calculations.
Two commonly used metrics are employed to assess the considered models' forecasting accuracy. Those are the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE). The former metric has the benefit of penalising large errors. Errors with larger absolute values are given more weight than those with smaller absolute values. This aspect makes it beneficial, especially when large errors are undesirable. The latter metric gives equal weight to all errors. Both are negatively-oriented scores, meaning that lower values are better. The two measures are defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i^2 - \sigma_i^2)^2} \quad \text{and} \quad \mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{\sigma}_i^2 - \sigma_i^2\right|,$$

where N is the number of forecasts, $\hat{\sigma}^2$ is the forecast volatility, and σ^2 is the “true” volatility.

4 RESULTS AND DISCUSSIONS

This section first summarises the four publications and presents their link to the thesis's aim as a whole, and second discusses the results. In line with the dissertation's objectives, results are presented and discussed as follows: First, ARCH effects testing outcomes are outlined, followed by a discussion of the findings in the case of return series with constant variance. That is the outcomes of their modelling and forecasting procedures. Next, in the case of return series with time-varying variance, their in-sample fit analysis and out-of-sample forecasting results are addressed. Finally, an outlook of the Finnish housing market in terms of three-year forecasts for both house price returns and volatilities is provided.

4.1 Publications Overview

The four publications in this study respond to the research questions posed earlier in Section 1.3 of the work. The published articles identify various aspects of the Finnish housing market related to the modelling and forecasting of its price returns. Table 2 presents each paper's title, its objectives, findings and the link of the respective article to the thesis's aim as a whole. In summary, article 1 investigates the evidence of long-term dependent behaviour in house price returns and volatilities using both semiparametric and parametric approaches. This investigation is crucial in determining the asset's predictability and developing adequate time series volatility forecasting models for the studied market. Article 2 examines the three volatility properties generally investigated in asset evaluation. Those are volatility clustering, risk-return association, and asymmetric effects. Paper 3 assesses the performances of four stochastic volatility models, namely SV, SV–2, SVt and SVl, in modelling as well as forecasting the Finnish house price volatility. Paper 4 compares the GARCH-type models, short- and long-memory based models in house price returns and volatility modelling and forecasting exercise.

4.2 ARCH effects testing

In article 2, using both Portmanteau and Lagrange Multiplier tests, results in Table 3 reveal that evidence of time-varying variance is found in the house price returns of the absolute majority of the studied cities/sub-areas. Plus precisely, in the one-room flats category, 74 per cent of the regions (equivalent to 28 out 38 areas) exhibit conditional variance. In the two-rooms flats category, it is 64 per cent of the cities/sub-areas (corresponding to 27 out of 42), while in the larger flats category, it is 79 per cent of the regions (equivalent to 31 out 39).

Table 2. Publications overview.

Paper	Title	Objectives	Findings	Link to the thesis
I	Long-range dependence in the returns and volatility of the Finnish housing market	Investigate the evidence of long-term dependent behaviour in house price returns and volatility	• A high degree of persistence in the house price returns • Long-range dependence behaviour in the house price volatility	The study outcomes were used to model and forecast the price returns and volatility dynamics of the studied types of dwellings
II	Volatility clustering, risk-return relationship and asymmetric adjustment in the Finnish housing market	Examine the three volatility properties generally investigated in asset evaluation, namely, volatility clustering, risk-return association, and asymmetric effects	• Clustering effects in over half of the studied regions • Mixed results on the sign of the significant risk-return association • Asymmetric impact of shocks on housing volatility in almost all the studied regions	The results of the volatility asymmetric effects were employed to forecast the house price volatility dynamics
III	Stochastic volatility forecasting of the Finnish housing market	Assess the performances of four stochastic volatility models in modelling as well as forecasting the Finnish house price volatility	The heavy-tailed SV model provided superior out-of-sample volatility forecasts in the majority of the studied regions	The study outcomes were used to forecast the volatility movements of the house prices

Paper	Title	Objectives	Findings	Link to the thesis
IV	Forecasting the Finnish house price returns and volatility: A comparison of Time Series Models	Compare the GARCH-type models, short- and long-memory based models in house price returns and volatility modelling and forecasting exercise	• No clear cut between competing models for house price returns modelling • The short-memory model stands out as the leading model for house price volatility modelling • The long-memory models outclass their short-memory peers in forecasting house price returns as well as volatility	The study findings were compared to their SV counterparts in forecasting the house price volatility of the considered regions

The ARCH effects evidence found in this study is more significant compared to, for instance, the US and UK housing markets. Milles (2008a) documented clustering effects in 56 per cent of the US states (28 out of 50). In his subsequent work, Milles (2011b), the author noted that 58 per cent of the UK regions (7 out of 12) indicated a time-varying variance. Moreover, Willcocks (2010) also reported a 54 per cent proportion of the UK regions (6 out of 13) showing a conditional variance. Furthermore, highly dense regions like Helsinki and least populated areas such as Kotka were found to demonstrate ARCH effects. A remark that contrasts to Lin and Fuerst's (2014) argument that the source of house price volatility clustering may be the region's population density. Therefore, identify the sources and determinants of these region's house price volatility warrants further research.
Additionally, regarding which city/sub-area manifested ARCH effects and which did not, in all flats categories, the Oulu and Lahti sub-markets2, and Joensuu city-level market did not demonstrate volatility clustering effects. Whereas, in both one-room and two-rooms flat types, no evidence of time-varying variance was found in Tampere city-level and Vaasa sub-market1. One possible interpretation of these constant variance series could be that these cities and sub-areas' house prices might be less active and thereby less volatile.

Table 3. ARCH effects tests results.

Regions	Cities/sub-areas	One room flats	Two rooms flats	Three rooms flats
		ARCH?	ARCH?	ARCH?
Helsinki	hki	Yes	Yes	Yes
	hki1	Yes	Yes	Yes
	hki2	Yes	Yes	No
	hki3	No	Yes	Yes
	hki4	Yes	Yes	Yes
Tampere	tre	No	No	Yes
	tre1	Yes	Yes	Yes
	tre2	No	Yes	Yes
	tre3	Yes	No	Yes
Turku	tku	Yes	Yes	Yes
	tku1	Yes	No	Yes
	tku2	Yes	Yes	Yes
	tku3	Yes	No	Yes
Oulu	oulu	Yes	No	Yes
	oulu1	Yes	No	Yes
	oulu2	No	No	No
Lahti	lti	Yes	Yes	Yes
	lti1	Yes	No	Yes
	lti2	No	No	No
Jyväskylä	jkla	Yes	Yes	Yes
	jkla1	Yes	Yes	Yes
	jkla2	Yes	Yes	Yes
Pori	pori	Yes	Yes	No
	pori1	Yes	Yes	Yes
	pori2	–	Yes	–
Kuopio	kuo	Yes	Yes	Yes
	kuo1	Yes	Yes	Yes
	kuo2	Yes	No	Yes
Joensuu	jnsu	No	No	No
	jnsu1	Yes	Yes	No
Seinäjoki	seoki	–	Yes	Yes
Vaasa	vaasa	No	Yes	Yes
	vaasa1	No	No	Yes
	vaasa2	–	–	Yes
Kouvola	kou	Yes	Yes	No
Lappeenranta	lrta	Yes	Yes	Yes
	lrta1	Yes	Yes	–
	lrta2	–	No	Yes
Hämeenlinna	hnlina	Yes	Yes	No
	hnlina1	No	Yes	Yes
Kotka	kotka	Yes	No	Yes
	kotka1	No	Yes	–
	kotka2	–	No	–

Notes: "Yes" denotes that a city/sub-area demonstrates ARCH effects, "No" indicates that a city or sub–area does not. The "–" sign denotes that no observations/data were available for that city or sub-area in that specific flat category.

4.3 Modelling and forecasting house price returns

In article 4, for cities and sub-areas with constant variance, the ARMA models' performance, in both in-sample fit and out-of-sample forecasting, is compared to their long-memory peers, the ARFIMA models. The motive of the long-memory models' choice is a high level of persistence found in house price returns during their examination of the long-range dependence behaviour in article 1. The long memory behaviour assessment results in these cities/sub-areas are as follows: In the one-room flats category, among ten regions with constant variance, eight demonstrated long-range dependence behaviour, whereas two were anti-persistent. The first group results imply that the estimated fractional differencing parameter d of this group's regions ranged from 0 to 0.5, while the second group mean that $d \in (-0.5, 0)$. In the two-room flats category, fourteen displayed long-range dependent behaviour among fifteen regions and one was anti-persistent. In the larger (over three rooms) flats category, seven among eight regions exhibited long memory behaviour, and one was anti-persistent. These long memory results; the d estimated parameters were incorporated in the ARFIMA models' estimation procedures. Table 4 reports for each city and sub-area, the best performing models both in-sample and out-of-sample.

Regarding the house price returns' modelling aspect, the two candidates models yield mixed results, with their performances alter across apartment categories. In the one-room flats group, the ARMA model takes the forefront in six out of eight cities/sub-areas. In the two-room flats group, the ARFIMA model leads in eleven out of fourteen cities/sub-areas. In the larger flats group, the two competing models share rankings with the short- and long-memory model standing out in, respectively, three and four cities/sub-areas out of seven. These results indicate that in most regions with constant variance, the current changes in house prices are explained by their past values behaviour (AR) and their past random shocks (MA), especially in the one-room apartment category and the half of the larger apartment category. Moreover, these findings reflect that the ARMA modelling framework, a leading major in modelling in various areas, also strongly contributes to examining housing markets; an observation also noted by Jadevicius and Huston (2015). Furthermore, the outcomes reveal the ARFIMA model's capability in capturing the high persistence feature found in house price returns, particularly in the two-room flats group.

Regarding the house price returns' forecasting aspect, the ARFIMA models, with their lowest error rates, outclass the ARMA models in all three flats categories. The long memory models provide superior out-of-sample returns forecasts in five out of eight cities/sub-areas in the one-room, in ten out fourteen in the two-rooms, and five out seven in the larger apartments group. These results confirm, once again, the ARFIMA model's capability of capturing these long-range dependences and the crucial role that this long memory feature plays in the returns forecasting procedure.

Table 4. House price returns - Best performing models.

Regions	Cities/Sub-areas	One room flats		
		In-sample		***Out-of-sample***
Helsinki	hki3	ARMA		ARMA
Tampere	tre	ARFIMA		ARFIMA
	tre2	ARMA		ARMA
Oulu	oulu2		Anti-persistent	
Lahti	lti2	ARFIMA		ARMA
Joensuu	jnsu	ARMA		ARFIMA
Vaasa	vaasa	ARMA		ARFIMA
	vaasa1		Anti-persistent	
Hämeenlinna	hnlina1	ARMA		ARFIMA
Kotka	kotka1	ARMA		ARFIMA
		Two rooms flats		
		In-sample		***Out-of-sample***
Tampere	tre	ARFIMA		ARFIMA
	tre3	ARFIMA		ARFIMA
Turku	tku1	ARFIMA		ARFIMA
	tku3	ARFIMA		ARFIMA
Oulu	oulu	ARFIMA		ARMA
	oulu1	ARMA		ARMA
	oulu2	ARMA		ARMA
Lahti	lti1	ARFIMA		ARMA
	lti2	ARFIMA		ARFIMA
Kuopio	kuo2	ARFIMA		ARFIMA
Joensuu	jnsu	ARFIMA		ARFIMA
Vaasa	vaasa1	ARMA		ARFIMA
Lappeenranta	ltra2	ARFIMA		ARFIMA
Kotka	kotka	ARFIMA		ARFIMA
	kotka2		Anti-persistent	
		Three rooms flats		
		In-sample		***Out-of-sample***
Helsinki	hki2	ARMA		ARFIMA
Oulu	oulu2	ARFIMA		ARFIMA
Lahti	lti2	ARMA		ARFIMA
Pori	pori	ARFIMA		ARMA
Joensuu	jnsu	ARFIMA		ARMA
	jnsu1		Anti-persistent	
Kouvola	kou	ARMA		ARFIMA
Hämeenlinna	hnlina	ARFIMA		ARFIMA

Notes: This table reports the house price returns best performing in-sample and out-of-sample models, for each city and sub-area, in each flat category. The “Anti-persistent” refers to the series with long-range negative dependence, meaning that their estimated fractional differencing parameter d ranged from -0.5 to 0.

Moreover, although the general view that the best performing in-sample fit model does not necessary produces accurate out-of-sample forecasts is observed in some regions, in most cities/sub-areas, especially in the two-rooms flats category, the contradictory is noted. In eleven out of fourteen cities/sub-areas, the same model provides the best in- and out-of-sample results. This observation is in line with the Jadevicius and Huston's (2015) study outcomes in the case of the Lithuanian housing market.

4.4 Modelling house price volatility

In article 3 and 4, the cities and sub-areas where their ARMA formats' residuals demonstrated evidence of the ARCH effects, the GARCH-type and the SV models are used to model their house price returns. In the former group, the ARMA mean process is expanded to incorporate a GARCH process to model the time-varying variance, and its evolution is modelled deterministically. In the latter group, the conditional variance is modelled probabilistically, meaning that the time-varying variance is treated as unobserved components that mimic a stochastic process. The considered GARCH-type models are the EGARCH model, which was found to outperform the GJR-GARCH model in the asymmetric volatility investigation in article 2. This short-memory GARCH model's performance is compared to the models that encompass long memory feature in the asset volatility. Those are the FIGARCH and CGARCH model. Similar to the house price returns, the ground of these long memory models' choice is a high level of persistence found in house price volatilities for most Finnish cities and sub-areas in article 1. The estimated long memory parameters d were incorporated in the FIGARCH estimations. The considered SV models include the vanilla SV, the SV whose latent volatility process mimics a stationary AR(2) process (SV–2), the heavy-tailed model (SVt), and the SV model with leverage effects (SVl) model. Table 5 reports for each city and sub-area, house price volatility modelling best performing models, GARCH-type model and SV model side-by-side.

The highlight of the results is the importance of the asymmetric volatility features (volatility feedback effects and leverage effects) in modelling asset returns. This observation is reflected in how the two groups' models that incorporate leverage effects rank as the best models for the Finnish house price volatility modelling in all three apartment categories. For the EGARCH model, in the one-room flats group, it comes on top in 17 out of 28 cities/sub-areas. In 19 out of 27 in the two-rooms flats group, 23 out of 31 in the larger flats group. For the SVl model, in the one-room flats group, it leads in 19 out of 28. In 24 out of 27 in the two-rooms flats group, 20 out of 31 in the larger flats group. This crucial role of asymmetric volatility has been well established in, for instance, the stock returns modelling. Thus, it is also mirrored in the case of house prices, whether one employs the deterministic or the probabilistic framework. Moreover, the long memory feature is also crucial, with the FIGARCH model taking the best performing models' second place in the GARCH-type group. Whilst the decomposition of volatility into permanent and transitory components appears to be insubstantial, with the CGARCH model standing out in only three regions. In the SV models group, the second place is taken by the SV–2 model, with the AR(2) component proving to be a valuable and worthwhile addition to the vanilla SV. The heavy-tailed model (SVt) and the vanilla SV share ranking in the rest of the regions.

Table 5. House price volatility modelling - Best performing models.

		One room flats		**Two rooms flats**		**Three rooms flats**	
Regions	Cities/Sub–areas	***GARCH-type***	***SV models***	***GARCH-type***	***SV models***	***GARCH-type***	***SV models***
Helsinki	hki	FIGARCH	SVl	FIGARCH	SVl	EGARCH	SV
	hki1	FIGARCH	SVl	EGARCH	SV–2	EGARCH	SVl
	hki2	FIGARCH	SVl	EGARCH	SVl	–	–
	hki3	–	–	FIGARCH	SVl	EGARCH	SVl
	hki4	EGARCH	SVl	EGARCH	SVl	EGARCH	SVl
Tampere	tre	–	–	–	–	EGARCH	SVl
	tre1	EGARCH	SVl	EGARCH	SVl	FIGARCH	SVl
	tre2	–	–	EGARCH	SVl	EGARCH	SV–2
	tre3	EGARCH	SV–2	–	–	FIGARCH	SVl
Turku	tku	EGARCH	SV–2	CGARCH	SVl	EGARCH	SVl
	tku1	EGARCH	SVl	–	–	EGARCH	SVl
	tku2	EGARCH	SV–2	EGARCH	SVl	EGARCH	SVl
	tku3	FIGARCH	SVl	–	–	EGARCH	SVl
Oulu	oulu	EGARCH	SVl	–	–	EGARCH	SVl
	oulu1	EGARCH	SVt	–	–	EGARCH	SVl
Lahti	lti	EGARCH	SVl	EGARCH	SVl	EGARCH	SVl
	lti1	EGARCH	SV–2	–	–	EGARCH	SVl
Jyväskylä	jkla	EGARCH	SVl	EGARCH	SVl	CGARCH	SV–2
	jkla1	FIGARCH	SV–2	EGARCH	SVl	FIGARCH	SVl
	jkla2	FIGARCH	SV–2	EGARCH	SVl	FIGARCH	SVl
Pori	pori	FIGARCH	SVl	EGARCH	SVl	–	–
	pori1	EGARCH	SVl	EGARCH	SVl	FIGARCH	SVl
	pori2	–	–	EGARCH	SVl	–	–
Kuopio	kuo	EGARCH	SVl	FIGARCH	SVl	EGARCH	SV–2
	kuo1	FIGARCH	SV–2	EGARCH	SVl	FIGARCH	SV–2
	kuo2	EGARCH	SV–2	–	–	EGARCH	SV–2
Joensuu	jnsu1	EGARCH	SVl	FIGARCH	SVl	–	–
Seinäjoki	seoki	–	–	FIGARCH	SV–2	FIGARCH	SV–2
Vaasa	vaasa	–	–	CGARCH	SVl	EGARCH	SVl
	vaasa1	–	–	–	–	EGARCH	SVt
	vaasa2	–	–	–	–	EGARCH	SV–2
Kouvola	kou	EGARCH	SVl	EGARCH	SVl	–	–
Lappeenranta	lrta	FIGARCH	SVl	EGARCH	SVl	EGARCH	SVl
	lrta1	FIGARCH	SVl	FIGARCH	SVl	–	–
	lrta2	–	–	–	–	EGARCH	SV–2
Hämeenlinna	hnlina	EGARCH	SVl	EGARCH	SV–2	–	–
	hnlina1	–	–	EGARCH	SVl	EGARCH	SV–2
Kotka	kotka	FIGARCH	SVl	–	–	EGARCH	SVl
	kotka1	–	–	EGARCH	SVl	–	–

Notes: This table reports the house price volatility modelling best performing models for each city and sub-area, in each flat category. "–" indicates that city or sub-area is with constant variance in that specific flat category.

4.5 Forecasting house price volatility

The core of the dissertation lies in the forecasting procedures of house price returns in both cases of constant and time-varying variances. For cities and sub-areas with significant ARCH effects, the forecasting performances of the models from two groups of time-series volatility forecastings models were compared separately. Those are the SV models: SV, SV–2, SVt, and SVl [Article 3], and GARCH-type models: EGARCH, FIGARCH, and CGARCH [Article 4]. For each city and sub-area, the best performing SV model is compared to its GARCH-type peer; Table 6 reports the model with the lowest error rate (RMSE or MAE). That is, the model that provides superior out-of-sample house price volatility forecasts for each city and sub-area.

A striking observation in the results is the outperformance of the GARCH-type models against their SV counterparts. The deterministic approach based models excel in more than 60 per cent of the studied regions. Plus precisely, in 17 out of 28 cities/sub-areas in the one-room flats category. They come on top in 17 out of 27 and 22 out of 31, in the two-rooms and larger flats category, respectively. Moreover, among the GARCH-type models, the ones accommodating the long memory feature outclass their short memory peers. In the larger apartments group, the CGARCH model leads in 9 cities/sub-areas; in the two-rooms apartments group, it is the FIGARCH model that comes on top in 7 areas. In the one-room apartments group, the two models are neck and neck, leading in 8 and 7 regions, respectively. These best performances outcomes of the long memory GARCH models, especially the CGARCH model, in forecasting house price volatility were also acknowledged by Milles (2011a) and C. L. Lee and Reed (2014). The former author found the CGARCH model to outperform the standard GARCH in the US house price volatility forecasting. The latter authors identified the CGARCH model as a better candidate to capture the Australian house price volatility's short and long-run movements. The authors also emphasised that the CGARCH's ability lies in its capabilities to break down the conditional variance into two components (transitory and permanent).

Furthermore, among the SV models, the SVl model stands out compared to the others, especially in the one-room flats category, where the model excels in 5 cities/sub-areas. This result highlights, once again, the vital role of asymmetric volatility not only in assets volatility modelling but also in their forecasting exercise, even by employing the probabilistic framework. The second place comes the heavy-tailed SV model, stressing the role that the fat tails stylised facts of the house price returns play in forecasting price volatility, with the heavy-tailed distributions models outperforming the standard errors based ones. Regarding whether the AR(2) process is a valuable addition component to the vanilla SV model when forecasting house price volatility, the two models share ranking with their performances varying across the three flats category. Overall, concerning the best performing model in each city/sub-area, whether being a GARCH-type or SV, no geographical or regions pattern is noted in all three flats type categories; the models' forecasting performances differ across sub-areas and cities and by apartment groups.

Table 6. House price volatility forecasting - Best performing models.

		One room flats		**Two rooms flats**		**Three rooms flats**	
Regions	Cities/Sub–areas	Model	RMSE/MAE	Model	RMSE/MAE	Model	RMSE/MAE
Helsinki	hki	EGARCH	0.0112	FIGARCH	0.0097	CGARCH	0.0136
	hki1	CGARCH	0.0174	FIGARCH	0.0132	EGARCH	0.0203
	hki2	EGARCH	0.0118	EGARCH	0.0087	–	–
	hki3	–	–	SV–2	0.0176	CGARCH	0.0174
	hki4	CGARCH	0.0174	CGARCH	0.0211	EGARCH	0.0184
Tampere	tre	–	–	–	–	EGARCH	0.0131
	tre1	SVl	0.0350	FIGARCH	0.0171	FIGARCH	0.0172
	tre2	–	–	FIGARCH	0.0215	CGARCH	0.0495
	tre3	SV–2	0.0570	–	–	CGARCH	0.0177
Turku	tku	FIGARCH	0.0163	EGARCH	0.0133	SV	0.0194
	tku1	CGARCH	0.0295	–	–	SVt	0.0289
	tku2	SVl	0.0332	SVt	0.0302	CGARCH	0.0308
	tku3	CGARCH	0.0391	–	–	SVt	0.0277
Oulu	oulu	CGARCH	0.0345	–	–	CGARCH	0.0138
	oulu1	SV	0.0494	–	–	EGARCH	0.0208
Lahti	lti	CGARCH	0.0541	SV	0.0176	CGARCH	0.0268
	lti1	FIGARCH	0.1572	–	–	FIGARCH	0.0284
Jyväskylä	jkla	SV–2	0.0332	CGARCH	0.0209	FIGARCH	0.0188
	jkla1	FIGARCH	0.0364	SV–2	0.0208	EGARCH	0.0222
	jkla2	SVl	0.0739	SVl	0.0647	EGARCH	0.0444
Pori	pori	FIGARCH	0.0612	EGARCH	0.0428	–	–
	pori1	FIGARCH	0.0473	SVl	0.0569	SVl	0.0765
	pori2	–	–	FIGARCH	0.0342	–	–
Kuopio	kuo	SVt	0.0271	CGARCH	0.0172	FIGARCH	0.0228
	kuo1	FIGARCH	0.0623	FIGARCH	0.0191	CGARCH	0.0349
	kuo2	SVl	0.0924	–	–	SV–2	0.0512
Joensuu	jnsu1	SV	0.0616	EGARCH	0.0209	–	–
Seinäjoki	seoki	–	–	EGARCH	0.0315	SVl	0.0423
Vaasa	vaasa	–	–	CGARCH	0.0174	CGARCH	0.0259
	vaasa1	–	–	–	–	SV	0.0392
	vaasa2	–	–	–	–	SVl	0.0299
Kouvola	kou	SVt	0.0549	SVl	0.0801	–	–
Lappeenranta	lrta	FIGARCH	0.0383	SVt	0.0245	SV	0.0348
	lrta1	CGARCH	0.0443	SVt	0.0295	–	–
	lrta2	–	–	–	–	FIGARCH	0.0070
Hämeenlinna	hnlina	SVl	0.0421	FIGARCH	0.0261	–	–
	hnlina1	–	–	CGARCH	0.0319	FIGARCH	0.0403
Kotka	kotka	CGARCH	0.0277	–	–	FIGARCH	0.0561
	kotka1	–	–	SV	0.0698	–	–

Notes: This table reports the house price volatility forecasting best performing models for each city and sub-area, in each flat category. MAE is the Mean Absolute Error and RMSE is Root Mean Squared Error. "–" indicates that city or sub-area is with constant variance in that specific flat category.

4.6 The Finnish housing market outlook

Once for each city and sub-area, the out-of-sample best performing model has been established, the model is re-estimated on the whole sample data $1988{:}Q1$ to $2018{:}Q4$. Next, the three-year house price returns and volatility forecasts are built, $2019{:}Q1$ to $2021{:}Q4$. In other words, the model is estimated through $2018{:}Q4$; then it is used to forecast 12 quarters-ahead. Tables 7 and 8 report the results of these forecasting procedures for both regions groups: with constant variances and time-varying variances. It is of paramount importance to express that these forecasts are subject to significant uncertainty as the forecasting period includes the starting phase of the ongoing health crisis, spring 2020. During this period, the housing market activities and the economy, in general, come to a halt due to the restrictive measures taken to combat the ongoing COVID-19 pandemic. Nevertheless, with the development of the pandemic and its effects on different parts of the economy still obscure, these forecasts can reflect the Finnish housing market's future movements and prospects. Moreover, with the pandemic's impacts on the Finnish housing market, a subject of future research, these results would serve as a key comparison in assessing the degree of impact of this health and economic crisis on the residential market.

4.6.1 Regions with constant variances

Table 7 reports the house price returns forecasts of regions with constant variances using the established best performing out-of-sample model for each city and sub-area. That is, either the short-memory ARMA model or the long memory ARFIMA model as specified in Table 4. The forecasts of the "Anti-persistent" series, meaning those series with long-range negative dependence, are carried out using the ARMA model of the respective series. The ex-ante forecasting results highlight several observations. First, in most cities/sub-areas, the house price return is predicted to reach its maximum (marked in bold) in the first three quarters of 2019; after that, it steadily decreases. Second, on average, in the 2019 - 2021 period, it is expected that the house price returns in the one-room flats category will continue to grow, especially in the Vaasa-area1, Helsinki-area3, and Lahti-area. These areas' return growth rate will be much higher than that achieved in the last thirty years. Those rates are respectively, as follows: 1.37 per cent compares to 1.00 per cent in the past three decades [see Table 1 in Article 2], 1.00 per cent (0.88 per cent), and 0.60 per cent (0.36 per cent). In the same period, in the two-rooms flats category, the regions which show an average positive growth in price returns are Turku-area1 and Turku-area 3, with both around 1 per cent per quarterly.

On the other hand, Lahti-area2 and Kuopio-area2 are predicted to take a downturn. This perspective is mainly due to a sharp decrease in house price returns experienced by these areas in the last three quarters of the estimation sample period. Lahti-area2 endured a 6 per cent decrease in the fourth quarter of 2018, whereas Kuopio-area2 experienced a 4 and 8 per cent price returns decline in the third and fourth quarters of 2018, respectively. Joensuu-city is on the flat line with an average growth return of 0.05 per cent per quarterly. Finally, in the larger flats category, the average price return in the Lahti-area2 is also expected to be in the negative territory. Again, it results from a declining trend noted in this sub-market; a negative house price return is observed since the fourth quarter of 2017. Thus, a close look at the house prices movements of this area and their determinants through the lenses of demographic and economic factors is a subject of future research. Kouvola-city is also predicted to have a negative average growth return; this is primarily due to the 11 per cent drop in house price return experienced by this city in the fourth quarter of 2018. Other regions in this group are expected to have continuous return growth.

Table 7. Ex-ante forecasts for the house price returns (%): 2019:$Q1$ - 2021:$Q4$.

		One room flats												
Regions	Cities/Sub-areas	**Average (2019 – 2021)**	2019:Q1	2019:Q2	2019:Q3	2019:Q4	2020:Q1	2020:Q2	2020:Q3	2020:Q4	2021:Q1	2021:Q2	2021:Q3	2021:Q4
Helsinki	hki3	1.00	-0.48	**3.63**	-1.21	3.15	-1.21	2.97	-1.12	2.85	-1.00	2.74	-0.89	2.63
Tampere	tre	0.64	**2.18**	1.22	0.17	-0.39	-0.42	-0.08	0.38	0.76	0.98	1.04	0.98	0.88
	tre2	0.95	-1.06	**2.96**	-1.06	2.96	-1.06	2.96	-1.05	2.96	-1.06	2.96	-1.06	2.96
Oulu	oulu2	0.66	-0.03	**0.89**	0.66	0.72	0.70	0.71	0.71	0.71	0.71	0.71	0.71	0.71
Lahti	lti2	0.60	**3.44**	-0.05	0.44	0.37	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38
Joensuu	jnsu	0.77	0.79	**2.34**	-0.70	1.68	-0.13	1.29	0.21	1.06	0.41	0.93	0.53	0.84
Vaasa	vaasa	0.33	**0.99**	-1.42	0.03	0.32	0.39	0.45	0.48	0.51	0.53	0.55	0.57	0.58
	vaasa1	1.37	**5.27**	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
Hämeenlinna	hnlina1	0.16	0.69	-1.42	-0.79	-0.32	0.01	0.24	0.39	0.50	0.58	0.63	0.67	**0.70**
Kotka	kotka1	0.34	-1.29	-0.69	**1.50**	0.09	0.31	0.73	0.47	0.54	0.63	0.59	0.62	0.64
		Two rooms flats												
		Average (2019 – 2021)	2019:Q1	2019:Q2	2019:Q3	2019:Q4	2020:Q1	2020:Q2	2020:Q3	2020:Q4	2021:Q1	2021:Q2	2021:Q3	2021:Q4
Tampere	tre	0.76	0.61	0.62	0.74	0.73	0.79	0.77	0.80	0.79	0.81	0.80	**0.82**	0.82
	tre3	0.75	0.02	0.81	**1.07**	1.02	0.91	0.82	0.77	0.75	0.74	0.73	0.73	0.72
Turku	tku1	1.03	**1.51**	1.04	1.08	0.99	0.99	0.97	0.97	0.96	0.96	0.96	0.95	0.95
	tku3	1.06	**2.22**	2.08	1.56	0.92	0.83	0.79	0.76	0.74	0.73	0.72	0.71	0.70
Oulu	oulu	0.66	0.45	**0.77**	0.63	0.69	0.66	0.67	0.67	0.67	0.67	0.67	0.67	0.67
	oulu1	0.65	0.28	0.58	0.62	0.65	0.67	0.69	0.70	0.72	0.73	0.73	**0.74**	0.74
	oulu2	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58
Lahti	lti1	0.95	**2.76**	0.45	1.36	0.62	0.94	0.69	0.80	0.73	0.76	0.74	0.75	0.74
	lti2	-0.59	-0.51	-1.42	-0.83	-0.78	-0.65	-0.57	-0.50	-0.45	-0.40	-0.36	-0.32	**-0.29**
Kuopio	kuo2	-0.78	-1.27	-2.41	-1.35	-1.19	-0.86	-0.68	-0.51	-0.39	-0.29	-0.21	-0.14	**-0.09**
Joensuu	jnsu	0.05	-0.99	-0.51	**0.66**	-0.17	0.08	0.25	0.12	0.21	0.24	0.23	0.26	0.28
Vaasa	vaasa1	0.39	0.11	-0.18	0.29	0.37	0.43	0.47	0.49	0.52	0.54	0.56	0.57	**0.59**
Lappeenranta	ltra2	0.40	-0.56	**0.69**	0.56	0.50	0.47	0.46	0.45	0.45	0.45	0.45	0.45	0.45
Kotka	kotka	0.39	-4.21	**2.88**	0.82	0.67	0.61	0.58	0.57	0.56	0.55	0.54	0.54	0.54
	kotka2	0.80	**3.51**	0.41	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57
		Three rooms flats												
		Average (2019 – 2021)	2019:Q1	2019:Q2	2019:Q3	2019:Q4	2020:Q1	2020:Q2	2020:Q3	2020:Q4	2021:Q1	2021:Q2	2021:Q3	2021:Q4
Helsinki	hki2	0.87	0.61	0.79	0.85	0.88	0.90	0.91	0.91	0.92	0.92	**0.93**	0.93	0.93
Oulu	oulu2	0.85	**3.69**	0.07	0.61	0.64	0.64	0.64	0.65	0.65	0.65	0.65	0.65	0.66
Lahti	lti2	-0.74	-2.25	-1.71	-0.98	-0.63	-0.57	-0.56	-0.51	-0.44	-0.38	-0.34	-0.30	**-0.27**
Pori	pori	0.56	1.39	-1.10	-1.02	0.43	**1.48**	1.47	0.88	0.44	0.43	0.67	0.86	0.86
Joensuu	jnsu	0.67	**1.06**	0.49	0.70	0.62	0.65	0.64	0.64	0.64	0.64	0.64	0.64	0.64
	jnsu1	0.76	**1.82**	0.25	0.87	0.62	0.72	0.68	0.70	0.69	0.69	0.69	0.69	0.69
Kouvola	kou	-0.29	**4.73**	-2.05	-0.95	-0.78	-0.70	-0.64	-0.60	-0.56	-0.53	-0.50	-0.48	-0.46
Hämeenlinna	hnlina	0.11	**2.56**	-0.39	-1.20	-0.29	-0.12	-0.03	0.03	0.08	0.12	0.15	0.18	0.20

Notes: This table reports in each apartment type, the house price returns forecasts of cities and sub-areas with constant variances. The bolded numbers are the maximum values in each row.

4.6.2 Regions with time-varying variances

Table 8 reports the 2019 – 2021 predicted average growth of both house price returns and volatility of regions with time-varying variances using the established best volatility forecasting model for each city and sub-area. That is, either the GARCH-type model or SV model as specified in Table 6. There are various observations to note from this prediction procedure. First, it is predicted that the growth triangle regions in Southern Finland will see a high quarterly average return growth in the 2019 – 2021 period across all three apartment types. It is expected that Helsinki and Turku regions will experience a quarterly growth return rate of at least 1.10 per cent, and Tampere will reach a 2.80 per cent return rate in the larger flats category and a 2.13 per cent in the two-rooms flats category. This perspective is essentially due to these regions' continuous growth in terms of population and economic development. The three areas currently count for 49 per cent of the Finnish population and 55 per cent of the total GDP. Hence, these regions are often at the top of the housing investors' list as their housing market activities are very active. Other areas projected to see a high quarterly house price return average growths are Lahi-area1 and Kuopio-area1, both in the one-room flats category with a 1.53 per cent and 1.39 per cent rate, respectively.

Second, in less densely populated regions, Hämeenlinna, Lappeenranta, and Kotka are predicted to face a decline in quarterly house price return on average in 2019 to 2021, especially in the two-rooms and larger flats categories. Both Lappeenranta-area2 and Hämeenlinna-area1 will experience the most significant drop of 0.30 per cent, and Kotka-area1 will see a decrease of 0.08 per cent. Other regions with the same trend include Pori-area1, whose average house price return growth is also expected to be in the negative territory in the both cited flats categories, and Seinäjoki-city as well in the larger flats category. This observation results from little or even sluggish demand for larger apartment types compared to other types. Currently, in Finland, urbanisation and a continuous decline in household's average size drive high demand for small and well-located flats. Therefore, more than three rooms flats category have experienced a significant downturn in their favour, leading to their drop in prices, especially in thinly populated regions.

Finally, regarding the risk measure – the volatility – it is predicted that most cities and sub-areas will experience a steady trend during the three years; no extreme swing is observed in most regions. The volatility measure is set to remain constant or even lower compared to the last three years [results not included in the table]. On average, the quarterly risk measure is set to vary between 1 per cent and 5 per cent in the vast majority of Finnish cities and sub-areas. The Helsinki-city market is expected to be less volatile across all three apartment types; this has also been the case for the last thirty years based on the unconditional standard deviation measure [See Table 1-3 in Article 2].

Table 8. Predicted average growth of house price returns and volatility (%).

Regions	Cities/Sub–areas	**One room flats**		**Two rooms flats**		**Three rooms flats**	
		Mean	Volatility	Mean	Volatility	Mean	Volatility
Helsinki	hki	1.23	1.02	1.34	1.56	0.97	1.65
	hki1	1.32	2.74	1.46	2.30	1.80	2.49
	hki2	1.16	1.19	1.36	1.73	–	–
	hki3	–	–	0.28	3.18	0.68	2.79
	hki4	0.84	3.34	0.75	2.15	1.04	3.11
Tampere	tre	–	–	–	–	2.80	2.86
	tre1	0.08	4.90	2.13	2.51	1.12	2.23
	tre2	–	–	0.77	3.40	0.79	4.34
	tre3	0.18	4.55	–	–	0.89	2.74
Turku	tku	0.75	3.02	1.07	2.45	0.21	3.78
	tku1	1.13	4.51	–	–	0.26	5.32
	tku2	0.33	6.84	0.22	4.88	1.09	4.13
	tku3	0.53	5.14	–	–	0.27	5.02
Oulu	oulu	0.69	4.11	–	–	0.52	2.26
	oulu1	0.09	4.71	–	–	0.50	1.80
Lahti	lti	0.67	6.71	-0.10	3.31	0.84	3.82
	lti1	1.53	13.21	–	–	0.97	3.91
Jyväskylä	jkla	-0.04	4.35	0.89	3.26	0.83	2.90
	jkla1	0.87	3.11	0.05	3.95	0.68	3.36
	jkla2	0.01	7.70	0.18	4.53	0.01	4.83
Pori	pori	0.04	7.00	0.93	3.75	–	–
	pori1	0.88	7.05	-0.07	5.75	-0.19	6.84
	pori2	–	–	0.85	4.51	–	–
Kuopio	kuo	0.04	4.64	0.68	2.55	0.60	2.58
	kuo1	1.39	5.62	0.88	2.56	0.91	4.93
	kuo2	0.24	7.66	–	–	0.34	5.25
Joensuu	jnsu1	0.34	5.43	0.29	4.66	–	–
Seinäjoki	seoki	–	–	0.81	5.59	-0.12	6.52
Vaasa	vaasa	–	–	0.49	3.15	0.20	3.78
	vaasa1	–	–	–	–	0.02	5.55
	vaasa2	–	–	–	–	0.52	6.69
Kouvola	kou	0.36	6.92	0.19	5.52	–	–
Lappeenranta	lrta	0.39	3.49	-0.16	4.13	-0.02	5.19
	lrta1	0.77	6.24	-0.29	5.41	–	–
	lrta2	–	–	–	–	-0.30	8.27
Hämeenlinna	hnlina	0.26	5.93	0.76	3.97	–	–
	hnlina1	–	–	0.35	3.64	-0.30	6.83
Kotka	kotka	0.51	5.58	–	–	0.81	5.64
	kotka1	–	–	-0.08	5.87	–	–

Notes: This table reports in each apartment type; the predicted average growth of house price returns and volatility of cities and sub-areas with time-varying variances.

Exceptions exist mostly in the one-room flats category, where few regions are expected to experience extreme volatility swing. The alarming case is for Lahti-area1 with a quarterly average risk measure of about 13 per cent, which is even high than the 11 per cent of the last three years. This area's housing market also recorded the highest risk measure in the past thirty years, and it seems that the highly volatile tendency will continue. Thus, a thorough investigation of this area's house price volatility determinants warrants future research. The same trend of regions with high house price fluctuations includes the Pori, Jyväskylä-area2, and Kuopio-area2, each with a quarterly average risk measure of about 7 per cent. Also, Lappeenranta-area1, Lahti and Kouvola cities, each with about 6 per cent risk measure per quarterly. Furthermore, in the larger flats category, three areas are also predicted to have a volatility upswing. Those are Pori-area1, Lappeenranta-area2, and Hämeenlinna-area1. It is interesting to note that these regions' high volatility is associated with a lower average quarterly return. One explanation of these areas' negative risk-return relationship would be the Glosten et al.'s (1993) argument that investors would accept a lower return during volatile periods, believing that the future would be more riskier.

5 CONCLUSIONS

In recent years, the housing market forecasting has been the theme of extensive research because of the vital role house price forecasts play in asset consumption, allocation, investment decision making, and predicting mortgage defaults. Moreover, the housing sector has been found to act as a leading indicator for the country's economy. Hene, future house price movements can provide policymakers with insights into the economy's overall trend, thereby support designing improved and more adapted policies. This study's four objectives were first to examine whether the Finnish house prices of fifteen main regions displayed a constant or a time-varying variance. Second, to investigate within the two groups, which time-series data generating processes can model the studied return series. Third, to establish which time-series forecasting model that provided superior out-of-sample forecasts for each region. Fourth, to offer the Finnish housing market development likelihood in terms of three-year forecasts in both cases of constant and time-varying variance. With these aims, several research questions served as a guide in the investigations, and four articles were published to address these raised issues. The following conclusions can be drawn from this study.

First, the evidence of time-varying variance was found in the house price returns of the absolute majority of the studied cities/sub-areas across the three studied flats categories. An observation which indicates that, in Finland, apartment types prices are heteroscedastic and clustered over time. Second, the two models candidates in modelling the house price returns of regions with constant variance, the ARMA and ARFIMA model, yield mixed results. Their performance is driven by the return series understudy and differs across flat categories. On the other hand, the ARFIMA models outclass their ARMA peers in forecasting these house price returns. This result highlights the essential role that long memory feature plays in returns forecasting procedure. Further, it confirms the ARFIMA model's capability to capture the long-range dependencies found in the studied house price returns. Additionally, by comparing the GARCH-types models and stochastic volatility models in modelling house price returns of regions with time-varying variances, the findings highlight how, in the two groups, the models' incorporating leverage effects rank as the best in-sample fit models. This aspect stresses, also in case of house prices, the importance of asymmetric volatility features in modelling asset returns.

Third, in forecasting the house price returns and volatilities of these regions with heteroscedastic errors, the GARCH-types models outperform their SV counterparts across the three apartment types. Moreover, among these best performing GARCH-types models, the ones accommodating the long memory feature come at the top. This observation emphasises, once again, the vital role of the long memory characteristic in forecasting not only asset returns but also asset volatilities. Finally, with regards to the outlook of the Finnish housing market; within the constant variance group regions, it is predicted that the house price returns in the one-room flats category will continue to grow with a 2019 – 2021 average growth rate of about 1 per cent. Two-rooms and larger flats price returns are expected to see the same trend; however, four regions are predicted to take a downturn. Those are Lahti-area2 and Kuopio-area2 in the former category and Lahti-area2 and Kouvola-city in the latter category. This perspective is mainly due to the sharp decline in house prices that these areas experienced in the three consecutive quarters of 2018; with Lahti-area 2 experiencing a steep decrease in larger apartment prices since the fourth quarter of 2017. Within the time-varying group regions, the growth triangle regions in Southern Finland, namely, Helsinki, Turku, and Tampere are expected to see a high quarterly average return growth in the 2019 – 2021 period across all three apartment types. On the other hand, in the same period, less densely populated areas such as Hämeenlinna, Lappeenranta, and Kotka, their quarterly house price returns are predicted to decline, especially in the two-rooms and larger flats categories. In the case of the risk measure, house price volatility, a steady trend is expected in most regions with the exception of a few areas in the one-room and larger flats categories, which will experience high price fluctuations.

These dissertation's outcomes have some crucial housing investment and policy implications. Due to the substantial evidence of clustering effects found in the studied regions' house prices, it is recommended that the housing consumers, investors, and policymakers track the asset volatility since it contains vital information. Moreover, accurate house price returns and volatility are central for housing investments and risk assessment and offer essential insights for portfolio allocation across Finland and apartment type. Furthermore, precise future house price movements support economists, policymakers and institutions' decision-making. This view is in connection with the linkages found between housing markets of various countries and their economic cycles.

5.1 Future research

As discussed above, some regions demonstrate a notable future house price returns and volatility movements. Thus, further study will focus on investigating the demographic and economic determinants of regions' house price returns and volatility trends. Moreover, the housing sector has been found to play a vital role in the country's macroeconomic. Therefore, the other component of further research would be to explore the relationship between the long-term and short-term dynamics of the key macroeconomic variables and the Finnish house prices and identify cross-country similarities and differences in this respect. The considered macroeconomic features would include interest rates, income, unemployment rate, inflation rate, lending rate, housing building approval rate, and Gross domestic product growth. The demographic factors, such as population growth would also be incorporated into the analysis. The information from these predictors can be further employed to improve the forecast accuracy of house prices. Furthermore, as Begiazi and Katsiampa (2019) pointed out, housing markets are usually exposed to major shocks that can lead to structural breaks. Hence, the other feature of further research would be to examine the existence of structural breaks in the studied housing markets, which could also improve the forecast accuracy of price returns and volatility. Additionally, the other facet of further research would be to investigate the impact of the ongoing health crisis – the COVID-19 pandemic - on the Finnish housing market through the lenses of the above-cited macroeconomic and demographic variables.

REFERENCES

Agnello, L., & Schuknecht, L. (2010). Booms and busts in housing markets: Determinants and implications. *Journal of Housing Economics*, *20*, 171-190.

Alvarez-Lacalle, E., Dorow, B., Eckmann, J.-P., & Moses, E. (2006). Hierarchical structures induce long-range dynamical correlations in written texts. *PNAS*, *103*, 7956–7961.

Anundsen, A. K. (2020). *Nordic house price bubbles?* (Working Paper Series 4). Housing Lab - Oslo Metropolitan University.

Apergis, N., & Payne, J. E. (2020). Modeling the time varying volatility of housing returns: Further evidence from the U.S. metropolitan condominium markets. *Review of Financial Economics*, *38*(1), 24-33. Retrieved from https://doi.org/10.1002/rfe.1063

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generarized autoregressive conditional heteroscedasticity. *Jouranl of Econometrics*, *74*(1), 3-30.

Baillie, R. T., Han, Y.-W., Myers, R. J., & Song, J. (2007). Long memory models for daily and high frequency commodity futures returns. *Journal of Futures Markets*, *27*, 643-688. Retrieved from https://doi.org/10.1002/fut.20267

Barari, M., Sarkar, N., Kundu, S., & Chowdhury, K. B. (2014). Forecasting House Prices in the United States with Multiple Structural Breaks. *International Econometric Review*, *6*(1), 1-23.

Barot, B., & Yang, Z. (2002). House prices and housing investment in Sweden and the UK: Econometric analysis for the period 1970-1998. *Review of Urban & Regional Development Studies*, *14*(2), 189-216.

Barros, C. P., Chen, Z., & Gil-Alana, L. A. (2013). Long memory in the housing price indices in China. *Asian Journal of Empirical Research*, *3*(7), 785-807.

Barros, C. P., Gil-Alana, L. A., & Payne, J. E. (2015). Modeling the long memory behavior in U.S housing price volatility. *Journal of Housing Research*, *24*(1), 87-106.

Baugnet, V., Butzen, P., Cheliout, S., Melyn, W., & Wibaut, Q. (2011). *End of the crisis in the housing markets? An international survey* (Economic Review).

National Bank of Belgium.

Bayer, P., Ellickson, B., & Ellickson, P. (2010). Dynamic asset pricing in a system of local housing markets. *American Economic Review*, *100*(2), 368-372.

Begiazi, K., & Katsiampa, P. (2019). Modelling UK House Prices with Structural Breaks and Conditional Variance Analysis. *The Journal of Real Estate Finance and Economics*, *58*, 290–309. Retrieved from https://doi.org/10.1007/s11146-018-9652-5

Beran, J., Feng, Y., Ghosh, S., & Kulik, R. (2013). *Long Memory Processes*. Heidelberg, Germany: Springer.

Berg, N., Jha, N., & Murdoch, J. C. (2012). *Risk in Housing Markets*. International Encyclopedia of Housing and Home, Elsevier. (pp. 193-203)

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *81*, 637-654.

Boitan, I. (2016). Residential property prices' modeling: evidence from selected European countries. *Journal of European Real Estate Research*, *9*(3), 273-285. Retrieved from https://doi.org/10.1108/JERER-01-2016-0001

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *31*(3), 307-327.

Bori, C., & Lowe, P. (2002). *Asset prices, financial and monetary stability: exploring the nexus* (Working Paper No. 114). BIS.

Bork, L., & Møller, S. (2015). Forecasting house prices in the 50 states using Dynamic Model Averaging and Dynamic Model Selection. *International Journal of Forecasting*, *31*(1), 63-78. Retrieved from https://sfx.aub.aau.dk/sfxaub?sid=pureportal&doi=10.1016/j.ijforecast.2014.05.005

Bork, L., & Møller, S. (2018). Housing price forecastability: A Factor Analysis. *Real Estate Economics*, *46*(3), 582-611.

Box, G., & Jenkins, G. (1970). *Time Series Analysis, Forecasting and Control*. Holden-Day: San Francisco, CA, USA.

Brown, J., Song, H., & McGillivray, A. (1997). Forecasting UK house prices: A time varying coefficient approach. *Economic Modelling*, *14*(4), 529-548.

Campbell, J. Y., & Cocco, J. F. (2003). *Household Risk Management and Optimal*

Mortgage Choice (Working Paper No. 1946). Harvard Institute of Economic.

Cannon, S., Miller, N., & Pandher, G. (2006). Risk and Return in the U.S. Housing Market: A Cross-Sectional Asset-Pricing Approach. *Real Estate Economics*, *34*, 519-552.

Case, K., Cotter, J., & Gabriel, S. (2011). Housing Risk and Return: Evidence of a Housing Asset-Pricing Model. *Journal of Portfolio Management*, *35*, 89-109.

Case, K., & Shiller, R. (1990). Forecasting prices and excess returns in the housing market. *Real Estate Economics*, *18*(3), 253-273.

Case, K., & Shiller, R. (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity*, *34*(2), 299-342.

Cecchetti, S., Feroli, M., Kashyap, A., Mann, C., & Schoenholtz, K. (2020). *Monetary policy in the next recession?* (Tech. Rep.). US Monetary Policy Forum. Retrieved from www.chicagobooth.edu/research/igm/events-forums/2020-us-monetary-policy-forum/paper

Chan, J. C., & Grant, A. L. (2016a). Modeling energy price dynamics: GARCH versus stochastic volatility. *Energy Economics*, *54*, 182-189.

Chan, J. C., & Grant, A. L. (2016b). On the Observed–Data Deviance Information Criterion for Volatility Modeling. *Journal of Financial Econometrics*, *14*(4), 772-802.

Cheridito, P. (2001). Mixed fractional Brownian motion. *Bernoulli*, *7*(6), 913-934. Retrieved from https://doi.org/10.2307/3318626

Cho, M. (1996). House price dynamics: A survey of theoretical and empirical issues. *Journal of Housing Research*, *7*(2), 145-172.

Christodoulou-Volos, C., & Siokis, F. M. (2006). Long range dependence in stock market returns. *Applied Financial Economics*, *16*(18), 1331-1338. Retrieved from https://doi.org/10.1080/09603100600829519

Cook, S., & Watson, D. (2017). Volatility in the Housing Market: Evidence on Risk and Return in the London Sub-market. *Quantitative Finance and Economics*, *1*(3), 272-287. Retrieved from https://doi.org/10.3934/QFE.2017.3.272

Coskun, Y., & Ertugrul, H. M. (2016). House price return volatility patterns in Turkey, Istanbul, Ankara and Izmir. *Journal of European Real Estate Research*, *8*(1), 26-51.

Crawford, G. W., & Fratantoni, M. C. (2003). Assessing the forecasting performance of Regime–Switching, ARIMA and GARCH models of house prices. *Real Estate Economics*, *31*(2), 223-243. Retrieved from https://doi.org/10.1111/1540-6229.00064

Crone, T., & Voith, R. (1999). Risk and Return Within the Single Family Housing Market. *Real Estate Economics*, *27*, 63-78.

Cunado, J., Gil-Alana, L. A., & Perez de Gracia, F. (2010). Persistence in some energy futures markets. *Journal of Futures Markets*, *30*, 490-507.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with unit root. *Journal of the American Statistical Association*, *74*, 427-31.

Dileep, A. D., & Gupta, S. (2020). Long range dependence in cloud servers: a statistical analysis based on google workload trace. *Computing*, *102*, 1031–1049. Retrieved from https://doi.org/10.1007/s00607-019-00779-4

Ding, Z., Granger, C., & Engle, R. (1993). A Long Memory Property of Stock Returns and a New Model. *Journal of Empirical Finance*, *1*(1), 83-106.

DiPasquale, D., & Wheaton, W. (1994). Housing market dynamics and the future of housing prices. *Journal of Urban Economics*, *35*(1), 1-27.

Dolde, W., & Tirtiroglu, D. (1997). Temporal and spatial information diffusion in real estate price changes and variances. *Real Estate Economics*, *25*(4), 539-565.

Dolde, W., & Tirtiroglu, D. (2002). Housing Price Volatility Changes and Their Effects. *Real Estate Economics*, *30*(1), 41-66.

Drake, L. (1993). Modelling UK house prices using cointegration: An application of the Johansen technique. *Applied Economics*, *25*(9), 1225-1228.

Duffie, D., & Singleton, K. J. (1993). Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica*, *61*(4), 929-952.

Dufitinema, J., Pynnönen, S., & Sottinen, T. (2020). Maximum likelihood estimators from discrete data modeled by mixed fractional Brownian motion with application to the Nordic stock markets. *Communications in Statistics - Simulation and Computation*, *Forthcoming*. Retrieved from https://doi.org/10.1080/03610918.2020.1764581

Elder, J., & Villupuram, S. (2012). Persistence in the return and volatility of home price indices. *Applied Financial Economics*, *22*, 1855-186.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica*, *50*(4), 987-100.

Engle, R. F., & Bollerslev, T. (1986). Modelling the peristence of conditional variances. *Econometrics Review*, *5*(1), 1-50.

Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating Time Varying Risk Premia in the Term Structure: The ARCH–M Model. *Econometrica*, *55*(”2), 391-407.

Englund, P., & Ioannides, Y. (1997). House price dynamics: An international empirical perspective. *Journal of Housing Economics*, *6*(2), 119-136.

Engsted, T., & Pedersen, T. Q. (2014). Housing market volatility in the OECD area: Evidence from VAR based return decompositions. *Journal of Macroeconomics*, *42*, 91–103. Retrieved from http://dx.doi.org/10.1016/j.jmacro.2014.07.005

Fama, E. (1968). Risk, Return and Equilibrium: Some Clarifying Comments. *The Journal of Finance*, *23*(1), 29-40.

Financial Accounts Data. (2018). *Financial Accounts of the United States* (Tech. Rep.). Retrieved from [2020-12-08]https://www.federalreserve.gov/releases/z1/20190920/html/b101h.htm

Fox, R., & Taqqu, M. S. (1986). Large-sample Properties of Parameter Estimation for Strongly Dependent Stationary Gaussian Time Series. *The Annals of Statistics*, *14*(2), 517-532.

French, C. (2003). The Treynor Capital Asset Pricing Model. *The Journal of Finance*, *1*(2), 60-72.

Geweke, J., & Porter-Hudak, S. (1983). The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, *4*(4), 221-238.

Ghalanos, A. (2020). rugarch: Univariate GARCH models [Computer software manual]. (R package version 1.4-4.)

Ghysels, E., Plazzi, A., Valkanov, R., & Torous, W. (2013). Forecasting real estate prices. In *Elliott, G. & Granger, C. & Timmermann, A., Handbook of Economic Forecasting* (1st ed., Vol. 2, p. 509-580). Amsterdam: Elsevier.

Glosten, L., Jagannathan, R., & Runkle, D. (1993). On the Relation between the Expected Value and the Volatility of the Norminal Excess Return on Stocks. *Journal of Finance*, *68*, 1179-1801.

Goetzmann, W., & Spiegel, M. (1997). A Spatial Model of Housing Returns and Neighbourhood Substitutability. *The Journal of Real Estate Finance and Economics*, *14*(1-2), 11-31.

Goodhart, C., & Hofmann, B. (2008). House prices, money, credit and the macroeconomy. *Oxford Review of Economic Policy*, *24*(1), 180-205.

Granger, C. W., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, *1*, 15–29.

Graves, T., Gramacy, R., Watkins, N., & Franzke, C. (2017). A Brief History of Long Memory: Hurst, Mandelbrot and the Road to ARFIMA, 1951–1980. *Entropy*, *19*(9), 437. Retrieved from https://doi.org/10.3390/e19090437

Guirguis, H. S., Giannikos, C. I., & Anderson, R. I. (2005). The U.S. housing market: Asset pricing forecasts using time varying coefficients. *The Journal of Real Estate Finance and Economics*, *30*, 33-53. Retrieved from https://doi.org/10.1007/s11146-004-4830-z

Guirguis, H. S., Giannikos, C. I., & Garcia, L. G. (2007). Price and Volatility Spillovers between large and small cities: A study of the Spanish market. *Journal of Real Estate Portfolio Management*, *13*, 311-316.

Guo, H., & Nelly, C. J. (2008). Investigating the intertemporal risk-return relation in international stock markets with the component GARCH model. *Economics Letters*, *99*, 371–374.

Guo, H., & Whitelaw, R. (2006). Uncovering the Risk-Return Relation in the Stock Market. *Journal of Finance*, *61*, 1443-1463.

Gupta, R., Kabundi, A., & Miller, S. M. (2011). Forecasting the US real house price index: Structural and non-structural models with and without fundamentals. *Economic Modelling*, *28*(4), 2013-2021.

Gupta, R., & Miller, S. M. (2012). The time-series properties on housing prices: A case study of the Southern California market. *Journal of Real Estate Finance and Economics*. Retrieved from https://dx.doi.org/10.2139/ssrn.1352768

Gupta, R., & Miller, S. M. (2012b). "Ripple effects" and forecasting home prices in Los Angeles, Las Vegas, and Phoenix. *The Annals of Regional Science*, *48*(3), 763-782.

Hadavandi, E., Ghanbari, A., Mohsen Mirjani, S., & Abbasian, S. (2011). An econometric panel data-based approach for housing price forecasting in Iran. *International Journal of Housing Markets and Analysis*, *4*(1), 70-83. Retrieved from https://doi.org/10.1108/17538271111111848

Han, L. (2013). Understanding the Puzzling Risk-Return Relationship for Housing. *Review of Financial Studies*, *26*(4), 877-928.

Harvey, A., Ruiz, E., & Shephard, N. (1994). Multivariate Stochastic Variance Models. *The Review of Economic Studies*, *61*(2), 247-264.

Hepşen, A., & Vatansever, M. (2011). Forecasting future trends in Dubai housing market by using Box-Jenkins autoregressive integrated moving average. *International Journal of Housing Markets and Analysis*, *4*(3), 210-223. Retrieved from https://doi.org/10.1108/17538271111153004

Hirata, H., Kose, M., Otrok, C., & Terrones, M. (2013). *Global house price fluctuations: Synchronization and determinants* (IMF Working Paper No. 13/38). IMF, Washington, DC.

Honkapohja, S. (2009). *The 1990's financial crises in Nordic countries* (Discussion Papers 7/2009). Bank of Finland Research.

Hosking, J. R. (1981). Fractional differencing. *Biometrika*, *68*(1), 165–176.

Hossain, B., & Latif, E. (2009). Determinants of housing price volatility in Canada: A dynamic analysis. *Applied Economics*, *41*(27), 3521-3531.

Hosszejni, D., & Kastner, G. (2020). Modeling Univariate and Multivariate Stochastic Volatility in R with stochvol and factorstochvol. *R package vignette*.

Huffman, F. E. (2003). Corporate real estate risk management and assessment. *Journal of Corporate Real Estate*, *5*, 31–41.

Hurst, H. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, *116*, 770-808.

International Monetary Fund. (April, 2003). *When bubles bust* (Tech. Rep.). Washington:IMF: World Economic Outlook – World Economic and Financial Survey.

Jacobsen, D. H., & Naug, B. E. (2005). *What drives house prices?* (Economic Bulletin No. 05 Q1). Norges Bank.

Jadevicius, A., & Huston, S. (2015). ARIMA modelling of lithuanian house price index. *International Journal of Housing Markets and Analysis*, *8*(1), 135-147. Retrieved from https://doi.org/10.1108/IJHMA-04-2014-0010

Kaleva, H. (2019). *The Finnish Property Market* (Technical Report). KTI Property Information Ltd.

Kaleva, H. (2020). *The Finnish Property Market* (Technical Report). KTI Property Information Ltd.

Karmeshu, D., & Krishnamachari, A. (2004). Sequence Variability and Long-Range Dependence in DNA: An Information Theoretic Perspective. In *Neural information processing* (Vol. 3316, p. 1354–1361). Berlin: Springer.

Karoglou, M., Morley, B., & Thomas, D. (2013). Risk and Structural Instability in US House Prices. *The Journal of Real Estate Finance and Economics*, *46*(3), 424-436.

Katsiampa, P., & Begiazi, K. (2019). An empirical analysis of the Scottish housing market by property type. *Scottish Journal of Political Economy*, *66*(4), 559-583. Retrieved from https://doi.org/10.1111/sjpe.12210

Kim, S., Shephard, N., & Chib, S. (1998). Stochastic Volatiltity: Likelihood Inference and Comparison with ARCH Models. *The Review of Economic Studies*, *65*(3), 361-393.

Kishor, N. K., & Marfatia, H. A. (2018). Forecasting house prices in OECD economies. *Journal of Forecasting*, *37*, 170-190.

Kivistö, J. (2012). *An assessment of housing price developments against various measures* (Bank of Finland Bulletin 3). Bank of Finland.

Kosonen, K. (1997). *House Price Dynamics in Finland* (Discussion Papers 137). Labour Institute for Economic Research.

KTI. (2020). *KTI Market Review* (Tech. Rep.). KTI Property Information Ltd.

Kuismanen, M., Laakso, S., & Loikkanen, H. A. (1999). *Demographic Factors and the Demand for Housing in the Helsinki Metropolitan Area* (Working Papers No. 191). Government Institute for Economic Research.

Laakso, S. (2000). *Regional housing markets in boom and bust: The experience of Finland* (Reports No. 169). Pellervo Economic Research Institute.

Lecat, R., & Mésonnier, J.-S. (2005). *What role do financial factors play in house* *price dynamics?* (Bulletin digest No. 134). Banque de France.

Lee, C. L. (2009). Housing price volatility and its determinants. *International Journal of Housing Markets and Analysis*, *2*(3), 293-308.

Lee, C. L. (2017). An examination of the risk-return relation in the Australian housing market. *International Journal of Housing Markets and Analysis*, *10*(3), 431-449.

Lee, C. L., & Reed, R. (2014). Volatility Decomposition of Australian Housing Prices. *Journal of Housing Research*, *23*(1), 21-44. Retrieved from http://www.jstor.org/stable/24862554

Lee, G. J., & Engle, R. F. (1999). *A permenant and transitory component model of stock return volatility*. Cointegration Causality and Forecasting A Festschrift in Honor of Clive WJ Granger, Oxford University Press.

Li, K.-W. (2012). A study on the volatility forecast of the US housing market in the 2008 crisis. *Applied Financial Economics*, *22*(22), 1869-1880.

Lin, P.-T., & Fuerst, F. (2014). Volatility clustering, risk-return relationship, and asymmertic adjustment in the Canadian housing market. *Journal of Real Estate Portifolio Management*, *20*(1), 37-46.

Lo, A. W. (2001). Fat tails, long memory, and the stock market since the 1960s. *Economic Notes, Banca Monte dei Paschi di Siena*.

Malpezzi, S. (1996). Housing prices, externalities, and regulation in US metropolitan areas. *Journal of Housing Research*, *7*(2), 209-241.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, *36*(4), 394-419.

Mandelbrot, B. (1965). Une classe de processus stochastiques homothetiques a soi; application a loi climatologique de H.E. Hurst. *Comptes Rendus de l'Académie des Sciences de Paris*, *240*, 3274–3277.

Mandelbrot, B. (1983). *The Fractal Geometry of Nature*. San Francisco: W.H. Freeman and Co.

Mandelbrot, B., & Van Ness, J. (1968). Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review*, *10*, 422–437.

Mandelbrot, B., & Wallis, J. (1968). Noah, Joseph and operational hydrology.

Water Resources Research, *4*, 909–918.

Mandelbrot, B., & Wallis, J. (1969). Computer experiments with fractional Gaussian noises. *Water Resources Research*, *5*, 228–267.

Mankiw, G. N., & Weil, D. N. (1989). The baby boom, the baby bust, and the housing market. *Regional and Urban Economics*, *19*(2), 235-258.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91.

Merton, R. C. (1973). An intertemporal capital asset pricing models. *Econometrica*, *41*, 867-887.

Meyer, R., & Wieand, K. (1996). Risk and Return to Housing, Tenure Choice and the Value of Housing in an Asset Pricing Context. *Real Estate Economics*, *24*, 113-131.

Miao, H., Ramchander, S., & Simpson, M. W. (2011). Return and volatility transimission in US housing markets. *Real Estate Economics*, *39*(4), 701-741.

Miller, N. G., & Peng, L. (2006). Exploring Metropolitan Housing Price Volatility. *Journal of Real Estate Finance and Economics*, *33*(1), 5-18. Retrieved from https://doi.org/10.2139/ssrn.468300

Milles, W. (2008a). Volatility clustering in US home prices. *Journal of Real Estate Research*, *30*(1), 73-90.

Milles, W. (2008b). Boom–Bust Cycles and the Forecasting Performance of Linear and Non-Linear Models of House Prices. *The Journal of Real Estate Finance and Economics*, *36*, 249-264. Retrieved from 10.1007/s11146-007-9067-1

Milles, W. (2010). Volatility transmission in U.K housing: A multivariate GARCH approach. *Journal of Real Estate Portfolio Management*, *16*(3), 241-248.

Milles, W. (2011a). Long-Range Dependence in U.S Home Price Volatility. *Journal of Real Estate Finance and Economics*, *42*, 329-347.

Milles, W. (2011b). Clustering in UK home prices volatility. *Journal of Housing Research*, *20*(1), 87-101.

Milles, W. (2015). Bubbles, busts and breaks in UK housing. *International Real Estate Review*, *18*(4), 455-471.

Moëc, G. (2004). *Y a-t-il un risque de bulle immobiliére en france?* (September

Bulletin No. 129). Banque de France.

Morley, B., & Thomas, D. (2011). Risk-return relationships and asymmetric adjustment in the UK housing market. *Applied Financial Economics*, *21*(10), 735-742.

Morley, B., & Thomas, D. (2016). An empirical analysis of UK house price risk variation by property type. *Review of Economics & Finance*, *6*(2), 45-56.

Munro, M., & Tu, Y. (1996). The dynamics of UK national and regional house prices. *Review of Urban & Regional Development Studies*, *8*(2), 186-201. Retrieved from https://doi.org/10.1111/j.1467-940X.1996.tb00117.x

Nagaraja, C. H., Brown, L. D., & Zhao, L. H. (2011). An autoregressive approach to house price modeling. *The Annals of Applied Statistics*, *5*(1), 124-149. Retrieved from https://projecteuclid.org/euclid.aoas/1300715185

Nakajima, J., & Omori, Y. (2012). Stochastic volatiltity model with leverage and asymmetrically heavy-tailed error using GH skew Student's *t*–distribution. *Computational Statistics & Data Analysis*, *56*(11), 3690-3704. Retrieved from https://doi.org/10.1016/j.csda.2010.07.012

Nellis, J., & Longbottom, J. (1981). An empirical analysis of the determination of house prices in the United Kingdom. *Urban Studies*, *18*, 9-21.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, *55*, 703-708.

Oikarinen, E. (2005). *The diffusion of housing price movements from centre to surrounding areas* (Discussion Papers No. 979). The Research Institute of the Finnish Economy (ETLA).

Oikarinen, E. (2007). *Studies on housing price dynamics* (PhD dissertation). Turku School of Economics.

Oikarinen, E., & Asposalo, H. (2004). Housing portfolio diversification potentials in the Helsinki metropolitan area in the short and long horizon. *Publications of the Turku School of Economics and Business Administration*, *Series: Discussion and Working Papers 11:2004*. Retrieved from https://urn.fi//URN:ISBN:951-564-223-X

Ólan, H. T. (2002). Long memory in stock returns: some international evidence. *Applied Financial Economics*, *12*(10), 725-729. Retrieved from https://doi.org/10.1080/09603100010025733

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*, 246-256.

Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335-46.

Poon, S.-H., & Granger, W. C. (2003). Forecasting Volatility in Financial Markets: A Review. *Journal of Economic Literature*, *41*(2), 478-539.

Quigley, J. (1999). Real estate prices and economic cycles. *International Real Estate Review*, *2*(1), 1-20.

R Core Team. (2020). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Rabemananjara, R., & Zakoian, J. M. (1993). Threshold Arch Models and Asymmetries in Volatility. *Journal of Applied Econometrics*, *8*(1), 31-49.

Rapach, D., & Strauss, J. (2009). Differences in housing price forecastability across US states. *International Journal of Forecasting*, *25*(2), 351-372.

Reen, T. A., & Razali, M. N. (2016). The dynamics of house price volatility in Malaysia. *Journal of Technology Management and Business*, *3*(2), 14-35.

Reisen, V., Abraham, B., & Lopes, S. (2001). Estimation of Parameters in ARFIMA Processes: A Simulation Study. *Communications in Statistics - Simulation and Computation*, *30*(4), 787-803.

Reserve Bank of Australia. (2004). *Financial stability review* (Tech. Rep.).

Robinson, P. (2003). *Time Series with Long Memory, Advanced Texts in Econometrics*. Oxford University Press.

Sadorsky, P. (2006). Modeling and forecasting petroleum futures volatility. *Energy Economics*, *28*, 467-488.

Samorodnitsky, G. (2006). Long Range Dependence. *Foundations and Trends in Stochastic Systems*, *1*(3), 163-257.

Savills. (2019). *Value of UK housing stock hits record high.* Retrieved from [2020-12-08]https://www.savills.com/insight-and-opinion/Blog

Savva, C. S., & Michail, N. A. (2017). Modelling house price volatility states in Cyprus with switching ARCH models. *Cyprus Economic Policy Review*, *11*(1), 69-82.

Scruggs, J. (1998). Resolving the Puzzling Intertemporal Relation Between the Market Risk Premium and Conditional Market Variance: A Two-Factor Approach. *Journal of Finance*, *53*, 575-603.

Segnon, M., Gupta, R., Lesame, K., & Wohar, M. E. (2020). High-frequency volatility forecasting of US housing markets. *The Journal of Real Estate Finance and Economics*. Retrieved from https://doi.org/10.1007/s11146-020-09745-w

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, *19*(3), 425-442.

Shiller, R. (1998). *Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks*. New York: Oxford University Press.

Statistics Finland. (2016). *Households' assets* (Tech. Rep.). Retrieved from [2020-12-08]http://www.stat.fi/til/vtutk/2016/vtutk-2016-2018-06-05-tie-001-en.html

Statistics Finland. (2019). *Building and dwelling production* (Tech. Rep.). Retrieved from [2020-12-23]http://www.stat.fi/til/ras/index-en.html

Statistics Finland Overview. (2018). *Household–dwelling units and housing conditions* (Tech. Rep.). Retrieved from [2020-11-08]http://www.stat.fi/til/asas/2018/011/asas-2018-01-2019-10-10-kat-002-en.html

Statistics Finland Overview. (2019). *Household–dwelling units and housing conditions* (Tech. Rep.). Retrieved from [2020-12-24]http://stat.fi/til/asas/2019/01/asas-2019-01-2020-10-14-kat-002-en.html

Sun, L., Wang, L., & Fu, P. (2018). Maximum likelihood estimators of a long–memory process from discrete observations. *Advances in Differential Equations*(154). Retrieved from https://doi.org/10.1186/s13662-018-1611-1

Sunega, P., Lux, M., & Zemčík, P. (2014). Housing Price Volatility and Econometrics. *Critical Housing Analysis*, *1*(2), 70-78. Retrieved from http://dx.doi.org/10.13060/23362839.2013.1.2.117

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes: A study of daily sugar prices 1961-75. *In OD Anderson (ed.), Time Series Analysis, Theory and Pratice*, 203–226. North–Holland, Amsterdam.

Taylor, S. J. (1986). *Modelling Financial Time Series*. Chichester: Wiley.

Tian, Y., & Gallagher, K. P. (2015). *Housing Price Volatility and the Capital Account in China* (GEGI Working Papers No. 2). Global Economic Governance Initiative.

Tsai, I.-C. (2014). Spillover effect between the regional and the national housing markets in the UK. *Regional Studies*, *49*(12), 1-20.

Tsai, I.-C., & Chen, M.-C. (2009). The asymmetric volatility of house prices in the UK. *Property Management*, *27*(2), 80-90.

Tsai, I.-C., Chen, M.-C., & Ma, T. (2010). Modelling house price volatility states in the UK by switching ARCH models. *Applied Economics*, *42*(9), 1145-1153.

Varotsos, D., & Kirk-Davidoff, D. (2006). Long-memory processes in global ozone and temperature variations. *Atmospheric Chemistry and Physics Discussions*, *6*, 4325–4340.

Webb, R. I., Yang, J., & Zhang, J. (2016). Price jump risk in the US housing market. *The Journal of Real Estate Finance and Economics*, *53*(1), 29-49.

Wei, Y., & Cao, Y. (2017). Forecasting house prices using dynamic model averaging approach: Evidence from China. *Economic Modelling*, *61*, 147-155. Retrieved from http://dx.doi.org/10.1016/j.econmod.2016.12.002

Whitelaw, R. (1994). Time variations and covariations in the expectation and volatility of stock market returns. *Journal of Finance*, *49*(2), 515-541.

Whittle, P. (1953). Estimation and Information in Stationary Time Series. *Arkiv for Matematik*, *3*, 423-434.

Wilhelmsson, M., & Zhao, J. (2018). Risk Assessment of Housing Market Segments: The Lender's Perspective. *Journal of Risk and Financial Management*, *11*(4). Retrieved from https://doi.org/10.3390/jrfm11040069

Willcocks, G. (2009). U.K. housing market: Time series processes with independent and identically distributed residuals. *The Journal of Real Estate Finance and Economics*, *39*, 403-414. Retrieved from https://doi.org/10.1007/s11146-020-09745-w

Willcocks, G. (2010). Conditional variances in UK regional house prices. *Spatial Economic Analysis*, *5*(3), 339-354.

Zakoian, J. M. (1994). Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, *18*(5), 931-955.

Zheng, X. (2015). Expectation, volatility and liquidity in the housing market. *Applied Economics*, *47*(37), 4020-4035. Retrieved from https://doi.org/10.1080/00036846.2015.1023943

Zhou, Z. (1997). Forecasting sales and price for existing single-family homes: A VAR model with error correction. *Journal of Real Estate Research*, *14*(2), 155-167.

Zhu, B., Fuss, R., & Rottke, N. B. (2013). Spatial linkages in returns and volatilities among US regional housing markets. *Real Estate Economics*, *41*(1), 29-64.

PUBLICATIONS

Publication I

Dufitinema, J. and Pynnönen, S.
Long-range dependence in the returns and volatility of the Finnish housing market

Journal of European Real Estate Research
Volume 13, Issue 1, pp. 29-50, 2020

This is the Author Accepted Manuscript version of the article. The original version can be found at DOI: https://doi.org/10.1108/JERER-07-2019-0019.

Long–range dependence in the returns and volatility of the Finnish Housing Market

Josephine Dufitinema[a] and Seppo Pynnönen[a]

[a]School of Technology and Innovation, Mathematics and Statistics Unit, University of Vaasa, Finland

Accepted 28 November 2019

Abstract

Purpose – The purpose of the paper is to examine the evidence of long–range dependence behaviour in both house price returns and volatility for fifteen main regions in Finland over the period of 1988:$Q1$ to 2018:$Q4$. These regions are divided geographically into forty–five cities and sub–areas according to their postcode numbers. The studied type of dwellings is apartments (block of flats) divided into one–room, two–rooms, and more than three rooms apartments types.
Design/methodology/approach – For each house price return series, both parametric and semiparametric long memory approaches are used to estimate the fractional differencing parameter d in an Autoregressive Fractional Integrated Moving Average (ARFIMA (p,d,q)) process. Moreover, for cities and sub–areas with significant clustering effects (ARCH effects), the semiparametric long memory method is used to analyse the degree of persistence in the volatility by estimating the fractional differencing parameter d in both squared and absolute price returns.
Findings – A higher degree of predictability was found in all three apartments types price returns with the estimates of the long memory parameter constrained in the stationary and invertible interval; implying that the returns of the studied types of dwellings are long–term dependent. This high level of persistence in the house price indices differs from other assets, such as stocks and commodities. Furthermore, the evidence of long–range dependence was discovered in the house price volatility with more than half of the studied samples exhibiting long memory behaviour.
Research limitations/implications – Investigating the long memory behaviour in both returns and volatility of the house prices is crucial for investment, risk, and portfolio management. One reason is that, the evidence of long–range dependence in the housing market returns suggests a high degree of predictability of the asset. The other reason is that, the presence of long memory in the housing market volatility aids in the development of appropriate time series volatility forecasting models in this market. The study outcomes will be used in modelling and forecasting the volatility dynamics of the studied types of dwellings. The quality of the data limits the analysis and the results of the study.
Originality/value – To the best of the authors' knowledge, this is the first research that assesses the long memory behaviour in the Finnish housing market. Also, it is the first study that evaluates the volatility of the Finnish housing market using data on both municipal and geographical level.
Keywords – House prices, Returns, Volatility, Long memory, Finland
Paper type – Research paper

1 Introduction

The Finnish property investment market is booming. It amounted up to EUR 69.5 billion at the end of 2018; that is an increase of 9.1 per cent compared to the previous year (Kaleva, 2019). In terms of property sector, currently, the residential properties are the largest sector in the Finnish property investment market. They represented 29 per cent of the total property investment market in 2018. The high demand for small and well–located apartments boosts this strong residential property investment as young or working–age population are moving towards urban areas. In 2018, up to 75 per cent of the newly constructed dwellings were for studios and one–bedroom flats (Statistics Finland, 2019). Moreover, according to the freshest statistics from 2016; housing consisted 50.3 per cent of the Finnish households' total wealth (Statistics Finland, 2016). Therefore, understanding the dynamics of the Finnish house prices, especially, investigating whether the returns and volatility of those types of dwellings preferred by investors exhibit long memory behaviour is crucial; for investment, risk, and portfolio management. One reason is that, the evidence of long–range dependence in the housing market returns suggests a high degree of predictability of the asset based on historical information. The other reason is that, the presence of long memory in the housing market volatility is the key element in the development of appropriate time series volatility forecasting models in this market; which can have substantial impacts of macroeconomic activity.

Previous research has examined the evidence of long memory in either returns or volatility of different assets classes; such as stocks (Hiemstra and Jones, 1997; Ólan, 2002; Christodoulou-Volos and Siokis, 2006), commodity futures (Baillie et al., 2007), and energy futures (Cunado et al., 2010). Moreover, the presence of persistence has been analysed in real state returns and volatility (Elder and Villupuram, 2012), and in individual housing markets (Milles, 2011; Feng and Baohua, 2015). While previous studies in different countries such as the United Kingdom and the United States have tested the evidence of long memory in housing markets using data sets at the state or metropolitan level of the family–home property type; for housing investment and portfolio allocation purposes; this study uses the Finnish house price indices data on both metropolitan and geographical level of the apartments in the block of flats property type which has increased its investors' attractiveness in the Finnish residential properties sector.

The general purpose of the study is to provide to the investors, risk managers, policymakers, and consumers the information regarding diversifying a housing investment portfolio across Finland and by apartment type; as many investors are often highly concentrated in narrow geographical regions such as Helsinki. In other words, the aim is to answer the following research question: "What type of apartments, geographically located in which area of Finland, should be included in the investment portfolio to acquire the best possible risk–return relationships?" This question is answered using appropriate modelling and forecasting approaches to understand the dynamics of the market. Thus, specifically, this article analyses the long–range behaviour in both returns and volatility of the Finnish housing market by the size of the apartments; that is, single–room apartments, two–rooms apartments, and apartments with more than three rooms. The study outcomes will be used in modelling and forecasting the volatility dynamics of the studied types of dwellings. Plus precisely, in an in–sample and out–of–sample forecasting test and performance comparison of different univariate time series models. The employed methodology is as follows. For each house price return series, both parametric and semiparametric long memory approaches are used to estimate the fractional differencing parameter d in an Autoregressive Fractional Integrated Moving Average (ARFIMA (p,d,q)) process. Moreover, for cities and sub–areas with significant clustering effects (ARCH effects), the semiparametric long memory method is used to analyse the degree of persistence in the volatility by estimating the fractional differencing parameter d in both squared and absolute price returns.

The study contributes to the literature by being the first attempt to assess the long memory behaviour in the Finnish housing market. Also, it is the first study that evaluates the volatility of the Finnish housing market using both municipal and geographical data level of the investors' favoured property type. Results reveal strong supportive evidence of the long memory behaviour in both returns and volatility of the studied apartment types. The high degree of persistence found in the house price returns differs from other assets, such as stocks and commodities. For house price volatility, the strong evidence of long memory is following other assets volatility dynamics. However, the degree of long–range dependence found is much higher.

The remainder of the article is organised as follows. Section 2 reviews the relevant literature. Section 3 describes the data and the methodology to be employed. Section 4 presents and discusses the results. Section 5 concludes the article.

2 Literature review

There has been extensive research on the housing market; whether the focus is on modelling the price dynamics, capturing the price volatility, or investigating the presence of substantial persistence in returns and/or volatility. The examination of these issues is done on individual housing markets or across different housing markets. For instance, Lin and Fuerst (2014) and Hossain and Latif (2009) examined Canadian house price volatility; Lee (2009), Lee (2017), and Lee and Reed (2013) studied Australian house price volatility. Guirguis et al. (2007) have investigated house price and volatility spillovers between two cities in the Spanish housing market, Madrid and Coslada; while Coskun and Ertugrul (2016) modelled the volatility properties of the house price of Turkey, Istanbul, Ankara, and Izmir. Apart from extensive studies on the house price volatility in developed countries; the dynamic of the housing market in small countries such as the Cyprus island (Savva and Michail, 2017) and developing countries such as Malaysia (Reen and Razali, 2016) have also been studied. However, the United States (US) and the United Kingdom (UK) are the two countries that have drawn more attention in terms of residential real state studies.

Regarding modelling house price volatility; authors who have studied US house prices include Dolde and Tirtiroglu (1997) who found evidence of time–varying volatility of the house price in the towns in Connecticut and San Francisco area. Moreover, Dolde and Tirtiroglu (2002) identified volatility shifts in the house price returns for four regions, and concluded that these shifts were due to regional conditions rather than national economic conditions. Miller and Peng (2006) and Milles (2008) investigated the evidence of ARCH effects in home prices at the metropolitan statistical area (MSA) level and state level, respectively. They found proof of ARCH effects in 34 MSAs out of 277 and 28 states out of 50. Furthermore, studies such as Miao et al. (2011), Karoglou et al. (2013), Webb et al. (2016), and Zhu et al. (2013) investigated different issues associated with shocks, price jump, and risk–return relationships in the various US areas and cities throughout different sample periods. The investigations of house price volatility in the UK include the works of Milles (2010), Milles (2011b), Milles (2015), Tsai (2014), Willcocks (2010), and Morley and Thomas (2011).

Regarding investigating whether house prices exhibit long–memory behaviour; Tsai et al. (2010) studied the volatility persistence in the UK housing market by older and newer homes. The authors employed the switch ARCH model and found a high magnitude of the high volatility regime for both the older and new housing market. In the US housing market, Milles (2011) found that over half of the 62 studied MSAs exhibit long memory in the conditional volatility, especially the West Coast MSAs. Elder and Villupuram (2012) examined the evidence of long–term behaviour of house price for 14 city indices and 10–city composite indices and found a higher degree of long–range dependence in both house price returns and volatility. Barros et al. (2015) evaluated the long–range dependence of house price volatility employing both data on state and metropolitan level. They found stationary long memory behaviour in the studied sample; to encompass each state and additional metropolitan areas; their analysis and results parallels to the Elder and Villupuram (2012) findings.

The outcomes of the above studies suggest evidence of volatility clustering in housing markets and long–range dependence in a limited number of countries. For Finland, there have been no investigations of house price volatility in general and of the long–term dependence behaviour in both returns and volatility in particular. Therefore, this paper aims to extend the current literature on countries' housing market volatility analysis. Moreover, in the literature, different studies employed data on the state, national, regional, or metropolitan level. However, few studies have been undertaken on house price volatility series using cross–level data for the seek of comparative analysis. Hence, this article attempts to fill that gap by using data on both metropolitan and geographical level for housing market investment and portfolio allocation purposes. Furthermore, as pointed out by Katsiampa and Begiazi (2019), few studies have attempted to analyse house price dynamics by property type level; hence to extend the extremely limited literature, this study uses data on apartments in the block of flats property type which has increased its investors' attractiveness in the Finnish residential properties sector.

3 Data and Methodology

Data

The study employs the Statistics Finland quarterly house price indices data of fifteen main regions in Finland; throughout 1988:$Q1$ to 2018:$Q4$, for a total of 124 observations. The studied regions are ranked according to their number of inhabitants. There are four regions with more than 250,000 inhabitants: Helsinki, Tampere, Turku, and Oulu; of which the three first make up the so–called growth triangle in Southern Finland, and Oulu is the growth center of Northern Finland. Seven regions with more than 100,000 inhabitants: Lahti, Jyväskylä, Kuopio, Pori, Seinäjoki, Joensuu, and Vaasa. Four regions with a population number between 80,000 – 90,000: Lappeenranta, Kouvola, Hämeenlinna, and Kotka. These regions are then divided geographically into cities and sub–areas according to their postcodes number (see Table 8 in Appendix A); to form a total of forty–five cities and sub–areas. The considered type of dwellings is apartments (block of flats) because they are the most homogenous assets in the housing market compared to other housing types, such as detached and terraced. Additionally, in Finland, flats are favored by investors. The apartments types are divided into single–room, two–rooms, and more than three rooms apartments.

Tables 1–3 provide the summary statistics of the quarterly house price returns for single–room, two–rooms, and more than three rooms flats respectively. Note that cities and sub–areas without available data for at least 20 years (80 observations) have been removed from the analysis. Over the studied period, Pori–area1 leads the one–room apartments type group with the highest average return (1.33 percent per quarterly). Kuopio–area1

follows with 1.32 percent per quarterly average return. Vaasa–area1, Lahti–area1, and Helsinki–area1 come in third place with an average return of at least 1.2 percent per quarterly. In terms of volatility dimension, Pori–area1 also recorded the highest risk measure (standard deviation), followed by Lahti–area1. The largest cities, such as Helsinki and Tampere, as well as Helsinki–area2, appear to be less volatile as they have the lowest risk level; suggesting a less significance of the ARCH effects in these cities and area.
The Two–rooms apartments type group appears to have less quarterly average returns, in general; compare to one–room and more than three rooms flats types. Helsinki–area1 scores the highest average return (1.30 percent per quarterly), followed by Helsinki–city, Helsinki–area2, Tampere–area1, and Turku-area1 with at least 1.0 percent per quarterly average return. Kotka–area2 leads the group in terms of risk measure. Same as in one–room apartments type group, the biggest cities (Helsinki, Tampere, Turku, and Oulu) and their surrounding areas seem to be less volatile. Helsinki–area1 also comes on top with 1.29 percent per quarterly average return in the more than three rooms apartments type group, followed by Lappeenranta–area2 and Tampere–area1. Hämeenlinna–area1, Joensuu–area1, and Seinäjoki–city are the more volatile areas of the group.

The house price movement of a sample of the three most volatile cities/sub–areas in each of the apartments categories over the studied period is shown in Figure 1. Those are Pori–area1, Pori–city, Jyväskylä–area2 in one–room apartments type group; Kotka–area2, Pori–area1, Kotka–area1 in two–rooms apartments type group; and Hämeenlinna–area1, Joensuu–area1, Seinäjoki–city in more than three rooms apartments type group. Initial evidence of volatility clustering effects is observed in all sample cities and sub–areas as they exhibit high fluctuations with certain time periods of high volatility followed by low volatility for other periods. A similar pattern is observed in all the graphs from the end of the 1980s until mid–1993, the period that Finland experienced financial market deregulation which induces a structural break in house price dynamics (Oikarinen, 2009a; Oikarinen, 2009b).

Methodology

The methodology employed in this study is presented as follows: first, we filter first order autocorrelations from the returns with an ARMA model of appropriate order determined by Akaike information criteria (AIC) and Bayesian information criteria (BIC). Thereafter, we test ARCH effects on the ARMA filtered returns. Next, an analysis of long memory behaviour in both returns and volatility is undertaken. That is, for each house price return series, both parametric and semiparametric long memory approaches are used to estimate the long memory parameter d of individual ARFIMA process. Lastly, for cities and sub–areas with significant clustering effects (ARCH effects), the semiparametric long memory method is used to analyse the degree of persistence in the volatility by estimating the fractional differencing parameter d in both squared and absolute price returns. All analysis was conducted in R (R Core Team, 2019).

Cities/Sub–areas	Abbrevations	Mean	Maximum	Minimum	Sd	nobs
Helsinki–city	hki	1.12	10.5	-9.1	3.5	124
Helsinki–area1	hki1	1.25	12.9	-8.7	4.1	124
Helsinki–area2	hki2	1.15	9.6	-9.0	3.6	124
Helsinki–area3	hki3	0.96	12.6	-12.6	4.1	124
Helsinki–area4	hki4	0.78	11.1	-12.0	4.3	124
Tampere–city	tre	1.01	11.6	-10.9	3.9	123
Tampere–area1	tre1	1.12	13.7	-13.8	4.9	123
Tampere–area2	tre2	1.13	15.8	-16.1	5.9	119
Tampere–area3	tre3	0.91	17.6	-11.9	5.0	123
Turku–city	tku	0.99	15.0	-9.6	4.4	124
Turku–area1	tku1	1.10	16.7	-11.7	5.5	124
Turku–area2	tku2	1.02	25.3	-19.3	6.9	111
Turku–area3	tku3	1.01	15.4	-23.0	6.4	114
Oulu–city	oulu	0.79	12.6	-10.3	4.3	124
Oulu–area1	oulu1	0.81	16.0	-12.0	5.1	124
Oulu–area2	oulu2	0.89	16.8	-16.7	5.7	116
Lahti–city	lti	0.80	17.6	-14.4	5.4	124
Lahti–area1	lti1	1.27	44.1	-24.6	8.1	109
Lahti–area2	lti2	0.55	17.9	-19.6	6.2	124
Jyväskylä–city	jkla	0.87	14.3	-10.1	4.7	124
Jyväskylä–area1	jkla1	0.99	15.7	-13.0	5.1	124
Jyväskylä–area2	jkla2	1.13	31.1	-18.5	7.4	91
Pori–city	pori	0.96	25.5	-23.5	7.6	124
Pori–area1	pori1	1.33	32.9	-23.6	8.7	100
Kuopio–city	kuo	0.95	17.9	-11.7	4.4	123
Kuopio–area1	kuo1	1.32	18.9	-18.6	5.9	111
Kuopio–area2	kuo2	1.12	16.6	-17.0	6.7	87
Joensuu–city	jnsu	0.88	17.1	-14.6	5.0	122
Joensuu–area1	jnsu1	0.93	18.7	-14.5	5.5	117
Vaasa–city	vaasa	1.01	15.7	-14.8	6.8	121
Vaasa–area1	vaasa1	1.29	18.8	-15.9	7.6	105
Kouvola–city	kou	0.39	16.5	-15.5	6.8	118
Lappeenranta–city	lrta	0.68	13.5	-12.9	4.9	124
Lappeenranta–area1	lrta1	1.01	18.6	-18.4	6.9	97
Hämeenlinna–city	hnlina	0.86	13.8	-15.7	5.9	124
Hämeenlinna–area1	hnlina1	1.07	13.3	-17.9	6.4	103
Kotka–city	kotka	0.71	18.2	-11.8	5.7	121
Kotka–area1	kotka1	1.11	17.5	-14.3	6.9	95

Notes: This table presents summary statistics on the one–room flats price index returns. Units are quarterly returns in percentage points. The sample is 1988:$Q1$ to 2018:$Q4$.

Table 1: One–room flats quarterly house price returns – Summary statistics (%).

Cities/Sub–areas	Abbrevations	Mean	Maximum	Minimum	Sd	nobs
Helsinki–city	hki	1.02	10.9	-8.8	3.2	124
Helsinki–area1	hki1	1.30	18.9	-14.2	4.8	124
Helsinki–area2	hki2	1.07	10.6	-8.3	3.3	124
Helsinki–area3	hki3	0.89	9.5	-10.7	3.7	124
Helsinki–area4	hki4	0.72	9.6	-9.7	3.6	124
Tampere–city	tre	0.93	10.5	-8.5	3.1	124
Tampere–area1	tre1	1.07	12.7	-7.9	3.7	123
Tampere–area2	tre2	0.85	10.4	-15.5	4.4	123
Tampere–area3	tre3	0.81	11.4	-11.8	3.6	123
Turku–city	tku	0.85	11.6	-8.3	3.4	124
Turku–area1	tku1	1.02	12.6	-11.6	4.2	124
Turku–area2	tku2	0.80	10.5	-12.2	4.6	124
Turku–area3	tku3	0.77	14.5	-8.3	4.7	124
Oulu–city	oulu	0.70	9.2	-6.2	3.1	124
Oulu–area1	oulu1	0.73	11.3	-6.9	3.6	124
Oulu–area2	oulu2	0.67	11.9	-9.8	4.2	124
Lahti–city	lti	0.58	10.9	-8.4	3.5	124
Lahti–area1	lti1	0.78	13.3	-10.8	4.5	124
Lahti–area2	lti2	0.46	11.9	-7.4	3.9	124
Jyväskylä–city	jkla	0.67	9.4	-8.6	3.3	124
Jyväskylä–area1	jkla1	0.79	12.4	-9.6	3.8	124
Jyväskylä–area2	jkla2	0.55	20.5	-18.6	4.7	124
Pori–city	pori	0.85	22.5	-15.5	5.2	124
Pori–area1	pori1	0.96	24.6	-17.4	6.4	124
Pori–area2	pori2	0.84	18.5	-15.9	6.3	122
Kuopio–city	kuo	0.75	12.9	-12.3	3.5	124
Kuopio–area1	kuo1	0.96	16.7	-15.5	4.8	123
Kuopio–area2	kuo2	0.60	11.5	-9.3	3.7	124
Joensuu–city	jnsu	0.76	14.8	-11.8	4.9	124
Joensuu–area1	jnsu1	0.80	16.8	-12.8	5.7	124
Seinajöki–city	seoki	0.88	20.9	14.4	6.0	118
Vaasa–city	vaasa	0.78	10.0	-8.6	4.1	123
Vaasa–area1	vaasa1	0.88	10.2	-9.3	4.3	121
Kouvoula–city	kou	0.42	27.0	-18.0	5.6	124
Lappeenranta–city	lrta	0.60	16.1	-10.8	3.9	124
Lappeenranta–area1	lrta1	0.72	21.1	-15.8	5.4	123
Lappeenranta–area2	lrta2	0.62	20.9	-17.9	5.7	122
Hämeenlinna–city	hnlina	0.73	12.1	-14.4	4.5	124
Hämeenlinna–area1	hnlina1	0.76	14.1	-16.8	5.2	124
Kotka–city	kotka	0.71	14.1	-10.5	5.1	124
Kotka–area1	kotka1	0.90	18.2	-16.1	6.4	121
Kotka–area2	kotka2	0.84	21.4	-23.7	8.1	96

Notes: This table presents summary statistics on the two–rooms flats price index returns. Units are quarterly returns in percentage points. The sample is 1988:$Q1$ to 2018:$Q4$.

Table 2: Two–rooms flats quarterly house price returns – Summary statistics (%).

Cities/Sub–areas	Abbrevations	Mean	Maximum	Minimum	Sd	nobs
Helsinki–city	hki	1.01	12.9	-9.7	3.6	124
Helsinki–area1	hki1	1.29	15.3	-14.5	5.1	124
Helsinki–area2	hki2	1.02	13.9	-8.5	3.7	124
Helsinki–area3	hki3	0.83	12.4	-9.0	3.9	124
Helsinki–area4	hki4	0.72	12.7	-11.2	3.8	124
Tampere–city	tre	0.94	11.7	-11.7	3.7	123
Tampere–area1	tre1	1.09	15.1	-14.6	4.7	123
Tampere–area2	tre2	1.07	12.4	-14.2	5.5	116
Tampere–area3	tre3	0.73	13.3	-12.0	3.5	123
Turku–city	tku	0.85	13.4	-10.2	3.9	124
Turku–area1	tku1	1.08	16.8	-15.8	5.3	124
Turku–area2	tku2	0.84	16.6	-14.8	4.9	124
Turku–area3	tku3	0.76	12.6	-10.5	4.5	124
Oulu–city	oulu	0.77	13.1	-12.4	3.8	124
Oulu–area1	oulu1	0.81	15.2	-14.6	4.6	123
Oulu–area2	oulu2	0.80	10.6	-13.5	4.5	123
Lahti–city	lti	0.66	12.3	-11.5	4.4	124
Lahti–area1	lti1	0.84	16.9	-13.8	5.7	124
Lahti–area2	lti2	0.51	10.6	-11.0	4.5	124
Jyväskylä–city	jkla	0.72	15.1	-9.3	4.4	124
Jyväskylä–area1	jkla1	0.79	17.0	-12.1	5.1	122
Jyväskylä–area2	jkla2	0.79	19.5	-17.4	6.3	122
Pori–city	pori	0.88	16.6	-16.7	5.7	124
Pori–area1	pori1	1.02	18.2	-18.3	6.6	116
Kuopio–city	kuo	0.69	14.61	-14.5	4.4	124
Kuopio–area1	kuo1	0.99	16.5	-27.9	6.9	115
Kuopio–area2	kuo2	0.62	16.7	-18.5	4.9	122
Joensuu–city	jnsu	0.85	19.3	-18.2	6.2	124
Joensuu–area1	jnsu1	0.98	22.6	-19.8	7.2	108
Seinajöki–city	seoki	1.06	27.5	-24.2	7.2	103
Vaasa–city	vaasa	0.81	15.8	-15.4	5.1	123
Vaasa–area1	vaasa1	0.97	19.1	-13.8	5.9	116
Vaasa–area2	vaasa2	1.09	14.8	-20.4	6.9	82
Kouvoula–city	kou	0.37	15.1	-13.9	6.7	121
Lappeenranta–city	lrta	0.59	12.7	-15.7	5.5	121
Lappeenranta–area2	lrta2	1.14	23.7	-22.2	6.9	80
Hämeenlinna–city	hnlina	0.80	22.0	-15.8	6.2	122
Hämeenlinna–area1	hnlina1	0.97	27.5	-16.2	7.4	108
Kotka–city	kotka	0.69	21.6	-17.5	6.4	120

Notes: This table presents summary statistics on the more than three rooms flats price index returns. Units are quarterly returns in percentage points. The sample is 1988:$Q1$ to 2018:$Q4$.

Table 3: More than three rooms flats quarterly house price returns – Summary statistics (%).

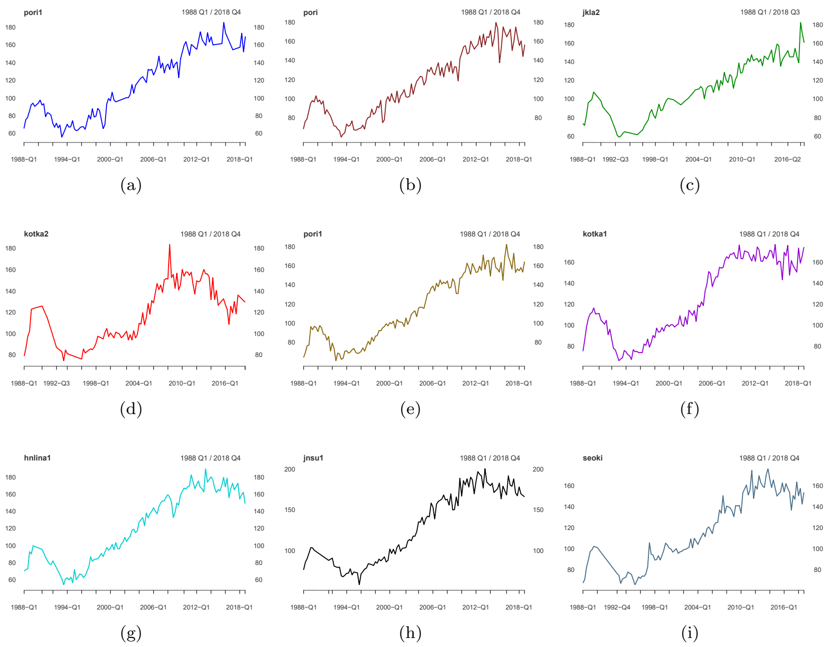


Figure 1: The house price movement of the most volatile cities/sub–areas.

Testing for ARCH effects

After filtering serial correlations from the returns series; the squared residual series are used to check the autoregressive conditional heteroscedasticity, also known as ARCH effects. If the null hypothesis of constant variance is rejected, then volatility modelling is required. Two tests are available. The first test, called Portmanteau *Q(m)*, is to examine whether the squares of the residuals are a sequence of white noise. It is the usual Ljung–Box test on the squared residuals, (see Mcleod and Li, 1983). The null hypothesis of the test statistic is that "there is no autocorrelation in the squared residuals up to lag m," that is, the first m lags of the autocorrelation function (ACF) of the squared residuals are zeros. A small p–value (smaller than the considered critical value) suggests the presence of autoregressive conditional heteroscedasticity (strong ARCH effects).

The second test is the Lagrange Multiplier test of Engle (1982), also known as ARCH–LM Engle's test. This test is to fit a linear regression model for the squared residuals and examine that the fitted model is significant. It is equivalent to the usual F statistic for testing $\gamma_i = 0$ $(i = 1, ..., m)$ in the linear regression

$$\hat{e}_t^2 = \gamma_0 + \gamma_1 \hat{e}_{t-1}^2 + ... + \gamma_m \hat{e}_{t-m}^2 + v_t, \quad t = m+1, ...N,$$

where $\hat{e}_t^2$ is the estimated residuals, v_t is the random error, m is a prespecified positive integer, and N is the sample size. The null hypothesis of the test is that "there are no ARCH effects," that is, $H_0 : \gamma_1 = ... = \gamma_m = 0$, and the alternative hypothesis is $H_1 : \gamma_i \neq 0$ (there are ARCH effects). Again, the null hypothesis is rejected if a p–value smaller than the considered critical value is obtained at the specified number of lags. The ARCH–LM tests were performed using the function ArchTest() from the *FinTs* package (Graves, 2019).

Testing for long–range dependence in returns

The methodology employed is based on the concept of long–term dependence, also called long memory or long–range persistence. This phenomenon describes time series processes whose autocorrelation function (ACF) decays slowly to 0 at a polynomial rate as the number of lag increases. One of the best–known classes of these processes, referred to as *the long–memory time series*, is the Autoregressive Fractionally Integrated Moving Average process of order (p,d,q), denoted by ARFIMA (p,d,q); proposed independently by Granger and Joyeux (1980) and Hosking (1981).
An ARFIMA (p,d,q) can be represented as

$$\Phi(B)(1-B)^d X_t = \Theta(B)u_t,\ t = 1, 2, ..., \tag{1}$$

where u_t is a white noise with $\mathbb{E}(u_t) = 0$, and variance σ_u^2. B is the lag operator or back–shift operator such that $BX_t = X_{t-1}$. $\Phi(B)$ and $\Theta(B)$ denotes finite polynomials of order p and q respectively with unit roots outside the unit circle. That is, $\Phi(B) = 1 - \phi_1 B - .... - \phi_p B^p$ and $\Theta(B) = 1 - \theta_1 B - ... - \theta_q B^q$. The studied discrete valued time series is denoted as X_t.

The estimation procedures of the fractional differencing parameter d can be classified into two groups: parametric and semiparametric approaches. In the former group where all the parameters (autoregressive, differencing, and moving average) are estimated simultaneously; the exact maximum likelihood estimation is used. The most commonly used methods within this group are those proposed by Fox and Taqqu (1986) and Sowell (1992). In the latter group, the most widely used estimator is the one developed by Geweke and Porter-Hudak (1983), usually referred to as the GPH estimator. Two steps are followed in the semiparametric estimation: first, the fractional parameter d is estimated alone, and other parameters are estimated in the second step.
The parametric approach involves the challenge of choosing the appropriate ARMA specification as it requires an explicit identification and estimation of the p and q values, parameters of $\Phi(B)$ and $\Theta(B)$ respectively. In the semiparametric method, however, the estimation of the long memory parameter d may be done without a full specification of the data generating process. Hence, in the long–range dependence analysis, different researchers have considered different semiparametric estimators (Elder and Villupuram, 2012; Christodoulou-Volos and Siokis, 2006) or a combination of both approaches (Barros et al., 2015). Additionally, parametric procedures have been found to require heavy computations, while semiparametric methods are easy to implement (Reisen et al., 2001). Reisen et al. (2001) conducted a simulation study on the estimation of parameters in ARFIMA processes; where they compared the performance of estimating all the ARFIMA parameters based on Hosking's algorithm (Hosking, 1981) and the parametric Whittle estimator proposed by Fox and Taqqu (1986). The semiparametric estimators used in the study are the Geweke and Porter-Hudak (1983) estimator, Smoothed Periodogram estimator by Reisen (1994), Robinson (Robinson, 1995a) estimator, and Robinson's estimator based on the smoothed periodogram. The results of the study indicated that regression methods (semiparametric) performed better than parametric Whittle's approach. In the light of the above research, this article employs both parametric and semiparametric

estimators to analyse the presence of long memory in the house price returns. The semiparametric estimator used is the GPH estimator, while the parametric one is the Whittle estimator. The GPH estimator, also known as the Periodogram estimator, is based on the regression equation using the periodogram function as an estimate of the spectral density. The Whittle estimator is due to Whittle (1953) with modifications suggested by Fox and Taqqu (1986). This estimator is also based on the periodogram, and it involves the spectral density function. The Whittle estimator is the value which minimises the spectral density function. For more details, (see Fox and Taqqu, 1986; Beran, 1994; Dahlhaus, 1989).

The outcome of the estimation is assessed as follows: if $d = 0$ in Equation 1, the process exhibits short memory; corresponding to stationary and invertible ARMA modelling. The process is described as "anti–persistence" if $d \in (-0.5, 0)$. If $d \in (0, 0.5)$, the process is said to manifest long–range positive dependence or long memory as the decay of the autocorrelation is hyperbolically slow. If $d \in [0.5, 1)$, the process is mean reverting, even though it is no longer covariance stationary. That is, shocks will disappear in the long run. Finally, if $d \geq 1$, the process is nonstationary without mean reversion. The estimations of the GPH estimators were performed using the function fdGPH() in the *fracdiff* package (Fraley et al., 2015), while the Whittle estimators were estimated using the function arfima.whittle() in the *afmtools* package (Contreras-Reyes and Palma, 2013).

Testing for long–range dependence in volatility

The presence of high persistence or long memory in the volatility of those cities and sub–areas with significant ARCH effects is analysed using the semiparametric approach to estimate the long memory parameter d in the squared and absolute price returns. For volatility series, the autocorrelation of the squared and absolute returns exhibit similar decay at high lags (Harvey, 1998); which justifies the use of the long memory parameter of any of these metrics. For our purposes, we use both squared and absolute returns for the seek of comparison of the estimated parameters; even though Wright (2002) claimed the strong Monte–Carlo evidence support of using absolute returns, as squared returns can cause a severe negative bias.

Previous studies have examined the persistence nature of the house price volatility employing different Generalized Autoregressive Conditional Heteroscedasticity (GARCH)–type models (Milles, 2011). Within the GARCH–family models, the most used ones which account for long memory in the conditional variance of the assets are Integrated GARCH (IGARCH) model of Engle and Bollerslev (1986), Component GARCH (CGARCH) model of Lee and Engle (1999), and Fractionally Integrated GARCH (FIGARCH) model of Baillie et al. (1996). FIGARCH and CGARCH have been applied more often of late than the IGARCH mainly because; in the IGARCH model, shocks persist forever, meaning that the model implies infinite persistence on the conditional variance; and hence, it is too restrictive (Tayefi and Ramanathan, 2012). The use of the above long memory GARCH–types models, however, to measure the degree of house price volatility persistence can be challenging to interpret as these models require to obtain convergent parameter estimates in the conditional variance equation. Moreover, the results obtained from these GARCH–types models are not directly comparable to the ones from the parametric or semiparametric estimators. Therefore, as we aim to estimate the degree of long–range dependence in the house price volatility; rather than modelling the process governing the studied types of dwellings volatility dynamics, this article employs the semiparametric method to examine the evidence of long memory in the volatility of the studied types of dwellings. The semiparametric estimator used is the GPH estimator described above. Again, the estimations of the GPH estimators were performed using the function fdGPH()

in the *fracdiff* package (Fraley et al., 2015).

4 Results and discussions

Testing for ARCH effects

Table 4 displays the p–values and their lag orders (in parentheses) of the two tests employed to investigate whether there is volatility clustering in each house price return series. Those tests are the Ljung–Box (LB) test and the Engle's Lagrange Multiplier (LM) test. The null hypothesizes of no serial correlation in squared residuals of the LB test and no ARCH effects of the LM test are rejected; in twenty–eight out of thirty–eight studied cities and sub–areas in the one–room flats category; in twenty–nine out forty–two in the two–rooms flats category; and in thirty–three out of forty in the more than three rooms flats category. Thus, strong evidence of volatility clustering (ARCH) effects is evident in over half of the cities and sub–areas in all three apartments types.
In some cases, one of the tests is inconclusive; for instance, in the case of Tampere–area1 (in the one–room flats category) and Turku–area2 (in the two–rooms flats category), the Portmanteau test is inconclusive (we fail to reject the null hypothesis because of the higher p–values); however, the Lagrange Multiplier values are statistically significant. Similarly, in the case of Lahti–area1 (in the more than three rooms flats category), this time, however, it is the Lagrange Multiplier test which is inconclusive. In these cases, the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the squared residuals (Figure 2) are used to show that there might be some autocorrelations left even though the significance might be small.

Testing for long–range dependence in returns

Table 5 gives the GPH estimates of the fractional differencing parameter d in the house price returns with their standard errors (in parentheses). Results reveal strong evidence of long–range dependence in most return series in all three apartment type categories. In approximately 68% (twenty–six out of thirty–eight) of the return series, in the one–room flats group; there is long–range positive dependence with values of d in the stationary and invertible interval (d is varying from 0.045 to 0.445). An anti–persistence behaviour is present in nine returns series; which implies a relatively quick dissipate of shocks in the house price returns. Three house price returns series (marked in bold) are mean reverting; however, they may no longer be covariance stationary as the estimates of the long memory parameter are greater than 0.5. Those return series are Turku–area3 ($d = 0.665$), Kuopio–area1 ($d = 0.529$), and Kotka–area1 ($d = 0.510$). In these three sub–areas, shocks will wear off in the long run. Approximately 83% (thirty–five out of forty–two) and 77.5% (thirty–one out of forty) of the return series exhibit stationary long memory behaviour in the two–rooms and the more than three rooms flats category, respectively. Further, in the two respective groups, seven out of forty–two cities/sub–areas and nine out of forty cities/sub–areas display long–range negative dependence or anti–persistence behaviour; which implies unpredictability of future returns based on historical returns.

Table 6 presents the Whittle estimates of the long memory parameter d with their p–values (in parentheses). The evidence of long–range dependence ($d \in (0, 0.5)$) is observed for most of the cities and sub–areas. Plus precisely, 80% (thirty out of thirty–eight), 83% (thirty–five out of forty–two), and 77.5% (thirty–one out of forty) return series exhibit long memory behaviour in the one–room, the two–rooms, and the more than three rooms flats

		One room flats		Two rooms flats		Three rooms flats	
Regions	Cities/sub–areas	LB p–values	LM p–values	LB p–values	LM p–values	LB p–values	LM p–values
	hki	0.01724**(1)	0.01873**(1)	0.00037***(1)	0.00046***(1)	0.01916**(6)	0.01009**(6)
	hki1	0.04039**(1)	0.04302**(1)	0.02057**(4)	0.05578*(4)	0.00343***(2)	0.00367***(2)
Helsinki	hki2	0.01778**(2)	0.01814**(2)	0.00361***(1)	0.00414***(1)	0.732	0.7356
	hki3	0.4441	0.5261	0.04761**(6)	0.03667**(12)	0.00744***(16)	0.03479**(16)
	hki4	0.01691**(2)	0.01927**(2)	0.08855*(3)	0.08347*(3)	0.03538**(1)	0.03665**(1)
	tre	0.4679	0.8395	0.7761	0.7786	0.07745*(1)	0.08072*(1)
Tampere	tre1	0.4878	0.0309**(11)	0.04128**(15)	0.07986*(15)	0.01222**(1)	0.0127**(1)
	tre2	0.992	0.7463	0.00801***(6)	0.00853***(10)	0.02761**(7)	0.05303*(7)
	tre3	0.09774*(1)	0.04147**(3)	0.2478	0.2204	0.0353**(11)	0.01721**(11)
	tku	0.00579**(10)	0.00826**(10)	0.06687*(3)	0.08406*(3)	0.04951**(1)	0.04706**(1)
Turku	tku1	0.06488*(5)	0.02419**(5)	0.5635	0.5692	0.07016*(1)	0.06882*(1)
	tku2	0.07601*(1)	0.08109*(1)	0.1173	0.0827*(16)	0.05191*(1)	0.05502*(1)
	tku3	0.08333*(15)	0.0791*(15)	0.09212*(2)	0.103	0.00029***(1)	0.00035***(1)
	oulu	0.00641**(1)	0.00589**(1)	0.3195	0.3242	0.00691***(4)	0.00947***(4)
Oulu	oulu1	0.08368*(1)	0.08781*(1)	0.5811	0.5853	0.06343*(9)	0.07595*(11)
	oulu2	0.03907**(1)	0.03648**(1)	0.08702*(11)	0.7267	0.25	0.2545
	lti	0.00397**(1)	0.00437**(1)	0.05585*(2)	0.08992*(13)	0.003756***(1)	0.003697***(1)
Lahti	lti1	0.00154**(1)	0.00185**(1)	0.4088	0.412	0.03861**(16)	0.3479
	lti2	0.9878	0.588	0.9908	0.9995	0.9251	0.9793
	jkla	0.02683**(2)	0.00648**(2)	0.01422**(1)	0.00958***(1)	0.07133*(1)	0.07546*(1)
Jyväskylä	jkla1	0.01705**(2)	0.00302**(2)	0.02106**(1)	0.02135**(1)	0.04016**(18)	0.08492*(18)
	jkla2	0.5235	0.03161**(9)	$5.42*10^{-7}$***(1)	$6.57*10^{-7}$***(1)	0.00879***(1)	0.00968***(1)
	pori	0.01178**(2)	0.01481**(2)	0.00783***(14)	0.00057***(11)	0.8245	0.8272
Pori	pori1	0.01847**(2)	0.02341**(2)	0.03601**(14)	0.02325**(7)	0.0135**(1)	0.01498**(1)
	pori2	–	–	0.03395**(1)	0.03511**(1)	–	–
	kuo	0.02197**(2)	0.02043**(2)	0.02826**(3)	0.01974**(3)	$9.37*10^{-5}$***(3)	$8.99*10^{-5}$***(3)
Kuopio	kuo1	0.0332**(2)	0.0343**(2)	0.00983***(3)	0.00361***(3)	0.00570***(1)	0.00651***(1)
	kuo2	0.01171**(1)	0.01355**(1)	0.7606	0.7552	0.05511*(3)	0.06738*(3)
Joensuu	jnsu	0.9552	0.8945	0.4816	0.4866	0.3902	0.3952
	jnsu1	0.02472**(1)	0.0229**(1)	0.06071*(15)	0.1541	0.09813*(17)	0.1872
Seinäjoki	seoki	–	–	0.02307**(2)	0.02324**(2)	0.07986*(14)	0.2042
	vaasa	0.7668	0.7694	0.06089*(1)	0.06419*(1)	0.09783*(1)	0.1031
Vaasa	vaasa1	0.8042	0.8075	0.2579	0.264	0.0145**(1)	0.01613**(1)
	vaasa2	–	–	–	–	0.00897***(2)	0.01832**(2)
Kouvola	kou	0.8457	0.8458	0.00059***(1)	0.00065***(1)	0.9867	0.9868
	lrta	0.01867**(1)	0.02028**(1)	0.00166***(1)	0.00196***(1)	0.06448*(2)	0.06807*(2)
Lappeenranta	lrta1	0.02763**(1)	0.03062**(1)	0.00393***(1)	0.00452***(1)	–	–
	lrta2	–	–	0.9146	0.916	0.00360***(1)	0.00450***(1)
Hämeenlinna	hnlina	0.00286**(1)	0.00331**(1)	0.08782*(6)	0.08461*(6)	0.8803	0.8821
	hnlina1	0.9694	0.9699	0.03765**(6)	0.03006**(6)	0.02362**(7)	0.01233**(4)
	kotka	0.117	0.08195*(3)	0.1482	0.109	0.0379**(3)	0.02734**(3)
Kotka	kotka1	0.8425	0.8342	0.4851	0.04158**(7)	0.04078**(1)	0.04221**(1)
	kotka2	–	–	0.1751	0.1834	–	–

Notes: This table reports the p–values from the Portmanteau (Ljung–Box) and Lagrange Multiplier tests. The values in parentheses are the lag orders of each test. *, **, and *** indicate respectively 10%, 5%, and 1% levels of significance.

Table 4: ARCH effects tests results.

group respectively. Considering the semiparametric approach (GPH estimator), an observation of the results reveals a geographical pattern regarding which cities and sub–areas exhibit long–term dependence in the returns series; densely populated regions Helsinki, Tampere, Turku, and Oulu display long memory behaviour in all three studied types of apartments. Except for Turku–area1, Oulu–city, and Oulu–area2 in the one–room flats category where the anti–persistence behaviour is observed and Turku–area3 with the long memory parameter greater than 0.5.

Similar results of high degrees of persistence in the house price returns were found in the United States house prices indices by Elder and Villupuram (2012) on the metropolitan level. Their estimates of the fractional differencing parameter d were restricted between 0 and 0.5, and even higher than 0.5 in some cities. Moreover, a similar conclusion to Elder and Villupuram's can be drawn in case of the Finnish housing market. That is, this high level of persistence in the house price indices differs from other assets, such as stocks, energy futures, and metal futures. Energy and metal futures assets classes generally display anti–persistence and modest long memory behaviour as documented by

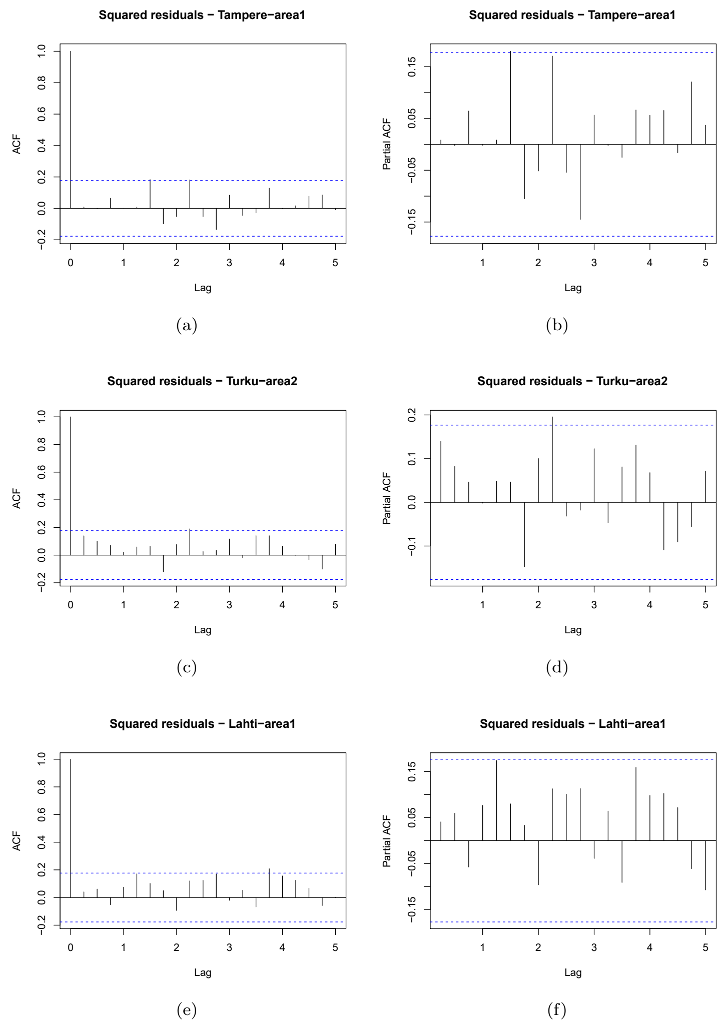


Figure 2: ACF and PACF of the squared residuals.

Barkoulas et al. (1999), Crato and Ray (2000), and Elder and Jin (2009). Also, stocks assets have been found to have the fractional differencing parameter d in the interval of -0.2 to 0.2, see Barkoulas and Baum (1996), Lo (1991), and Hiemstra and Jones (1997). Therefore, as stressed by Elder and Villupuram (2012), the long memory in real estates index returns is notable, and it is a relevant feature for issues such as constructing hedge ratios for risk management. It is worth mentioning again the challenge of specifying the appropriate ARMA *(p,q)* order in the parametric approach as different lag orders lead to varying estimations of the long memory parameter.

Testing for long–range dependence in volatility

Table 7 displays the estimates of the fractional differencing parameter in both squared and absolute returns of house prices in all three apartment types with their standard errors (in parentheses). As with house price returns, results indicate very persistent long memory for house price volatility. In the one–room flats group, the d estimates in squared and absolute returns place respectively, nineteen and eighteen out of twenty–eight cities/sub–areas into the stationary and invertible interval ($d \in (0, 0.5)$); which implies stationary long memory behaviour. An anti–persistence behaviour is present in four (for d in squared returns) and five (for d in absolute returns) cities/sub–areas; which implies a relatively quick disappearance of shocks in house price volatility. Both squared and absolute returns estimate equally place five cities/sub–areas into the interval of d between 0.5 and 1; where shocks in the house price volatility will disappear in the long run.

In the two–rooms flats category, for d estimates in both squared and absolute returns, twenty–five out of twenty–nine cities/sub–areas have long–range positive dependence with values of d in the stationary and invertible interval. Two and three cities/sub–areas, for d estimates in squared and absolute returns respectively, display long–range negative dependence. Further, in the two respective estimates, the house price volatility of two and one cities/sub–areas are mean–reverting but no longer covariance stationary. There is an exception in the more than three rooms flats group, the d estimate in two sub–areas (marked in bold) is higher than one; implying that the house price volatility process in these two sub–areas is nonstationary without mean reversion. Those sub–areas are Vaasa–area2 and Kotka–area1. Otherwise, up to twenty–four cities/sub–areas in this category exhibit stationary long memory behaviour; two cities/sub–areas are anti-persistent; and up two seven cities/sub–areas are mean–reverting. Regarding overall comparison, an inspection of the results of both squared and absolute returns fails to reveal any regional pattern as regards to the degree of persistence in volatility. However, there is some evidence to suggest that regions such as Helsinki–city, Jyväskylä–city, Kuopio–area1, and Hämeenlinna–area1 come on top in terms of house prices volatility persistence in at least two out three types of apartments. Moreover, the estimated d parameters for squared and absolute returns are quite close in most of the cases. However, there are cases where the two estimated parameters place the corresponding city or sub–area into two different intervals. For instance, Turku–area2 (in the one–room flats), d estimate for absolute returns puts it into the stationary and invertible range while the one for squared returns puts it into the mean–reverting, no longer covariance stationary interval. Therefore, as the estimated parameters will be considered in modelling and forecasting house price volatility of the studied types of apartments; both estimates will be used in the recognised model and assessing the best one will be based on the information criteria or other tools for models specification.

In summary, in twenty–eight, twenty–nine, and thirty–three cities/sub–areas which exhibited significant ARCH effects in the one–room, two–rooms, and more than three rooms flats group respectively; over half exhibit long memory in the house price volatility. Moreover, contrasting with other asset classes; such as stocks, for example, Cotter and Stevenson (2008) reported an estimate of the parameter d equalling 0.42 for volatility persistence of the S&P 500 index. Therefore, the higher estimates of long–run dependence found in our study suggest that persistence in house price volatility is much stronger than for stock prices.

Regions	Cities/Sub–areas	d estimates (GPH) One room flats	Two rooms flats	Three rooms flats
Helsinki	hki	0.140 (0.274)	0.155 (0.274)	0.121 (0.274)
	hki1	0.123 (0.273)	0.150 (0.274)	0.107 (0.274)
	hki2	0.140 (0.273)	0.178 (0.274)	0.117 (0.274)
	hki3	0.120 (0.273)	0.094 (0.274)	0.147 (0.274)
	hki4	0.232 (0.274)	0.166 (0.274)	0.146 (0.274)
Tampere	tre	0.169 (0.273)	0.245 (0.274)	0.245 (0.274)
	tre1	0.164 (0.273)	0.160 (0.274)	0.091 (0.274)
	tre2	0.271 (0.293)	0.252 (0.274)	0.221 (0.293)
	tre3	0.066 (0.274)	0.273 (0.274)	0.426 (0.274)
Turku	tku	0.071 (0.274)	0.118 (0.274)	0.209 (0.274)
	tku1	-0.008 (0.274)	0.035 (0.274)	0.190 (0.274)
	tku2	0.264 (0.293)	0.168 (0.274)	0.265 (0.274)
	tku3	**0.665** (0.293)	0.110 (0.274)	0.142 (0.274)
Oulu	oulu	-0.017 (0.274)	0.282 (0.274)	0.313 (0.274)
	oulu1	0.058 (0.274)	0.378 (0.274)	0.264 (0.274)
	oulu2	-0.605 (0.293)	0.081 (0.274)	0.187 (0.274)
Lahti	lti	0.436 (0.274)	0.233 (0.274)	0.321 (0.274)
	lti1	0.445 (0.293)	0.147 (0.274)	0.267 (0.274)
	lti2	0.208 (0.274)	0.331 (0.274)	0.347 (0.274)
Jyväskylä	jkla	0.045 (0.274)	0.095 (0.274)	0.341 (0.274)
	jkla1	-0.005 (0.274)	0.102 (0.274)	0.499 (0.274)
	jkla2	-0.509 (0.317)	0.087 (0.274)	0.390 (0.274)
Pori	pori	-0.124 (0.274)	-0.063 (0.274)	0.098 (0.274)
	pori1	0.059 (0.317)	-0.280 (0.274)	-0.272 (0.293)
	pori2	–	-0.074 (0.274)	–
Kuopio	kuo	-0.107 (0.274)	0.037 (0.274)	0.190 (0.274)
	kuo1	**0.529** (0.293)	-0.166 (0.274)	-0.198 (0.293)
	kuo2	-0.154 (0.318)	0.176 (0.274)	0.215 (0.274)
Joensuu	jnsu	0.056 (0.274)	0.291 (0.274)	0.230 (0.274)
	jnsu1	0.311 (0.293)	0.287 (0.274)	-0.065 (0.293)
Seinäjoki	seoki	–	-0.358 (0.293)	-0.551 (0.293)
Vaasa	vaasa	0.101 (0.293)	0.188 (0.274)	0.077 (0.274)
	vaasa1	-0.327 (0.293)	0.174 (0.293)	-0.160 (0.293)
	vaasa2	–	–	-0.296 (0.318)
Kouvola	kou	0.053 (0.293)	0.401 (0.274)	0.370 (0.293)
Lappeenranta	lrta	0.174 (0.274)	0.186 (0.274)	0.089 (0.293)
	lrta1	0.424 (0.317)	0.247 (0.274)	–
	lrta2	–	-0.027 (0.274)	-0.686 (0.347)
Hämeenlinna	hnlina	0.281 (0.274)	0.425 (0.274)	0.197 (0.274)
	hnlina1	0.068 (0.293)	0.401 (0.274)	-0.171 (0.293)
Kotka	kotka	0.248 (0.293)	0.116 (0.274)	0.0813 (0.293)
	kotka1	**0.510** (0.317)	0.216 (0.293)	-0.307 (0.347)
	kotka2	–	-0.329 (0.317)	–

Notes: This table reports the GPH estimates of the long memory parameter d in the house price returns. The values in parentheses are their standard errors. If $d \in (-0.5, 0)$, the series is described as "anti–persistence". If $d \in (0, 0.5)$, the process manifests the long–range dependence. If $d \in [0.5, 1)$, the process is mean reverting, even though it is no longer covariance stationary.

Table 5: Estimates of fractional differencing parameter (GPH).

Regions	Cities/Sub–areas	*d* estimates (Whittle) One room flats	Two rooms flats	Three rooms flats
Helsinki	hki	0.194 (0.188)	-0.281 (0.057)	0.256 (0.084)
	hki1	0.037 (0.801)	0.179 (0.224)	0.362 (0.014)
	hki2	0.315 (0.032)	0.296 (0.045)	0.198 (0.180)
	hki3	0.265 (0.072)	0.499 (0.000)	0.377 (0.011)
	hki4	0.468 (0.001)	-0.166 (0.260)	0.499 (0.001)
Tampere	tre	-0.345 (0.020)	-0.195 (0.187)	0.426 (0.004)
	tre1	0.033 (0.820)	0.382 (0.010)	0.251 (0.091)
	tre2	0.247 (0.101)	0.044 (0.762)	0.305 (0.046)
	tre3	-0.377 (0.011)	0.195 (0.189)	-0.344 (0.021)
Turku	tku	-0.523 (0.000)	0.307 (0.037)	0.499 (0.001)
	tku1	0.185 (0.210)	0.327 (0.027)	0.151 (0.305)
	tku2	0.055 (0.721)	0.078 (0.594)	-0.369 (0.013)
	tku3	0.178 (0.247)	0.499 (0.000)	0.090 (0.543)
Oulu	oulu	0.235 (0.112)	0.309 (0.036)	0.406 (0.006)
	oulu1	0.034 (0.816)	0.324 (0.028)	-0.333 (0.024)
	oulu2	0.102 (0.502)	-0.016 (0.909)	0.036 (0.810)
Lahti	lti	0.116 (0.432)	0.499 (0.000)	0.091 (0.539)
	lti1	0.014 (0.929)	0.405 (0.006)	0.149 (0.312)
	lti2	0.092 (0.534)	0.183 (0.217)	0.404 (0.006)
Jyväskylä	jkla	0.091 (0.535)	0.209 (0.158)	0.136 (0.356)
	jkla1	0.025 (0.865)	0.248 (0.093)	0.277 (0.063)
	jkla2	-0.108 (0.533)	-0.207 (0.161)	0.147 (0.323)
Pori	pori	-0.045 (0.758)	0.312 (0.035)	-0.002 (0.984)
	pori1	0.064 (0.695)	0.245 (0.097)	-0.054 (0.722)
	pori2	–	0.187 (0.208)	–
Kuopio	kuo	0.111 (0.455)	0.437 (0.003)	0.041 (0.777)
	kuo1	0.241 (0.122)	0.043 (0.769)	0.005 (0.974)
	kuo2	-0.040 (0.819)	0.147 (0.319)	0.299 (0.045)
Joensuu	jnsu	0.169 (0.257)	0.281 (0.057)	0.026 (0.857)
	jnsu1	0.209 (0.170)	0.311 (0.035)	0.001 (0.990)
Seinäjoki	seoki	–	0.029 (0.846)	-0.583 (0.000)
Vaasa	vaasa	-0.008 (0.952)	-0.132 (0.373)	0.255 (0.085)
	vaasa1	-0.306 (0.056)	0.115 (0.441)	-0.021 (0.886)
	vaasa2	–	–	-0.127 (0.484)
Kouvola	kou	0.189 (0.212)	0.074 (0.614)	0.244 (0.102)
Lappeenranta	lrta	0.073 (0.620)	0.093 (0.527)	0.201 (0.181)
	lrta1	0.017 (0.914)	0.065 (0.659)	–
	lrta2	–	0.125 (0.401)	-0.153 (0.409)
Hämeenlinna	hnlina	0.293 (0.047)	-0.387 (0.008)	0.241 (0.106)
	hnlina1	0.294 (0.069)	0.354 (0.016)	0.021 (0.891)
Kotka	kotka	0.279 (0.063)	0.355 (0.016)	0.300 (0.046)
	kotka1	0.259 (0.125)	0.015 (0.918)	0.171 (0.365)
	kotka2	–	0.032 (0.848)	–

Notes: This table reports the Whittle estimates of the long memory parameter d in the house price returns. The values in parentheses are their standard errors. If $d \in (-0.5, 0)$, the series is described as "anti–persistence". If $d \in (0, 0.5)$, the process manifests the long–range dependence. If $d \in [0.5, 1)$, the process is mean reverting, even though it is no longer covariance stationary.

Table 6: Estimates of fractional differencing parameter (Whittle).

Regions	Cities/sub–areas	One room flats		Two rooms flats		Three rooms flats	
		d – Squared returns	d – Absolute returns	d – Squared returns	d – Absolute returns	d – Squared returns	d – Absolute returns
Helsinki	hki	0.696 (0.274)	0.536 (0.274)	0.327 (0.274)	0.381 (0.274)	0.334 (0.274)	0.410 (0.274)
	hki1	0.381 (0.274)	0.439 (0.274)	0.278 (0.274)	0.229 (0.274)	0.685 (0.274)	0.654 (0.274)
	hki2	0.493 (0.274)	0.535 (0.274)	0.652 (0.274)	0.646 (0.274)	–	–
	hki3	–	–	0.113 (0.274)	0.017 (0.274)	0.396 (0.274)	0.565 (0.274)
	hki4	0.742 (0.274)	0.631 (0.274)	-0.056 (0.274)	-0.0411 (0.274)	0.024 (0.274)	0.029 (0.274)
Tampere	tre	–	–	–	–	0.501 (0.274)	0.353 (0.274)
	tre1	0.378 (0.274)	0.182 (0.274)	0.178 (0.274)	0.306 (0.274)	0.236 (0.274)	0.296 (0.274)
	tre2	–	–	0.374 (0.274)	0.225 (0.274)	0.166 (0.293)	0.241 (0.293)
	tre3	0.152 (0.274)	0.247 (0.274)	–	–	0.157 (0.274)	0.257 (0.274)
Turku	tku	0.106 (0.274)	-0.044 (0.274)	0.168 (0.274)	0.262 (0.274)	0.345 (0.274)	0.331 (0.274)
	tku1	0.083 (0.274)	-0.116 (0.274)	–	–	0.366 (0.274)	0.405 (0.274)
	tku2	0.721 (0.293)	0.347 (0.293)	0.137 (0.274)	0.085 (0.274)	0.511 (0.274)	0.393 (0.274)
	tku3	0.217 (0.293)	0.184 (0.293)	0.156 (0.274)	0.20 (0.274)	0.571 (0.274)	0.541 (0.274)
Oulu	oulu	-0.146 (0.274)	-0.193 (0.274)	–	–	0.480 (0.274)	0.365 (0.274)
	oulu1	0.017 (0.274)	0.019 (0.274)	–	–	0.635 (0.274)	0.423 (0.274)
	oulu2	0.006 (0.293)	0.449 (0.293)	0.013 (0.274)	0.125 (0.274)	–	–
Lahti	lti	0.025 (0.274)	0.087 (0.274)	0.125 (0.274)	0.163 (0.274)	0.036 (0.274)	-0.048 (0.274)
	lti1	0.354 (0.293)	0.853 (0.293)	–	–	0.123 (0.274)	0.005 (0.274)
Jyväskylä	jkla	0.151 (0.274)	0.257 (0.274)	0.320 (0.274)	0.343 (0.274)	0.305 (0.274)	0.414 (0.274)
	jkla1	-0.072 (0.274)	-0.049 (0.274)	0.559 (0.274)	0.409 (0.274)	0.312 (0.274)	0.396 (0.274)
	jkla2	0.304 (0.317)	0.391 (0.317)	0.073 (0.274)	0.240 (0.274)	-0.261 (0.274)	0.253 (0.274)
Pori	pori	-0.142 (0.274)	-0.218 (0.274)	0.078 (0.274)	0.188 (0.274)	–	–
	pori1	0.135 (0.317)	0.184 (0.317)	0.0009 (0.274)	-0.021 (0.274)	0.362 (0.293)	0.252 (0.293)
	pori2	–	–	0.355 (0.274)	0.150 (0.274)	–	–
Kuopio	kuo	0.363 (0.274)	0.315 (0.274)	0.147 (0.274)	0.243 (0.274)	0.202 (0.274)	0.296 (0.274)
	kuo1	0.309 (0.293)	0.288 (0.293)	0.316 (0.274)	0.391 (0.274)	0.358 (0.293)	0.411 (0.293)
	kuo2	0.757 (0.318)	0.626 (0.318)	–	–	0.159 (0.274)	0.137 (0.274)
Joensuu	jnsu1	-0.123 (0.293)	0.008 (0.293)	-0.133 (0.274)	0.003 (0.274)	-0.421 (0.293)	-0.184 (0.293)
Seinäjoki	seoki	–	–	0.249 (0.293)	0.428 (0.293)	0.255 (0.293)	0.218 (0.293)
Vaasa	vaasa	–	–	0.338 (0.274)	0.327 (0.274)	0.140 (0.274)	0.105 (0.274)
	vaasa1	–	–	–	–	0.331 (0.293)	0.109 (0.293)
	vaasa2	–	–	–	–	**1.05** (0.318)	**1.22** (0.318)
Kouvola	kou	–	–	0.360 (0.274)	0.432 (0.274)	–	–
Lappeenranta	lrta	0.539 (0.274)	0.426 (0.274)	0.158 (0.274)	0.242 (0.274)	0.199 (0.293)	0.299 (0.293)
	lrta1	0.403 (0.317)	0.422 (0.317)	0.096 (0.274)	-0.332 (0.274)	–	–
	lrta2	–	–	–	–	0.884 (0.347)	0.722 (0.347)
Hämeenlinna	hnlina	0.072 (0.274)	0.146 (0.274)	0.274 (0.274)	0.089 (0.274)	–	–
	hnlina1	–	–	0.408 (0.274)	0.362 (0.274)	0.459 (0.293)	0.470 (0.293)
Kotka	kotka	0.157 (0.293)	0.235 (0.293)	–	–	0.933 (0.293)	0.732 (0.293)
	kotka1	–	–	0.365 (0.293)	0.390 (0.293)	**1.26** (0.347)	0.679 (0.347)

Notes: This table reports the estimates of the long memory parameter d in both squared and absolute returns of house prices. The values in parentheses are their standard errors. If $d \in (-0.5, 0)$, the series is described as "anti–persistence". If $d \in (0, 0.5)$, the process manifests the long–range dependence. If $d \in [0.5, 1)$, the process is mean reverting, even though it is no longer covariance stationary. If $d \geq 1$, the process is nonstationary without mean reversion.

Table 7: Estimates of d in the Squared and Absolute returns house prices.

5 Conclusion

The presence of long memory in the asset returns implies that the considered asset returns may be predictable at long horizons; which is why investigating this issue is crucial in the development of appropriate time series forecasting models in the financial market. With this motivation, this study examines the persistence or long memory behaviour of the house price returns and volatility for fifteen main regions in Finland. The study employs both parametric and semiparametric long memory approaches to estimate the degree of long–range dependence in both returns and volatility. The results reveal strong supportive evidence of long memory in the returns; suggesting that the house price return series, contrary to other asset classes such as stocks, are strongly autocorrelated and hence highly forecastable. Moreover, in the majority of the cities and sub–areas with significant clustering effects, the long memory behaviour was found in the volatility using either squared or absolute returns. The evidence of high degree of persistence found in the house price volatility is essential higher than that exhibited by other assets categories.

In the standpoint of developing appropriate time series volatility forecasting models in this housing market; for further research, these results will be used in modelling and forecasting the volatility dynamics of the studied types of dwellings. That is, for cities and sub–areas with no significant ARCH effects, meaning those cities with constant mean and variance, and with long–range dependence in the returns; a short memory ARMA (p,q) model and a long memory ARFIMA (p,d,q) model will be used to examine which model leads to the best results in modelling house price returns. The long memory parameter d estimated in the house price returns will be incorporated in the ARFIMA (p,d,q) estimation procedure. However, as the model which fits better does not necessarily mean it will forecast well, an in–sample and out–of–sample forecasting performance of both univariate models will be assessed. Furthermore, for those cities and sub–areas with significant ARCH effects, and exhibiting long memory behaviour in the volatility; a short memory GARCH model will be employed to capture the house price volatility dynamics, and it will be compared to the other GARCH–type models which account for long memory in the conditional variance such as the Fractionally Integrated GARCH (FIGARCH) model, and the Component GARCH (CGARCH) model. Again, long memory parameter d estimated in the house price volatility will be incorporated in the FIGARCH estimation procedure, and a forecasting test will be performed to provide information regarding which forecasting methods delivers superior volatility forecasts of the studied types of apartments.

Appendices

A

Regional division of quarterly house price index data		
Cities/Sub–areas	Abbreviations for cities and sub–areas	Postcode numbers
Helsinki	hki	City area
Helsinki–area1	hki1	100, 120, 130, 140, 150, 160, 170, 180, 220, 260
Helsinki–area2	hki2	200, 210, 250, 270, 280, 290, 300, 310, 320, 330, 340, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 610, 810, 850, 990
Helsinki–area2	hki3	240, 350, 360, 370, 400, 430, 440, 440, 620, 650, 660, 670, 680, 690, 730, 780, 790, 800, 830, 840, 950
Helsinki–area4	hki4	Other postcodes
Tampere	tre	City area
Tampere–area1	tre1	33100, 33180, 33200, 33210, 33230, 33240, 33250, 33500, 33540
Tampere–area2	tre2	33270, 33400, 33530, 33560, 33610, 33700, 33730, 33820, 33900, 34240
Tampere–area3	tre3	Other postcodes

Regional division of quarterly house price index data		
Cities/Sub–areas	Abbreviations for cities and sub–areas	Postcode numbers
Turku	tku	City area
Turku–area1	tku1	20100, 20500, 20700, 20810, 20900
Turku–area2	tku2	20200, 20250, 20300, 20380, 20400, 20520, 20720, 20880, 20960
Turku–area3	tku3	Other postcodes
Oulu	oulu	City area
Oulu–area1	oulu1	90100,90120, 90130, 90140, 90230, 90400, 90410, 90420, 90510
Oulu–area2	oulu2	Other postcodes
Lahti	lti	City area
Lahti–area1	lti1	15100, 15110, 15140, 15160, 15320, 15340, 15610, 15850, 15900
Lahti–area2	lti2	Other postcodes
Jyväskylä	jkla	City area
Jyväskylä–area1	jkla1	40100, 40200, 40500, 40520, 40530, 40600, 40700, 40720
Jyväskylä–area2	jkla2	Other postcodes
Pori	pori	City area
Pori–area1	pori1	28100, 28130, 28300, 28430, 28540, 28660, 28900
Pori–area2	pori2	Other postcodes
Kuopio	kuo	City area
Kuopio–area1	kuo1	70100, 70110, 70300, 70600, 70800, 70840
Kuopio–area2	kuo2	Other postcodes
Joensuu	jnsu	City area
Joensuu–area1	jnsu1	80100, 80110, 80200, 80220
Joensuu–area2	jnsu2	Other postcodes
Seinajöki	seoki	City area
Vaasa	vaasa	City area
Vaasa–area1	vaasa1	65100, 65170, 65200, 65410
Vaasa–area2	vaasa2	Other postcodes
Kouvola	kou	City area
Lappeenranta	lrta	City area
Lappeenranta–area1	lrta1	53100, 53130, 53500, 53600, 53900, 55330
Lappeenranta–area2	lrta2	Other postcodes
Hämeenlinna	hnlina	City area
Hämeenlinna–area1	hnlina1	13100, 13130, 13200, 13220, 13270
Hämeenlinna–area2	hnlina2	Other postcodes
Kotka	kotka	City area
Kotka–area1	kotka1	48100, 48210, 48310, 48710
Kotka–area2	kotka2	Other postcodes

Source: Statistics Finland

Table 8: Regional division by postcode numbers.

References

Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generarized autoregressive conditional heteroscedasticity, *Journal of Econometrics* **74**(1): 3–30.

Baillie, R. T., Han, Y.-W., Myers, R. J. and Song, J. (2007). Long memory models for daily and high frequency commodity futures returns, *Journal of Futures Markets* **27**: 643–688.
URL: *https://doi.org/10.1002/fut.20267*

Barkoulas, J. T. and Baum, C. F. (1996). Long-term dependence in stock returns, *Economics Letters* **53**: 253–359.

Barkoulas, J. T., Labys, W. C. and Onochie, J. I. (1999). Long memory in futures prices, *Financial Review* **34**: 91–100.

Barros, C. P., Gil-Alana, L. A. and Payne, J. E. (2015). Modeling the long memory behavior in U.S housing price volatility, *Journal of Housing Research* **24**(1): 87–106.

Beran, J. (1994). *Statistics for Long–Memory Processes*, New York: Chapman and Hall.

Christodoulou-Volos, C. and Siokis, F. M. (2006). Long range dependence in stock market returns, *Applied Financial Economics* **16**(18): 1331–1338.
URL: *https://doi.org/10.1080/09603100600829519*

Contreras-Reyes, J. E. and Palma, W. (2013). Statistical analysis of autoregressive fractionally integrsted moving average models in R, *Comput Stat* .
URL: *DOI: 10.1007/s00180-013-0408-7*

Coskun, Y. and Ertugrul, H. M. (2016). House price return volatility patterns in Turkey, Istanbul, Ankara and Izmir, *Journal of European Real Estate Research* **8**(1): 26–51.

Cotter, J. and Stevenson, S. (2008). Modeling long memory in REITs, *Real Estate Economics* **36**: 533–54.

Crato, N. and Ray, B. K. (2000). Memory in returns and volatilities of futures contracts, *Journal of Futures Markets* **20**: 525–543.

Cunado, J., Gil-Alana, L. A. and de Gracia, F. P. (2010). Persistence in some energy futures markets, *Journal of Futures Markets* **30**: 490–507.

Dahlhaus, R. (1989). Efficient Parameter Eestimation for Self–Similar Processes, *The Anals of Statistics* **17**(4): 1749–1766.

Dolde, W. and Tirtiroglu, D. (1997). Temporal and spatial information diffusion in real estate price changes and variances, *Real Estate Economics* **25**(4): 539–565.

Dolde, W. and Tirtiroglu, D. (2002). Housing price volatility changes and their effects, *Real Estate Economics* **30**: 41–66.

Elder, J. and Jin, H. (2009). Fractional integration in commodity futures returns, *Financial Review* **44**: 583–602.

Elder, J. and Villupuram, S. (2012). Persistence in the return and volatility of home price indices, *Applied Financial Economics* **22**: 1855–1868.
URL: *http://dx.doi.org/10.1080/09603107.2012.687095*

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation, *Econometrica* **50**(4): 987–1007.

Engle, R. F. and Bollerslev, T. (1986). Modelling the peristence of conditional variances, *Econometrics Review* **5**(1): 1–50.

Feng, L. and Baohua, C. (2015). Research on the Long-term Memory of Commodity Housing Price Volatility Based on the FIGARCH Model, *Advanced Materials Research* **1079–1080**: 1194–1198.

Fox, R. and Taqqu, M. S. (1986). Large – sample Properties of Parameter Estimation for Strongly Dependent Stationary Gaussian Time Series, *The Annals of Statistics* **14**(2): 517–532.

Fraley, S., Leisch, F., Maechler, M., Reisen, V. and Lemonte, A. (2015). *Fractionally differenced ARIMA aka ARFIMA(p,d,q) models*. R package version 1.4-2.

Geweke, J. and Porter-Hudak, S. (1983). The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis* **4**(4): 221–238.

Granger, C. W. and Joyeux, R. (1980). An introduction to long–memory time series models and fractional differencing, *Journal of Time Series Analysis* **1**: 15–29.

Graves, S. (2019). *Companion to Tsay (2005) Analysis of Financial Time Series*. R package version 0.4-6.

Guirguis, H. S., Giannikos, C. I. and Garcia, L. G. (2007). Price and Volatility Spillovers between large and small cities:A study of the Spanish market, *Journal of Real Estate Portfolio Management* **13**: 311–316.

Harvey, A. C. (1998). *Long memory in stochastic volatility, in Forecasting Volatility in the Financial Markets, 3rd edn (Eds) J*, Knight and S. Satchell, Butterworth - Heinemann, Oxford, pp. 351-64.

Hiemstra, C. and Jones, J. D. (1997). Another look at long memory in common stock returns, *Journal of Empirical Finance* **4**(4): 373–401.
URL: *https://doi.org/10.1016/S0927-5398(96)00016-3*

Hosking, J. R. (1981). Fractional differencing, *Biometrika* **68**(1): 165–176.

Hossain, B. and Latif, E. (2009). Determinants of housing price volatility in Canada: a dynamic analysis, *Applied Economics* **41**(27): 3521–3531.

Kaleva, H. (2019). The Finnish Property Market Report, *Technical report*, KTI Property Information Ltd.

Karoglou, M., Morley, B. and Thomas, D. (2013). Risk and structual instability in US house prices, *The Journal of Real Estate Finance and Economics* **46**(3): 424–436.

Katsiampa, P. and Begiazi, K. (2019). An empirical analysis of the Scottish housing market by property type, *Scottish Journal of Political Economy* .
URL: *DOI: 10.1111/sjpe.12210*

Lee, C. L. (2009). Housing price volatility and its determinants, *International Journal of Housing Markets and Analysis* **2**(3): 293–308.

Lee, C. L. (2017). An examination of the risk–return relation in the Australian housing market, *International Journal of Housing Markets and Analysis* **10**(3): 431–449.

Lee, C. L. and Reed, R. (2013). Volatility decomposition of Australian housing prices, *Journal of Housing Research* **23**(1): 21–43.

Lee, G. J. and Engle, R. F. (1999). A permenant and transitory component model of stock return volatility, *Cointegration Causality and Forecasting A Festschrift in Honor of Clive WJ Granger* pp. 475–497.

Lin, P.-T. and Fuerst, F. (2014). Volatility clustering, risk–return relationship, and asymmertic adjustment in the Canadian housing market, *Journal of Real Estate Portifolio Management* **20**(1): 37–46.

Lo, A. W. (1991). Long-term memory in stock market prices, *Econometrica* **59**: 1279–1313.

Mcleod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared – residual autocorrelations, *Journal of Time Series Analysis* **4**(4): 269–273.

Miao, H., Ramchander, S. and Simpson, M. W. (2011). Return and volatility transimission in US housing markets, *Real Estate Economics* **39**(4): 701–741.

Miller, N. and Peng, L. (2006). Exploring metropolitan housing price volatility, *The Jouarnal of Real Estate Finance and Economics* **33**(1): 5–18.

Milles, W. (2008). Volatility clustering in US home prices, *Journal of Real Estate Research* **30**(1): 73–90.

Milles, W. (2010). Volatility transmission in U.K housing: A multivariate GARCH approach, *Journal of Real Estate Portfolio Management* **16**(3): 241–248.

Milles, W. (2011). Long–Range Dependence in U.S Home Price Volatility, *Journal of Real Estate Finance and Economics* **42**: 329–347.

Milles, W. (2011b). Clustering in UK home prices volatility, *Journal of Housing Research* **20**(1): 87–101.

Milles, W. (2015). Bubbles, busts and breaks in UK housing, *International Real Estate Review* **18**(4): 455–471.

Morley, B. and Thomas, D. (2011). Risk–return relationships ans asymmetric adjustment in the UK housing market, *Applied Financial Economics* **21**(10): 735–742.

Oikarinen, E. (2009a). Household borrowing and metropolitan house price dynamics – Empirical evidence from Helsinki, *Journal of Housing Economics* **18**(2): 126–139.

Oikarinen, E. (2009b). Interaction between housing prices and household borrowing: The Finnish case, *Journal of Banking & Finance* **33**(4): 747–756.

Ólan, H. T. (2002). Long memory in stock returns: some international evidence, *Applied Financial Economics* **12**(10): 725–729.
URL: *https://doi.org/10.1080/09603100010025733*

R Core Team (2019). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

Reen, T. A. and Razali, M. N. (2016). The dynamics of house price volatility in Malaysia, *Journal of Technology Management and Business* **3**(2): 14–35.

Reisen, V. A. (1994). Estimation of the Fractional Difference Parameter in the ARIMA (p,d,q) Model Using the Smoothed Periodogram, *Journal of Time Series Analysis* **15**(3): 335–350.

Reisen, V., Abraham, B. and Lopes, S. (2001). Estimation of Parameters in ARFIMA Processes: A Simulation Study, *Communications in Statistics – Simulation and Computation* **30**(4): 787–803.

Robinson, P. M. (1995a). Log – Periodogram Regression of Time Series with Long Range Dependence, *The Anals of Statistics* **23**(3): 1048–1072.

Savva, C. S. and Michail, N. A. (2017). Modelling house price volatility states in Cyprus with switching ARCH models, *Cyprus Economic Policy Review* **11**(1): 69–82.

Sowell, F. (1992). Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, *Journal of Econometrics* **53**: 165–188.

Statistics Finland (2016). Households' assets, *Technical report*. Last checked: 08/07/2019.
URL: *http://www.stat.fi/til/vtutk/2016/vtutk-2016-2018-06-05-tie-001-en.html*

Statistics Finland (2019). Building and dwelling production, *Technical report*. Last checked: 25/06/2019.
URL: *http://www.stat.fi/til/ras/index-en.html*

Tayefi, M. and Ramanathan, T. V. (2012). An Overview of FIGARCH and Related Time Series Models, *Austrian Jouranl of Statistics* **41**(3): 175–196.

Tsai, I.-C. (2014). Spillover effect between the regional and the national housing markets in the UK, *Regional Studies* **49**(12): 1–20.

Tsai, I.-C. T., Chen, M.-C. and Ma, T. (2010). Modelling house price volatility states in the UK by switching ARCH models, *Applied Economics* **42**(9): 1145–1153.

Webb, R. I., Yang, J. and Zhang, J. (2016). Price jump risk in the US housing market, *The Journal of Real Estate Finance and Economics* **53**(1): 29–49.

Whittle, P. (1953). Estimation and Information in Stationary Time Series, *Arkiv for Matematik* **2**: 423–434.

Willcocks, G. (2010). Conditional variances in UK regional house prices, *Spatial Economic Analysis* **5**(3): 339–354.

Wright, J. H. (2002). Log – periodogram estimation of long memory volatility dependencies with conditionally heavy tailed returns, *Econometric Reviews* **21**: 397–417.

Zhu, B., Fuss, R. and Rottke, N. B. (2013). Spatial linkages in returns and volatilities among US regional housing markets, *Real Estate Economics* **41**(1): 29–64.

Publication II

Dufitinema, J.

Volatility clustering, risk-return relationship and asymmetric adjustment in the Finnish housing market

International Journal of Housing Markets and Analysis
Volume 13, Issue 4, pp. 661-688, 2020

This is the Author Accepted Manuscript version of the article. The original version can be found at DOI: https://doi.org/10.1108/IJHMA-12-2019-0125.

Volatility clustering, risk–return relationship, and asymmetric adjustment in the Finnish housing Market

Josephine Dufitinema

School of Technology and Innovation, Mathematics and Statistics Unit, University of Vaasa, Finland

Accepted 03 February 2020

Abstract

Purpose – The purpose of the paper is to examine whether the house prices in Finland share financial characteristics with assets such as stocks. The studied regions are fifteen main regions in Finland over the period of 1988:$Q1$ to 2018:$Q4$. These regions are divided geographically into forty–five cities and sub–areas according to their postcode numbers. The studied type of dwellings is apartments (block of flats) divided into one–room, two–rooms, and more than three rooms apartment types.
Design/methodology/approach – Both Ljung–Box and Lagrange Multiplier tests are used to test for clustering effects (ARCH effects). For cities and sub–areas with significant clustering effects, the Generalized Autoregressive Conditional Heteroscedasticity–in–mean (GARCH–M) model is used to determine the potential impact that the conditional variance may have on returns. Moreover, the Exponential GARCH (EGARCH) model is employed to examine the possibility of asymmetric effects of shocks on house price volatility. For each apartment type, individual models are estimated; enabling different house price dynamics, and variation of signs and magnitude of different effects across cities and sub–areas.
Findings – Results reveal that clustering effects exist in over half of the cities and sub–areas in all studied types of apartments. Moreover, mixed results on the sign of the significant risk–return relationship are observed across cities and sub–areas in all three apartment types. Furthermore, the evidence of the asymmetric impact of shocks on housing volatility is noted in almost all the cities and sub–areas housing markets. These studied volatility properties are further found to differ across cities and sub–areas, and by apartment types.
Research limitations/implications – The existence of these volatility patterns has essential implications, such as investment decision making and portfolio management. The study outcomes will be used in a forecasting procedure of the volatility dynamics of the studied types of dwellings. The quality of the data limits the analysis and the results of the study.
Originality/value – To the best of the author's knowledge, this is the first study that evaluates the volatility of the Finnish housing market in general, and by using data on both municipal and geographical level, particularly.
Keywords – Finland, House prices, Returns, Volatility, GARCH–M, EGARCH
Paper type – Research paper

1 Introduction

The housing market is a vital factor in the economy of most developed countries. In Finland, housing consisted 50.3 per cent of the Finnish households' total wealth according to the freshest statistics from Statistics Finland (2016). Moreover, the residential properties are the largest sector in the Finnish property investment market. They represented 29 per cent of the total property investment in 2018 (Kaleva, 2019). The main booster of this strong residential property investment is the high demand for small and well–located apartments as young or working–age populations are moving towards urban areas. The reason why the Finnish housing construction has been most active in apartment buildings while the construction for other house types has decreased. In 2018, up to 75 per cent of newly constructed dwellings were studios and one–bedroom flats (Statistics Finland, 2019). In 2019, up to 4,700 apartments will be completed in the Helsinki region, and up to 2,600 in other major areas; and all these flats are being developed 100 per cent for investment market or rental market. This housing development is also boosted by a strong investment demand; currently, foreign investors hold some 15,000 rental flats. Their share in this housing construction varied between 31 and 38 per cent between 2015 and 2018. Some 40 per cent is split between domestic and individual investors; meaning that between 2015 and 2018, up to some 60 per cent of all the Finnish housing construction is targeted for investment market (KTI, Autumn, 2019). Furthermore, in Finland, living in blocks of flats is growing in popularity in comparison to other house types such as attached or terraced and detached houses. At the end of 2018, 46 per cent of all occupied dwellings were in block of flats, 39 per cent in detached and semi–detached houses, 14 per cent in attached houses, and around 1 per cent in other buildings (Statistics Finland Overview, 2018).

The Finnish residential property investment is also fueled by the fact that Finland continues to experience a period of extreme low–interest rates of mortgages, which started in 2012. The average for the last 20 years is 2.37 per cent, and since 2008, the interest rate has stayed under 2 per cent (Bank of Finland, 2018). Therefore, understanding the dynamics of the house price volatility of these types of dwellings preferred by investors is crucial for investment, risk, and portfolio management. In other words, it is essential to investigate whether the studied apartment types display volatility clustering – a volatility pattern which is often observed in stock indices. This clustering or Autoregressive Conditional Heteroscedasticity (ARCH) effect is the characteristic of a series that exhibits certain periods of higher volatility followed by lower volatility for other periods. This aspect is important because, if a process exhibits clustering effects, there is a much higher risk of large losses than standard mean–variance would suggest [see Milles (2008) and Milles (2011b) for the case of the United States and United Kingdom home prices respectively]. Further, as investors are concerned not only about the rate of return on their investment but also the risk associated with the investment; it is essential to analyse the relationship between housing risk and housing return. Finally, it is crucial to examine the possibility of asymmetric effects of shocks on the studied house price volatility. This phenomenon has two popular explanations; the leverage effect and the volatility feedback effect.

Previous research has examined house price volatility of various housing markets and highlighted the importance of testing and analysing ARCH effects in the housing markets (Milles, 2011b); investigating the relation between volatility (risk) and return (Lee, 2017); and exploring whether asymmetric effects of shocks are observable in the housing markets (Lin and Fuerst, 2014). While previous studies in different countries such the United States, United Kingdom, Australia, and Canada have tested the above issues using data sets at the state, metropolitan, and/or provincial level of the family–home property type;

for housing investment and portfolio allocation purposes, this study uses the Finnish house price indices data on both metropolitan and geographical level of the apartments in the block of flats property type which has increased its investors' and consumers' attractiveness in the Finnish residential properties sector. Moreover, the growing literature on housing volatility studies has been focusing on the countries mentioned above. There has yet to be an analysis of the conditional variance in the Finnish housing market in general, and by property type in particular. Thus, this paper aims to fill that gap by being the first study that comprehensively explores the volatility of the Finnish housing market.

The purpose of the study is, to test volatility clustering in the Finnish housing market by the size of apartments; that is, single–room apartments, two–rooms apartments, and apartments with more than three rooms, for fifteen main regions divided geographically into forty–five cities and sub–areas according to their postcode numbers. Next, for cities and sub–areas exhibiting ARCH effects, individual Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models are estimated. These models are used to answer the following questions: What is the nature of the time–varying volatility? What is the relation between housing returns and house price risk? In other words, does conditional variance affect average returns? Is there asymmetric effect in the volatility of the studied apartment types? These questions are answered by allowing different responses for each time series unit under study. That is, enabling different house price dynamics, and variation of signs and magnitude of different effects across cities and sub–areas; as various studies have found house price dynamics to be heterogeneous across different areas and property types (Milles, 2011b; Katsiampa and Begiazi, 2019). Thus, instead of imposing one single GARCH–type model on the entire data set; an individual model is estimated for each of the cities and sub–areas exhibiting ARCH effects. Results reveal that, clustering effects exist in over half of the cities and sub–areas in all three apartment types. Moreover, mixed results on the sign of the significant risk–return impact are observed across cities and sub–areas in all three apartment types. Furthermore, the evidence of the asymmetric impact of shocks on housing volatility is noted in almost all the cities and sub–areas housing markets. These studied volatility properties are further found to differ across cities and sub–areas, and by apartment types.

The remainder of the article is organised as follows. The next section reviews the relevant literature. The third describes the data and the methodology to be employed. The fourth presents and discusses the results. The fifth concludes the article.

2 Literature review

Asset volatility has been acknowledged as the most widely used risk measure in many areas of finance since it holds essential information (Bollerslev et al., 1992). In the housing literature, recent studies have underlined the importance of analysing and understanding house price volatility for housing investment, proper portfolio management, and policy decision–making (Milles, 2011; Lee and Reed, 2014a). One emphasis of the research on the housing market volatility in different countries has been to investigate whether a specific housing market exhibits volatility clustering (time–varying) effect. The United States housing market has been widely examined with authors such as Dolde and Tirtiroglu (1997; 2002), Miller and Peng (2006), and Milles (2008). Their results suggest the evidence of ARCH effects in the US housing market at the metropolitan statistical area (MSA) and state level. Moreover, similar evidence has also been documented in other housing markets such as the United Kingdom with the works of Tsai et al. (2010), Willcocks (2010), and Milles (2011b). Lee (2009) and Lee and Reed (2014b) studied Australian house price volatility; Hossain and Latif (2009) and Lin and Fuerst (2014) examined Canadian house price volatility; while Coskun and Ertugrul (2016) modelled the volatility properties of the house price of Turkey, Istanbul, Ankara, and Izmir. More recently, Katsiampa and Begiazi (2019) studied the Scottish market by property types (detached, semi–detached, terraced, and flats). All these studies confirm the evidence of clustering effects in different housing markets.

Another aspect which is crucial in asset valuation is the risk–return relationship. While many studies have investigated the risk–return interaction for other assets such as stocks (see among others, Guo and Nelly, 2008); the housing risk–return relation is somewhat under–researched. Again, starting from the widely studied housing market, the US housing market, Dolde and Tirtiroglu (1997) analysed the US housing risk–return tradeoff on the municipality level for towns in Connecticut and San Francisco area. They found mixed evidence of a significantly positive risk–return effect in San Francisco and a negative effect in Connecticut. Mixed results were also found by Miller and Peng (2006), who employed GARCH and Vector Autoregressions (VAR) models to analyse the impacts between volatility and house price variables at the MSA level. The abovementioned studies were, however, criticised by Milles (2008). The pointed out limitations were that first, Dolde and Tirtiroglu's (1997) investigation was at the municipal level, and the authors did not first formal conduct tests for the evidence of the ARCH effects in the studied areas. Second, Miller and Peng's (2006) study covered the MSA level data; however, Milles (2008) underlined that for investors, real state risk could be at a broader region than just a municipality or metropolitan area. Specifically, the author examined at the state level, the effect of conditional variance on mean returns by using the GARCH–in–mean model. The analysis was done on twenty–eight states, which exhibited ARCH effects; and similar to the above studies, mixed results of positive and negative risk–return impact were found across eight states. Contrary, on the national level, a positive relationship between house price returns and volatility was found by Cannor et al. (2006), who analysed ZIP code–level housing data by using a cross–sectional asset–pricing approach.

In the UK housing market, the return–volatility relationship was investigated by Milles (2011b). By first conducting clustering effect tests in twelve UK regions, the author found that two out of seven regions with significant ARCH effects exhibited significant risk–return impacts; those regions were Wales and East Midlands with positive and negative impacts, respectively. Moreover, Morley and Thomas (2011) found evidence of a positive return–volatility effect in most of the English regions and Wales except for the South West area. These heterogeneous impacts of housing volatility on returns across provinces were also found by Lin and Fuerst (2014) in the Canadian housing market. Their results indicated a positive risk–return relationship in the regions of Ontario and Quebec, while a negative one is observed in British Colombus. More recently, Lee (2017) analysed this issue in the Australian housing market by using an asymmetric Component–GARCH–in–Mean (C–GARCH–M) model and found a strong positive risk–return relationship in the whole Australia as well as in capital cities such as Sydney, Melbourne, Adelaide, Perth, and Brisbane. Lee's (2017) results differ from earlier studies' results by being conclusive and homogenous across the studied areas. The author attributed the difference to the enhancement of the C–GARCH–M model in which the model decomposes the volatility into a short and long–run component and simultaneously incorporates the asymmetric effect of the shocks in the housing market under analysis

Another characteristic of volatility, which is also under–explored in the housing literature, is that of the asymmetric volatility. This volatility phenomenon has been examined in different asset classes, and has two popular explanations; the leverage effect and volatility feedback effect. The former effect follows from Black's (1976) and Christie's (1982) finding that a negative shock has more effect on the equity price volatility increase than a positive shock. The latter effect is based on the observed positive correlations between asset volatility and returns (French et al., 1987; Campbell and Hentschell, 1992; Bekaert and Wu, 2000). In the case of the housing market, the literature is still limited. Similar to the two volatility properties discussed above; the asymmetric effects of the shocks were studied in the US housing market by Milles (2008); in the UK by Tsai and Chen (2009), Milles (2011b), and Morley and Thomas (2011); in the Australian market by Lee (2009) and Lee (2017); in the Canadian market by Lin and Fuerst (2014), and more recently in the Scottish market by Katsiampa and Begiazi (2019). However, Katsiampa and Begiazi's (2019) study differs from the others by considering house prices by property types (detached, semi–detached, terraced, and flats), while the other studies generally consider family–home property type. The US, Australian, and Canadian housing markets seem to be in line in terms of asymmetric effects. In all three housing markets, significant leverage effects were found; implying that the three housing markets are more sensitive to bad news than good news. In the UK housing market, however, results were mixed; on the one hand, no significant leverage effects were observed by Milles (2011b) during the studied sample period. On the other hand, Tsai and Chen (2009) and Morley and Thomas (2011) found respectively significant negative asymmetric effects and little evidence of positive asymmetry in their considered areas.

Collectively, although there is growing research on the housing market volatility, whether the focus is on an investigation of one aspect or a combination of more; the emphasis has been on a limited number of countries. No particular empirical research has been undertaken for the Finnish housing market; even though housing consisted 50.3 per cent of the Finnish households' total wealth (Statistics Finland, 2016). Therefore, this article aims to fill that gap and extend the current literature on the countries' house price volatility analysis by analysing the three volatility characteristics that are commonly investigated in asset valuation; namely, volatility clustering, risk–return relationship, and asymmetric effects. Moreover, this study uses cross–level housing data; that is data on both metropolitan and geographical level, for cross–comparative analysis of the housing investment and portfolio allocation in different Finnish cities and sub–areas. Contrary to the previous studies which employed data on the state, national, regional, or metropolitan level. Furthermore, in the same standpoint of housing investment; this article uses data on apartments in the block of flats property type, which has increased its investors' and consumers' attractiveness in the Finnish residential property sector.

3 Data and Methodology

Data

The study employs the Statistics Finland quarterly house price indices data of fifteen main regions in Finland; throughout 1988:$Q1$ to 2018:$Q4$, for a total of 124 observations. The studied regions are ranked according to their number of inhabitants. There are four regions with more than 250,000 inhabitants: Helsinki, Tampere, Turku, and Oulu; of which the three first make up the so–called growth triangle in Southern Finland, and Oulu is the growth centre of Northern Finland. Seven regions with more than 100,000 inhabitants: Lahti, Jyväskylä, Kuopio, Pori, Seinäjoki, Joensuu, and Vaasa. Four regions with a population number between 80,000 – 90,000: Lappeenranta, Kouvola, Hämeenlinna, and Kotka. These regions are then divided geographically into cities and sub–areas according to their postcode numbers (see Table 8 in Appendix); to form a total of forty–five cities and sub–areas. The considered type of dwellings is apartments (block of flats) because they are the most homogenous assets in the housing market compared to other housing types, such as detached and terraced. Additionally, in Finland, flats are favoured by investors. The apartment types are divided into single–room, two–rooms, and more than three rooms apartments.

Tables 1–3 provide the summary statistics of the quarterly house price returns for single–room, two–rooms, and more than three rooms flats respectively. Note that cities and sub–areas without available data for at least 20 years (80 observations) have been removed from the analysis. Over the studied period, Helsinki–area1 leads the one–room apartments type group with the highest average return (1.16 per cent per quarterly). Kuopio–area1 follows with 1.14 per cent per quarterly average return. Helsinki–area2, Helsinki–city, and Vaasa–area1 come in third place with an average return of at least 1.0 per cent per quarterly. In terms of volatility dimension, Pori–area1 recorded the highest risk measure (standard deviation), followed by Lahti–area1. The largest cities, such as Helsinki and Tampere, as well as Helsinki–area2, appear to be less volatile as they have the lowest risk level; suggesting a less significance of the ARCH effects in these cities and area.

The two–rooms apartments type group appears to have less quarterly average returns, in general; compare to the one–room and more than three rooms flat types. Helsinki–area1 also scores the highest average return (1.18 per cent per quarterly), followed by Helsinki–area2 and Tampere–area1 with at least 1.0 per cent per quarterly average return. Kotka–area2 leads the group in terms of risk measure. Same as in the one–room apartments type group, the biggest cities (Helsinki, Tampere, Turku, and Oulu) and their surrounding areas seem to be less volatile. Helsinki–area1 also comes on top with 1.15 per cent per quarterly average return in the more than three rooms apartments type group, followed by Tampere–area1 and Helsinki–area2. Hämeenlinna–area1, Joensuu–area1, and Seinäjoki–city are the more volatile areas of the group.

The house price movement of a sample of the three most volatile cities/sub–areas in each of the apartments categories over the studied period is shown in Figure 1. Those are Pori–area1, Pori–city, Jyväskylä–area2 in the one–room apartments type group; Kotka–area2, Pori–area1, Kotka–area1 in the two–rooms apartments type group; and Hämeenlinna–area1, Joensuu–area1, Seinäjoki–city in the more than three rooms apartments type group. Initial evidence of volatility clustering effects is observed in all sample cities and sub–areas as they exhibit high fluctuations with certain time periods of high volatility followed by low volatility for other periods. A similar pattern is observed in all the graphs from the end of the 1980s until mid–1993; the period that Finland experienced financial market deregulation which induces a structural break in the house price dynamics (Oikarinen, 2009a; Oikarinen, 2009b).

Cities/Sub–areas	Abbrevations	Mean	Maximum	Minimum	Sd	nobs
Helsinki–city	hki	1.05	10.0	-9.6	3.5	124
Helsinki–area1	hki1	1.16	12.2	-9.1	4.1	124
Helsinki–area2	hki2	1.08	9.2	-9.4	3.6	124
Helsinki–area3	hki3	0.88	11.9	-13.4	4.1	124
Helsinki–area4	hki4	0.68	10.5	-12.8	4.3	124
Tampere–city	tre	0.93	10.9	-11.6	3.9	123
Tampere–area1	tre1	0.99	12.9	-14.9	4.9	123
Tampere–area2	tre2	0.95	14.7	-17.6	5.9	119
Tampere–area3	tre3	0.78	16.2	-12.6	4.9	123
Turku–city	tku	0.88	14.0	-10.1	4.4	124
Turku–area1	tku1	0.95	15.4	-12.5	5.4	124
Turku–area2	tku2	0.77	22.5	-21.5	6.9	111
Turku–area3	tku3	0.79	14.3	-26.1	6.5	114
Oulu–city	oulu	0.69	11.8	-10.8	4.3	124
Oulu–area1	oulu1	0.68	14.9	-12.8	5.1	124
Oulu–area2	oulu2	0.73	15.5	-18.3	5.7	116
Lahti–city	lti	0.66	16.2	-15.6	5.4	124
Lahti–area1	lti1	0.96	36.5	-28.3	7.8	109
Lahti–area2	lti2	0.36	16.5	-21.8	6.1	124
Jyväskylä–city	jkla	0.76	13.4	-10.6	4.7	124
Jyväskylä–area1	jkla1	0.85	14.6	-13.9	5.1	124
Jyväskylä–area2	jkla2	0.86	27.0	-20.4	7.3	91
Pori–city	pori	0.67	22.7	-26.8	7.7	124
Pori–area1	pori1	0.95	28.4	-26.9	8.6	100
Kuopio–city	kuo	0.86	16.5	-12.5	4.3	123
Kuopio–area1	kuo1	1.14	17.3	-20.6	5.9	111
Kuopio–area2	kuo2	0.89	15.4	-18.6	6.7	87
Joensuu–city	jnsu	0.75	15.8	-15.8	5.0	122
Joensuu–area1	jnsu1	0.78	17.2	-15.7	5.4	117
Vaasa–city	vaasa	0.78	14.5	-16.0	6.8	121
Vaasa–area1	vaasa1	1.00	17.3	-17.4	7.6	105
Kouvola–city	kou	0.17	15.3	-16.9	6.8	118
Lappeenranta–city	lrta	0.55	12.6	-13.8	4.9	124
Lappeenranta–area1	lrta1	0.77	17.0	-20.3	6.9	97
Hämeenlinna–city	hnlina	0.68	12.9	-17.0	6.0	124
Hämeenlinna–area1	hnlina1	0.85	12.4	-19.8	6.5	103
Kotka–city	kotka	0.55	16.7	-12.5	5.6	121
Kotka–area1	kotka1	0.87	16.1	-15.4	6.9	95

Notes: This table presents summary statistics on the one–room flats price index returns. Units are quarterly returns in percentage points. The sample is 1988:$Q1$ to 2018:$Q4$.

Table 1: One–room flats quarterly house price returns – Summary statistics (%).

Cities/Sub–areas	Abbrevations	Mean	Maximum	Minimum	Sd	nobs
Helsinki–city	hki	0.96	10.4	-9.1	3.2	124
Helsinki–area1	hki1	1.18	17.3	-15.2	4.7	124
Helsinki–area2	hki2	1.01	10.1	-8.7	3.3	124
Helsinki–area3	hki3	0.81	9.1	-11.3	3.7	124
Helsinki–area4	hki4	0.65	9.2	-10.2	3.6	124
Tampere–city	tre	0.88	10.0	-8.8	3.1	124
Tampere–area1	tre1	1.00	11.9	-8.3	3.6	123
Tampere–area2	tre2	0.75	9.9	-16.8	4.5	123
Tampere–area3	tre3	0.75	10.7	-12.5	3.6	123
Turku–city	tku	0.79	11.0	-8.7	3.4	124
Turku–area1	tku1	0.93	11.9	-12.4	4.1	124
Turku–area2	tku2	0.70	9.9	-13.0	4.6	124
Turku–area3	tku3	0.66	13.5	-8.7	4.7	124
Oulu–city	oulu	0.65	8.8	-6.4	3.1	124
Oulu–area1	oulu1	0.67	10.7	-7.1	3.5	124
Oulu–area2	oulu2	0.58	11.3	-10.3	4.2	124
Lahti–city	lti	0.52	10.4	-8.8	3.4	124
Lahti–area1	lti1	0.68	12.4	-11.4	4.5	124
Lahti–area2	lti2	0.34	11.2	-7.7	3.9	124
Jyväskylä–city	jkla	0.61	9.0	-8.9	3.3	124
Jyväskylä–area1	jkla1	0.71	11.7	-10.1	3.8	124
Jyväskylä–area2	jkla2	0.44	18.6	-20.6	4.7	124
Pori–city	pori	0.72	20.3	-16.8	5.2	124
Pori–area1	pori1	0.76	22.0	-19.1	6.3	124
Pori–area2	pori2	0.64	16.9	-17.3	6.3	122
Kuopio–city	kuo	0.69	12.1	-13.1	3.5	124
Kuopio–area1	kuo1	0.84	15.4	-16.9	4.8	123
Kuopio–area2	kuo2	0.53	10.9	-9.7	3.7	124
Joensuu–city	jnsu	0.64	13.8	-12.6	4.9	124
Joensuu–area1	jnsu1	0.64	15.5	-13.7	5.6	124
Seinajöki–city	seoki	0.70	18.9	-15.6	5.9	118
Vaasa–city	vaasa	0.70	9.6	-9.0	4.0	123
Vaasa–area1	vaasa1	0.78	9.8	-9.8	4.3	121
Kouvoula–city	kou	0.27	23.9	-19.9	5.5	124
Lappeenranta–city	lrta	0.52	14.9	-11.5	3.9	124
Lappeenranta–area1	lrta1	0.57	19.1	-17.1	5.4	123
Lappeenranta–area2	lrta2	0.46	18.9	-19.8	5.7	122
Hämeenlinna–city	hnlina	0.63	11.4	-15.6	4.5	124
Hämeenlinna–area1	hnlina1	0.62	13.1	-18.4	5.2	124
Kotka–city	kotka	0.58	13.2	-11.0	5.0	124
Kotka–area1	kotka1	0.70	16.7	-17.5	6.3	121
Kotka–area2	kotka2	0.52	19.4	-27.0	8.1	96

Notes: This table presents summary statistics on the two–rooms flats price index returns. Units are quarterly returns in percentage points. The sample is 1988:$Q1$ to 2018:$Q4$.

Table 2: Two–rooms flats quarterly house price returns – Summary statistics (%).

Cities/Sub–areas	Abbrevations	Mean	Maximum	Minimum	Sd	nobs
Helsinki–city	hki	0.94	12.2	-10.2	3.6	124
Helsinki–area1	hki1	1.15	14.2	-15.6	5.1	124
Helsinki–area2	hki2	0.95	13.0	-8.9	3.7	124
Helsinki–area3	hki3	0.75	11.7	-9.5	3.9	124
Helsinki–area4	hki4	0.65	11.9	-11.8	3.8	124
Tampere–city	tre	0.87	11.0	-12.4	3.7	123
Tampere–area1	tre1	0.98	14.0	-15.8	4.7	123
Tampere–area2	tre2	0.91	11.7	-15.3	5.5	116
Tampere–area3	tre3	0.66	12.5	-12.8	3.5	123
Turku–city	tku	0.77	12.6	-10.8	3.9	124
Turku–area1	tku1	0.93	15.5	-17.2	5.3	124
Turku–area2	tku2	0.71	15.3	-16.0	4.9	124
Turku–area3	tku3	0.66	11.8	-11.1	4.5	124
Oulu–city	oulu	0.70	12.3	-11.0	3.7	124
Oulu–area1	oulu1	0.70	14.2	-15.7	4.6	123
Oulu–area2	oulu2	0.70	10.1	-14.4	4.5	123
Lahti–city	lti	0.56	11.6	-12.2	4.4	124
Lahti–area1	lti1	0.68	15.7	-14.9	5.6	124
Lahti–area2	lti2	0.41	10.1	-11.7	4.5	124
Jyväskylä–city	jkla	0.62	14.1	-9.8	4.4	124
Jyväskylä–area1	jkla1	0.66	15.7	-12.9	5.0	122
Jyväskylä–area2	jkla2	0.59	17.9	-19.1	6.3	122
Pori–city	pori	0.72	15.3	-18.3	5.8	124
Pori–area1	pori1	0.80	16.7	-20.2	6.6	116
Kuopio–city	kuo	0.59	13.6	-15.7	4.4	124
Kuopio–area1	kuo1	0.74	15.2	-32.7	7.1	115
Kuopio–area2	kuo2	0.49	15.4	-20.5	4.9	122
Joensuu–city	jnsu	0.66	17.6	-20.1	6.2	124
Joensuu–area1	jnsu1	0.72	20.4	-22.0	7.2	108
Seinajöki–city	seoki	0.80	24.3	-27.7	7.2	103
Vaasa–city	vaasa	0.68	14.7	-16.7	5.1	123
Vaasa–area1	vaasa1	0.79	17.5	-14.9	5.9	116
Vaasa–area2	vaasa2	0.84	13.8	-22.9	7.0	82
Kouvoula–city	kou	0.14	14.1	-15.0	6.7	121
Lappeenranta–city	lrta	0.45	11.9	-17.1	5.5	121
Lappeenranta–area2	lrta2	0.89	21.2	-25.1	7.0	80
Hämeenlinna–city	hnlina	0.61	19.9	-17.1	6.1	122
Hämeenlinna–area1	hnlina1	0.70	24.3	-17.7	7.3	108
Kotka–city	kotka	0.49	19.6	-19.2	6.4	120

Notes: This table presents summary statistics on the more than three rooms flats price index returns. Units are quarterly returns in percentage points. The sample is 1988:$Q1$ to 2018:$Q4$.

Table 3: More than three rooms flats quarterly house price returns – Summary statistics (%).

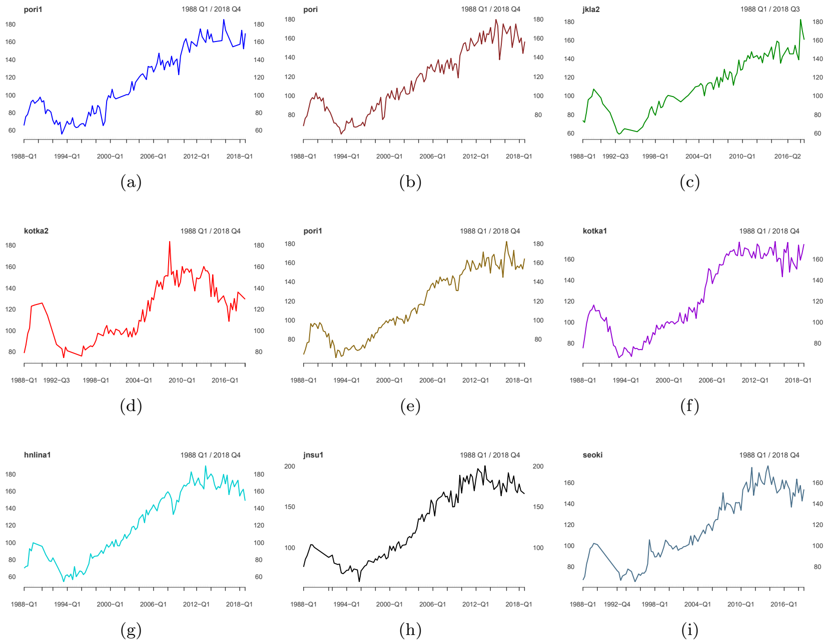


Figure 1: The house price movement of the most volatile cities/sub–areas.

Methodology

The methodology employed in this study is presented as follows: first, we define returns as log differences of the price indices in percentages (multiplied by 100) for each city and sub–area in each of the apartment type. Thereafter we filter first order autocorrelations from the returns with an ARMA model of appropriate order determined by Akaike information criteria (AIC) and Bayesian information criteria (BIC). Next, we test ARCH effects on the ARMA filtered returns. Then, for cities and sub–areas with significant clustering effects, individual GARCH models are used to model the nature of the time–varying volatility. Moreover, GARCH–in–mean (GARCH–M) model is employed to investigate the relationship between housing returns and volatility. Lastly, the Exponential GARCH (EGARCH) model is used to analyse the asymmetric effects of the shocks. Although the Glosten, Jagannthan, and Runkle GARCH (GJR–GARCH) model was also employed; the EGARCH proved to be effective in modelling the studied house prices. All analysis was conducted in *R* (R Core Team, 2019).

Testing for ARCH effects

After filtering serial correlations from the returns series, the squared residual series are used to check the autoregressive conditional heteroscedasticity, also known as ARCH effects. If the null hypothesis of constant variance is rejected, then volatility modelling is required.

Two tests are available. The first test, called Portmanteau *Q(m)*, is to examine whether the squares of the residuals are a sequence of white noise. It is the usual Ljung–Box test on the squared residuals (see Mcleod and Li, 1983). The null hypothesis of the test statistic is that "there is no autocorrelation in the squared residuals up to lag m;" that is, the first m lags of the autocorrelation function (ACF) of the squared residuals are zeros. A small p–value (smaller than the considered critical value) suggests the presence of autoregressive conditional heteroscedasticity (strong ARCH effects).

The second test is the Lagrange Multiplier test of Engle (1982), also known as ARCH–LM Engle's test. This test is to fit a linear regression model for the squared residuals and examine that the fitted model is significant. It is equivalent to the usual F statistic for testing $\gamma_i = 0$ $(i = 1, ..., m)$ in the linear regression

$$\hat{e}_t^2 = \gamma_0 + \gamma_1 \hat{e}_{t-1}^2 + ... + \gamma_m \hat{e}_{t-m}^2 + v_t, \quad t = m+1, ...N,$$

where $\hat{e}_t^2$ is the estimated residuals, v_t is the random error, m is a prespecified positive integer, and N is the sample size. The null hypothesis of the test is that "there are no ARCH effects;" that is, $H_0 : \gamma_1 = ... = \gamma_m = 0$, and the alternative hypothesis is $H_1 : \gamma_i \neq 0$ (there are ARCH effects). Again, the null hypothesis is rejected if a p–value smaller than the considered critical value is obtained at the specified number of lags. The ARCH–LM tests were performed using the function ArchTest() from the *FinTs* package (Graves, 2019).

Volatility modelling

The dynamics of the time–varying conditional variance of the house price returns of those cities and sub–areas with significant ARCH effects are analysed using the Generalized Autoregressive Conditional Heteroscedasticity (GARCH)–type models.

Let R_t be the log–return of an asset at time t. The conditional volatility model is of the following form:

$$R_t = u_t + e_t, \quad e_t \sim \mathcal{N}(0, \sigma_t^2),$$
$$\sigma_t^2 = \omega + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + ... + \alpha_q e_{t-q}^2,$$

where u_t is the conditional mean, σ_t is the conditional standard deviation, e_t is the error term, and $\omega > 0$ is the intercept.

In the GARCH model proposed by Bollerslev (1986), the conditional variance σ_t^2 depends on previous squared errors and past volatility. That is,

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i e_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

where $\alpha_i \geq 0$ (coefficients of e_{t-i}) are referred to as the ARCH parameters and $\beta_j \geq 0$ (coefficients of σ_{t-j}^2) as the GARCH parameters.

It is worth mentioning that it is not plausible to assume a normal distribution when analysing asset returns with GARCH–type models. A realistic distribution needs to accommodate asset returns characteristics such as fat tails and skewness. Thus, univariate distributions such as Generalized Error ('GED'), Student t ('Std'), and their skew variants ('sGED', 'sStd') can be used. For more details, see Tsay (2013).

The estimation of individual GARCH models were performed using the *rugarch* package (Ghalanos, 2019), and followed the following steps:

- Model specification by ugarchspec() function,

- Model estimation by ugarchfit() function,
- Model adequacy checking using the standardised residuals tests.

Regarding the distribution of the error term (e_t) for each city and sub–area; based on Akaike information criteria (AIC) and Bayesian information criteria (BIC), the selected distributions are as follows:

- In the one–room flats category, nine cities/sub–areas follow a normal distribution. Four follow a Generalized Error Distribution ("GED"). Three follow a Student t distribution ("Std"). Eleven follow a skew normal distribution ("snorm"), and one follows a skew Generalized Error Distribution ("sGED").
- In the two–rooms flats category, fourteen cities/sub–areas follow a normal distribution. Three follow a GED; six follow a snorm; three follow a sGED, and one follows a skew Student t distribution ("sStd").
- In the more than three rooms flats category, eighteen cities/sub–areas follow a normal distribution. Four follows a GED; three follow a Std; five follow a snorm, and one follows a sGED.

Testing for GARCH–M effects

In addition to testing and estimating GARCH models, and for modelling and forecasting purposes; there are other crucial volatility issues to note, that GARCH models are employed to investigate. One is the determination of the potential impact that the conditional variance may have on returns. This phenomenon was initially pointed out by Engle et al. (1987); the authors developed an innovation GARCH model known as GARCH–in–mean (GARCH–M) model and applied to the Treasury bond returns.
The model setup consists of specifying the conditional mean as a linear function of the conditional variance. That is,

$$R_t = u_t + \lambda\sigma_t^2 + e_t, \quad e_t \sim \mathcal{N}(0,\, \sigma_t^2),$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i e_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

where the parameter λ describes the nature of the relationship between asset return and volatility. A significant positive λ implies that the asset return is positively related to its volatility. In other words, an increased risk leads to an increased mean return. This technique has been applied to other asset markets, such as stocks and equities. However, even though intuitively a positive λ is expected; studies such as Glosten et al. (1993) highlighted reasons that could indicate a negative λ, as there is no theoretical restriction on the sign of λ. Thus, different signs of λ have been documented in different markets.
In the housing market, the GARCH–M model has been applied by Milles (2008) to the US market and Milles (2011b) to the UK market; mixed results of positive and negative risk–return impact were found across the studied areas. Following these studies, this article employs the GARCH–M model to investigate the relationship between housing returns and volatility in the Finnish market. Moreover, an individual estimation is done on each city and sub–area to allow the impact of volatility on returns to differ from region to region and by apartment type.

Testing for Asymmetric volatility

Another issue in GARCH modelling is the presence of asymmetric volatility. That is the possibility of asymmetric effects of shocks on conditional variance. An observation that a negative return gives rise to subsequent volatility than a positive return is termed the leverage effect; while the contradicting observation, the volatility feedback effect is based on the observed positive correlations between asset volatility and returns. Two GARCH models are mostly used to examine these effects; those are the Glosten, Jagannthan, and Runkle GARCH (GJR–GARCH) model by Glosten et al. (1993) and the Exponential GARCH (EGARCH) model by Nelson (1991).

The former model allows negative shocks to have a different effect on volatility than positive shocks, and it is specified as follows:

$$R_t = u_t + e_t, \quad e_t \sim \mathcal{N}(0,\, \sigma_t^2),$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q}(\alpha_i + \gamma_i I_{t-i})e_{t-i}^2 + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^2,$$

where α_i and $\alpha_i+\gamma_i$ represent the impact of good news and bad news respectively, and the indicator function I_{t-i} is equal to one if $e_{t-1} < 0$ and zero otherwise. That is, a negative shock e_{t-i} has a more significant impact $(\alpha_i+\gamma_i)e_{t-i}^2$ with $\gamma_i > 0$, whereas a positive shock e_{t-i} have less effect $\alpha_i e_{t-i}^2$ to σ_t^2. An estimated $\gamma \neq 0$ implies asymmetry, and leverage effects exist if $\gamma > 0$, meaning that if γ is positive significant, negative innovations affect volatility than positive shocks.
The latter model applies the logged conditional variance so that the volatility is always positive. Its specification is as follows:

$$R_t = u_t + e_t, \quad e_t \sim \mathcal{N}(0,\, \sigma_t^2),$$

$$\log(\sigma_t^2) = \omega + \sum_{i=1}^{q}(\alpha_i|z_{t-i}| + \gamma_i z_{t-i}) + \sum_{j=1}^{p}\beta_j \log\sigma_{t-j}^2,$$

where $z_t = e_t/\sqrt{\sigma_t^2}$ is the standardized shock. Again, $\gamma \neq 0$ implies an asymmetric impact. However, leverage effect exist if $\gamma < 0$; following Black (1976)'s argument that negative shocks have more effect on the rise of the conditional volatility than positive innovations of equal magnitude.

In the housing market, both GJR–GARCH and EGARCH models have been employed in different markets to investigate asymmetric volatility. The GJR–GARCH model has been used by Milles (2008) in the US; Tsai and Chen (2009) and Milles (2011b) in the UK market. The EGARCH model which is the widely used one has been employed by Morley and Thomas (2011) in the UK; Lee (2009) in Australian; and Lin and Fuerst (2014) in the Canadian market. The reason behind the acknowledgement of the EGARCH model is the evidence of its good performance in stock and real estate markets provided by Engle and Ng (1993) and Stevenson (2002). More recently Katsiampa and Begiazi (2019) applied both models on the Scottish housing market. To that end, and for models' estimation comparison purposes; this study also employs both GJR–GARCH and EGARCH models to examine the presence of asymmetric volatility in the Finnish housing market. Same as the above–cited studies, the EGARCH model proved to be effective in modelling the studied house prices; confirming once more its good performance (GJR–GARCH model's results available from the author upon request). Again, an individual estimation is done for each city and sub–area in each apartment type to allow different effects. Following Lee

(2009), high–order models are compared to the EGARCH(1,1) model using Akaike information criteria (AIC) and Bayesian information criteria (BIC) to determine the optimal specification for an EGARCH model for each city and sub–area.

4 Results and discussions

Testing for ARCH effects

Tables 4 and 5 display the p–values of the two tests employed to investigate whether there is volatility clustering in each housing return series. Those tests are the Ljung–Box (LB) test and the Engle's Lagrange Multiplier (LM) test. The null hypothesizes of no serial correlation in squared residuals and no ARCH effects, of the LB test and LM test respectively are rejected in twenty–eight out of thirty–eight studied cities and sub–areas in the one–room flats category; in twenty–seven out forty–two in the two–rooms flats category; and in thirty–one out of thirty–nine in the more than three rooms flats category. Thus, strong evidence of volatility clustering (ARCH) effects is evident in over half of the cities and sub–areas in all three apartments types; suggesting that the Finnish house price volatility contains critical information that should be addressed by housing investors, buyers, and policymakers to assess the risk of housing investment.

In some cases, one of the tests is inconclusive, for instance, in the case of Tampere–area1 (in the one–room flats category) and Turku–area2 (in the two–rooms flats category), the Ljung–Box test is inconclusive (we fail to reject the null hypothesis because of the higher p–values); however, the Lagrange Multiplier values are statistically significant. Similarly, in the case of Lahti–area1 (in the more than three rooms flats category), this time, however, it is the Lagrange Multiplier test, which is inconclusive. In these cases, the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the squared residuals (figures available from the author upon request) may be used to demonstrate that there might be some autocorrelations left even though the significance might be low.

The evidence of the volatility clustering effect found in this study is stronger compared to other housing markets such as the US and the UK, where Milles (2008) and Milles (2011b) respectively documented less significant ARCH effects. More precisely, twenty–eight out of fifty US states and seven out of twelve UK regions were found to exhibit ARCH effects. Strong ARCH effects were also found in the Australian housing market by Lee (2017), where all the studied cities displayed time–varying volatility; however, this strong effect was attributed to the use of higher frequency data (daily data), while other studies employed quarterly or monthly data. Moreover, we found a little pattern to which cities and sub–areas had ARCH effects and which did not. Regions such us Oulu–area2, Lahti–area2, and Joensuu–city lack volatility clustering in all three apartment types, while no clustering effects were discovered in Tampere–city and Vaasa–area1 in both one–room and two–rooms apartment types. Furthermore, densely populated regions such as Helsinki as well as thinly populated cities such as Kotka were found to exhibit volatility clustering effects. An observation which is, on the one hand, in contrast with Lin and Fuerst's (2014) claim that the density of the population in the region may be the source of volatility clustering. On the other hand, our results are consistent with Lee's (2017) analysis outcome, who found significant ARCH effects even in small Australian cities such as Adelaide. Therefore, the identification of possible sources of volatility is the subject of further research.

In summary, ARCH effects discovered in different cities and sub–areas in all three apartment types, suggest that a housing portfolio concentrated in one or several regions may be risky; however, diversifying a portfolio across Finland and by apartment type can lower risk. This finding is important as many investors are often highly concentrated in narrow geographical regions such as Helsinki.

		One room flats				Two rooms flats			
Regions	Cities/sub–areas	ARMA	LB	LM	ARCH?	ARMA	LB	LM	ARCH?
Helsinki	hki	ARMA(2,1)	0.0204**	0.0221**	yes	ARMA(2,1)	0.0004***	0.0005***	yes
	hki1	MA(2)	0.0296**	0.0317**	yes	ARMA(2,1)	0.0443**	0.0186**	yes
	hki2	ARMA(2,1)	0.0162**	0.0193**	yes	AR(1)	0.0025**	0.0029**	yes
	hki3	ARMA(2,1)	0.5206	0.4721	no	AR(2)	0.0256**	0.0319**	yes
	hki4	AR(2)	0.0255**	0.0305**	yes	ARMA(1,1)	0.0921*	0.0905*	yes
Tampere	tre	ARMA(1,1)	0.3557	0.8176	no	ARMA(2,1)	0.829	0.831	no
	tre1	ARMA(2,2)	0.4351	0.0105**	yes	AR(2)	0.0315**	0.0752*	yes
	tre2	ARMA(1,1)	0.978	0.7062	no	ARMA(0,0)	0.0051**	0.0108**	yes
	tre3	AR(2)	0.0913*	0.0258**	yes	ARMA(2,2)	0.2858	0.2683	no
Turku	tku	ARMA(2,2)	0.0039**	0.0126**	yes	ARMA(2,2)	0.0432**	0.0494**	yes
	tku1	ARMA(1,1)	0.0369**	0.0179**	yes	AR(2)	0.5196	0.5257	no
	tku2	AR(1)	0.0635*	0.0681*	yes	ARMA(0,0)	0.2146	0.0997*	yes
	tku3	AR(1)	0.2593	0.0019**	yes	MA(3)	0.1846	0.187	no
Oulu	oulu	ARMA(1,1)	0.0116**	0.0111**	yes	AR(2)	0.333	0.3378	no
	oulu1	AR(1)	0.0716*	0.0754*	yes	ARMA(1,2)	0.5143	0.5191	no
	oulu2	AR(1)	0.1065	0.1049	no	ARMA(0,0)	0.1008	0.7517	no
Lahti	lti	AR(2)	0.0092**	0.0099**	yes	AR(2)	0.0882*	0.0881*	yes
	lti1	AR(1)	0.0026**	0.0031**	yes	AR(2)	0.5808	0.5838	no
	lti2	AR(1)	0.6409	0.5526	no	ARMA(1,2)	0.989	0.9995	no
Jyväskylä	jkla	ARMA(1,1)	0.0296**	0.0111**	yes	ARMA(2,2)	0.0108**	0.007**	yes
	jkla1	ARMA(1,1)	0.0255**	0.0075**	yes	MA(3)	0.0123**	0.0125**	yes
	jkla2	ARMA(0,0)	0.3453	0.0326**	yes	ARMA(1,2)	$1.22{*}10^{-5}$***	$1.43{*}10^{-5}$***	yes
Pori	pori	MA(1)	0.0257**	0.0323**	yes	MA(3)	0.0027**	0.0017**	yes
	pori1	AR(2)	0.0522*	0.0699*	yes	MA(3)	0.0124**	0.0126**	yes
	pori2	–	–	–	–	ARMA(2,2)	0.0568*	0.0589*	yes
Kuopio	kuo	ARMA(0,0)	0.0324**	0.0305**	yes	AR(2)	0.0224**	0.0195**	yes
	kuo1	MA(2)	0.0842*	0.0891*	yes	ARMA(0,0)	0.0030**	0.0017**	yes
	kuo2	ARMA(0,0)	0.0167**	0.0191**	yes	AR(2)	0.6987	0.69	no
Joensuu	jnsu	MA(3)	0.9751	0.961	no	AR(3)	0.3522	0.3578	no
	jnsu1	MA(3)	0.0502**	0.0457**	yes	AR(3)	0.0424**	0.1197	yes
Seinäjoki	seoki	–	–	–	–	AR(1)	0.0109**	0.0103**	yes
Vaasa	vaasa	MA(1)	0.8133	0.8154	no	ARMA(1,2)	0.0634*	0.0668*	yes
	vaasa1	MA(1)	0.9039	0.9057	no	MA(2)	0.3136	0.3199	no
Kouvola	kou	AR(1)	0.0161**	0.0105**	yes	ARMA(1,2)	0.0012**	0.0013**	yes
Lappeenranta	lrta	AR(1)	0.0548*	0.0581*	yes	MA(3)	0.0009***	0.0011**	yes
	lrta1	MA(1)	0.0086**	0.0361**	yes	ARMA(2,2)	0.0021**	0.0025**	yes
	lrta2	–	–	–	–	AR(1)	0.9971	0.997	no
Hämeenlinna	hnlina	MA(3)	0.0032**	0.0038**	yes	ARMA(0,0)	0.0703*	0.0700*	yes
	hnlina1	MA(3)	0.8936	0.8955	no	ARMA(1,2)	0.0168**	0.0166**	yes
Kotka	kotka	MA(1)	0.0304**	0.0244**	yes	MA(3)	0.1588	0.1269	no
	kotka1	MA(3)	0.8673	0.8621	no	MA(2)	0.5307	0.0577*	yes
	kotka2	–	–	–	–	MA(2)	0.1854	0.194	no

Notes: This table reports the ARMA model for each city and sub–area, and the p–values from the Ljung–Box and Lagrange Multiplier tests. *, **, and *** represent respectively 10%, 5%, and 1% levels of significance. "yes" indicates that a city/sub–area exhibits ARCH effects, "no" means that a city/sub–area does not.

Table 4: ARCH effects tests results.

		Three rooms flats			
Regions	Cities/sub–areas	ARMA	LB	LM	ARCH?
Helsinki	hki	AR(1)	0.0063**	0.0037**	yes
	hki1	AR(2)	0.0034**	0.0038**	yes
	hki2	AR(1)	0.6742	0.6784	no
	hki3	AR(2)	0.0079**	0.0524*	yes
	hki4	AR(2)	0.0366**	0.0382**	yes
Tampere	tre	ARMA(2,2)	0.0386**	0.0371**	yes
	tre1	ARMA(2,2)	0.0346**	0.0362**	yes
	tre2	ARMA(2,2)	0.0301**	0.0329**	yes
	tre3	ARMA(1,1)	0.0911*	0.0238**	yes
Turku	tku	ARMA(2,2)	0.0255**	0.0244**	yes
	tku1	AR(1)	0.0596*	0.0599*	yes
	tku2	ARMA(2,2)	0.0784*	0.0825*	yes
	tku3	ARMA(0,0)	0.0009***	0.0011**	yes
Oulu	oulu	ARMA(1, 2)	0.0172**	0.0249**	yes
	oulu1	ARMA(1,2)	0.0465**	0.0745*	yes
	oulu2	MA(3)	0.4018	0.407	no
Lahti	lti	ARMA(2,2)	0.0034**	0.0034**	yes
	lti1	MA(3)	0.0141**	0.2274	yes
	lti2	ARMA(2,2)	0.9306	0.9865	no
Jyväskylä	jkla	ARMA(1,2)	0.0393**	0.0423**	yes
	jkla1	ARMA(2,2)	0.0369**	0.0859*	yes
	jkla2	ARMA(1,2)	0.0199**	0.0216**	yes
Pori	pori	ARMA(2,2)	0.75	0.7537	no
	pori1	MA(1)	0.0312**	0.0338**	yes
Kuopio	kuo	ARMA(0,0)	$3.25{*}10^{-5}$***	$3.72{*}10^{-5}$***	yes
	kuo1	MA(1)	0.0506*	0.0544*	yes
	kuo2	ARMA(1,2)	0.9288	0.0023**	yes
Joensuu	jnsu	AR(1)	0.7833	0.7856	no
	jnsu1	AR(1)	0.2663	0.3491	no
Seinäjoki	seoki	MA(3)	0.0281**	0.0130**	yes
Vaasa	vaasa	ARMA(1,2)	0.0938*	0.0990*	yes
	vaasa1	MA(1)	0.0156**	0.0173**	yes
	vaasa2	ARMA(0,0)	0.0038**	0.0094**	yes
Kouvola	kou	MA(3)	0.8182	0.8207	no
Lappeenranta	lrta	MA(3)	0.0645*	0.0680*	yes
	lrta2	ARMA(0,0)	0.0004***	0.0006***	yes
Hämeenlinna	hnlina	MA(3)	0.9513	0.9519	no
	hnlina1	AR(1)	0.0031**	0.0009***	yes
Kotka	kotka	ARMA(2,2)	0.0604*	0.0498*	yes

Notes: This table reports the ARMA model for each city and sub–area, and the p–values from the Ljung–Box and Lagrange Multiplier tests. *, **, and *** represent respectively 10%, 5%, and 1% levels of significance. ”yes” indicates that a city/sub–area exhibits ARCH effects, ”no” means that a city/sub–area does not.

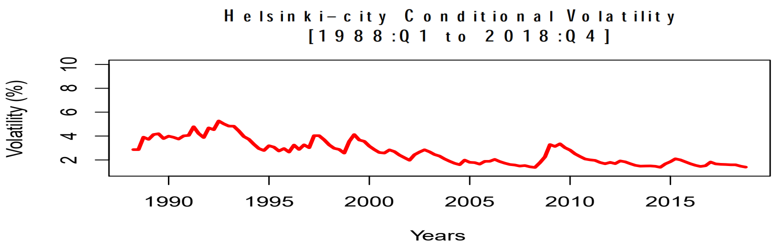
Table 5: ARCH effects tests results – Continued.

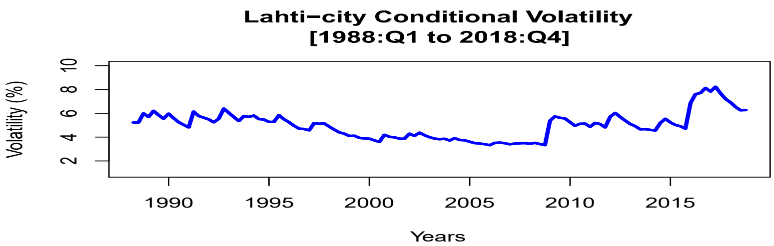
Volatility modelling

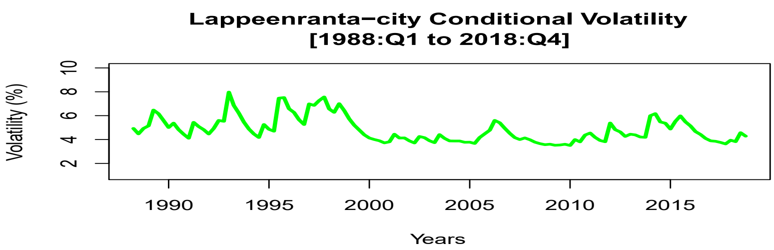
For cities and sub–areas displaying significant volatility clustering, the historical conditional volatility dynamic is investigated from individuals GARCH models, to allow different regions volatility behaviour, and analyse how these regions adjust to the financial shocks.

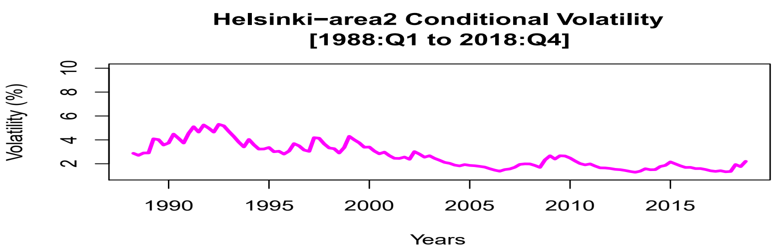
Based on the number of inhabitants, a sample of three cities/sub–areas over the studied period is drawn in each apartment category and shown in Figure 2. Those are Helsinki, Lahti, and Lappeenranta cities in the one–room flats category. Helsinki–area2, Vaasa–city, and Kotka–area1 in the two–rooms flats category. Tampere–city, Kuopio–are2, and Hämeenlinna–area1 in the more than three rooms flats category. As it can be observed in almost all sample regions, the conditional variance was high at the end of the 1980s due to the housing boom that Finland experienced between 1987 and mid–1989; mainly as a result of the financial market deregulation which induces house prices to rise to more than 60 per cent (André and García, 2012). After the bursting of the housing bubble, the Finnish housing market experienced some downturns, upturns, and some steady trend in some regions until around mid–1993 where the extreme volatility swing is observed in almost all the cities/sub–areas especially in Lappeenranta and Vaasa. The moderate growth path noticed in the Finnish housing market in the 1990s is mainly due to the reformation of the state support for housing done around the same time; with the creation of the Housing Fund of Finland (ARA), and Finland's entry into the European Union which causes the drop of inflation (Kivistö, 2012). This moderate housing trend can be very well observed in all three samples from the more than three rooms apartment type after the year 1995; meaning that the house prices of this flat category have been somehow steady, no extremely volatility swings.

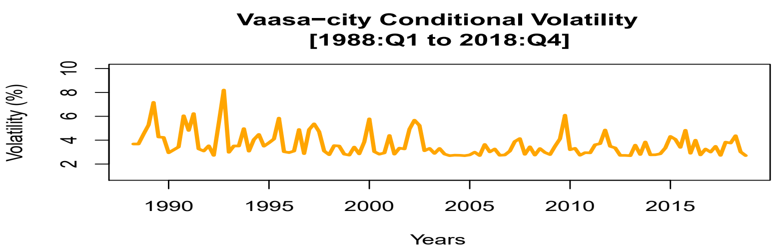
An increase in conditional volatility is observed during the Global Financial Crisis from 2007 to 2008; however, the adjustment to such financial shock differed across city and sub–area market. In the Helsinki and Lahti city, the highest housing volatility occurred in 2008, while in Lappeenranta–city, it began in 2007. In Vaasa–city, the most significant swing was observed in 2009. For Lahti–city and Kotka–area1, the peak level of conditional volatility is at the end of the sample around the year 2016–2017.

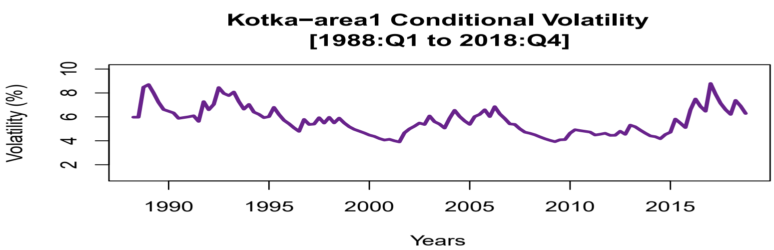





Lappeenranta−city Conditional Volatility
[1988:Q1 to 2018:Q4]
Volatility (%)
2
4
6
8
10
1990
1995
2000
2005
2010
2015
Years


Helsinki−area2 Conditional Volatility
[1988:Q1 to 2018:Q4]
Volatility (%)
2
4
6
8
10
1990
1995
2000
2005
2010
2015
Years


Vaasa−city Conditional Volatility
[1988:Q1 to 2018:Q4]
Volatility (%)
2
4
6
8
10
1990
1995
2000
2005
2010
2015
Years


Kotka−area1 Conditional Volatility
[1988:Q1 to 2018:Q4]
Volatility (%)
2
4
6
8
10
1990
1995
2000
2005
2010
2015
Years

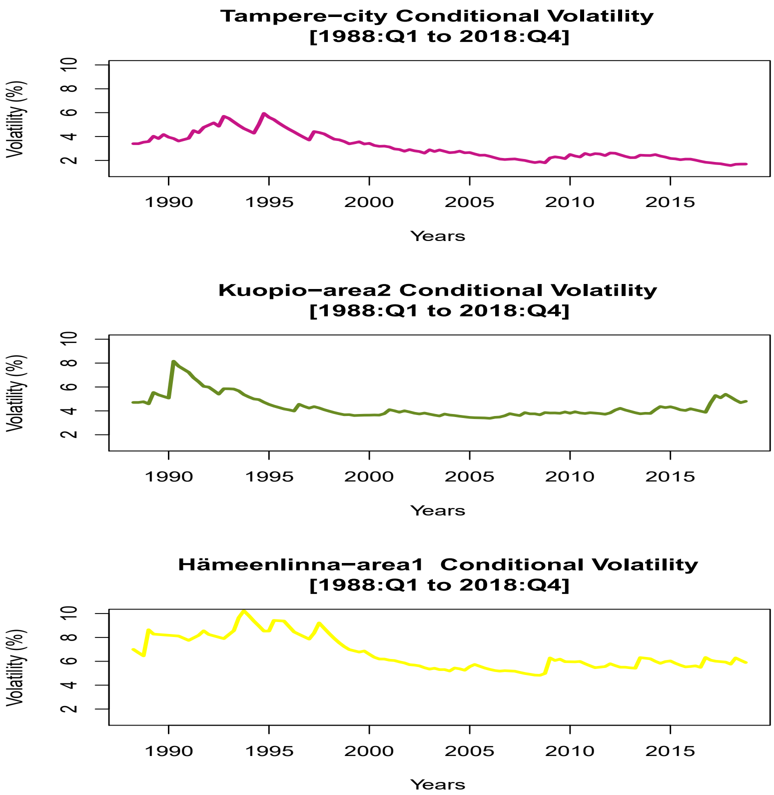


Figure 2: Conditional variance of housing volatility – Sample cities/sub–areas.

Testing for GARCH–M effects

Next, an analysis is conducted to examine the potential impact that the house price volatility may have on house price returns by estimating individual GARCH–M models for cities and sub–areas with significant ARCH effects.

Table 6 presents the λ estimated coefficients with their standard errors (in parentheses). Mixed results on the sign of the significant risk–return impact are observed across cities and sub–areas in all three apartment types. Plus precisely, in the one–room flats category, out of twenty–eight cities/sub–areas exhibiting clustering effects, twenty–two have GARCH–M effects: twelve areas with a positive sign and ten with a negative sign. In the two–rooms flats category, out of twenty–seven cities/sub–areas with significant ARCH effects, nineteen have GARCH–M effects: fifteen areas with a positive sign and four areas with a negative sign. In the more than three rooms flats category, out of thirty–one cities/sub–areas exhibiting volatility clustering, twenty–five have GARCH–M effects: seventeen areas with a positive sign and eight areas with a negative sign. Note that no GARCH–M model could be fitted which yielded convergent estimates for Tutku–city in the two–rooms flats category.

As discussed earlier, a significant positive λ implies that an increased risk leads to an increased mean return. Results confirm that there is evidence of a positive risk–return relationship in the regions of Oulu, Pori, Joensuu, Lappeenranta, and Kotka in the one–room flats category. In the two–rooms flats category, the risk–return impact is positive in the regions of Lahti, Pori, Kuopio, Joensuu, Kovoula, and Lappeenranta. In the more than three rooms flats, the risk–return relationship is positive in the regions of Oulu, Jyväskylä, Seinäjoki, Lappeenranta, and Hämeenlinna. In all three apartment types, the risk–return impact is positive in Lahti and Lappeenranta cities. These results imply that these regions' investors are risk–averse and would require a higher return as a reward for the increased risk. For densely populated areas such as Helsinki, Tampere, and Turku, their housing citywide and sub–markets display mostly a negative risk–return effect; indicating that a higher conditional variance lowers returns in these regions. In other words, in these areas, periods with higher volatility may cause investors to believe that the future is riskier, which would give rise to higher savings and lower the risk premium.

These results are in line with the findings of Milles (2008), Milles (2011b), Morley and Thomas (2011), and Lin and Fuerst (2014), who documented mixed results on the sign of the risk–return relationship in the US, UK, and Canadian housing markets, respectively. Moreover, the results in most of the cities and sub–areas, especially in the two–rooms apartments type support Merton's (1973) hypothesis that investors (in this case Finnish housing investors) would demand a higher return to compensate the high risk. Furthermore, in some regions, especially in the one–room flats category, the results provide support to Glosten et al.'s (1993) argument that during volatile periods, investors may accept a lower return as the future would be more riskier. Lee (2017) did not find evidence to support the latter argument in the Australian housing market. The author argued that as the Australian housing market is dominated by home buyers and home owner–occupiers instead of housing investors (RBA, 2015); if the market is too volatile, investors may not accept a lower return. Instead, they may hold off their investment plans; the reason behind the strong positive relationship between risk and return in the whole Australian housing market.

However, this is not the case for the Finnish housing market, as discussed above. The plausible reason could be the fact that the Finnish housing market is dominated by investors; foreign as well as domestic. According to the Finnish Landlord Association's (2017) survey, Finnish residential property investments are no longer exclusively for professional investors. Ordinary people in Finland, especially younger residents, are increasingly buying flats as an investment rather than consumption (a place to live). Among the interviewed, 63 per cent own one or two apartments for rent; 33 per cent own three to ten apartments, and over three per cent own more than ten flats. Thus, when the housing market is highly volatile in one region, for instance, Kouvola area; the investor may accept a lower return as the future would be riskier, to invest in a positive risk–return market region such as Pori. Hence, the evidence of a negative risk–return relationship in some Finnish areas. Furthermore, these negative risk–return relationship results can also be supported by Hibbert et al.'s (2008) findings that in the stock market, the main explanation of the negative risk-return relation is the behaviour of traders (heterogeneous preferences).

In summary, the two–rooms apartments type have more cities/sub–areas with a positive housing risk–return relationship. Only four areas have a negative one; those are Helsinki–area2, Helsinki–area3, Tampere–area1, and Jyväskylä–area1. This finding suggests that the two–rooms flats investors would require a higher return to compensate for the high risk.

Moreover, there is no evidence of a geographical pattern with regards to the regions with a positive or a negative risk–return relationship; mixed results in the sign is observed across regions in all three apartments types. A finding which confirms, again as above, that a housing portfolio concentrated in one or several areas may be risky; however, diversifying a portfolio across Finland and by apartment type can lower risk.

Regions	Cities/Sub–areas	One room flats		Two rooms flats		Three rooms flats	
Helsinki	hki	-0.011	(0.053)	0.011	(0.089)	0.022***	(0.000)
	hki1	-0.183***	(0.003)	0.022**	(0.010)	-0.033	(0.046)
	hki2	0.043**	(0.017)	-0.055***	(0.000)	–	
	hki3	–		-0.033***	(0.000)	0.097***	(0.028)
	hki4	-0.009	(0.048)	0.478**	(0.189)	0.306**	(0.132)
Tampere	tre	–		–		-0.077***	(0.000)
	tre1	-0.178***	(0.054)	-0.077***	(0.016)	0.013***	(0.000)
	tre2	–		0.144***	(0.028)	-0.142**	(0.055)
	tre3	0.422***	(0.058)	–		-0.133***	(0.029)
Turku	tku	-0.028***	(0.003)	**no convergence**		-0.075***	(0.000)
	tku1	0.265***	(0.063)	–		-0.044	(0.055)
	tku2	-0.005***	(0.000)	-0.050	(0.040)	-0.068**	(0.031)
	tku3	-0.085***	(0.022)	–		-0.023	(0.035)
Oulu	oulu	0.286***	(0.058)	–		0.503***	(0.065)
	oulu1	0.258***	(0.000)	–		0.113**	(0.048)
Lahti	lti	0.207***	(0.061)	0.359***	(0.110)	0.165**	(0.053)
	lti1	-0.007	(0.016)	–		-0.123***	(0.032)
Jyväskylä	jkla	-0.074***	(0.000)	-0.040	(0.051)	0.215**	(0.086)
	jkla1	-0.136***	(0.017)	-0.001***	(0.000)	0.115***	(0.018)
	jkla2	-0.092*	(0.054)	0.577***	(0.144)	0.335***	(0.027)
Pori	pori	0.128***	(0.028)	0.066**	(0.031)	–	
	pori1	0.099***	(0.011)	0.047**	(0.021)	-0.047	(0.034)
	pori2	–		0.143***	(0.039)	–	
Kuopio	kuo	-0.048	(0.071)	0.278***	(0.072)	0.165***	(0.000)
	kuo1	-0.012	(0.032)	0.011***	(0.000)	-0.034	(0.024)
	kuo2	-0.055	(0.048)	–		-0.006***	(0.000)
Joensuu	jnsu1	0.193**	(0.068)	0.087***	(0.025)	–	
Seinäjoki	seoki	–		-0.018	(0.016)	0.120***	(0.026)
Vaasa	vaasa	–		0.031	(0.038)	0.200***	(0.008)
	vaasa1	–		–		0.052***	(0.011)
	vaasa2	–		–		-0.048	(0.049)
Kouvola	kou	-0.064***	(0.000)	0.425***	(0.031)	–	
Lappeenranta	lrta	0.181**	(0.071)	0.276***	(0.071)	0.119***	(0.024)
	lrta1	0.071*	(0.043)	0.083**	(0.026)	–	
	lrta2	–		–		0.122***	(0.036)
Hämeenlinna	hnlina	-0.130***	(0.000)	-0.037	(0.124)	–	
	hnlina1	–		0.349***	(0.086)	0.074***	(0.009)
Kotka	kotka	0.083***	(0.000)	–		-0.078**	(0.044)
	kotka1	–		-0.002	(0.041)	–	

Notes: This table reports the estimates of the parameter λ describing the nature of the relationship between house price return and volatility in individual GARCH–M model for each city and sub–area. The values in parentheses are the standard errors of the estimated parameters. *, **, and *** represent respectively 10%, 5%, and 1% levels of significance.

Table 6: Risk–return estimated coefficients.

Testing for Asymmetric volatility

Last, for cities and sub–areas exhibiting volatility clustering, the potentially asymmetric effects of shocks are investigated by estimating individual EGARCH models.

Table 7 displays the optimal specification of the EGARCH models, and their γ estimated coefficients. The evidence of asymmetric impact of shocks on housing volatility is noted in all the cases; with significant threshold term in all three apartment types across all cities and sub–areas. The exception is Kuopio–area1 (highlighted) in the one–room flats category. Following Black's (1976) argument that negative shocks have more effect on the rise of the conditional volatility than positive innovations of equal magnitude; across all three apartment types, the leverage effects ($\gamma < 0$) exist in thirteen out of twenty–eight cities and sub–areas in the one–room flats category. In ten out of twenty–seven in the two–rooms flats category, and twelve out of thirty–one in the more than three rooms flats category. In these areas, leverage effects can be interpreted as follows: a drop in apartment type price will induce the rise of debt to housing equity, and thereby causes an increase in the house price volatility. Note again that no EGARCH model could be fitted which yielded convergent estimates for Tutku–city in the two–rooms flats category.

A striking geographical pattern is observed in the Helsinki region, where a significant positive sign of the asymmetric term is found in the whole region in both one–room and two–rooms apartments types, and in the part of the region (sub–market 3 and 4) in the more than three rooms apartments type. A result which is consistent with other studies on equity markets (Koutmos et al., 1993; Kassimatis, 2002; Apergis and Eleptherou, 2001). Although, this positive sign suggests that there is no leverage effect, however, it may reflect a speculative housing bubble as pointed out by Morley and Thomas (2011); which means that in a highly populated region such as Helsinki, high demand of these apartments types would lead to a dramatic rise in house prices, and hence, the housing bubble. Therefore, following Lin and Fuerst (2014) who found similar results in the populated Canadian provinces: Alberta, Quebec, and Saskatchewan; the same recommendation of designing local housing policies with the aim to counterbalance developing housing bubbles can be given to Finnish housing policymakers of the Helsinki region. A similar pattern in all three apartments types is also observed in Turku–area2 and Jyväskylä–area2. One key solution would be to implement a supply response to these types of apartments in these populated areas.

The above–discussed results are consistent with the findings of Milles (2008), Morley and Thomas (2011), Lee (2009), and Lin and Fuerst (2014) who documented asymmetric effects in the US, UK, Australian, and Canadian housing markets, respectively. The results also prove and confirm once more the effective performance of the EGARCH model in modelling asymmetric effects in the housing markets.

		One room flats		Two rooms flats		Three rooms flats	
Regions	Cities/sub–areas	Model	γ	Model	γ	Model	γ
Helsinki	hki	(1,3)	0.762***	(2,3)	1.062***	(2,2)	-1.027***
	hki1	(2,2)	0.502***	(2,3)	1.695***	(3,1)	-0.800***
	hki2	(1,1)	0.100***	(2,3)	1.257***	–	–
	hki3	–	–	(2,3)	0.694***	(2,3)	1.111***
	hki4	(2,3)	0.324***	(1,3)	1.255***	(2,2)	0.372***
Tampere	tre	–	–	–	–	(2,3)	0.432***
	tre1	(3,2)	-0.342***	(1,3)	-0.336***	(3,1)	-0.339***
	tre2	–	–	(3,3)	-0.298***	(3,1)	1.566***
	tre3	(3,2)	-0.218***	–	–	(2,2)	0.747***
Turku	tku	(2,3)	0.367***	**No convergence**	–	(3,2)	-0.074***
	tku1	(3,2)	-1.180***	–	–	(2,3)	0.294***
	tku2	(2,3)	0.896***	(2,3)	0.532***	(1,3)	0.366**
	tku3	(3,3)	0.001***	–	–	(2,3)	1.035***
Oulu	oulu	(3,2)	-0.335***	–	–	(1,3)	0.797***
	oulu1	(2,3)	-0.586***	–	–	(1,2)	-0.324***
Lahti	lti	(2,3)	-0.221***	(2,2)	-0.282***	(1,1)	-0.377***
	lti1	(2,3)	0.415***	–	–	(3,1)	-0.829***
Jyväskylä	jkla	(3,2)	0.722***	(3,1)	-0.068***	(3,3)	-0.402***
	jkla1	(3,1)	-0.626***	(2,2)	0.489***	(1,2)	0.384***
	jkla2	(3,3)	0.151***	(2,1)	0.557***	(2,3)	0.062***
Pori	pori	(2,3)	0.544***	(2,3)	1.275***	–	–
	pori1	(1,2)	-0.883***	(3,3)	-1.480***	(3,1)	0.686***
	pori2	–	–	(1,1)	0.972***	–	–
Kuopio	kuo	(3,1)	-1.432***	(3,2)	-0.985***	(3,3)	-0.952***
	kuo1	(2,1)	**0.188**	(3,1)	-0.968***	(1,1)	0.183***
	kuo2	(3,3)	1.638***	–	–	(2,2)	-0.454***
Joensuu	jnsu1	(2,2)	-0.454***	(1,3)	-0.552***	–	–
Seinäjoki	seoki	–	–	(1,1)	0.557**	(2,3)	0.701***
Vaasa	vaasa	–	–	(3,1)	0.926***	(2,1)	-0.738***
	vaasa1	–	–	–	–	(2,2)	-0.344***
	vaasa2	–	–	–	–	(3,3)	0.039***
Kouvola	kou	(1,3)	0.170***	(2,1)	0.861***	–	–
Lappeenranta	lrta	(2,2)	-0.254***	(3,2)	-0.202***	(2,2)	0.330***
	lrta1	(3,3)	0.486***	(3,2)	0.763***	–	–
	lrta2	–	–	–	–	(3,1)	0.312***
Hämeenlinna	hnlina	(3,2)	-0.866***	(2,3)	0.887***	–	–
	hnlina1	–	–	(2,3)	0.085***	(3,3)	0.979***
Kotka	kotka	(2,1)	-0.648***	–	–	(1,2)	0.568***
	kotka1	–	–	(2,3)	-0.002***	–	–

Note: This table reports the optimal specification of the EGARCH model for each city and sub–area and their γ estimated coefficients. A significant γ implies asymmetry, and leverage effects exist if $\gamma < 0$. *, **, and *** represent respectively 10%, 5%, and 1% levels of significance.

Table 7: Asymmetric estimated coefficients.

5 Conclusion and implications

The literature on the assessment of volatility clustering, risk–return relationship, and asymmetric effects of the house prices is quite limited; compared to other assets such as stocks, despite housing having a dual role of investment and consumption. This article examines whether these financial properties are valid in the Finnish housing market, using quarterly house price indices from 1988:$Q1$ to 2018:$Q4$, for fifteen main regions in Finland.

The study has various major findings. First, strong evidence of volatility clustering (ARCH) effects is evident in over half of the cities and sub–areas in all three apartments types; indicating that apartments types prices in Finland are time–varying and clustered over time. Second, mixed results on the sign of the significant risk–return impact are observed across cities and sub–areas in all three apartments types; indicating that, in some regions, the Finnish housing investors are risk–reverse and would require a higher return to compensate the high risk. In other regions, during volatile periods, investors may accept a lower return as they think that the future would be riskier. Last, the evidence of the asymmetric impact of shocks on housing volatility is noted in almost all the cities and sub–areas housing markets. However, across all three apartments types, the leverage effects exist in few regions; indicating that a drop in apartment price will raise uncertainty, and hence house price volatility. Moreover, a geographical pattern is noticeable in the whole Helsinki region in both one–room and two–rooms flats categories, and in the part of the area in the more than three rooms flats category. These results in the Helsinki region serve as an "early warning" of a housing bubble.

The findings have some housing investment and policymakers implications. First, the strong evidence of ARCH effects suggests that housing investors, policymakers, and consumers should be aware of higher risks of substantial losses in house price returns during volatile periods. Moreover, they should monitor the asset volatility as it holds essential information. Second, mixed results in the sign of the risk–return relation observed across regions in all three apartments types recommend diversification of a housing portfolio across Finland and by apartment type; as the one concentrated in one or several areas may be risky. Last, the evidence of the asymmetric adjustment to shocks suggests that the Finnish housing market is more sensitive to bad news than good news. The Helsinki region' results indicate a speculative housing bubble, and it is recommended to the housing policymakers of these regions to address this issue in their housing policy designs.

In the standpoint of developing appropriate time series volatility forecasting models of this housing market; the study outcomes will be used in a forecasting procedure of the volatility dynamics of the studied types of dwellings. That is, as the model which fits better does not necessarily mean it will forecast well, an in–sample and out–of–sample forecasting performance of these short memory GARCH–types model will be compared and assessed to the long memory GARCH models such as the Fractionally Integrated GARCH (FIGARCH) model and the Component GARCH (CGARCH) model; to provide information regarding which forecasting methods deliver superior volatility forecasts of the studied types of apartments.

Appendices

Cities/Sub–areas	Abbreviations for cities and sub–areas	Postcode numbers
Helsinki	hki	City area
Helsinki–area1	hki1	100, 120, 130, 140, 150, 160, 170, 180, 220, 260
Helsinki–area2	hki2	200, 210, 250, 270, 280, 290, 300, 310, 320, 330, 340, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 610, 810, 850, 990
Helsinki–area2	hki3	240, 350, 360, 370, 400, 430, 440, 440, 620, 650, 660, 670, 680, 690, 730, 780, 790, 800, 830, 840, 950
Helsinki–area4	hki4	Other postcodes
Tampere	tre	City area
Tampere–area1	tre1	33100, 33180, 33200, 33210, 33230, 33240, 33250, 33500, 33540
Tampere–area2	tre2	33270, 33400, 33530, 33560, 33610, 33700, 33730, 33820, 33900, 34240
Tampere–area3	tre3	Other postcodes
Turku	tku	City area
Turku–area1	tku1	20100, 20500, 20700, 20810, 20900
Turku–area2	tku2	20200, 20250, 20300, 20380, 20400, 20520, 20720, 20880, 20960
Turku–area3	tku3	Other postcodes
Oulu	oulu	City area
Oulu–area1	oulu1	90100,90120, 90130, 90140, 90230, 90400, 90410, 90420, 90510
Oulu–area2	oulu2	Other postcodes
Lahti	lti	City area
Lahti–area1	lti1	15100, 15110, 15140, 15160, 15320, 15340, 15610, 15850, 15900
Lahti–area2	lti2	Other postcodes
Jyväskylä	jkla	City area
Jyväskylä–area1	jkla1	40100, 40200, 40500, 40520, 40530, 40600, 40700, 40720
Jyväskylä–area2	jkla2	Other postcodes
Pori	pori	City area

Cities/Sub–areas	Abbreviations for cities and sub–areas	Postcode numbers
Pori–area1	pori1	28100, 28130, 28300, 28430, 28540, 28660, 28900
Pori–area2	pori2	Other postcodes
Kuopio	kuo	City area
Kuopio–area1	kuo1	70100, 70110, 70300, 70600, 70800, 70840
Kuopio–area2	kuo2	Other postcodes
Joensuu	jnsu	City area
Joensuu–area1	jnsu1	80100, 80110, 80200, 80220
Joensuu–area2	jnsu2	Other postcodes
Seinajöki	seoki	City area
Vaasa	vaasa	City area
Vaasa–area1	vaasa1	65100, 65170, 65200, 65410
Vaasa–area2	vaasa2	Other postcodes
Kouvola	kou	City area
Lappeenranta	lrta	City area
Lappeenranta–area1	lrta1	53100, 53130, 53500, 53600, 53900, 55330
Lappeenranta–area2	lrta2	Other postcodes
Hämeenlinna	hnlina	City area
Hämeenlinna–area1	hnlina1	13100, 13130, 13200, 13220, 13270
Hämeenlinna–area2	hnlina2	Other postcodes
Kotka	kotka	City area
Kotka–area1	kotka1	48100, 48210, 48310, 48710
Kotka–area2	kotka2	Other postcodes

Source: Statistics Finland

Table 8: Regional division by postcode numbers.

References

André, C. and García, C. (2012). Housing Price and Investment Dynamics in Finland, *Technical report*, OECD Economics Department Working Papers, No.962, OECD Publishing.
URL: *http://dx.doi.org/10.1787/5k98rwldjr44-en*

Apergis, N. and Elieptherou, S. (2001). Stock Returns and Volatility Evidence from the Athens Stock Market Index, *Journal of Economics and Finance* **25**: 50–61.

Bank of Finland (2018). *Helibor and Euribor rates.*
URL: *https://www.suomenpankki.fi/en/Statistics/interest-rates/tables/*

Bekaert, G. and Wu, G. (2000). Asymmetric volatility and risk in equity markets, *Review of Financial Studies* **13**: 1–42.

Black, F. (1976). Studies of Stock Market Volatility Changes, *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section*, pp. 177–181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics* **31**(3): 307–327.

Bollerslev, T., Chou, R. Y. and Kroner, K. F. (1992). ARCH modeling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics* **52**: 1–2, 5–59.

Campbell, J. Y. and Hentschell, L. (1992). No News is Good News: An asymmetric model of changing volatility is stock returns, *Journal of Financial Economics* **31**: 281–318.

Cannor, S., Miller, N. and Pandher, G. S. (2006). Risk and return in US housing market: a cross–sectional asset–pricing approach, *Real State Economics* **34**(4): 519–552.

Christie, A. A. (1982). The Stochastic Behavior of Common Stock Variances–Value, Leverage and Interest Rate Effects, *Journal of Financial Economics* **10**: 407–432.

Coskun, Y. and Ertugrul, H. M. (2016). House price return volatility patterns in Turkey, Istanbul, Ankara and Izmir, *Journal of European Real Estate Research* **8**(1): 26–51.

Dolde, W. and Tirtiroglu, D. (1997). Temporal and spatial information diffusion in real estate price changes and variances, *Real Estate Economics* **25**(4): 539–565.

Dolde, W. and Tirtiroglu, D. (2002). Housing price volatility changes and their effects, *Real Estate Economics* **30**: 41–66.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation, *Econometrica* **50**(4): 987–1007.

Engle, R. F., Lilien, D. M. and Robins, R. P. (1987). Estimating Time Varying Risk Premia in the Term Structure: The ARCH–M Model, *Econometrica* **55**(2): 391–407.

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility, *Journal of Finance* **48**(5): 1749–1778.

Finnish Landlord Association (2017). *A survey among Finnish private landlords.* **URL:** *https://vuokranantajat.fi/asunntosijoittaminen/tutkimusket/vuokranantaja-2017/*

French, K. R., Schwert, G. W. and Stambaugh, R. (1987). Expected stock returns and volatility, *Journal of Financial Economics* **19**: 3–29.

Ghalanos, A. (2019). *rugarch: Univariate GARCH models.* R package version 1.4-1.

Glosten, L., Jagannathan, R. and Runkle, D. (1993). On the Relation between the Expected Value and the Volatility of the Norminal Excess Return on Stocks, *Journal of Finance* **68**: 1179–1801.

Graves, S. (2019). *Companion to Tsay (2005) Analysis of Financial Time Series.* R package version 0.4-6.

Guo, H. and Nelly, C. J. (2008). Investigating the intertemporal risk–return relation in international stock markets with the component GARCH model, *Economics Letters* **99**: 371–374.

Hibbert, A. M., Daigler, R. T. and Dupoyet, B. (2008). A behaviour explanation for the negative asymmetric return–volatility relation, *Journal of Banking & Finance* pp. 322254–322266.

Hossain, B. and Latif, E. (2009). Determinants of housing price volatility in Canada: a dynamic analysis, *Applied Economics* **41**(27): 3521–3531.

Kaleva, H. (2019). The Finnish Property Market Report, *Technical report*, KTI Property Information Ltd.

Kassimatis, K. (2002). Financial Liberalisation and Stock Market Volume in Selected Developing Markets, *Applied Financial Economics* **12**: 389–394.

Katsiampa, P. and Begiazi, K. (2019). An empirical analysis of the Scottish housing market by property type, *Scottish Journal of Political Economy* .
URL: *DOI: 10.1111/sjpe.12210*

Kivistö, J. (2012). An assessment of housing price developments against various measures, *Technical report*, Bank of Finland Bulletin.

Koutmos, G., Negakis, C. and Theodossiou, P. (1993). Stochastic Behaviour of the Athens Stock Exchange, *Applied Financial Economics* **3**: 119–126.

KTI (Autumn, 2019). KTI Market Review, *Technical report*, KTI Property Information Ltd.

Lee, C. L. (2009). Housing price volatility and its determinants, *International Journal of Housing Markets and Analysis* **2**(3): 293–308.

Lee, C. L. (2017). An examination of the risk–return relation in the Australian housing market, *International Journal of Housing Markets and Analysis* **10**(3): 431–449.

Lee, C. L. and Reed, R. (2014a). The relationship between housing market intervention for the first–time buyers and house price volatility, *Housing Studies* **29**(8): 1073–1095.
URL: *doi: 10.1080/02673037.2014.927420*

Lee, C. L. and Reed, R. (2014b). Volatility decomposition of Australian housing prices, *Journal of Housing Research* **23**(1): 21–43.

Lin, P.-T. and Fuerst, F. (2014). Volatility clustering, risk–return relationship, and asymmertic adjustment in the Canadian housing market, *Journal of Real Estate Portifolio Management* **20**(1): 37–46.

Mcleod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared – residual autocorrelations, *Journal of Time Series Analysis* **4**(4): 269–273.

Merton, R. C. (1973). An intertemporal capital asset pricing models, *Econometrica* pp. 41867–41887.

Miller, N. and Peng, L. (2006). Exploring metropolitan housing price volatility, *The Jouarnal of Real Estate Finance and Economics* **33**(1): 5–18.

Milles, W. (2008). Volatility clustering in US home prices, *Journal of Real Estate Research* **30**(1): 73–90.

Milles, W. (2011). Long–Range Dependence in U.S Home Price Volatility, *Journal of Real Estate Finance and Economics* **42**: 329–347.

Milles, W. (2011b). Clustering in UK home prices volatility, *Journal of Housing Research* **20**(1): 87–101.

Morley, B. and Thomas, D. (2011). Risk–return relationships ans asymmetric adjustment in the UK housing market, *Applied Financial Economics* **21**(10): 735–742.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* **55**: 703–708.

Oikarinen, E. (2009a). Household borrowing and metropolitan house price dynamics – Empirical evidence from Helsinki, *Journal of Housing Economics* **18**(2): 126–139.

Oikarinen, E. (2009b). Interaction between housing prices and household borrowing: The Finnish case, *Journal of Banking & Finance* **33**(4): 747–756.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

RBA (2015). Submission to the Inquiry into Home Ownership, *Technical report*, Reserve Bank of Australia, Sydney.
URL: *www.rba.gov.au/publications/submissions/housing-and-housing-finance/inquiry-into-home-ownership/pdf/inquiry-into-home-ownership.pdf*

Statistics Finland (2016). Households' assets, *Technical report*. Last checked: 10/09/2019.
URL: *http://www.stat.fi/til/vtutk/2016/vtutk-2016-2018-06-05-tie-001-en.html*

Statistics Finland (2019). Building and dwelling production, *Technical report*. Last checked: 10/09/2019.
URL: *http://www.stat.fi/til/ras/index-en.html*

Statistics Finland Overview (2018). 2018 overview, household–dwelling units and housing conditions, *Technical report*. Last checked: 28/01/2020.
URL: *http://www.stat.fi/til/asas/2018/011/asas-2018-01-2019-10-10-kat-002-en.html*

Stevenson, S. (2002). An examination of volatility spillovers in REIT returns, *Journal of Real Estate Portfolio Management* **8**(3): 229–238.

Tsai, I.-C. and Chen, M.-C. (2009). The asymmetric volatility of house prices in the UK, *Property Management* **27**(2): 80–90.

Tsai, I.-C., Chen, M.-C. and Ma, T. (2010). Modelling house price volatility states in the UK by switching ARCH models, *Applied Economics* **42**(9): 1145–1153.

Tsay, R. S. (2013). *An introduction to analysis of financial data with R*, John Wiley & Sons, Inc., Hoboken, New Jersey.

Willcocks, G. (2010). Conditional variances in UK regional house prices, *Spatial Economic Analysis* **5**(3): 339–354.

Publication III

Dufitinema, J.

Stochastic volatility forecasting of the Finnish housing market

Reprinted with permission from
Applied Economics
Volume 53, Issue 1, pp. 98-114, 2021

 DOI: https://doi.org/10.1080/00036846.2020.1795074

APPLIED ECONOMICS
2021, VOL. 53, NO. 1, 98–114
https://doi.org/10.1080/00036846.2020.1795074

Stochastic volatility forecasting of the Finnish housing market

Josephine Dufitinema

School of Technology and Innovation, Mathematics and Statistics Unit, University of Vaasa, Vaasa, Finland

ABSTRACT

The purpose of the article is to assess the in-sample fit and the out-of-sample forecasting performances of four stochastic volatility (SV) models in the Finnish housing market. The competing models are the vanilla SV, the SV model where the latent volatility follows a stationary AR(2) process, the heavy-tailed SV and the SV with leverage effects. The models are estimated using Bayesian technique, and the results reveal that the SV with leverage effects is the best model for modelling the Finnish house price volatility. The heavy-tailed SV model provides accurate out-of-sample volatility forecasts in most of the studied regions. Additionally, the models' performances are noted to vary across almost all cities and sub-areas, and by apartment types. Moreover, the AR (2) component substantially improves the in-sample fit of the standard SV, but it is unimportant for the out-of-sample forecasting performance. The study outcomes have crucial implications, such as portfolio management and investment decision-making. To establish suitable time-series volatility forecasting models of this housing market, these study outcomes will be compared to the performances of their GARCH models counterparts.

KEYWORDS

Stochastic volatility; Bayesian estimation; forecasting; Finland; house prices

JEL CLASSIFICATION

C11; C22; C53

I. Introduction

Volatility modelling and forecasting are a vital task in financial markets. As the asset volatility holds critical information; it has been recognized as the most risk measure broadly used in many areas of finance (Bollerslev, Chou, and Kroner 1992). In the housing market, as housing assets have a dual role of consumption and investment; understanding price volatility plays an essential role in the housing investment decision-making and the asset allocation (Milles 2008a). Moreover, housing is a crucial factor for the country's economy; in particular, in Finland, Statistics Finland (2016) reported that housing made up to 50.3% of the Finnish households' total wealth. Thus, housing affects the country's economy through wealth effects (Case, Quigley, and Shiller 2013) as well as through influences on many parties exposed to housing and mortgage activity. Therefore, better housing modelling and forecasting would be beneficial for consumers, mortgage market, mortgage insurance and mortgage-backed securities (Segnon et al. 2020). Furthermore, as pointed out by Zhou and Haurin (2010), insights into house price volatility are the key input in designing housing policies. In the light of the abovementioned points, understanding the dynamics of the house price volatility is crucial for portfolio management, risk assessment and investment decision-making.

An increasing amount of studies have attempted to model and/or forecast the house price volatility of individual markets. However, the literature has mainly focussed on the use of different Generalized Autoregressive Conditional Heteroscedasticity (GARCH)-type models. Under this approach, the volatility evolution is modelled deterministically; a framework which has its roots from the Engle's (1982) and Bollerslev's (1986) groundbreaking works. Taylor (1982), on the other hand, provided an alternative way; to model volatility probabilistically, meaning that volatility is treated as an unobserved component that follows a stochastic process. The specification is known as the Stochastic Volatility (SV) models. Even though SV models are theoretically attractive and there is some empirical evidence in their favour over GARCH models (Jaquier, Polson, and Rossi 1994; Gysels, Harvey, and Renault 1996; Kim et al., 1998; Nakajima and Omori 2012); they have drawn little attention among practitioners. The challenges

CONTACT Josephine Dufitinema josephine.dufitinema@uwasa.fi School of Technology and Innovation, Mathematics and Statistics Unit, University of Vaasa, Vaasa, Finland

pointed out by Bos (2012) are highly non-linear estimations and lack of standard software packages implementing these methods. In response to these challenges, Chan and Grant (2016b) provided the means for the Bayesian estimation of not only the vanilla SV model but also the heavy-tailed SV model and the SV model with leverage effects. Specifically, this study uses Chan and Grant (2016b) approach to model and forecast the studied housing market. To the best of the author' knowledge, in the housing markets, there has yet to be empirical modelling and forecasting using the SV framework. Hence, this is the first study that models and forecasts the Finnish housing market volatility using the SV framework in general, and incorporating both non-Gaussianity and asymmetry effects in particular.

Moreover, the emphasis of the housing market volatility modelling and/or forecasting has been on a limited number of countries such as the United States, United Kingdom, Australia and Canada. Regarding housing market volatility modelling without the forecasting aspect, the authors (to cite few) who have employed GARCH-type models to study US house prices include Dolde and Tirtiroglu (1997, 2002), Miller and Peng (2006), Milles (2008b), and more recently, Apergis and Payne (2020). The UK house price volatility investigation consists of the work of Willcocks (2010), Tsai, Chen, and Ma (2010), Milles (2011b), and more recently, Begiazi and Katsiampa (2019). The Australian house price volatility has been examined by Lee (2009) and Lee and Reed (2014b); while Hossain and Latif (2009) and Lin and Fuerst (2014) studied the Canadian house price volatility. For Finland, Dufitinema (2020) has recently explored different aspects of the Finnish housing market volatility. Regarding the housing market volatility forecasting, the US housing market is the widely studied housing market. Beginning with the work of Crawford and Fratantoni (2003), followed by Milles (2008a), Li (2012), more recently, Segnon et al. (2020). For Finland, there has yet to be an empirical forecasting of the Finnish housing market, even though Statistics Finland (2016) reported that housing made up to 50.3% of the Finnish households' total wealth. Therefore, this article aims to fill that gap by being the first study that forecasts the Finnish housing market volatility and further extends the ongoing literature on the countries' house price volatility forecasting.

Furthermore, in contrast to previous studies which employed the data sets of the family–home property type; the studied type of dwellings in the article at hand is apartments (block of flats) categorize by the number of rooms. That is one-room, two-room and more than three rooms apartment types. One reason is that, according to Statistics Finland Overview, at the end of 2018, among all occupied dwellings, 46% were in apartments; which reflects how living in flats is growing in popularity in Finland, compared to other house types. Detached and semi-detached houses occupied 39%, terraced 14%, while 1% were in other buildings. The other reason is that apartments property type has not only increased its attractiveness in consumers but also in the Finnish residential property investors. Currently, foreign investors own some 15,000 rental flats, and between 2015 and 2018, in the Finnish housing development which has been very active in apartment buildings (Statistics Finland 2019); the share of foreign investors was up to 38%, and domestic and individual investors together hold some 40% (KTI, Autumn, 2019). Additionally, in the same standpoint of housing investment, this study uses data on both metropolitan and geographical level, to analysis and cross-compare housing investment in different cities and sub-areas, and portfolio allocation across Finland.

The purpose of the study is to assess the in-sample fit and the out-of-sample forecasting performance of four stochastic volatility models in the Finnish housing market. The competing models are the vanilla SV, the SV model where the latent volatility follows a stationary AR(2) process, the heavy-tailed SV and the SV with leverage effects. In other words, the goal of this model comparison exercise is to examine, in the SV framework, which volatility model tends to fit better the dynamics of the Finnish house prices and which one provides superior out-of-sample forecasts. Additionally, these models are used to answer the following questions: Are leverage effects and heavy-tailed distributions crucial in modelling and forecasting the Finnish house price volatility? Is the AR(2) component a useful addition to the vanilla SV model? The study assesses the Finnish housing market by apartment types categorize by the number of rooms. That is, single-room, two-room and apartments with more than three rooms. These apartment type prices are for 15 main regions divided geographically, according to their postcode numbers, into 45 cities and sub-areas. Each model is estimated for each city and sub-area with significant clustering effects. For the assessment of the out-of-sample forecasting performance of the four models, the data is split into two parts: the training set used for the estimation and prediction, and the test set used for the evaluation of the forecast built by the fitted model. Results reveal that, for the in-sample fit analysis, in all three apartment types, the stochastic volatility model with leverage effect ranks as the best model for modelling the Finnish house price volatility. For the out-of-sample forecasting assessment, in most of the regions, the heavy-tailed stochastic volatility model excels in forecasting the house price volatility of the studied types of apartments. Additionally, the models' performances are noted to vary across almost all cities and sub-areas, and by apartment types – no geographical pattern is observed. Moreover, for the in-sample fit analysis, the AR(2) component is found to be a valuable addition to the vanilla SV, whereas, for the out-of-sample forecasting assessment, the vanilla SV model outperforms the SV-2 in most of the regions.

The remainder of the article is as follows. Section 2 describes the data and outlines the methodology to be employed. Section 3 presents and discusses the results. Section 4 concludes the article.

II. Data and methodology

Data

The study uses quarterly house price indices of 15 main regions in Finland estimated by Statistics Finland using the so-called hedonic method. The studied period is from 1988:Q1 to 2018:Q4, and the type of dwellings is apartments categorize by the number of rooms, that is, one-room, two-room and more than three rooms apartment types. The studied regions are Helsinki, Tampere, Turku, Oulu, Lahti, Jyväskylä, Kuopio, Pori, Seinäjoki, Joensuu, Vaasa, Lappeenranta, Kouvola, Hämeenlina and Kotka. Additionally, these regions are divided geographically, according to their postcode numbers, into 45 cities and sub-areas. The data regions' ranking according to their number of inhabitants and regional division by postcode numbers is described in detailed in Dufitinema (2020).

For a sample of three cities in each of the apartments categories, a house price movement is graphed in Figure 1. Those are Helsinki, Tampere, Turku in the one-room flats group; Pori, Joensuu, Vaasa in the two-room flats group; Lappeenranta, Hämeenlina, Kotka in the more than three rooms flats group. A similar pattern is observed in all sample graphs from the end of 1980s to mid-1993. During this period, house prices in Finland experienced a structural break due to the financial market deregulation (Oikarinen 2009a, 2009b). Moreover, as it can be noted since the bursting of the bubble, one-room apartment prices have been increasing. Two-room apartments experienced downturns in the 2010s, same as large apartments; however, large apartments prices continue to decrease especially in less densely populated regions such as Kotka-city.

Methodology

The methodology used in this study is as follows: For each city and sub-area in each apartment type, we transform house price indices into continuous compound returns. Next, by employing the Akaike and Bayesian information criteria, we determine the ARMA model of appropriate order that filters out the first autocorrelations from the returns. Then, we test the clustering effects or Autoregressive Conditional Heteroscedasticity (ARCH) effects from the ARMA filtered returns. Lastly, for cities and sub-areas exhibiting ARCH effects, the four SV models' in-sample estimations are performed, and the out-of-sample volatility forecasting performances are evaluated using the stochastic volatility framework.

Testing for ARCH effects

Two tests are employed to test clustering effects; those are Ljung-Box (LB) and Lagrange Multiplier (LM). An extensive discussion is given in Dufitinema (2020) and results are outlined in Table 1. In summary, both tests found significant clustering effects in over half of the cities/sub-areas

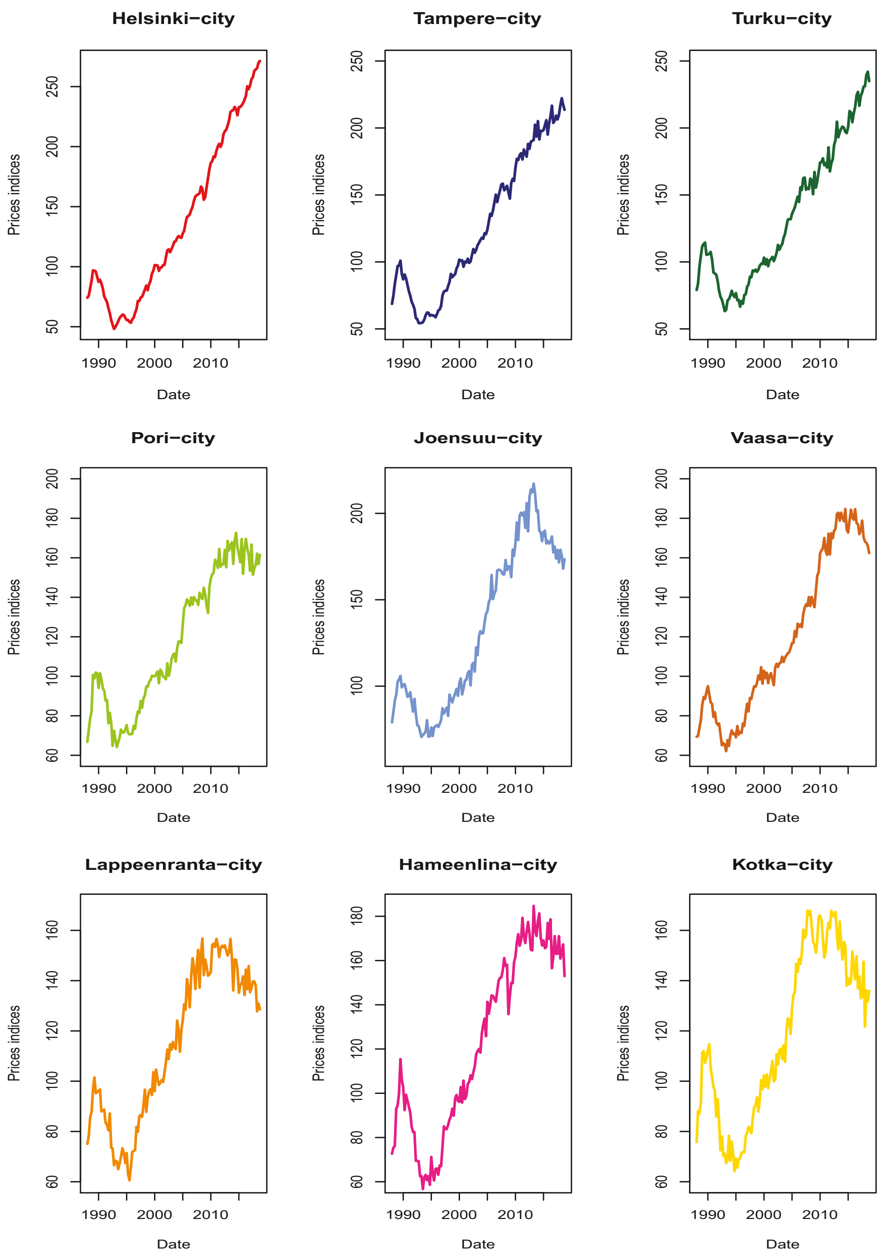


Figure 1. The house price movement – Sample cities.

in all three studied types of apartments. Plus precisely, in the one–room flats category, ARCH effects were found in 28 out of 38 cities/sub-areas. In two-room flats category, they were significant in 27 out of 42;, and in the more than three rooms flats category, they were found in 31 out of 39.

Table 1. ARCH effects tests results.

		One room flats		Two rooms flats		Three rooms flats	
Regions	Cities/sub-areas	ARMA	ARCH?	ARMA	ARCH?	ARMA	ARCH?
Helsinki	hki	ARMA(2,1)	Yes	ARMA(2,1)	Yes	AR(1)	Yes
	hki1	MA(2)	Yes	ARMA(2,1)	Yes	AR(2)	Yes
	hki2	ARMA(2,1)	Yes	AR(1)	Yes	AR(1)	No
	hki3	ARMA(2,1)	No	AR(2)	Yes	AR(2)	Yes
	hki4	AR(2)	Yes	ARMA(1,1)	Yes	AR(2)	Yes
Tampere	tre	ARMA(1,1)	No	ARMA(2,1)	No	ARMA(2,2)	Yes
	tre1	ARMA(2,2)	Yes	AR(2)	Yes	ARMA(2,2)	Yes
	tre2	ARMA(1,1)	No	ARMA(0,0)	Yes	ARMA(2,2)	Yes
	tre3	AR(2)	Yes	ARMA(2,2)	No	ARMA(1,1)	Yes
Turku	tku	ARMA(2,2)	Yes	ARMA(2,2)	Yes	ARMA(2,2)	Yes
	tku1	ARMA(1,1)	Yes	AR(2)	No	AR(1)	Yes
	tku2	AR(1)	Yes	ARMA(0,0)	Yes	ARMA(2,2)	Yes
	tku3	AR(1)	Yes	MA(3)	No	ARMA(0,0)	Yes
Oulu	oulu	ARMA(1,1)	Yes	AR(2)	No	ARMA(1, 2)	Yes
	oulu1	AR(1)	Yes	ARMA(1,2)	No	ARMA(1,2)	Yes
	oulu2	AR(1)	No	ARMA(0,0)	No	MA(3)	No
Lahti	lti	AR(2)	Yes	AR(2)	Yes	ARMA(2,2)	Yes
	lti1	AR(1)	Yes	AR(2)	No	MA(3)	Yes
	lti2	AR(1)	No	ARMA(1,2)	No	ARMA(2,2)	No
Jyväskylä	jkla	ARMA(1,1)	Yes	ARMA(2,2)	Yes	ARMA(1,2)	Yes
	jkla1	ARMA(1,1)	Yes	MA(3)	Yes	ARMA(2,2)	Yes
	jkla2	ARMA(0,0)	Yes	ARMA(1,2)	Yes	ARMA(1,2)	Yes
Pori	pori	MA(1)	Yes	MA(3)	Yes	ARMA(2,2)	No
	pori1	AR(2)	Yes	MA(3)	Yes	MA(1)	Yes
	pori2	–	–	ARMA(2,2)	Yes	–	–
Kuopio	kuo	ARMA(0,0)	Yes	AR(2)	Yes	ARMA(0,0)	Yes
	kuo1	MA(2)	Yes	ARMA(0,0)	Yes	MA(1)	Yes
	kuo2	ARMA(0,0)	Yes	AR(2)	No	ARMA(1,2)	Yes
Joensuu	jnsu	MA(3)	No	AR(3)	No	AR(1)	No
	jnsu1	MA(3)	Yes	AR(3)	Yes	AR(1)	No
Seinäjoki	seoki	–	–	AR(1)	Yes	MA(3)	Yes
Vaasa	vaasa	MA(1)	No	ARMA(1,2)	Yes	ARMA(1,2)	Yes
	vaasa1	MA(1)	No	MA(2)	No	MA(1)	Yes
	vaasa2	–	–	–	–	ARMA(0,0)	Yes
Kouvola	kou	AR(1)	Yes	ARMA(1,2)	Yes	MA(3)	No
Lappeenranta	lrta	AR(1)	Yes	MA(3)	Yes	MA(3)	Yes
	lrta1	MA(1)	Yes	ARMA(2,2)	Yes	–	–
	lrta2	–	–	AR(1)	No	ARMA(0,0)	Yes
Hämeenlinna	hnlina	MA(3)	Yes	ARMA(0,0)	Yes	MA(3)	No
	hnlina1	MA(3)	No	ARMA(1,2)	Yes	AR(1)	Yes
Kotka	kotka	MA(1)	Yes	MA(3)	No	ARMA(2,2)	Yes
	kotka1	MA(3)	No	MA(2)	Yes	–	–
	kotka2	–	–	MA(2)	No	–	–

Notes: This table reports, for each city and sub-area, the ARMA model and the outcomes of the two tests of ARCH effects. "Yes" indicates that a city/sub-area exhibits ARCH effects, "No" means that a city/sub-area does not.

In-sample fit analysis

For cities and sub-areas exhibiting clustering effects, the in-sample fit is performed using the stochastic volatility approach. That is, in contrast to the GARCH-type framework where the conditional variance is assumed to follow a deterministic process; a stochastic volatility (SV) model treats the time-varying volatility as an unobserved component that mimics a stochastic process. The most popular SV model is the vanilla SV model with normal distribution errors proposed and developed by Taylor (1982, 1986). However, several authors have pointed out that a normal distribution assumption is not plausible when analysing asset returns with SV framework as well as GARCH-type models (Tsay 2013; Harvey and Shephard 1996; Omari et al. 2007; Nakajima and Omori 2012). A suitable distribution requires to accommodate the characteristics of asset returns such as skewness and fat tails. Therefore, for each city and sub-area in each apartment type, the in-sample estimations of the vanilla SV model and the SV model with additional AR(2) component are compared to the SV model with Student's *t* errors (heavy-tailed SV) and SV model with leverage effects. The models are estimated on the whole sample data from 1988:Q1 to 2018:Q4.

Vanilla SV model

Let y_t denotes the demeaned return process $y_t = \log(S_t/S_{t-1}) - \mu_t$. A basic stochastic volatility model is of the following form:

$$y_t = \sigma_t \epsilon_t, \quad t = 1, 2, \ldots .T,$$

where the $\log \sigma_t^2$ follows an AR(1) process. To adopt the convention often used in literature, we write for $h_t = \log \sigma_t^2$,

$$\begin{aligned} y_t &= \sigma_t \epsilon_t, \quad t = 1, 2, \ldots .n \\ \sigma_t^2 &= \exp(h_t) \\ h_t &= \mu + \phi h_{t-1} + \sigma_\eta \eta_t, \end{aligned} \tag{1}$$

where h_t is the latent stochastic process (more precisely, the log-variance process), μ is a constant or the level of the log-variance process, ϕ is a parameter representing persistence in the log-variance process, σ_η is the volatility or the standard deviation of the log-variance process (also called *volvol*), and η_t is the random shocks in the log-variance process; a white noise uncorrelated with ϵ_t. $\theta = (\mu, \phi, \sigma_\eta)^T$ is referred to as the SV parameter vector.

Equation (1) can be expressed in hierarchical form. In its centred parameterization form, it is written as:

$$\begin{aligned} & y_t | h_t \sim \mathcal{N}(0, \exp(h_t)), \\ & h_t | h_{t-1}, \theta \sim \mathcal{N}(\mu + \phi(h_{t-1} - \mu), \sigma_\eta^2), \end{aligned}$$

where $\mathcal{N}(\mu, \sigma_\eta^2)$ denotes the normal distribution with mean μ and variance σ_η^2.

The SV model with additional AR(2) component, which is referred to as the **SV-2**, is the model where the observation is the same as in Equation (1); however, the log-variance h_t mimics a stationary AR(2) process.

SV with student's t errors (SVt)

As discussed above, the non-normal conditional residual distributions are recommended when analysing asset returns. The proposed distributions include, for instance, the Student's t distribution by Harvey, Ruiz, and Shephard (1994); the (semi-)parametric residuals by Jensen and Maheu (2010) and Delatola and Griffin (2011); the extended generalized Inverse Gaussian by (Silva, Lopes, and Migon 2006); and the generalized hyperbolic skew Student's t errors by Nakajima and Omori (2012).

The SV model with Student's t errors is described as:

$$\begin{aligned} & y_t | h_t, \nu \sim t_\nu(0, \exp(h_t/2)), \\ & h_t | h_{t-1}, \theta \sim \mathcal{N}(\mu + \phi(h_{t-1} - \mu), \sigma_\eta^2). \end{aligned} \tag{2}$$

The observations now follow a conditionally Student's t distribution $t_\nu(a, b)$ with ν degrees of freedom, mean a and scale b. The parameter vector of the SVt model is $\theta = (\mu, \phi, \sigma_\eta, \nu)^T$.

SV with leverage effects (SVl)

It has been argued that the returns of financial variables have three major distribution characteristics. Those are heavy-tailedness, skewness and volatility clustering with leverage effects. The leverage effect emerged from Black's (1976) and Christie's (1982) studies outcome that a drop in return (a negative chock) has more impact on asset price volatility increase than a rise in return (a positive chock). Various extensions of the vanilla SV model with normal errors have been proposed to model this effect. The proposed asymmetric innovations include, for instance, the distributions featuring correlation and variance by Harvey and Shephard (1996), and Jaquier, Polson, and Rossi (2004); the skewed distributions by Nakajima and Omori (2012) and the non-parametric distributions by Jensen and Maheu (2014).

The SV model with leverage effects is described as:

$$\begin{aligned} & y_t | h_t, \theta \sim \mathcal{N}(0, \Sigma), \\ & h_t | h_{t-1}, \theta \sim \mathcal{N}(\mu + \phi(h_{t-1} - \mu), \Sigma), \\ & \Sigma = \begin{pmatrix} \exp(h_t) & \rho\sigma_\eta \exp(h_t/2) \\ \rho\sigma_\eta \exp(h_t/2) & \sigma_\eta^2 \end{pmatrix}. \end{aligned} \tag{3}$$

The vector $\theta = (\mu, \phi, \sigma_\eta, \rho)^T$ collects the SVl parameters. The parameter ρ measures the correlation between the residuals of the observations (ϵ_t) and the innovations of the log-variance process (η_t). Leverage effects exist when $\rho < 0$.

Model comparison

As the latent volatility process (h_t) enters the models in a non-linear fashion, the maximum likelihood estimation framework is not a straightforward task as in the GARCH-type models' case. The reason being that for the SV models, the likelihood function does not have a closed form (Gysels, Harvey, and Renault 1996). Hence, the estimation of the SV models is done through Bayesian parameter estimation technique via Markov Chain Monte Carlo (MCMC) methods (Kim et al., 1998). The estimation of the four SV models was performed by following Chan and Grant's approach, which is outlined in Chan and Grant (2016b, Appendix A). In estimating the SV models, the vital step is the joint sampling of the log volatilities. The novelty of Chan and Grant's approach is that instead of using the conventional Kalman Filter to achieve this key step; the algorithm employs the fast band matrix routines (Chan and Jeliazkov 2009; Chan 2013).

The four models performances are compared using two popular Bayesian model comparison criteria, namely, deviance information criterion (DIC) and Bayes factor. The deviance information criterion (DIC) proposed by Spiegelhalter et al. (2002) is a trade-off between the model's goodness of fit and its corresponding complexity. The fit is measured by the *deviance*, defined as

$$D(\theta) = -2\log\mathcal{L}(y|\theta),$$

where $\mathcal{L}(y|\theta)$ is the likelihood function. The complexity is measured by an estimate of the *effective number of parameters* p_{D}, defined as

$$p_{\mathrm{D}} = \overline{D} - D(\bar{\theta}).$$

That is, the difference between the posterior mean deviance and the deviance evaluated at the posterior mean of parameters. Thus, the DIC is the sum between the Monte Carlo estimated posterior mean deviance and the effective number of parameters:

$$\mathrm{DIC} = \overline{D} + p_{\mathrm{D}}.$$

The smaller the DIC, the better the model supports the data. The widely used version of DIC is the one obtained by conditioning on the latent variables, that is, the DIC based on conditional likelihood. However, studies such as Li, Zeng, and Yu (2012) have warned against using this DIC version on the grounds of being non-regular and thus invalidates the needed justification of DIC – the standard asymptotic arguments. Moreover, Millar (2009) and Chan and Grant (2016a) provided Monte Carlo evidence that this DIC version always favours the most complex and overfitted model. To overcome this issue, Chan and Grant (2016a) proposed importance sampling algorithms to compute DIC by integrating out the latent variables; that is, the DIC based on the observed data likelihood. The authors showed in a Monte Carlo study that indeed the observed data DIC was able to select the correct model. Following Chan and Grant (2016a) approach, this article carries out the four models comparison exercise using the observed data DICs.

Another popular metric for Bayesian model comparison is the Bayes factor; it is defined as a ratio of marginal likelihoods. That is, given the likelihood function $\mathcal{L}(y|\theta_k, M_k)$ of a model M_k and its prior density $\mathcal{L}(\theta_k|M_k)$, the Bayes factor in favour of Model M_i against M_j is

$$\mathrm{BF}_{ij} = \frac{\mathcal{L}(y|M_i)}{\mathcal{L}(y|M_j)} > 1, \qquad \text{where}$$

$$\mathcal{L}(y|M_k) = \int \mathcal{L}(y|\theta_k, M_k)\mathcal{L}(\theta_k|M_k)d\theta_k \qquad (4)$$

is the marginal likelihood under model M_k, $k = i, j$.

The interpretation of the marginal likelihood is that of the density forecast of the data under model M_k evaluated at the actual observed data y. Therefore, the more likely the observed data are to be under the model, the 'larger' the corresponding marginal likelihood would be. Furthermore, the Bayes factor is a consistent model selection creation (Kass and Raftery 1995). However, one potential drawback of the marginal likelihoods is that they are relatively sensitive to the prior distribution. In addition, their computation is non-trivial; the integral in Equation (4) does not have an analytical solution as it is often high-dimensional. Chan and Grant (2016b) provided an improved approach to compute the marginal likelihoods using an adaptive importance sampling method called the cross-entropy method. It is an importance sampling estimator based on independent draws from convenient distributions. This article employs Chan and Grant (2016b) approach; the model selection criterion results are available from the author upon request.

Out-of-sample volatility forecasting

For the out-of-sample forecasting performance comparison of the four used models, the data is split into two parts: the training set which includes 25-year sample data (estimation sample: 1988: Q1–2013:Q4) and 5-year sample data for the test set or validation test (5-year forecast: 2014: Q1–2018:Q4). The prediction procedure starts with the estimation of each model using the training data set. Next, the estimated models are used to build the one-step-ahead (quarter) volatility forecasts. Finally, the predicted volatility ($\hat{\sigma}^2$) is compared to the proxy of the true volatility (σ^2).

By nature, true volatility is unobserved, and its appropriate proxy to use in the evaluation of the forecasting performance of different models remains the centre of active ongoing debate. Although most studies such as Brailsford and Faff (1996), Brooks and Persands (2002), and Sadorsky (2006) have employed the squared return as a proxy of σ^2, the realized volatility (RV) has been recognized as the natural benchmark against which to quantify volatility forecasts since it provides a consistent non-parametric estimate of the variability of the asset price over a given discrete period. The point which was first pointed out by Andersen and Bollerslev, in their (1998a)'s work which was further developed by Andersen, Bollerslev, and Lange (1999); (2003); 2004) and Patton (2007). Recently, in the stock market, the use of available intraday data and realized daily volatility had been praised for providing better forecast accuracy (Xingyi and Zakamulin 2018). In the housing market, σ^2 is also proxied by realized volatility calculated from the asset returns, as employed by Zhou and Kang (2011). Following this study, a proxy of the true volatility used in this article is the released volatility constructed as a rolling sample. Moreover, following other studies on conditional volatility forecasting, the forecasting accuracy of the studied models is measured using two popular measures; the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE). The two criteria are defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\sigma}_i^2 - \sigma_i^2\right)^2} \quad \text{and MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{\sigma}_i^2 - \sigma_i^2\right|,$$

where N is the number of forecasts, $\hat{\sigma}^2$ is the forecast volatility, and σ^2 is the true volatility.

III. Results and discussions

In-sample fit analysis

For cities and sub-areas with significant clustering effects in each apartment category, all four stochastic models are estimated using the Bayesian approach. The estimated observed data DICs and their standard errors are reported in Table 2–4. Various conclusions can be drawn from this model comparison exercise.

Overall, in all three apartment types, the SVl model ranks as the best model for modelling the Finnish house price volatility. In the one–room flats category, out of 28 cities/sub-areas exhibiting ARCH effects, SVl model comes on top in 19. In two–room flats category, SVl model leads in 24 cities/sub–areas out of 27;, and in the more than three rooms flats category, SVl comes on top in 20 cities/sub–areas out of 31. These results are in line with the general finding that asymmetric volatility (leverage effect and volatility feedback effect) is a crucial component in modelling assets returns. The results are also consistent with the findings of Dufitinema (2020) who documented, using the GARCH-type framework, the evidence of leverage effects in the price volatility of the studied types of apartment.

Next, the SV-2 model interchanges with the SVl and takes the first place. This pattern is observed in eight cities/sub-areas in the one-room flats category, in three cities/sub-areas in the two-room flats category, and in nine cities/sub-areas in the more than three rooms flats category. The exceptions of this general pattern are Oulu-area1 in the one-room apartments, Helsinki-city and Vassa-area1 in the more than three rooms apartments. In both sub–areas (Oulu and Vassa), the heavy-tailed model (SVt) performs better, followed by

Table 2. Estimated DICs – One room flats.

Regions	Cities/Sub-areas	One room flats				
		SV	SV-2	SVt	SVI	The best model
Helsinki	hki	602.6 (0.94)	603.9 (0.57)	601.6 (0.16)	600.6 (0.19)	SVI
	hki1	687.5 (0.09)	686.0 (0.49)	688.0 (0.30)	685.3 (0.25)	SVI
	hki2	627.7 (0.73)	628.5 (0.53)	627.1 (0.12)	626.0 (0.32)	SVI
	hki4	697.7 (0.23)	700.8 (0.59)	697.9 (0.16)	693.7 (0.27)	SVI
Tampere	tre1	735.6 (0.39)	736.0 (0.75)	734.9 (0.13)	728.3 (0.29)	SVI
	tre3	726.1 (0.82)	718.8 (1.16)	725.7 (0.30)	722.5 (1.27)	SV-2
Turku	tku	711.5 (0.25)	705.3 (1.05)	711.7 (0.12)	708.1 (0.38)	SV-2
	tku1	764.7 (0.29)	764.8 (1.59)	764.9 (0.23)	757.1 (0.44)	SVI
	tku2	728.1 (0.32)	717.6 (2.42)	727.6 (0.21)	724.2 (0.43)	SV-2
	tku3	749.5 (0.38)	742.3 (1.35)	749.3 (0.71)	741.1 (0.52)	SVI
Oulu	oulu	699.8 (0.57)	705.2 (0.19)	702.5 (0.46)	698.4 (0.69)	SVI
	oulu1	748.9 (0.37)	749.2 (0.10)	747.1 (0.86)	759.2 (11.19)	SVt
Lahti	lti	757.9 (0.64)	760.4 (0.26)	757.0 (0.36)	750.0 (0.76)	SVI
	lti1	720.2 (0.20)	717.4 (1.47)	719.8 (0.19)	719.9 (0.37)	SV-2
Jyväskylä	jkla	730.1 (0.24)	729.4 (1.77)	731.9 (0.91)	724.7 (0.17)	SVI
	jkla1	753.6 (0.70)	748.5 (0.71)	753.0 (0.20)	750.4 (0.51)	SV-2
	jkla2	614.9 (0.40)	599.5 (1.02)	614.6 (0.44)	607.5 (0.41)	SV-2
Pori	pori	853.9 (0.39)	851.9 (0.16)	852.8 (0.71)	845.2 (0.54)	SVI
	pori1	717.8 (1.74)	711.2 (0.21)	716.1 (0.72)	710.3 (0.49)	SVI
Kuopio	kuo	695.3 (0.20)	691.7 (0.79)	695.5 (0.09)	687.7 (0.55)	SVI
	kuo1	689.0 (0.07)	682.7 (0.71)	689.3 (0.32)	686.2 (0.25)	SV-2
	kuo2	573.7 (0.25)	570.2 (0.86)	573.6 (0.10)	571.1 (0.54)	SV-2
Joensuu	jnsu1	724.4 (0.94)	722.5 (0.27)	723.7 (0.27)	719.3 (0.73)	SVI
Kouvola	kou	777.3 (0.44)	774.3 (0.52)	778.7 (0.72)	764.4 (0.49)	SVI
Lappeenranta	lrta	725.0 (0.30)	722.0 (0.89)	724.2 (0.41)	718.8 (0.31)	SVI
	lrta1	635.5 (0.59)	632.0 (1.43)	635.8 (0.30)	631.1 (0.27)	SVI
Hämeenlinna	hnlina	787.1 (0.21)	786.4 (0.40)	788.0 (0.64)	780.1 (0.52)	SVI
Kotka	kotka	756.7 (1.29)	755.3 (1.54)	755.8 (0.83)	748.6 (0.60)	SVI

Notes: This table reports, for each city and sub-area, the estimated observed data DICs – the information criterion for model comparison. The preferred model is the one with the minimum DIC value. The standard errors are in parentheses.

Table 3. Estimated DICs – Two-room flats.

Regions	Cities/Sub-areas	Two rooms flats				
		SV	SV-2	SVt	SVI	The best model
Helsinki	hki	583.5 (0.47)	585.7 (1.06)	583.4 (0.42)	581.8 (0.45)	SVI
	hki1	698.8 (0.35)	697.4 (1.05)	698.5 (0.10)	697.9 (0.28)	SV-2
	hki2	601.1 (0.07)	604.3 (1.12)	602.3 (0.13)	599.9 (0.36)	SVI
	hki3	645.7 (0.23)	644.3 (0.43)	646.0 (0.12)	638.1 (0.28)	SVI
	hki4	643.7 (0.27)	645.4 (1.48)	643.9 (0.17)	636.0 (0.36)	SVI
Tampere	tre1	637.0 (0.21)	635.8 (0.70)	637.2 (0.43)	631.2 (0.30)	SVI
	tre2	712.1 (0.58)	710.7 (1.98)	711.0 (0.37)	708.5 (0.30)	SVI
Turku	tku	629.3 (0.43)	630.8 (1.69)	628.6 (0.24)	627.0 (0.18)	SVI
	tku2	713.9 (0.29)	714.7 (1.53)	714.6 (0.30)	710.8 (0.35)	SVI
Lahti	lti	638.8 (0.29)	640.4 (1.36)	639.5 (0.23)	631.4 (0.44)	SVI
Jyväskylä	jkla	630.5 (0.72)	631.4 (1.01)	629.2 (0.30)	622.3 (0.42)	SVI
	jkla1	661.7 (0.25)	661.5 (1.71)	662.7 (0.32)	655.2 (0.30)	SVI
	jkla2	704.6 (0.41)	701.7 (0.42)	703.3 (0.28)	693.5 (0.50)	SVI
Pori	pori	743.2 (0.57)	739.2 (2.03)	743.0 (0.19)	733.8 (0.50)	SVI
	pori1	802.4 (0.43)	801.2 (2.28)	802.8 (0.43)	789.1 (0.35)	SVI
	pori2	787.1 (0.45)	785.8 (0.24)	786.9 (0.31)	780.2 (0.75)	SVI
Kuopio	kuo	640.1 (0.23)	641.4 (1.98)	640.3 (0.40)	638.0 (0.60)	SVI
	kuo1	722.4 (0.14)	719.1 (0.92)	722.4 (0.12)	716.6 (0.46)	SVI
Joensuu	jnsu1	761.2 (0.16)	758.5 (2.15)	761.5 (0.27)	757.7 (0.10)	SVI
Seinäjoki	seoki	750.8 (0.39)	743.0 (1.22)	751.0 (0.40)	746.9 (0.54)	SV-2
Vaasa	vaasa	689.0 (0.34)	689.7 (0.52)	690.5 (0.57)	685.4 (0.25)	SVI
Kouvola	kou	767.6 (0.29)	761.7 (1.93)	765.7 (0.36)	759.8 (0.57)	SVI
Lappeenranta	lrta	680.1 (0.43)	680.6 (0.22)	679.4 (0.59)	677.1 (0.20)	SVI
	lrta1	756.4 (0.33)	753.5 (0.15)	756.2 (1.05)	750.9 (0.34)	SVI
Hämeenlinna	hnlina	714.3 (0.26)	709.1 (1.07)	715.1 (0.22)	709.2 (0.39)	SV-2
	hnlina1	745.2 (0.91)	741.1 (1.32)	744.1 (0.35)	739.1 (0.36)	SVI
Kotka	kotka1	786.6 (0.88)	784.0 (2.77)	786.9 (0.35)	778.1 (0.16)	SVI

Notes: This table reports, for each city and sub-area, the estimated observed data DICs – the information criterion for model comparison. The preferred model is the one with the minimum DIC value. The standard errors are in parentheses.

Table 4. Estimated DICs – More than three rooms flats.

		Three rooms flats				
Regions	Cities/Sub-areas	SV	SV-2	SVt	SVl	The best model
Helsinki	hki	627.4 (0.15)	631.2 (0.84)	628.0 (0.28)	628.2 (0.79)	SV
	hki1	728.5 (0.14)	728.9 (0.73)	729.2 (0.19)	727.7 (0.55)	SVl
	hki3	649.8 (0.14)	648.3 (0.63)	649.9 (0.12)	646.9 (0.50)	SVl
	hki4	665.4 (0.43)	661.6 (1.81)	664.6 (0.11)	661.2 (0.38)	SVl
Tampere	tre	629.9 (0.65)	631.7 (0.81)	630.1 (0.22)	628.2 (0.20)	SVl
	tre1	713.2 (0.15)	713.2 (1.25)	713.4 (0.17)	710.5 (0.44)	SVl
	tre2	721.9 (0.60)	714.8 (1.93)	722.8 (0.49)	717.5 (0.52)	SV-2
	tre3	617.4 (0.18)	619.4 (1.09)	617.3 (0.42)	614.2 (0.26)	SVl
Turku	tku	676.3 (0.12)	673.7 (1.12)	676.6 (0.34)	671.0 (0.34)	SVl
	tku1	757.3 (0.15)	756.1 (2.07)	757.7 (0.23)	753.3 (0.52)	SVl
	tku2	725.3 (0.40)	725.9 (2.41)	724.3 (0.39)	722.2 (0.43)	SVl
	tku3	706.1 (0.25)	706.6 (0.90)	706.9 (0.21)	702.3 (0.74)	SVl
Oulu	oulu	658.7 (0.17)	658.0 (1.00)	659.8 (0.15)	656.0 (0.31)	SVl
	oulu1	716.9 (0.20)	715.0 (1.18)	717.7 (0.23)	713.8 (0.62)	SVl
Lahti	lti	710.3 (0.16)	711.7 (0.16)	710.6 (0.72)	701.8 (0.51)	SVl
	lti1	769.6 (0.40)	767.7 (0.12)	770.8 (0.23)	762.2 (0.58)	SVl
Jyväskylä	jkla	709.5 (0.59)	703.8 (0.25)	710.5 (0.20)	706.2 (0.26)	SV-2
	jkla1	730.1 (0.40)	725.4 (2.05)	730.7 (0.45)	725.0 (0.63)	SVl
	jkla2	787.1 (0.44)	785.2 (0.11)	787.2 (0.35)	781.6 (0.28)	SVl
Pori	pori1	768.6 (0.94)	762.1 (2.18)	769.9 (0.64)	761.1 (0.41)	SVl
Kuopio	kuo	703.2 (0.08)	700.0 (0.42)	703.4 (0.16)	701.1 (0.34)	SV-2
	kuo1	754.2 (0.22)	746.2 (0.83)	754.9 (0.39)	751.1 (0.49)	SV-2
	kuo2	719.2 (0.33)	714.7 (1.32)	717.8 (0.17)	716.5 (0.22)	SV-2
Seinäjoki	seoki	697.2 (0.31)	686.6 (0.19)	697.1 (0.52)	691.0 (0.48)	SV-2
Vaasa	vaasa	744.2 (0.55)	743.5 (0.20)	744.9 (0.39)	737.9 (0.30)	SVl
	vaasa1	737.3 (1.04)	737.6 (0.13)	737.2 (0.90)	740.2 (0.41)	SVt
	vaasa2	544.1 (0.26)	536.9 (1.88)	544.7 (0.25)	542.8 (0.25)	SV-2
Lappeenranta	lrta	749.5 (0.38)	747.8 (0.10)	749.7 (0.33)	743.3 (0.29)	SVl
	lrta2	511.7 (0.19)	500.1 (1.20)	511.0 (0.10)	510.8 (0.12)	SV-2
Hämeenlinna	hnlina1	727.9 (0.51)	718.5 (0.15)	727.7 (0.20)	720.8 (0.49)	SV-2
Kotka	kotka	778.1 (0.14)	773.0 (1.03)	778.9 (0.33)	770.9 (0.38)	SVl

Notes: This table reports, for each city and sub-area, the estimated observed data DICs – the information criterion for model comparison. The preferred model is the one with the minimum DIC value. The standard errors are in parentheses.

the Vanilla SV; whereas in the Helsinki–city the model performance rank is the other way around.

Finally, to further investigate the features that are vital in modelling the Finnish house price volatility dynamics; the vanilla SV and SV-2 model are compared. In doing so, the question of whether the AR(2) component is a useful addition to the vanilla SV model is also answered. As it can be observed in the one–room flats category where the SV-2 model outperforms the vanilla SV in 20 out 28 cities/sub–areas; the richer AR(2) volatility process provides significant benefits. In the two–room flats category, the SV-2 performs better than SV in 17 cities/sub-areas out of 27, and in 22 out of 31 in the more than three rooms flats category. Although the SV-2 general excel in comparison to the vanilla SV, cautions should be taken when modelling house prices volatility of individual regions. As it can be noted, the performance of the two models differs across cities and sub-areas, and by apartment types – no geographical pattern is observed. Therefore, retaining the standard specification of an AR(1) volatility process or adding a component depends on the house price dataset under study.

In summary, the stochastic volatility model with leverage effect is the best model for modelling the house prices volatility of most of the Finnish cities and sub-areas. In the rest of the regions, the SVl swaps places with the SV model where the latent volatility follows a stationary AR (2) process. In a few cases, the second place is less clear-cut; the vanilla and the heavy-tailed SV models share the ranking. However, again as above, the model performance differs from region to region. Therefore, when modelling house price, even by employing the SV framework, one has to enable different house price dynamics across cities and sub-areas; rather than imposing one SV model on the whole dataset. As it has been stressed in various studies, such as Milles (2011b) and Begiazi and Katsiampa (2019) that house prices present a heterogeneous dynamics across different areas and property types.

Out-of-sample volatility forecasting

Since the model that performs better in-sample does not necessarily imply that it will provide accurate forecasts, the out-of-sample forecast performance of the four competing models is investigated. The procedure starts by estimating the models using the training dataset, build 5-year volatility forecasts in terms of one step ahead, and validate the constructed predictions using the test dataset. For each city and sub-area in each apartment category, Tables 5–7 report the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE); the measures used in assessing the forecasting accuracy for each model. The lower the value of the two criteria, the better the model's forecasting performance.

Overall, in all three apartment types, both evaluation criteria rank the heavy-tailed stochastic volatility model (SVt) as the best model. Especially in the two rooms and more than three rooms flats categories, where the SVt model provides the best forecasts in, respectively, 17 out on 27 and 18 out of 31 cities/sub-areas. In the one-room flats category, the SVt and SVl models are neck and neck; they forecast best in, respectively, 9 and 10 out of 28 cities/sub-areas. These results confirm again, the importance of the heavy-tailed distributions not only in modelling but also in forecasting assets volatility. Moreover, as it has been found in other assets such as stocks (Nakajima and Omori 2009; Chan and Grant 2016a), even in the SV framework, when the heavy-tailed distribution is employed, it provides the model with extra flexibility against misspecification and outlier. The same conclusion can also be drawn in the case of house prices, where the SVt outperforms the SV model with standard errors.

A geographical pattern is observed in some regions where, in all three apartment types, the same model performs well in producing accurate forecasts. In Helsinki-city, Helsinki-area1 and Kuopio-city, the SVt is the first–ranked model across all apartment types, whereas the SVl comes on top in Pori-area1. These results imply that, in addition to the volatility clustering, the returns distributions of the former regions in all three apartments types are characterized by skewness and heavy-tailedness. While in the latter area, the returns' major characteristic is leverage effect; a drop in apartment price causes an increase in house price volatility.

Regarding, the forecasting performance of the vanilla SV in comparison to the SV-2 model, unlike in the in-sample fit analysis where the SV-2 general excel; for the out-of-sample forecasting assessment, the vanilla SV model outperforms the SV-2 in most of the regions. Plus precisely, the vanilla SV does better in approximately 64% (18 out of 28) in the one-room apartments category; in 59% (16 out of 27) in the two-room apartments category; and in 52% (16 out of 31) in the more than three rooms apartments category. Thus, for forecasting the house prices at least, one can feel comfortable retaining the standard specification of an AR(1) volatility process. However, as there is no geographical pattern observed, the same as discussed above, cautions should be taken when forecasting house prices volatility of individual regions.

In summary, indeed, a model that performs well in the in-sample analysis may not provide accurate out-of-sample forecasts. The heavy-tailed stochastic volatility model is the best model for forecasting the house prices volatility of most of the Finnish cities and sub-areas. On the second place comes the stochastic volatility model with leverage effect, while the vanilla SV and SV-2 models share the last two rankings. Moreover, apart from a few areas (two cities and two sub–areas), no geographical pattern is observed in all three apartment types; the models' forecasting performances vary across cities and sub-areas, and by apartment types.

IV. Conclusions, implications and further research

Volatility forecasting is one of the most fundamental methodologies in financial economics as it is a vital tool for asset allocation in general, and specifically for investors who implement volatility targeting. This article assesses the in-sample fit and the out-of-sample forecasting performance of four stochastic volatility models in the Finnish housing market. The competing models are the vanilla SV, the SV model where the latent volatility follows a stationary AR(2) process, the heavy-tailed SV and the SV with leverage effects. The study uses

Table 5. The results of RMSE and MAE – One room flats.

Regions	Cities/Sub-areas	One room flats SV		SV-2		SVt		SVI		
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	The best model
Helsinki	hki	0.0152	0.0142	0.0167	0.0156	0.0148	0.0138	0.0157	0.0147	SVt
	hki1	0.0228	0.0195	0.0232	0.0198	0.0218	0.0187	0.0227	0.0194	SVt
	hki2	0.0169	0.0153	0.0178	0.0161	0.0166	0.0149	0.0173	0.0156	SVt
	hki4	0.0186	0.0152	0.0189	0.0154	0.0182	0.0148	0.0191	0.0155	SVt
Tampere	tre1	0.0351	0.0304	0.0351	0.0302	0.0352	0.0304	0.0350	0.0300	SVI
	tre3	0.0580	0.0435	0.0570	0.0420	0.0583	0.0435	0.0572	0.0441	SV-2
Turku	tku	0.0277	0.0249	0.0256	0.0231	0.0264	0.0238	0.0278	0.0249	SV-2
	tku1	0.0384	0.0332	0.0385	0.0334	0.0381	0.0324	0.0384	0.0333	SVt
	tku2	0.0338	0.0253	0.0335	0.0252	0.0333	0.0254	0.0332	0.0252	SVI
	tku3	0.0412	0.0362	0.0418	0.0367	0.0411	0.0359	0.0417	0.0367	SVt
Oulu	oulu	0.0357	0.0242	0.0360	0.0241	0.0359	0.0240	0.0357	0.0239	SVI
	oulu1	0.0494	0.0359	0.0497	0.0360	0.0501	0.0362	0.0495	0.0359	SV
Lahti	lti	0.0550	0.0394	0.0554	0.0394	0.0555	0.0394	0.0548	0.0393	SVI
	lti1	0.1657	0.1277	0.1664	0.1281	0.1674	0.1288	0.1655	0.1275	SVI
Jyväskylä	jkla	0.0337	0.0281	0.0332	0.0280	0.0339	0.0282	0.0337	0.0281	SV-2
	jkla1	0.0372	0.0336	0.0372	0.0338	0.0369	0.0332	0.0373	0.0337	SVt
	jkla2	0.0739	0.0579	0.0740	0.0580	0.0744	0.0581	0.0739	0.0578	SVI
Pori	pori	0.0619	0.0527	0.0615	0.0524	0.0624	0.0529	0.0617	0.0526	SV-2
	pori1	0.0484	0.0388	0.0483	0.0388	0.0497	0.0388	0.0481	0.0386	SVI
Kuopio	kuo	0.0271	0.0201	0.0272	0.0203	0.0271	0.0198	0.0272	0.0198	SVt
	kuo1	0.0672	0.0423	0.0691	0.0429	0.0693	0.0430	0.0673	0.0424	SV
	kuo2	0.0927	0.0739	0.0929	0.0740	0.0943	0.0751	0.0924	0.0738	SVI
Joensuu	jnsu1	0.0616	0.0372	0.0619	0.0373	0.0623	0.0374	0.0626	0.0379	SV
Kouvola	kou	0.0549	0.0411	0.0549	0.0411	0.0549	0.0407	0.0552	0.0412	SVt
Lappeenranta	lrta	0.0388	0.0316	0.0387	0.0314	0.0390	0.0315	0.0389	0.0319	SV-2
	lrta1	0.0459	0.0397	0.0461	0.0398	0.0464	0.0398	0.0461	0.0399	SV
Hämeenlinna	hnlina	0.0422	0.0311	0.0424	0.0312	0.0428	0.0313	0.0421	0.0310	SVI
Kotka	kotka	0.0289	0.0240	0.0290	0.0241	0.0292	0.0243	0.0288	0.0240	SVI

Notes: This table reports the performance of the four competing models in forecasting the house price volatility. The training set is 1988:*Q*1–2013:*Q*4, while the test set is 2014:*Q*1–2018:*Q*4.
RMSE is Root Mean Squared Error and MAE is the Mean Absolute Error.

Table 6. The results of RMSE and MAE – Two-room flats.

Regions	Cities/Sub-areas	Two rooms flats SV		SV-2		SVt		SVI		
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	The best model
Helsinki	hki	0.0107	0.0091	0.0109	0.0093	0.0106	0.0090	0.0112	0.0096	SVt
	hki1	0.0182	0.0144	0.0183	0.0145	0.0179	0.0142	0.0191	0.0151	SVt
	hki2	0.0106	0.0093	0.0105	0.0092	0.0103	0.0090	0.0106	0.0092	SVt
	hki3	0.0182	0.0156	0.0176	0.0153	0.0179	0.0155	0.0183	0.0156	SV-2
	hki4	0.0227	0.0204	0.0222	0.0200	0.0220	0.0198	0.0223	0.0201	SVt
Tampere	tre1	0.0227	0.0213	0.0233	0.0219	0.0224	0.0210	0.0219	0.0204	SVI
	tre2	0.0247	0.0216	0.0241	0.0209	0.0239	0.0207	0.0252	0.0221	SVt
Turku	tku	0.0144	0.0126	0.0145	0.0127	0.0136	0.0118	0.0147	0.0129	SVt
	tku2	0.0309	0.0284	0.0308	0.0283	0.0302	0.0276	0.0306	0.0281	SVt
Lahti	lti	0.0176	0.0153	0.0178	0.0153	0.0179	0.0153	0.0177	0.0153	SV
Jyväskylä	jkla	0.0219	0.0146	0.0218	0.0146	0.0222	0.0148	0.0215	0.0143	SVI
	jkla1	0.0210	0.0158	0.0208	0.0158	0.0209	0.0157	0.0211	0.0160	SV-2
	jkla2	0.0648	0.0400	0.0653	0.0398	0.0652	0.0397	0.0647	0.0401	SVI
Pori	pori	0.0443	0.0339	0.0444	0.0340	0.0444	0.0340	0.0442	0.0339	SVI
	pori1	0.0572	0.0432	0.0574	0.0433	0.0578	0.0435	0.0569	0.0429	SVI
	pori2	0.0396	0.0356	0.0398	0.0358	0.0383	0.0345	0.0399	0.0358	SVt
Kuopio	kuo	0.0176	0.0151	0.0179	0.0154	0.0175	0.0151	0.0181	0.0155	SVt
	kuo1	0.0224	0.0197	0.0226	0.0198	0.0222	0.0196	0.0225	0.0198	SVt
Joensuu	jnsu1	0.0288	0.0256	0.0285	0.0253	0.0270	0.0239	0.0286	0.0255	SVt
Seinäjoki	seoki	0.0376	0.0321	0.0374	0.0319	0.0373	0.0318	0.0375	0.0320	SVt
Vaasa	vaasa	0.0192	0.0159	0.0199	0.0168	0.0188	0.0156	0.0194	0.0162	SVt
Kouvola	kou	0.0802	0.0474	0.0802	0.0474	0.0807	0.0474	0.0801	0.0474	SVI
Lappeenranta	lrta	0.0255	0.0223	0.0251	0.0220	0.0245	0.0214	0.0256	0.0224	SVt
	lrta1	0.0301	0.0270	0.0300	0.0269	0.0295	0.0260	0.0302	0.0272	SVt
Hämeenlinna	hnlina	0.0278	0.0246	0.0279	0.0247	0.0274	0.0237	0.0277	0.0244	SVt
	hnlina1	0.0328	0.0284	0.0330	0.0288	0.0324	0.0277	0.0329	0.0285	SVt
Kotka	kotka1	0.0698	0.0579	0.0699	0.0581	0.0705	0.0584	0.0702	0.0583	SV

Notes: This table reports the performance of the four competing models in forecasting the house price volatility. The training set is 1988:*Q*1–2013:*Q*4, while the test set is 2014:*Q*1–2018:*Q*4.
RMSE is Root Mean Squared Error and MAE is the Mean Absolute Error.

Table 7. The results of RMSE and MAE – More than three rooms flats.

		Three rooms flats								
Regions	Cities/Sub-areas	SV		SV-2		SVt		SVl		
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	The best model
Helsinki	hki	0.0205	0.0185	0.0206	0.0186	0.0199	0.0180	0.0207	0.0187	SVt
	hki1	0.0243	0.0197	0.0241	0.0194	0.0237	0.0191	0.0247	0.0199	SVt
	hki3	0.0178	0.0153	0.0181	0.0157	0.0176	0.0151	0.0183	0.0159	SVt
	hki4	0.0201	0.0174	0.0194	0.0166	0.0196	0.0168	0.0204	0.0176	SV-2
Tampere	tre	0.0216	0.0198	0.0213	0.0194	0.0208	0.0189	0.0217	0.0199	SVt
	tre1	0.0251	0.0224	0.0255	0.0228	0.0241	0.0212	0.0253	0.0226	SVt
	tre2	0.0506	0.0426	0.0504	0.0423	0.0503	0.0421	0.0505	0.0425	SVt
	tre3	0.0186	0.0152	0.0182	0.0146	0.0181	0.0145	0.0191	0.0161	SVt
Turku	tku	0.0194	0.0148	0.0195	0.0150	0.0195	0.0152	0.0196	0.0150	SV
	tku1	0.0290	0.0257	0.0291	0.0258	0.0289	0.0256	0.0291	0.0259	SVt
	tku2	0.0311	0.0263	0.0312	0.0263	0.0310	0.0261	0.0311	0.0264	SVt
	tku3	0.0358	0.0279	0.0359	0.0282	0.0357	0.0277	0.0358	0.0279	SVt
Oulu	oulu	0.0179	0.0157	0.0177	0.0155	0.0173	0.0152	0.0181	0.0159	SVt
	oulu1	0.0255	0.0228	0.0255	0.0229	0.0247	0.0220	0.0262	0.0235	SVt
Lahti	lti	0.0277	0.0237	0.0276	0.0236	0.0275	0.0234	0.0278	0.0238	SVt
	lti1	0.0317	0.0270	0.0311	0.0263	0.0309	0.0261	0.0320	0.0273	SVt
Jyväskylä	jkla	0.0212	0.0183	0.0218	0.0189	0.0210	0.0181	0.0213	0.0184	SVt
	jkla1	0.0281	0.0238	0.0278	0.0235	0.0268	0.0227	0.0282	0.0238	SVt
	jkla2	0.0526	0.0389	0.0522	0.0387	0.0536	0.0397	0.0524	0.0388	SV-2
Pori	pori1	0.0766	0.0536	0.0767	0.0536	0.0772	0.0538	0.0765	0.0535	SVl
Kuopio	kuo	0.0278	0.0246	0.0277	0.0246	0.0271	0.0237	0.0279	0.0247	SVt
	kuo1	0.0358	0.0329	0.0353	0.0322	0.0354	0.0323	0.0355	0.0324	SV-2
	kuo2	0.0514	0.0400	0.0512	0.0400	0.0521	0.0405	0.0513	0.0400	SV-2
Seinäjoki	seoki	0.0425	0.0350	0.0435	0.0361	0.0438	0.0366	0.0423	0.0346	SVl
Vaasa	vaasa	0.0341	0.0275	0.0339	0.0271	0.0340	0.0271	0.0341	0.0276	SV-2
	vaasa1	0.0392	0.0299	0.0396	0.0300	0.0398	0.0301	0.0395	0.0301	SV
	vaasa2	0.0300	0.0279	0.0301	0.0277	0.0301	0.0281	0.0299	0.0275	SVl
Lappeenranta	lrta	0.0348	0.0296	0.0350	0.0297	0.0349	0.0297	0.0349	0.0297	SV
	lrta2	0.0178	0.0157	0.0185	0.0164	0.0206	0.0189	0.0171	0.0150	SVl
Hämeenlinna	hnlina1	0.0424	0.0370	0.0423	0.0370	0.0422	0.0368	0.0424	0.0371	SVt
Kotka	kotka	0.0573	0.0387	0.0572	0.0386	0.0578	0.0391	0.0574	0.0386	SV-2

Notes: This table reports the performance of the four competing models in forecasting the house price volatility. The training set is 1988:*Q*1–2013:*Q*4, while the test set is 2014:*Q*1–2018:*Q*4.
RMSE is Root Mean Squared Error and MAE is the Mean Absolute Error.

quarterly house price indices from 1988:Q1 to 2018:Q4, for 15 main regions in Finland.

The study has various findings. First, in all three apartment types, the stochastic volatility model with leverage effect ranks as the best model for modelling the Finnish house price volatility; indicating that leverage effect is a crucial component in modelling house price returns. Second, in most of the regions, the heavy-tailed stochastic volatility model excels in forecasting the house price volatility of the studied types of apartments, indicating that the skewness and the heavy-tailedness characteristics are vital components in forecasting house price volatility. Moreover, results suggest that the *t* innovations component is a useful addition to the vanilla SV model. Third, for the in-sample fit analysis, the AR(2) component is found to be a valuable addition to the vanilla SV, whereas, for the out-of-sample forecasting assessment, the vanilla SV model outperforms the SV-2 in most of the regions. Last, except for two cities and two sub-areas, no geographical pattern is observed for the models' out-of-sample forecasting performances in all three apartment types. Their performances vary across cities and sub-areas, and by apartment types.

The findings have some housing investment implications. As housing investors, policymakers and consumers are recommended to monitor the asset volatility; accurate forecasts help to improve portfolio diversifications across Finland and by apartment type. In addition, in the viewpoint of volatility as a measure of risk, precise predictions are the key to assessing investment risks; an essential decision-making factor for foreign as well as domestic investors who dominate the Finnish housing market.

In the standpoint of establishing suitable time-series volatility forecasting models of this housing market; these study findings – the performance of the four stochastic models – will be weighed up to their GARCH models counterparts. One reason is that Dufitinema and Pynnönen (2020) have found, in all three apartment types, evidence of long-range dependence in the returns and volatility for the majority of cities and sub-areas. The long memory present in the housing market returns suggests that the asset is forecastable on a long horizon, whereas the evidence of long-range dependence in the housing market volatility is the key to establish suitable time-series volatility forecasting models for the market. The other reason is that Dufitinema (2020) employed the Exponential GARCH (EGARCH) model to investigate whether the asymmetric effects of shocks are noted in the Finnish house price volatility. The author found that, indeed, these asymmetric impacts of shocks are observed in all three studied apartment types. Therefore, to assess whether the deterministic conditional variance under GARCH or the unobserved time-varying volatility under SV is more favoured by the house price data; these study outcomes will be compared to the performance of the short memory and long memory GARCH-type models. Namely, the EGARCH model, the Component GARCH (CGARCH) model and the Fractionally Integrated GARCH (FIGARCH) model. The aim is to provide to the investors, risk managers and consumers enlightenments with regards to which forecasting approach delivers accurate and superior volatility forecasts of the apartment types under study.

Moreover, it would also be of interest to incorporate, in a multivariate analysis, macroeconomic factors such as interest rates and unemployment rates; as the interaction between these variables and house prices is often of interest. Additionally, several studies have referred to the importance of spatial dependence in regional housing markets known as 'the ripple effect'. The phenomenon refers to the house prices' tendency to rise first in the part of the country during an upswing and to gradually spread out or 'ripple out' across the country. Meen, 1999 was the fisrt to provide convincing economic explanations for the ripple effect, and by utilizing different approaches, many studies have contributed to the discussions of the spatial interaction of regional house prices. Among the methods used to detect the ripple effect includes tests of cointegration (Alexander and Barrow 1994), the concept of absolute and conditional convergence (Chow, Fung, and Cheng 2016), a measure of the regional–national return spillover indices through Vector Autoregressive (VAR) model (Tsai 2015), and use of time-series volatility models (Morley and Thomas 2011; Lin and Fuerst 2014). Therefore, following Morley and Thomas and Lin and Fuerst, and using the current study outcomes, the analysis of the spatial spillover in the Finnish housing market is also subjected to future research. That is, as the stochastic volatility model with leverage effect (SVl) has been ranked as the best model for modelling the Finnish house price volatility of most of the regions. The ripple effects will be allowed in the model by incorporating house prices of the most populated area – the Helsinki region – as highlighted by the above-cited studies that the most populated area in a country may be a leading factor to influence the rest of the housing markets.

Furthermore, it would be worth investigating the structural breaks in the studied housing market. For instance, as discussed earlier, during the period of the end of 1980s to mid-1993, house prices in Finland experienced a structural break due to the financial market deregulation. By examining the occurrence of structural breakpoints, the full sample data can be divided into subsamples based on the estimated break dates, and hence improve forecast accuracy.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This work was supported by the Foundation for Economic Education (Liikesivistyrahasto) [16-8653].

ORCID

Josephine Dufitinema http://orcid.org/0000-0001-6716-1229

References

Alexander, C., and M. Barrow. 1994. "Seasonality and Cointegration of Regional House Prices in the UK".

Urban Studies. 31: 1667–1689. doi:10.1080/00420989420081571.

Andersen, T. G., and T. Bollerslev. 1998a. "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts". *International Economic Review*. 39 (4):885–905. doi:10.2307/2527343.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. 2003. "Modeling and Forecasting Realized Volatility". *Econometrica*. 71:579–625. doi:10.1111/1468-0262.00418.

Andersen, T. G., T. Bollerslev, and N. Meddahi. 2004. "Analytic Evaluation of Volatility Forecasts". *International Economic Review*. 45:1079–1110. doi:10.1111/j.0020-6598.2004.00298.x.

Andersen, T. G., T. Bollerslev, and S. Lange. 1999. "Forecasting Financial Market Volatility: Sample Frequency Vis–á–vis Forecast Horizon". *Journal of Empirical Finance*. 6:457–477. doi:10.1016/S0927-5398(99)00013-4.

Apergis, N., and J. E. Payne. 2020. "Modeling the Time Varying Volatility of Housing Returns: Further Evidence from the U.S Metropolitan Condominium Markets". *Review of Financial Economics*. 38(1):24–33. doi:10.1002/rfe.1063.

Begiazi, K., and P. Katsiampa. 2019. "Modelling U.K House Prices with Structual Breaks and Conditional Variance Analysis". *Journal of Real Estate Finance and Economics*. 58:290–309. doi:10.1007/s11146-018-9652-5.

Black, F. 1976. "Studies of Stock Market Volatility Changes", *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section*, Boston, MA; pp. 177–181.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroscedasticity". *Journal of Econometrics*. 31 (3):307–327. doi:10.1016/0304-4076(86)90063-1.

Bollerslev, T., R. Y. Chou, and K. F. Kroner. 1992. "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence". *Journal of Econometrics*. 52:1–2, 5–59. doi:10.1016/0304-4076(92)90064-X.

Bos, C. S. 2012. "Relating Stochastic Volatility Estimation Methods", *In* L Bauwens, C Hafner, S Laurent (*Ed.s*), *Handbook of Volatility Models and Their Applications*, 147–174. John Wiley & Sons. doi:10.1002/9781118272039.ch6.

Brailsford, T. J., and R. W. Faff. 1996. "An Evaluation of Volatility Forecasting Techniques". *Journal of Banking and Finance*. 20:307–327. doi:10.1016/0378-4266(95)00015-1.

Brooks, C., and G. Persands. 2002. "Model Choice and Value–at–risk Performance". *Financial Analysts Journal*. 58:87–97. doi:10.2469/faj.v58.n5.2471.

Case, K. E., J. M. Quigley, and R. J. Shiller. 2013. "Wealth Effects Revisited 1975–2012". *Critical Finance Review*. 2:101–128. doi:10.1561/104.00000009.

Chan, J. C. 2013. "Moving Average Stochastic Volatility Models with Application to Inflation Forecast". *Journal of Econometrics*. 176(2):162–172. doi:10.1016/j.jeconom.2013.05.003.

Chan, J. C., and A. L. Grant. 2016b. "Modeling Energy Price Dynamics: GARCH versus Stochastic Volatility". *Energy Economics*. 54:182–189. doi:10.1016/j.eneco.2015.12.003.

Chan, J. C., and I. Jeliazkov. 2009. "Efficient Simulation and Integrated Likelihood Estimation in State Space Models". *International Journal of Mathematical Modelling and Numerical Optimisation*. 1(1):101–120. doi:10.1504/IJMMNO.2009.030090.

Chan, J. C. C., and A. L. Grant. 2016a. "On the Observed–Data Deviance Information Criterion for Volatility Modeling". *Journal of Financial Econometrics*. 14(4):772–802. doi:10.1093/jjfinec/nbw002.

Chow, W. W., M. K. Fung, and A. C. S. Cheng. 2016. "Convergence and Spillover of House Prices in Chinese Cities". *Applied Economics*. 48(51):4922–4941. doi:10.1080/00036846.2016.1167829.

Christie, A. A. 1982. "The Stochastic Behavior of Common Stock Variances–Value, Leverage and Interest Rate Effects". *Journal of Financial Economics*. 10:407–432. doi:10.1016/0304-405X(82)90018-6.

Crawford, G. W., and M. C. Fratantoni. 2003. "Assessing the Forecasting Performance of Regime–Switching, ARIMA and GARCH Models of House Prices". *Real Estate Economics*. 31(2):223–243. doi:10.1111/1540-6229.00064.

Delatola, E. I., and J. E. Griffin. 2011. "Bayesian Nonparametric Modeling of the Return Distribution with Stochastic Volatility". *Bayesian Analysis*. 6: 901–926. doi:10.1214/11-BA632.

Dolde, W., and D. Tirtiroglu. 1997. "Temporal and Spatial Information Diffusion in Real Estate Price Changes and Variances". *Real Estate Economics*. 25(4):539–565. doi:10.1111/1540-6229.00727.

Dolde, W., and D. Tirtiroglu. 2002. "Housing Price Volatility Changes and Their Effects". *Real Estate Economics*. 30:41–66. doi:10.1111/1540-6229.00029.

Dufitinema, J. 2020. "Volatility Clustering, Risk–return Relationship and Asymmetric Adjustment in the Finnish Housing Market". *International Journal of Housing Markets and Analysis*. 13(4):661–688. doi:10.1108/IJHMA-12-2019-0125.

Dufitinema, J., and S. Pynnönen. 2020. "Long–range Dependence in the Returns and Volatility of the Finnish Housing Market". *Journal of European Real Estate Research*. 13(1):29–50. doi:10.1108/JERER-07-2019-0019.

Engle, R. F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation". *Econometrica*. 50(4):987–1007. doi:10.2307/1912773.

Gysels, E., A. C. Harvey, and E. Renault. 1996. "Stochastic Volatiltiy", *In* GS Maddala, CR Rao (*Eds.*), *Handbook of Statistics*, 14, 119–191: Elsevier. doi:10.1016/s0169-7161(96)14007-4.

Harvey, A. C., E. Ruiz, and N. Shephard. 1994. "Multivariate Stochastic Variance Models". *The Review of Economic Studies*. 61(2):247–264. doi:10.2307/2297980.

Harvey, A. C., and N. Shephard. 1996. "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns". *Journal of Business & Economic Statistics*. 14 (4):429–434. doi:10.1080/07350015.1996.10524672.

Hossain, B., and E. Latif. 2009. "Determinants of Housing Price Volatility in Canada: A Dynamic Analysis". *Applied*

Economics. 41(27):3521–3531. doi:10.1080/00036840701522861.

Jaquier, E., N. G. Polson, and P. E. Rossi. 1994. "Bayesian Analysis of Stochastic Volatility Models". *Journal of Business & Economic Statistics*. 20(1):69–87. doi:10.1080/07350015.1994.10524553.

Jaquier, E., N. G. Polson, and P. E. Rossi. 2004. "Bayesian Analysis of Stochastic Volatility Models with Fat–Tails and Correlated Errors". *Journal of Econometrics*. 122(1):185–212.

Jensen, M. J., and J. M. Maheu. 2010. "Bayesian Semiparametric Stochastic Volatility Modeling". *Journal of Econometrics*. 157(2):306–316.

Jensen, M. J., and J. M. Maheu. 2014. "Estimating a Semiparametric Asymmetric Stochastic Volatility Model with a Dirichlet Process Mixture". *Journal of Econometrics*. 178(3):523–538.

Kass, R. E., and A. E. Raftery. 1995. "Bayes Factors". *Journal of the American Statistical Association*. 90(430):773–794.

Kim, S., Shephard, N. And Chib, S. 1998. "Stochastic Volatility: Likelihood Inference and Comparison with ARCH models". *The Review of Economic Studies*. 65 (3):361–393 doi:10.1111/1467-937x.00050.

KTI Autumn, 2019. "KTI Market Review", *Technical report*, KTI Property Information .

Lee, C. L. 2009. "Housing Price Volatility and Its Determinants". *International Journal of Housing Markets and Analysis*. 2 (3):293–308. doi:10.1108/17538270910977572.

Lee, C. L., and R. Reed. 2014b. "Volatility Decomposition of Australian Housing Prices". *Journal of Housing Research*. 23(1):21–43. doi:10.1080/10835547.2013.12092084.

Li, K.-W. 2012. "A Study on the Volatility Forecast of the US Housing Market in the 2008 Crisis". *Applied Financial Economics*. 22(22):1869–1880. doi:10.1080/09603107.2012.687096.

Li, Y., T. Zeng, and Y. Yu 2012. "Robust Deviance Information Criterion for Latent Variable Models", *SMU Economics and Statistics Working Paper Series*.

Lin, P.-T., and F. Fuerst. 2014. "Volatility Clustering, Risk–return Relationship, and Asymmetric Adjustment in the Canadian Housing Market". *Journal of Real Estate Portfolio Management*. 20(1):37–46.

Meen, G. 1999. "Regional House Prices and the Ripple Effect: A New Interpretation". *Housing Studies*. 14(6):733–753. doi:10.1080/02673039982524.

Millar, R. B. 2009. "Comparison of Hierarchical Bayesian Models for Overdispersed Count Data Using DIC and Bayes Factors". *Biometrics*. 65:962–969. doi:10.1111/j.1541-0420.2008.01162.x.

Miller, N., and L. Peng. 2006. "Exploring Metropolitan Housing Price Volatility". *The Journal of Real Estate Finance and Economics*. 33(1):5–18. doi:10.1007/s11146-006-8271-8.

Milles, W. 2008a. "Boom–bust Cycles and the Forecasting Performance of Linear and Non-linear Models of House Prices". *Journal of Real Estate Finance and Economics*. 36:249–264. doi:10.1007/s11146-007-9067-1.

Milles, W. 2008b. "Volatility Clustering in US Home Prices". *Journal of Real Estate Research*. 30(1):73–90.

Milles, W. 2011b. "Clustering in UK Home Prices Volatility". *Journal of Housing Research*. 20(1):87–101. doi:10.1080/10835547.2011.12092031.

Morley, B., and D. Thomas. 2011. "Risk–return Relationships and Asymmetric Adjustment in the UK Housing Market". *Applied Financial Economics*. 21(10):735–742. doi:10.1080/09603107.2010.535782.

Nakajima, J., and Y. Omori. 2009. "Leverage, Heavy–tails and Correlated Jumps in Stochastic Volatility Models". *Computational Statistics and Data Analysis*. 53:2535–2553. doi:10.1016/j.csda.2008.03.015.

Nakajima, J., and Y. Omori. 2012. "Stochastic Volatility Model with Leverage and Asymmetrically Heavy–tailed Error Using GH Skew Student's *T* –distribution". *Computational Statistics & Data Analysis*. 56(11):3690–3704.

Oikarinen, E. 2009a. "Household Borrowing and Metropolitan House Price Dynamics – Empirical Evidence from Helsinki". *Journal of Housing Economics*. 18(2):126–139. doi:10.1016/j.jhe.2009.04.001.

Oikarinen, E. 2009b. "Interaction between Housing Prices and Household Borrowing: The Finnish Case". *Journal of Banking & Finance*. 33(4):747–756. doi:10.1016/j.jbankfin.2008.11.004.

Omari, Y., S. Chib, N. Shephard, and J. Nakajima. 2007. "Stochastic Volatility with Leverage: Fast and Efficient Likelihood Inference". *Journal of Econometrics*. 140 (2):425–449.

Patton, A. J. 2007. "Volatility Forecast Comparison Using Imperfect Volatility Proxies". *Working paper, Oxford University*.

Sadorsky, P. 2006. "Modeling and Forecasting Petroleum Futures Volatility". *Energy Economics*. 28:467–488. doi:10.1016/j.eneco.2006.04.005.

Segnon, M., R. Gupta, K. Lesame, and M. E. Wohar. 2020. "High–frequency Volatility Forecasting of US Housing Markets". *Journal of Real Estate Finance & Economics*. doi: 10.1007/s11146-020-09745-w.

Silva, R. S., H. F. Lopes, and H. S. Migon. 2006. "The Extended Generalized Inverse Gaussian Distribution for Log–Linear and Stochastic Volatility Models". *Brazilian Journal of Probability and Statistics*. 20(1):67–91. *https://www.jstor.org/stable/43601074*

Spiegelhalter, D., N. Best, B. Carlin, and A. V. der Linde. 2002. "Bayesian Measures of Model Complexity and Fit(with Discussion)". *Journal of the Royal Statistical Society Series B*. 64:583–639. doi:10.1111/1467-9868.00353.

Statistics Finland. 2016. "Households' Assets", *Technical report*. Last checked: 10/ 02/2020.*http://www.stat.fi/til/vtutk/2016/vtutk-2016-2018-06-05-tie-001-en.html*

Statistics Finland (2019). "Building and Dwelling Production", *Technical report*. Last checked: 10/ 02/2020 *http://www.stat.fi/til/ras/index-en.html*

Statistics Finland Overview (2018). 2018 "Overview, Household–dwelling Units and Housing Conditions", *Technical report*.

Last checked: 10/ 02/2020. *http://www.stat.fi/til/asas/2018/011/asas-2018-01-2019-10-10-kat-002- en.html*

Taylor, S. J. 1982. "Financial Returns Modelled by the Product of Two Stochastic Processes: A Study of Daily Sugar Prices 1961-75". In O. D. Anderson, edited by. *Time Series Analysis, Theory and Practice*. Amsterdam: North–Holland; p. 203–226.

Taylor, S. J. 1986. *Modelling Financial Time Series*. Chichester: Wiley.

Tsai, I.-C. 2015. "Spillover Effect between the Regional and the National Housing Markets in the UK". *Regional Studies*. 49 (12):1957–1976.

Tsai, I.-C., M.-C. Chen, and T. Ma. 2010. "Modelling House Price Volatility States in the UK by Switching ARCH Models". *Applied Economics*. 42(9):1145–1153. doi:10.1080/00036840701721133.

Tsay, R. S. 2013. *An Introduction to Analysis of Financial Data with R*. Hoboken, New Jersey: John Wiley & Sons.

Willcocks, G. 2010. "Conditional Variances in UK Regional House Prices". *Spatial Economic Analysis*. 5(3):339–354. doi:10.1080/17421772.2010.493951.

Xingyi, L., and V. Zakamulin (2018). "Forecasting Stock Volatility: The Gains Form Using Intraday Data." Available at SSRN: https//ssrn.com/abstract=2847059; http://dx.doi.10.2139/ssrn.2847059.

Zhou, J., and Z. Kang. 2011. "A Comparison of Alternative Forecast Models of REIT Volatility". *The Journal of Real Estate Finance and Economics*. 42:275–294 . doi: 10.1007/s11146-009-9198-7.

Zhou, Y., and D. R. Haurin. 2010. "On the Determinants of House Value Volatility". *The Journal of Real Estate Research*. 32:377–396.

Publication IV

Dufitinema, J.
Forecasting the Finnish house price returns and volatility: A comparison of Time Series Models

Reprinted with permission from
International Journal of Housing Markets and Analysis, 2021

DOI: https://doi.org/10.1108/IJHMA-12-2020-0145

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1753-8270.htm

Forecasting the Finnish house price returns and volatility: a comparison of time series models

Josephine Dufitinema
School of Technology and Innovation, Mathematics and Statistics Unit, University of Vaasa, Vaasa, Finland

Received 2 December 2020
Revised 19 December 2020
Accepted 3 January 2021

Abstract

Purpose – The purpose of this paper is to compare different models' performance in modelling and forecasting the Finnish house price returns and volatility.

Design/methodology/approach – The competing models are the autoregressive moving average (ARMA) model and autoregressive fractional integrated moving average (ARFIMA) model for house price returns. For house price volatility, the exponential generalized autoregressive conditional heteroscedasticity (EGARCH) model is competing with the fractional integrated GARCH (FIGARCH) and component GARCH (CGARCH) models.

Findings – Results reveal that, for modelling Finnish house price returns, the data set under study drives the performance of ARMA or ARFIMA model. The EGARCH model stands as the leading model for Finnish house price volatility modelling. The long memory models (ARFIMA, CGARCH and FIGARCH) provide superior out-of-sample forecasts for house price returns and volatility; they outperform their short memory counterparts in most regions. Additionally, the models' in-sample fit performances vary from region to region, while in some areas, the models manifest a geographical pattern in their out-of-sample forecasting performances.

Research limitations/implications – The research results have vital implications, namely, portfolio allocation, investment risk assessment and decision-making.

Originality/value – To the best of the author's knowledge, for Finland, there has yet to be empirical forecasting of either house price returns or/and volatility. Therefore, this study aims to bridge that gap by comparing different models' performance in modelling, as well as forecasting the house price returns and volatility of the studied market.

Keywords ARMA, GARCH models, House prices, Forecasting, Finland

Paper type Research paper

1. Introduction

Forecasting house price returns and volatility is vital for numerous sectors such as consumers, policymakers, investors and risk managers. The reasons being, firstly, the housing assets' dual

The author express my gratitude to Professor Seppo Pynnönen for his constructive comments and advice.

The funding from the foundation for economic education (Liikesivistyrahasto) is highly acknowledged.

International Journal of Housing Markets and Analysis
Emerald Publishing Limited
1753-8270
DOI 10.1108/IJHMA-12-2020-0145

role of investment and consumption; thus, accurate forecasting of house price dynamics plays a crucial role in asset allocation and investment decision-making. Secondly, housing is a substantial component of the country's economy. Notably, in Finland, over half of the households' total wealth (50.3%) is in the form of housing (Statistics Finland, 2016). In the USA, housing is the largest component of household wealth; it represented, respectively, 28.3 and 24.6% of the total households' net worth and households' asset (Financial Accounts Data, 2018). In the UK, Savills (2019) estimated the housing stock total value to £7.29tn, highlighting an essential part that housing and its market have in the sustainability of the economy. Thirdly, housing affects the country's economy by influencing many parties involved in housing and mortgage activities. Therefore, accurate house price forecasting would benefit consumers and mortgage parties (Segnon *et al.*, 2020). Last, insights into house price dynamics provide recommendations to the housing policymakers and they are the fundamental inputs in outlining housing plans and policies, as stressed by Zhou and Haurin (2010).

Having noted the importance of the housing market, house price analysis of individual markets has been the subject of an increasing amount of studies. However, the focus has been on a restricted number of countries, namely, the USA, UK, Canada and Australia (Apergis and Payne, 2020). For Finland, even though over half of the households' total wealth is in the form of housing, as reported by Statistics Finland (2016), there has yet to be empirical forecasting of either house price returns or/and volatility. Therefore, this study aims to bridge that gap by comparing different models' performance in modelling as well as forecasting the house price returns and volatility. Thereby providing the information on the accurate model for modelling and forecasting the Finnish housing market, moreover extending the ongoing literature on the analysis of the housing market of various countries.

The purpose of the study is to find the most suitable and accurate model for Finnish house price returns and volatility modelling and forecasting. The number of rooms is used to categorise the studied dwellings, that is, one-room, two-rooms and larger (over three rooms) apartments. The 15 studied regions are distributed into 45 cities and sub-areas following their Zone Improvement Plan (ZIP)-code or postcode numbers. The competing models are the autoregressive moving average (ARMA) model and autoregressive fractional integrated moving average (ARFIMA) model for house price returns. The exponential GARCH (EGARCH) model, the fractionally integrated GARCH (FIGARCH) model and the component GARCH (CGARCH) model for house price volatility. The models' choice derives from Dufitinema and Pynnönen's (2020) and Dufitinema's (2020) studies outcomes. After testing for ARCH effects, the former article found grounds of long-range dependence in the house price returns and volatility for a greater number of the Finnish cities and sub-areas. The latter article used the EGARCH model and found that shocks' asymmetric impact on housing volatility was recorded in nearly all the Finnish cities and sub-markets. Therefore, to develop time-series models suitable for this housing market forecasting exercise, for cities and sub-areas with no ARCH effects, the short memory ARMA model's forecasting performances and long memory ARFIMA model are compared. For cities and sub-areas with substantial clustering effects, a short memory GARCH model, in this case, the EGARCH model's forecasting performance is weighed up to the GARCH models, which accommodate the long memory in the conditional variance; those are FIGARCH and CGARCH models. To assess the models' out-of-sample forecasting performances, the data is split into training and test sets. The former set is used to estimate the model and build predictions; the latter is used to evaluate the model produced forecasts. Results reveal that the house price return understudy drives the models' performance for the in-sample fit examination. While the EGARCH model is the best-ranked model for house price volatility modelling. The long memory models outclass their short memory peers in the out-of-sample forecasting for house price returns and volatility. Additionally, the models' in-sample fit performances vary from region to region, while in some areas, the models manifest a geographical pattern in their out-of-sample forecasting performances.

The remainder of the paper is organised as follows. The data and methodology used are described in Section 2; results are presented and discussed in Section 3. Section 4 concludes and presents further research.

2. Related literature

The housing market is a fundamental factor of the economy of various developed countries and it has been found to hold strong interlinkages with business cycles. Therefore, it is of great importance to understand and forecast house price dynamics. However, in the housing literature, whether the focal point is house price returns and volatility modelling and/or forecasting, a restricted number of countries has been targeted. These include the USA, UK, Canada and Australia. Moreover, the emphasis has been on the house price dynamics modelling while, apart from the USA housing market, research on forecasting individual housing markets is quite limited. Regarding modelling house prices of the above-cited countries, Apergis and Payne (2020) provide an extensive literature review with a striking dominance of the USA and UK studies. The reviewed studies also confirm the evidence of Autoregressive Conditional Heteroscedasticity (ARCH) effects in different housing markets. Further, the studies use various Generalised Autoregressive Conditional Heteroscedasticity (GARCH)-type models to investigate house price returns and volatility dynamics.

Regarding forecasting house prices, as mentioned above, the widely studied market is the US housing market. Crawford and Fratantoni's (2003) work paved the way; the authors investigated the performance of three types of models in forecasting the US home prices for the state of Texas, FL, OH, CA and Massachusetts. The three used models were Autoregressive Integrated Moving Average (ARIMA), Regime-Switching and GARCH. The authors found that the Regime-Switching models performed better in-sample fit, while the ARIMA models delivered superior out-of-sample forecasts. However, Milles (2008) criticised Crawford and Fratantoni's (2003) study by pointing out that, in a Monte Carlo study, Bessec and Bouabdallah (2005) found the Regime-Switching model to provide poor out-of-sample forecasts and it was recommended to use other nonlinear approaches. Specifically, the author used the Generalised AR (GAR) model and found that the GAR outperformed GARCH and ARIMA models in the out-of-sample forecasting. Li (2012) carried out in-sample and out-of-sample evaluation performance of the GARCH, Asymmetric Power ARCH (PARCH) and RiskMetrics model on the US housing market pre- and post-2008 financial crisis. The author's empirical results revealed that for the in-sample estimation, the benchmark model, the RiskMetrics performed satisfactorily, while all models achieved poor post-crisis out-of-sample forecasts. Recently, Segnon *et al.* (2020) introduced and used the Markov-Switching Multifractal (MSM) process to model and forecast the US house price volatility for 10 major cities, namely, Miami, Boston, New York, Chicago, San Diego, WA DC, Los Angeles, San Francisco, Denver and Las Vegas. The authors tested the MSM's forecasting abilities in comparison to the GARCH-type models; their results suggested that improved forecast accuracy is achieved through MSM and FIGARCH frameworks.

Broadly, despite the housing market analysis growing literature, whether the focus is on modelling house prices, forecasting their dynamics or a combination of two; special attention has been given to a limited number of countries. No particular empirical forecasting of either house price returns and/or volatility has been undertaken for the Finnish housing market, even though more than half of the households' total wealth is in the form of housing (Statistics Finland, 2016). Therefore, this article aims to fill that gap by comparing different models' performance in modelling as well as forecasting the house price returns and volatility. Furthermore, previous studies used the family-home property type data sets; the article at hand, however, uses apartments (also referred to as, a block of flats) type data. The number of rooms categorises the studied dwellings: one-room, two-rooms and larger apartments (over three rooms) types. The reasons for using flats property type data are their fast-growing popularity as a place to live in Finland and their increased attractiveness to both consumers and investors. At the end of 2018, Statistics Finland Overview reported that apartments counted for nearly half of all occupied dwellings, they represented 46%. Detached and semi-detached was the second favourable house type, with 39%, followed by terraced with 14%. Regarding the investment aspect, apartments continue to strengthen their position in the Finnish residential property market with foreign, domestic as well as individual investors continue to increase their portfolios across the country (KTI, 2019). In addition, in the same viewpoint of housing investment and portfolio allocation, this analysis uses metropolitan as well as ZIP-code level data for cross-examination and comparison of housing investment on the city and sub-market levels.

3. Data and methodology

The data used in this study are quarterly house price indices, retrieved from Statistics Finland's PxWeb databases (2020). The number of rooms categorises the studied types of dwellings: one-room, two-rooms and larger (over three rooms) apartment types. The considered period spans from the first quarter (*Q*1) of 1988 to the fourth quarter (*Q*4) of 2018 and the 15 considered regions are Helsinki, Oulu, Tampere, Lahti, Pori, Turku, Seinäjoki, Jyväskylä, Lappeenranta, Kuopio, Hämeenlina, Vaasa, Kotka, Joensuu and Kouvola. The regions of Helsinki, Turku and Tampere form an important and growing area, called the growth triangle in Southern Finland. Currently, the area accounts for, respectively, 49 and 55.5% of the Finnish population and total gross domestic product (GDP). The Oulu region, called the Northern Finland growth centre, is also amongst the well-performing region with substantial economic development and population growth. The other regions also show significant expansion and economic performance. These regions are then divided into 45 cities and sub-areas according to their ZIP-code or postcode numbers. Dufitinema (2020) details the regions' ranking and division. The number of inhabitants ranks regions and postcode numbers divide them.

The methodology used in this study is an extension of Dufitinema's (2020). That is, house price indices are transformed into log-returns. The process is done for each city and sub-area in every apartment type. Next, first-order autocorrelations are filtered out from the returns. The task is done by determining the appropriate order of the ARMA model using the Akaike and Bayesian information criteria (respectively, AIC and BIC). Then, from the transformed returns, ARCH effects are tested. Thereafter, the current study extends this methodology by examining the ARMA and ARFIMA models' forecasting performances for cities and sub-areas with no substantial ARCH effects. The EGARCH model's forecasting abilities are compared to the FIGARCH and CGARCH models for cities and sub-areas with substantial clustering effects.

Regarding testing for ARCH effects, details are given and results are described in Dufitinema (2020). In a nutshell, both used tests Lagrange Multiplier (LM) and Ljung-Box (LB) found, in all three considered types of apartments, that clustering effects were significant in the majority of the cities/sub-areas. Specifically, the results are as follows: in the one-room flats category, the evidence of clustering effects was found in 28 out of 38 cities/sub-areas. In 27 out of 42 and 31 out of 39 in, respectively, the two-rooms and larger (over three rooms) flats category. Moreover, as in forecasting the house price dynamics of the considered types of dwellings, short memory and long memory time series models are compared, we make use of Dufitinema and Pynnönen's (2020) study outcomes. The results summary is as follows: in those cities/sub-areas with no significant clustering effects, in the one-room apartment type category, 8 out of 10 exhibited long memory behaviour. Meaning that their Geweke and Porter-Hudak (1983) (GPH) estimates of the fractional differencing parameter d varied from 0 to 0.5. The two returns series were anti-persistent $[d \in (-0.5, 0)]$. In both two-room and larger (over three rooms) apartment categories, one sub-area displayed anti-persistence behaviour while the rest 14 and 7 returns series exhibited long-range dependence behaviour in the respective groups. These results are used as hyperparameters of the ARFIMA models in the estimation procedure.

The same applies to Dufitinema and Pynnönen's (2020) findings on the long-range dependence in those cities/sub-areas with substantial ARCH effects. In squared as well as absolute house price returns, in all three apartment types, the fractional differencing parameter d was estimated and the outcomes indicated a very persistent long memory behaviour in the house price volatility. Both metrics results are used as hyperparameters of the FIGARCH models in the estimation procedure and the best model is assessed based on different model selection tools. This approach of tuning the parameter d, that is, estimate the long memory parameter first and get the other parameters estimations using these d estimates, is at the core of most semiparametric estimation approaches (Lopes and Mendes, 2006; Härdle and Mungo, 2008). Furthermore, as pointed out by different researchers such as Tsay (2013), when GARCH-type models are used to assess asset returns, an assumption of a normal distribution is not tenable. An appropriate distribution must accommodate asset returns characteristics, for instance, skewness and fat tails. Therefore, based on AIC and BIC, appropriate distribution is selected, for each city and sub-area in every apartment type, amongst univariate distributions, namely, Student t ("Std"), Generalised Error ("GED") and their skew variants ("sStd" and "sGED").

3.1 Models for forecasting house price returns

House prices returns are predicted for cities/sub-areas with no substantial clustering effects, meaning those regions with both constant mean and variance. The types of models tested relate to this constant mean/variance specification of the series. The ARMA models fulfil this property; however, they do not capture the long-memory behaviour that house price returns of these cities/sub-areas exhibit. Therefore, their forecasting performances are compared to the models that accommodate the high persistence present in the returns series; those are ARFIMA models.

3.1.1 Autoregressive moving average model. ARMA models have been a leading major of modelling and forecasting in numerous areas of finance and economics. In the housing market, we refer to Jadevicius and Huston (2015) and the references therein. Jadevicius and Huston assess the ARMA's application for forecasting the Lithuanian housing market in particular and extend their findings to the global housing market. The ARMA model is a combination of AR and MA processes (Box *et al.*, 1994). Its standard specification is as follows:

$$r_t = \varphi_0 + \sum_{i=1}^{p} \varphi_i r_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i},$$

where $\sum_{i=1}^{p} \varphi_i r_{t-i}$ represents the AR portion of the model and $\sum_{i=1}^{q} \theta_i a_{t-i}$ represents the model's MA portion. By assumption, r_t is stationary, for a collect specification of the ARMA model; otherwise, the series has a unit root and it is termed as AR Integrated MA (ARIMA) process. However, Dufitinema and Pynnönen (2020) have conducted unit root tests on the studied house prices returns and concluded that the null hypothesis of a unit root in all return series in all the three apartment types was rejected at least at the 5% level. Hence, stationarity was ensured across all cities and sub-areas, in all apartment types.

3.1.2 Autoregressive fractional integrated moving average model. ARFIMA models are the extension of the ARIMA models to accommodate the time series's long-memory behaviour. They were independently put forward by Granger and Joyeux (1980) and Hosking (1981). The standard specification of an ARFIMA model is as follows:

$$\Phi(L)(1-L)^d Y_t = \Theta(L)\epsilon_t,\ t = 1, 2, \ldots,$$

where Y_t denotes the discrete-valued studied time series, d is the fractional differencing parameter and ϵ_t is a white noise with $E(\epsilon_t) = 0$ and variance σ^2_ϵ. L is the lag operator or back-shift operator such that $LY_t = Y_{t-1}$. $\Phi(L)$ and $\Theta(L)$ are the AR and MA polynomials in the lag operator, respectively. That is, $\Phi(L) = 1 - \phi_1 L - \ldots. - \phi_p L^p$ and $\Theta(L) = 1 - \theta_1 L - \ldots - \theta_q L^q$.

The value of d – the long memory parameter – dictates the properties and the interpretations of the ARFIMA model. If $d = 0$, ARFIMA reduces to ARIMA and the process is stated to exhibit short memory. If $d \in (-0.5, 0)$, it is characterised as anti-persistence or long-range negative dependence. The process is said to manifest long memory or long-range positive dependence if $d \in (0, 0.5)$ and it is non-stationary with mean reversion if $d \in [0.5, 1)$, whereas it becomes non-stationary without mean reversion if $d \geq 1$.

3.2 Models for forecasting house price volatility

For regions with time-varying variance, meaning those cities and sub-areas with evidence of ARCH effects, GARCH-type models are used to forecast house price volatility. Motivated by the persistence or long memory behaviour found in these cities/sub-areas' house price volatility, short memory GARCH models are compared to the GARCH models that accommodate the long memory property. The EGARCH model is selected amongst the short memory GARCH models, over the standard GARCH. The grounds of the EGARCH selection are the evidence of asymmetric effects of shocks on housing volatility recorded in the studied types of dwellings and its effective performance over the Glosten *et al.*'s (1993) GJR-GARCH model in modelling the studied house prices' asymmetric volatility (Dufitinema, 2020). Amongst the GARCH models that accommodate the long memory in the assets' conditional variance, the selected ones are the FIGARCH and CGARCH models. The FIGARCH model allows a slower hyperbolic rate decay of shocks, making it the best candidate for explaining and capturing the high degree of autocorrelation in financial market volatility. The CGARCH model investigates the conditional variance's long- and short-run movement by decomposing the conditional variance into permanent and transistor components. Both models have been applied more often of late compare to, for instance, the Integrated GARCH (IGARCH) model (Engle and Bollerslev, 1986). The reason is that Tayefi and Ramanathan (2012) have found the IGARCH model to be too restrictive as it implicates on the conditional variance, an infinite persistence and consequently, shocks persist forever.

There is an extensive collection of studies on the FIGARCH and CGARCH applicabilities to model and/or forecast different assets' volatility. In the housing markets, Milles (2011) used the CGARCH model to investigate whether there is long-range dependence in the US home price volatility. The author found that housing markets of over half of the US metropolitan areas exhibited persistent volatility. For those regions, the CGARCH model provided better forecasts than the standard GARCH model. The Milles's choice of the CGARCH was based on Maheu's (2005) Monte Carlo study, which showed that the CGARCH captured long-range dependence better than FIGARCH in equity markets. On the other hand, Feng and Baohua (2015) discovered that the FIGARCH model could well catch the long memory of the Zhengzhou house price volatility. To that end and for the models' cross-check assessment, this article uses both FIGARCH and CGARCH models to forecast house price volatility of the considered types of dwellings.

3.2.1 Exponential generalised autoregressive conditional heteroscedasticity model. Let R_t denotes the asset log-return at time t. The standard form of the conditional volatility model is as follows:

$$R_t = v_t + \epsilon_t, \quad \epsilon_t \sim N\left(0,\, \sigma_t^2\right),$$

where v_t is the conditional mean, σ_t is the conditional standard deviation and ϵ_t is the error term. Given that many financial assets exhibited volatility clustering, instead of modelling the variance of the innovation ϵ_t as a constant, Bollerslev (1986) proposed a GARCH process where the conditional variance σ_t^2 is a function of past volatility and previous squared errors. That is,

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \qquad (1)$$

where $\omega > 0$ is the intercept, $\alpha_i \geq 0$ (coefficients of e_{t-i}) and $\beta_j \geq 0$ (coefficients of σ_{t-j}^2) are referred to, respectively, as the ARCH and GARCH parameters.To investigate the potential asymmetric effects of shocks on conditional variance, Nelson (1991) proposed the EGARCH model. The model enables negative shocks to have a distinct impact on conditional variance than positive shocks, an observation which is termed to leverage effects. Its standard specification is as follows:

$$R_t = v_t + \epsilon_t, \quad \epsilon_t \sim N\left(0,\, \sigma_t^2\right),$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} (\alpha_i + \gamma_i I_{t-i}) \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

where α_i and $\alpha_i + \gamma_i$ indicate, respectively, the effects of good and bad news. I_{t-i} is the indicator function and it equals to one if $\epsilon_{t-1} < 0$ and zero otherwise. Implying a more sizable influence $(\alpha_i + \gamma_i)\epsilon_{t-i}^2$ with $\gamma_i > 0$ of a negative shock ϵ_{t-i}, while a positive shock ϵ_{t-i} have little influence $\alpha_i \epsilon_{t-i}^2$ to σ_t^2.

3.2.2 Fractionally integrated generalised autoregressive conditional heteroscedasticity model. The evidence of slow decay in correlations of squared and absolute returns of financial assets gave rise to the FIGARCH model, first introduced by Baillie *et al.* (1996). The model adds the fractional differences in the standard GARCH process, thereby explaining and capturing the high degree of autocorrelation in financial market volatility.

The GARCH process in equation (1) can be written as:

$$\sigma_t^2 = \omega + \alpha(B)\epsilon_t^2 + \beta(B)\sigma_t^2,$$

where B is the lag operator such that $\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \ldots + \alpha_q B^p$ and $\beta(B) = \beta_1 B + \beta_2 B^2 + \ldots + \beta_p B^p$. Its equivalent ARMA type representation is given by:

$$[1 - \alpha(B) - \beta(B)]\epsilon_t = \omega + [1 - \beta(B)]u_t,$$

where $u_t = \epsilon_t^2 - \sigma_t^2$. From this formulation, Engle and Bollerslev (1986) presented the IGARCH model by allowing the presence of unit root in $1 - \alpha(B) - \beta(B)$ as follows:

$$[1 - \alpha(B) - \beta(B)](1 - B)\epsilon_t = \omega + [1 - \beta(B)]u_t. \tag{2}$$

However, as discussed above, the IGARCH model is too restrictive as shocks persist forever. Hence, the introduction of the FIGARCH model, where the fractional differencing operator $(1 - B)^d$ with $0 < d < 1$ replaces the first difference operator $(1 - B)$ in equation (2). The general form of the FIGARCH model is as follows:

$$[1 - \alpha(B) - \beta(B)](1 - B)^d \epsilon_t = \omega + [1 - \beta(B)]u_t.$$

If $d = 0$, the FIGARCH model reduces to the standard GARCH, while if $d = 1$, it turns into an IGARCH model.

3.2.3 Component generalised autoregressive conditional heteroscedasticity model. Lee and Engle (1999) developed the CGARCH model by decomposing the conditional variance into permanent and transitory components, thereby investigating the long- and short-run volatility movements. Unlike in the GARCH process where the conditional variance reverts to a long-run constant mean ω in equation (1), the CGARCH model allows a time-varying mean reversion of the conditional variance. Its specification is as follows:

$$\sigma_t^2 = q_t + \sum_{i=1}^{q} \alpha_i \left(\epsilon_{t-i}^2 - q_{t-i}\right) + \sum_{j=1}^{p} \beta_j \left(\sigma_{t-j}^2 - q_{t-j}\right), \tag{3}$$

$$q_t = \omega + \rho q_{t-1} + \phi\left(\epsilon_{t-1}^2 - \sigma_{t-1}^2\right). \tag{4}$$

Equation (4) represents the long-run (permanent) component of the volatility; the time-varying mean reversion of the conditional variance. It describes how the GARCH model's intercept is now time-varying following first-order autoregressive type dynamics, and thus, captures the long memory portion of volatility. Equation (3) describes the short-term (transitory) component of the volatility, which is the difference between the conditional variance and its trend $(\sigma_t^2 - q_t)$. To ensure the stationarity conditions, the sum of (α, β) coefficients must be less than 1 and $\rho < 1$ for the persistence of the transitory and permanent components. If $\rho = \phi = 0$, the CGARCH model reduces to the standard GARCH.

3.3 Forecast evaluation

To test and compare the prediction abilities of the above-mentioned models; the data is divided into training and test set. The training set, which consists of 25 years of sample data, is used to build the models (estimation sample: 1988:Q1-2013:Q4). The test set is used to evaluate the models' predictive accuracy; it consists of 5 years of sample data (forecasting sample: 2014:Q1-2018:Q4). The forecasting process starts by estimating each model on the training data set. Thereafter, the one-step-ahead (quarter) volatility forecasts are built using the estimated model. Finally, the predicted volatility ($\hat{\sigma}^2$) and the proxy of the true volatility (σ^2) are compared.

When evaluating volatility forecasts, one has to deal with the problem that the true volatility σ^2 is unobserved. Various studies have proposed the appropriate proxy of σ^2 such as the squared returns (Brooks and Persands, 2002; Sadorsky, 2006). Patton (2011) discussed that squared returns are a rather noisy proxy for the true conditional variance and that a conditionally unbiased estimator of the conditional variance, the realised volatility (RV), is a more efficient estimator than the squared returns. Recently, Xingyi and Zakamulin (2018) pointed out that the usage of realised daily volatility and available intraday data provided better forecast accuracy in the stock market. In the housing market, Zhou and Kang (2011) also used realised volatility calculated from assets returns as σ^2 proxy. Following this study, in this article, the true volatility is also proxied by realised volatility built as a rolling sample. Furthermore, in line with other studies on volatility forecasting, two popular metrics, namely, the root mean squared error (RMSE) and the mean absolute error (MAE), is used to evaluate the studied models' forecasting accuracy. The former metric has the benefit of penalising large errors as it gives errors with larger absolute values more weight than errors with smaller absolute values, which makes it useful when large errors are particularly undesirable. The latter metric gives the same weight to all errors. Both are negatively-oriented scores, meaning that lower values are better. The two measures are defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\sigma}_i^2 - \sigma_i^2\right)^2} \text{ and MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{\sigma}_i^2 - \sigma_i^2|,$$

where N is the number of forecasts, $\hat{\sigma}^2$ is the forecast volatility and σ^2 is the true volatility.

4. Results and discussions

4.1 Forecasting house price returns

The ARMA and ARFIMA models' performances are compared, in each apartment category, for cities and sub-areas with no substantial clustering effects, meaning those regions with both constant mean and variance. Recall that in the one-room apartment category, there are 10 cities/sub-areas and eight of them exhibited long memory behaviour. In the two-room and larger (over three rooms) apartment categories, there are 15 and 8 cities/sub-areas, respectively. In total, 14 and 7 returns series exhibited long-range dependence behaviour in each apartment category, respectively. Table 1 reports the house price returns' best performing in-sample and out-of-sample models for each city and sub-area, in each apartment type. In Appendix, Table A1 details the Akaike information criteria (AIC) of each model, while Table A2 presents the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE); the used metrics in evaluating the forecasting accuracy of every model. A lower criteria value describes a better model's performance.

To investigate which feature (short or long memory) is crucial in the Finnish house price returns modelling, results are mixed; the two models' performances differ by apartment types and across cities and sub-areas. Firstly, in the one-room flat category, the ARMA model ranks as the leading in-sample performing model in six out of eight cities/sub-areas. Secondly, in the two-room flat category, it is the ARFIMA model, which excels in 11 out of 14 cities/sub-areas. Last, in larger (over three rooms) flat type, both models split the ranking as the ARMA model fits the house price returns best in three cities/sub-areas, while

One room flats				
Regions	Cities/sub-areas	*In-sample*		*Out-of-sample*
Helsinki	hki3	ARMA		ARMA
Tampere	tre	ARFIMA		ARFIMA
	tre2	ARMA		ARMA
Oulu	oulu2		Anti-persistent	
Lahti	lti2	ARFIMA		ARMA
Joensuu	jnsu	ARMA		ARFIMA
Vaasa	vaasa	ARMA		ARFIMA
	vaasa1		Anti-persistent	
Hämeenlinna	hnlina1	ARMA		ARFIMA
Kotka	kotka1	ARMA		ARFIMA
		Two rooms flats		
		In-sample		*Out-of-sample*
Tampere	tre	ARFIMA		ARFIMA
	tre3	ARFIMA		ARFIMA
Turku	tku1	ARFIMA		ARFIMA
	tku3	ARFIMA		ARFIMA
Oulu	oulu	ARFIMA		ARMA
	oulu1	ARMA		ARMA
	oulu2	ARMA		ARMA
Lahti	lti1	ARFIMA		ARMA
	lti2	ARFIMA		ARFIMA
Kuopio	kuo2	ARFIMA		ARFIMA
Joensuu	jnsu	ARFIMA		ARFIMA
Vaasa	vaasa1	ARMA		ARFIMA
Lappeenranta	ltra2	ARFIMA		ARFIMA
Kotka	kotka	ARFIMA		ARFIMA
	kotka2		Anti-persistent	
		Three rooms flats		
		In-sample		*Out-of-sample*
Helsinki	hki2	ARMA		ARFIMA
Oulu	oulu2	ARFIMA		ARFIMA
Lahti	lti2	ARMA		ARFIMA
Pori	pori	ARFIMA		ARMA
Joensuu	jnsu	ARFIMA		ARMA
	jnsu1		Anti-persistent	
Kouvola	kou	ARMA		ARFIMA
Hämeenlinna	hnlina	ARFIMA		ARFIMA

Notes: This table reports the house price returns best performing in-sample and out-of-sample models, for each city and sub-area, in each apartment type. The "anti-persistent" refers to the series with long-range negative dependence, meaning that their estimated fractional differencing parameter *d* varied from −0.5 to 0

Table 1.
House price returns – best performing models

ARFIMA performs well in four out of seven cities/sub-areas. These results are in line with Jadevicius and Huston's (2015) study outcomes and Hepsen and Vatansever's (2011) recommendations. Jadevicius and Huston highlighted that the ARIMA modelling approach strongly contributes to examining housing markets. Hepsen and Vatansever pointed out that house price modelling with ARIMA provides perceptions for a range of stakeholders. Moreover, the ARFIMA model's ability to capture the long memory feature of the house price returns, notably in the two-room flat category; stresses the high persistence of house prices (Dufitinema and Pynnönen, 2020).

The out-of-sample forecast performance of the two models is investigated by estimating the models on the training data set, generating 5-year returns forecasts and validating the constructed predictions using the test set. Generally, in all three apartment types, the ARFIMA model outperforms the ARMA in most regions. The ARFIMA model provides the best returns forecasts in 5 out of 8, 10 out 14 and 5 out of 7 cities/sub-areas in the one-room, two-room and larger (over three rooms) flats categories, respectively. Given the strong evidence of long memory found in the Finnish house price returns by Dufitinema and Pynnönen (2020), these results confirm again the long memory models' ability to capture these long-range dependencies and their superiority in forecasting house price returns. In the two-room apartment category, an interesting observation emerges, the best in-sample performing model also produces accurate out-of-sample forecasts. This remark is noted in 11 out of 14 cities/sub-areas. On the one hand, it contradicts previous studies, which expressed that a better in-sample fit does not automatically suggest a superior forecasting performance (Newell *et al.*, 2002; Stevenson and McGrath, 2003). On the other hand, the remark aligned with Jadevicius and Huston's (2015) findings that the same model [ARIMA(3,0,3)] provided superior in- and out-of-sample modelling results for the Lithuanian housing market.

In summary, regarding modelling the Finnish house price returns, the short or long memory model's performance is driven by the house price data set under study. Therefore, across cities and sub-areas, one must enable different house price dynamics instead of imposing one model on the full data set. With respect to forecasting house price returns, the long memory models outclass their short memory peers. This result highlights the advantage of long memory models in forecasting different asset prices.

4.2 Forecasting house price volatility

For regions with time-varying variance, meaning those cities and sub-areas with substantial ARCH effects, short and long memory GARCH models are compared. Those are the EGARCH, FIGARCH and CGARCH models. Table 2 reports the house price volatility' best-performing in-sample and out-of-sample models for each city and sub-area, in each apartment type. In the Appendix, the models' in-sample fits are detailed in Table A3 and their RMSE and MAE forecasting accuracies in Table A4.

Mostly, the best-ranked model for the Finnish house price volatility modelling, in all three apartment types, is the EGARCH model. It comes on top in 17 out of 28 cities/sub-areas exhibiting clustering effects in the one-room flat category. It leads in 19 out 27 and 23 out 31 cities/sub-areas in, respectively, two-room and larger (over three rooms) flat categories. These outcomes are in line with Dufitinema's (2021) findings, who underlined, using the Stochastic Volatility framework, that the stochastic volatility model with leverage effects was also the leading in-sample performing model for the studied type of dwellings. The results also highlight, once more, the importance of asymmetric volatility features in modelling house price volatility. In the rest of the regions, the FIGARCH model alternatives with EGARCH and takes the lead. This pattern is noted in 11, 6 and 7 cities/sub-areas in the respective flat categories. The exceptions of this general pattern are Turku and Vaasa cities in the two-room apartments and Jyväskylä-city in the category of larger (over three rooms) apartments, where the CGARCH model excels in comparison to the other two models.

The out-of-sample forecasting performance of the three models is examined. The forecasting exercise starts with an estimation of the models on the training set. Next, using the estimated models, 5-years volatility forecasts are generated in the form of one-step ahead. Finally, the built predictions are validated on the test set. Mostly, the long memory GARCH models overcome their short memory counterparts in all three apartment types. The CGARCH model provides the superior forecasts in, respectively, 14 out of 28, 11 out of

Regions	Cities/sub-areas	One room flats *In-sample*	One room flats *Out-of-sample*	Two rooms flats *In-sample*	Two rooms flats *Out-of-sample*	Three rooms flats *In-sample*	Three rooms flats *Out-of-sample*
Helsinki	hki	FIGARCH	EGARCH	FIGARCH	FIGARCH	EGARCH	CGARCH
	hki1	FIGARCH	CGARCH	EGARCH	FIGARCH	EGARCH	EGARCH
	hki2	FIGARCH	EGARCH	EGARCH	EGARCH	–	–
	hki3	–	–	FIGARCH	CGARCH	EGARCH	CGARCH
	hki4	EGARCH	CGARCH	EGARCH	CGARCH	EGARCH	EGARCH
Tampere	tre	–	–	–	–	EGARCH	EGARCH
	tre1	EGARCH	FIGARCH	EGARCH	FIGARCH	FIGARCH	FIGARCH
	tre2	–	–	EGARCH	FIGARCH	EGARCH	CGARCH
	tre3	EGARCH	EGARCH	–	–	FIGARCH	CGARCH
Turku	tku	EGARCH	FIGARCH	CGARCH	EGARCH	EGARCH	CGARCH
	tku1	EGARCH	CGARCH	–	–	EGARCH	FIGARCH
	tku2	EGARCH	EGARCH	EGARCH	CGARCH	EGARCH	CGARCH
	tku3	FIGARCH	CGARCH	–	–	EGARCH	CGARCH
Oulu	oulu	EGARCH	CGARCH	–	–	EGARCH	CGARCH
	oulu1	EGARCH	CGARCH	–	–	EGARCH	EGARCH
Lahti	lti	EGARCH	CGARCH	EGARCH	CGARCH	EGARCH	CGARCH
	lti1	EGARCH	FIGARCH	–	–	EGARCH	FIGARCH
Jyväskylä	jkla	EGARCH	CGARCH	EGARCH	CGARCH	CGARCH	FIGARCH
	jkla1	FIGARCH	FIGARCH	EGARCH	EGARCH	FIGARCH	EGARCH
	jkla2	FIGARCH	CGARCH	EGARCH	FIGARCH	FIGARCH	EGARCH
Pori	pori	FIGARCH	FIGARCH	EGARCH	EGARCH	–	–
	pori1	EGARCH	FIGARCH	EGARCH	CGARCH	FIGARCH	FIGARCH
	pori2	–	–	EGARCH	FIGARCH	–	–
Kuopio	kuo	EGARCH	FIGARCH	FIGARCH	CGARCH	EGARCH	FIGARCH
	kuo1	FIGARCH	FIGARCH	EGARCH	FIGARCH	FIGARCH	CGARCH
	kuo2	EGARCH	CGARCH	–	–	EGARCH	EGARCH
Joensuu	jnsu1	EGARCH	CGARCH	FIGARCH	EGARCH	–	–
Seinäjoki	seoki	–	–	FIGARCH	EGARCH	FIGARCH	CGARCH
Vaasa	vaasa	–	–	CGARCH	CGARCH	EGARCH	CGARCH
	vaasa1	–	–	–	–	EGARCH	EGARCH
	vaasa2	–	–	–	–	EGARCH	CGARCH
Kouvola	kou	EGARCH	CGARCH	EGARCH	FIGARCH	–	–
Lappeenranta	lrta	FIGARCH	FIGARCH	EGARCH	CGARCH	EGARCH	FIGARCH
	lrta1	FIGARCH	CGARCH	FIGARCH	EGARCH	–	–
	lrta2	–	–	–	–	EGARCH	FIGARCH
Hämeenlinna	hnlina	EGARCH	FIGARCH	EGARCH	FIGARCH	–	–
	hnlina1	–	–	EGARCH	CGARCH	EGARCH	FIGARCH
Kotka	kotka	FIGARCH	CGARCH	–	–	EGARCH	FIGARCH
	kotka1	–	–	EGARCH	CGARCH	–	–

Note: This table reports the house price volatility best performing in-sample and out-of-sample models for each city and sub-area, in each apartment type

Table 2.
House price volatility – best performing models

27 and 13 out of 31 cities/sub-areas in the one-room, two-room and larger (over three rooms) flats categories. The FIGARCH model follows with superior performance in 10, 9 and 10 cities/sub-areas in the respective flat categories. These findings are consistent with Milles's (2011), who concluded that the CGARCH provided better forecasts than the standard GARCH for the US home price volatility. Moreover, Lee and Reed (2014), in regard to the Australian housing market, also acknowledged the CGARCH model's ability to decompose the price volatility into "permanent" and "transitory" components. And thereby, be a better candidate to capture the short- and long-run movements of volatility.

A regional pattern is noted in few regions where the same model produces better out-of-sample forecasts in all three apartment types. In Tampere-area1, the FIGARCH is the leading model throughout all apartment types, while the CGARCH model stands out in Lahti-city. These results suggest that the house price volatility of the former region is characterised by a significant degree of autocorrelation. While the conditional variance of the latter city includes two components (permanent and transitory).

In summary, for a larger number of Finnish cities and sub-areas, the EGARCH model is the best model for modelling their house price volatilities. In the remaining regions, the EGARCH switches places with the FIGARCH model. However, no geographical is noted; the performance of the model varies from region to region. Hence, again as above, when modelling house price volatility, one must enable different house price dynamics across cities and sub-areas and types of apartment. Regarding the models' out-of-sample forecasting performances, the long memory models (CGARCH and FIGARCH) take the lead, dominating their short-memory counterparts. Apart from few regions (one city and one sub-area), the models' forecasting performances vary across cities and sub-areas and by type of apartment – no geographical or regional pattern is noted.

5. Conclusions, implications and further research

Over recent years, housing market forecasting has been the theme of extensive research due to the vital role of house price forecasts in asset allocation, consumption, investment, policy decision-making and also in predicting mortgage defaults. This article determines, in the Finnish housing market, which model is best able to forecast movements of both house price returns and volatility. The two competing models are the ARMA model and ARFIMA model for house price returns. For house price volatility, the EGARCH model is competing with the FIGARCH and CGARCH models. The study uses quarterly house price indices for 15 main regions in Finland, spanning from the first quarter (*Q*1) of 1988 to the fourth quarter (*Q*4) of 2018.

There are several important findings. Firstly, to investigate whether the short or long memory feature captures the house price returns movements, the models' performance is driven by the house price data set under investigation. In contrastingly, the ARFIMA model tops in the house price returns forecasting; it outperforms the ARMA model in most regions. This result indicates that the long-range dependencies that house price exhibits are a crucial component in their forecasting. Secondly, the EGARCH model ranks as the leading model for the Finnish house price volatility modelling, highlighting the importance of asymmetric volatility in the house price volatility modelling. The long memory GARCH models (CGARCH and FIGARCH) outperforms the EGARCH in forecasting the house price volatility, indicating the long term dependence in house price volatility and the ability of long memory models to capture and predict this property of house price volatility. Last, in all three apartment types, no geographical or regional pattern is noted for models' in-sample fit; each model's performance varies from region to region for both house price returns and volatility. For the out-of-sample analysis, however, some interesting observations emerge. For house price returns, especially in the two-room flat category, the same model provides the best in- and out-of-sample forecasts. While for the house price volatility, in two regions, the same model comes on top across all apartment types.

These outcomes have some vital housing investment and policy implications. For consumers, investors and policymakers, who monitor the house price volatility and whose decisions are based on future house price movements, accurate forecasts help their decision-making. Moreover, precise predictions are essential for housing investment risk assessment and are more significant insights for portfolio allocation across Finland and apartment type. Additionally, as interlinkages have been found between housing markets and the economic cycle of various developed countries, a view into house prices outlook would be beneficial for economists and policy institutions. Also, as pointed out by Balcilara *et al.* (2015), forecasting housing market movements plays a significant role in monetary policy authorities and their willingness to "lean against the wind".

Furthermore, as housing has been found to play a crucial role in macroeconomic factors fluctuations (Kishor and Marfatia, 2018), it would be of interest to investigate the interaction between house prices and the variables such as unemployment rates and interest rates from region to region. The information from these macroeconomic predictors can be further used to improve the forecast accuracy. In the same viewpoint, the existence of the structural break in the studied housing market merits an examination. In this aspect, the data can be split into subsamples supported by the break dates and thereby improving forecast accuracy.

References

Apergis, N. and Payne, J.E. (2020), "Modeling the time varying volatility of housing returns: further evidence from the US metropolitan condominium markets", *Review of Financial Economics*, Vol. 38 No. 1, pp. 24-33, doi: 10.1002/rfe.1063.

Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O. (1996), "Fractionally integrated generarized autoregressive conditional heteroscedasticity", *Journal of Econometrics*, Vol. 74 No. 1, pp. 3-30.

Balcilara, M., Guptab, R. and Millerc, S.M. (2015), "The out-of-sample forecasting performance of nonlinear models of regional housing prices in the US", *Applied Economics*, Vol. 47 No. 22, pp. 2259-2277, doi: 10.1080/00036846.2015.1005814.

Bessec, M. and Bouabdallah, O. (2005), "What causes the forecasting failure of markov–switching modesls?A monte carlo study", *Studies in Nonlinear Dynamics and Econometrics*, Vol. 9.

Bollerslev, T. (1986), "Generalized autoregressive conditional heteroscedasticity", *Journal of Econometrics*, Vol. 31 No. 3, pp. 307-327.

Box, G., Jenkins, G. and Reinsel, G. (1994), *Time Series Analysis: Forecasting and Control*, 3rd, ed., Prentice Hall, Englewood Clifs, NJ.

Brooks, C. and Persands, G. (2002), "Model choice and value–at–risk performance", *Financial Analysts Journal*, Vol. 58 No. 5, pp. 87-97.

Crawford, G.W. and Fratantoni, M.C. (2003), "Assessing the forecasting performance of regime–switching, ARIMA and GARCH models of house prices", *Real Estate Economics*, Vol. 31 No. 2, pp. 223-243, doi: 10.1111/1540-6229.00064.

Dufitinema, J. (2020), "Volatility clustering, risk–return relationship and asymmetric adjustment in the finnish housing market", *International Journal of Housing Markets and Analysis*, Vol. 13 No. 4, pp. 661-688, doi: 10.1108/IJHMA-12-2019-0125.

Dufitinema, J. (2021), "Stochastic volatility forecasting of the Finnish housing market", *Applied Economics*, Vol. 53 No. 1, pp. 98-114, doi: 10.1080/00036846.2020.1795074.

Dufitinema, J. and Pynnönen, S. (2020), "Long–range dependence in the returns and volatility of the finnish housing market", *Journal of European Real Estate Research*, Vol. 13 No. 1, pp. 29-50, doi: 10.1108/JERER-07-2019-0019.

Engle, R.F. and Bollerslev, T. (1986), "Modelling the peristence of conditional variances", *Econometric Reviews*, Vol. 5 No. 1, pp. 1-50.

Feng, L. and Baohua, C. (2015), "Research on the long-term memory of commodity housing price volatility based on the FIGARCH model", *Advanced Materials Research*, Vol. 1079–1080, pp. 1194-1198.

Financial Accounts Data (2018), "Financial accounts of the United States", available at: www.federalreserve.gov/releases/z1/20190920/html/b101h.htm (accessed 8 June 2020).

Geweke, J. and Porter-Hudak, S. (1983), "The estimation and application of long memeory time series models", *Journal of Time Series Analysis*, Vol. 4 No. 4, pp. 221-238.

Glosten, L., Jagannathan, R. and Runkle, D. (1993), "On the relation between the expected value and the volatility of the norminal excess return on stocks", *Journal of Finance*, Vol. 68, pp. 1179-1801.

Granger, C.W. and Joyeux, R. (1980), "An introduction to long–memory time series models and fractional differencing", *Journal of Time Series Analysis*, Vol. 1 No. 1, pp. 15-29.

Härdle, W.K. and Mungo, J. (2008), "Long memory persistence in the factor of implied volatility dynamics", *International Research Journal of Finance and Economics*, Vol. 18, pp. 213-230.

Hepsen, A. and Vatansever, M. (2011), "Forecasting future trends in dubai housing market by Box-Jenkins autoregressive integrated moving average", *International Journal of Housing Markets and Analysis*, Vol. 4 No. 3, pp. 210-223.

Hosking, J.R. (1981), "Fractional differencing", *Biometrika*, Vol. 68 No. 1, pp. 165-176.

Jadevicius, A. and Huston, S. (2015), "ARIMA modelling of lithuanian house price index", *International Journal of Housing Markets and Analysis*, Vol. 8 No. 1, pp. 135-147.

Kishor, N.K. and Marfatia, H.A. (2018), "Forecasting house prices in OECD economies", *Journal of Forecasting*, Vol. 37 No. 2, pp. 170-190, doi: 10.1002/for.2483.

KTI (2019), "KTI market review", *Technical Report*, KTI Property Information Ltd.

Lee, G.J. and Engle, R.F. (1999), "A permenant and transitory component model of stock return volatility", *Cointegration Causality and Forecasting a Festschrift in Honor of Clive WJ Granger*, Oxford University Press.

Lee, C.L. and Reed, R. (2014), "Volatility decomposition of australian housing prices", *Journal of Housing Research*, Vol. 23 No. 1, pp. 21-43, doi: 10.1080/10835547.2013.12092084.

Li, K.W. (2012), "A study on the volatility forecast of the US housing market in the 2008 crisis", *Applied Financial Economics*, Vol. 22 No. 22, pp. 1869-1880.

Lopes, S.R.C. and Mendes, B.V.M. (2006), "Bandwidth selection in classical and robust estimation of long memory", *International Journal of Statistics and Systems*, Vol. 1, pp. 107-190.

Maheu, J. (2005), "Can GARCH models capture the long–range dependence?", *Studies in Nonlinear Dynamics and Econometrics*, Vol. 9 No. 4, pp. 1-43.

Milles, W. (2008), "Boom–bust cycles and the forecasting performance of linear and non-linear models of house prices", *Journal of Real Estate Finance and Economics*, Vol. 36, pp. 249-264.

Milles, W. (2011), "Long-Range dependence in U.S home price volatility", *Journal of Real Estate Finance and Economics*, Vol. 42, pp. 329-347.

Nelson, D.B. (1991), "Conditional heteroskedasticity in asset returns: a new approach", *Econometrica*, Vol. 59 No. 2, pp. 703-708.

Newell, G., Acheampong, P. and Karantonis, A. (2002), *The Accuracy of Property Forecasting*, Pacific Rim Real Estate Society Conference, Christchurch, p.11.

Patton, A.J. (2011), "Volatility forecast comparison using imperfect volatility proxies", *Journal of Econometrics*, Vol. 160 No. 1, pp. 246-256.

Sadorsky, P. (2006), "Modeling and forecasting petroleum futures volatility", *Energy Economics*, Vol. 28 No. 4, pp. 467-488.

Savills (2019), "Value of UK housing stock hits record high", available at: www.savills.com/blog/article/274512/residential-property/value-of-uk-housing-stock-hits-record-high.aspx (accessed 8 June 2020).

Segnon, M., Gupta, R., Lesame, K. and Wohar, M.E. (2020), "High–frequency volatility forecasting of US housing markets", *Journal of Real Estate Finance and Economics*, doi: 10.1007/s11146-020-09745-w.

Statistics Finland (2016), "Households' assets", Technical report. Last checked on 08 June 2020. available at: www.stat.fi/til/vtutk/2016/vtutk-2016-2018-06-05-tie-001-en.html

Statistics Finland's PxWeb databases (2020), "Price indices of old dwellings in housing companies", available at: http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin-asu-ashi-nj/statfin-ashi-pxt-112t.px/

Stevenson, S. and McGrath, O. (2003), "A comparison of alternative rental forecasting models: empirical tests on the london office market", *Journal of Property Research*, Vol. 20 No. 3, pp. 235-260.

Tayefi, M. and Ramanathan, T.V. (2012), "An overview of FIGARCH and related time series models", *Austrian Jouranl of Statistics*, Vol. 41 No. 3, pp. 175-196.

Tsay, R.S. (2013), *An Introduction to Analysis of Financial Data with R*, John Wiley and Sons, Inc., Hoboken, NJ.

Xingyi, L. and Zakamulin, V. (2018), "Forecasting stock volatiltiy: the gains form using intraday data", available at SSRN: https//ssrn.com/abstract=2847059 or http://dx.doi.org/10.2139/ssrn.2847059

Zhou, Y. and Haurin, D.R. (2010), "On the determinants of house value volatility", *The Journal of Real Estate Research*, Vol. 32, pp. 377-396.

Zhou, J. and Kang, Z. (2011), "A comparison of alternative forecast models of REIT volatility", *The Journal of Real Estate Finance and Economics*, Vol. 42 No. 3, pp. 275-294, doi: 10.1007/s11146-009-9198-7.

Further reading

Statistics Finland Overview (2018), "2018 Overview, household–dwelling units and housing conditions", Technical report, Last checked on 09 June 2020. available at: www.stat.fi/til/asas/2018/011/asas-2018-01-2019-10-10-kat-002-en.html

Corresponding author
Josephine Dufitinema can be contacted at: josephine.dufitinema@uwasa.fi

Appendix

Comparison of time series models

Regions	Cities/sub-areas	One room flats ARMA Order (p,q)	AIC	ARFIMA Order (p,d,q)	AIC
Helsinki	hki3	(2,1)	**685.503**	(2,0.14,1)	687.823
Tampere	tre	(1,1)	678.811	(2,0.20,2)	**662.259**
	tre2	(1,1)	**747.802**	(0,0.31,2)	752.563
Oulu	oulu2	(1,0)	723.337	Anti-persistent	–
Lahti	lti2	(1,0)	798.635	(1,0.24,2)	**794.762**
Joensuu	jnsu	(0,3)	**730.946**	(1,0.05,2)	732.678
Vaasa	vaasa	(0,1)	**785.643**	(0,0.15,3)	786.159
	vaasa1	(0,1)	702.467	Anti-persistent	–
Hämeenlinna	hnlina1	(0,3)	**662.039**	(1,0.09,2)	663.959
Kotka	kotka1	(0,3)	**625.391**	(2,0.46,0)	634.882
		Two rooms flats ARMA Order (p,q)	AIC	ARFIMA Order (p,d,q)	AIC
Tampere	tre	(2,1)	587.509	(2,0.27,1)	**585.939**
	tre3	(2,2)	631.758	(2,0.31,2)	**630.768**
Turku	tku1	(2,0)	699.340	(3,0.06,0)	**696.621**
	tku3	(0,3)	721.061	(0,0.15,3)	**703.969**
Oulu	oulu	(2,0)	627.435	(0,0.30,3)	**626.219**
	oulu1	(1,2)	**658.029**	(0,0.39,3)	659.520
	oulu2	(0,0)	**705.876**	(0,0.13,2)	707.335
Lahti	lti1	(2,0)	712.556	(2,0.16,0)	**709.631**
	lti2	(1,2)	677.356	(1,0.36,0)	**676.637**
Kuopio	kuo2	(2,0)	662.183	(2,0.20,1)	**659.772**
Joensuu	jnsu	(3,0)	727.037	(2,0.29,0)	**725.219**
Vaasa	vaasa1	(0,2)	**673.098**	(0,0.16,2)	675.471
Lappeenranta	ltra2	(1,0)	761.701	(1,0.01,2)	**751.964**
Kotka	kotka	(0,2)	737.003	(0,0.16,2)	**725.713**
	kotka2	(0,2)	659.653	Anti-persistent	–
		Three rooms flats ARMA Order (p,q)	AIC	ARFIMA Order (p,d,q)	AIC
Helsinki	hki2	(1,0)	**653.996**	(1,0.14,0)	654.658
Oulu	oulu2	(0,3)	708.763	(0,0.19,2)	**706.500**
Lahti	lti2	(2,2)	**707.073**	(2,0.37,2)	710.338
Pori	pori	(2,2)	770.727	(1,0.12,2)	**765.959**
Joensuu	jnsu	(1,0)	783.782	(0,0.27,2)	**780.175**
	jnsu1	(1,0)	712.655	Anti-persistent	–
Kouvola	kou	(0,3)	**778.805**	(0,0.41,2)	779.629
Hämeenlinna	hnlina	(0,3)	776.563	(0,0.26,3)	**771.045**

Notes: This table records, for every city and sub-area, the estimated Akaike information criteria (AICs) for model comparison. The favourable model is the one witd the minimum AIC value. The "anti-persistent" refers to the series with long-range negative dependence, meaning that their estimated fractional differencing parameter d varied from −0.5 to 0. The best model's values are marked in bold

Table A1.
In-sample fit – returns models

Regions	Cities/sub-areas	ARMA RMSE	ARMA MAE	ARFIMA RMSE	ARFIMA MAE	The best model
One room flats						
Helsinki	hki3	**0.0393**	**0.0341**	0.0404	0.0346	ARMA
Tampere	tre	0.0344	0.0265	**0.0336**	**0.0265**	ARFIMA
	tre2	**0.0642**	**0.0495**	0.0676	0.0530	ARMA
Oulu	oulu2	0.0695	0.0507	Anti-persistent	–	–
Lahti	lti2	**0.0713**	**0.0500**	0.0714	0.0507	ARMA
Joensuu	jnsu	0.0595	0.0485	**0.0588**	**0.0471**	ARFIMA
Vaasa	vaasa	0.0831	0.0703	**0.0814**	**0.0678**	ARFIMA
	vaasa1	0.0879	0.0751	Anti-persistent	–	–
Hämeenlinna	hnlina1	0.0558	0.0544	**0.0548**	**0.0537**	ARFIMA
Kotka	kotka1	0.0548	0.0548	**0.0393**	**0.0393**	ARFIMA
Two rooms flats						
		ARMA RMSE	ARMA MAE	ARFIMA RMSE	ARFIMA MAE	The best model
Tampere	tre	0.0133	0.0102	**0.0131**	**0.0099**	ARFIMA
	tre3	0.0285	0.0219	**0.0278**	**0.0214**	ARFIMA
Turku	tku1	0.03622	0.02978	**0.03623**	**0.02977**	ARFIMA
	tku3	0.0335	0.0231	**0.0330**	**0.0223**	ARFIMA
Oulu	oulu	**0.0295**	**0.0237**	0.0297	0.0239	ARMA
	oulu1	**0.0405**	**0.0354**	0.0406	0.0354	ARMA
	oulu2	**0.0451**	**0.0327**	0.0451	0.0329	ARMA
Lahti	lti1	**0.0551**	**0.0441**	0.0552	0.0442	ARMA
	lti2	0.0298	0.0217	**0.0290**	**0.0212**	ARFIMA
Kuopio	kuo2	0.0389	0.0311	**0.0372**	**0.0296**	ARFIMA
Joensuu	jnsu	0.0344	0.0284	**0.0334**	**0.0272**	ARFIMA
Vaasa	vaasa1	0.0322	0.0261	**0.0321**	**0.0261**	ARFIMA
Lappeenranta	ltra2	0.0526	0.0454	**0.0526**	**0.0453**	ARFIMA
Kotka	kotka	0.0587	0.0489	**0.0584**	**0.0488**	ARFIMA
	kotka2	0.1010	0.0894	Anti-persistent	–	–
Three rooms flats						
		ARMA RMSE	ARMA MAE	ARFIMA RMSE	ARFIMA MAE	The best model
Helsinki	hki2	0.0117	0.0101	**0.0116**	**0.0099**	ARFIMA
Oulu	oulu2	0.0461	0.0392	**0.0455**	**0.0382**	ARFIMA
Lahti	lti2	0.0454	0.0382	**0.0439**	**0.0351**	ARFIMA
Pori	pori	**0.0776**	**0.0577**	0.0779	0.0578	ARMA
Joensuu	jnsu	**0.0675**	**0.0550**	0.0678	0.0554	ARMA
	jnsu1	0.0667	0.0578	Anti-persistent	–	–
Kouvola	kou	0.0681	0.0558	**0.0668**	**0.0546**	ARFIMA
Hämeenlinna	hnlina	0.0527	0.0405	**0.0524**	**0.0399**	ARFIMA

Notes: This table records the root mean squared error (RMSE) and the mean absolute error (MAE) values of the two competing models in forecasting the house price returns. The estimation sample is 1988:*Q*1–2013:*Q*4, whereas the forecasting sample is 2014:*Q*1–2018:*Q*4. The "anti-persistent" refers to the series with long-range negative dependence, meaning that their estimated fractional differencing parameter d varied from −0.5 to 0. The best model's values are marked in bold

Table A2.
Results of RMSE and MAE – return models

Comparison of time series models

		One room flats					
		EGARCH		FIGARCH		CGARCH	
Regions	Cities/sub-areas	Order (q,p)	AIC	Order (q,d,p)	AIC	Order (q,p)	AIC
Helsinki	hki	(1,3)	4.781	(1,0.58,3)	**4.745**	(2,1)	4.758
	hki1	(2,2)	5.608	(1,0.47,1)	**5.529**	(1,2)	5.584
	hki2	(1,1)	4.966	(2,0.58,3)	**4.844**	(2,1)	4.939
	hki4	(2,3)	**5.500**	(3,0.72,3)	5.562	(2,3)	5.622
Tampere	tre1	(3,2)	**5.694**	(3,0.54,2)	5.845	(1,2)	5.945
	tre3	(3,2)	**5.812**	(1,0.20,1)	5.923	(1,1)	5.961
Turku	tku	(2,3)	**5.487**	(2,0.15,1)	5.587	(1,1)	5.572
	tku1	(3,2)	**5.992**	(1,0.17,1)	6.202	(1,1)	6.203
	tku2	(2,3)	**6.423**	(1,0.54,1)	6.666	(1,1)	6.701
	tku3	(3,3)	6.444	(3,0.23,3)	**6.432**	(1,1)	6.505
Oulu	oulu	(2,3)	**5.662**	(1,-0.20,1)	5.690	(1,1)	5.763
	oulu1	(2,3)	**5.874**	(1,0.02,1)	6.033	(1,1)	6.060
Lahti	lti	(3,2)	**6.123**	(1,0.07,1)	6.151	(1,2)	6.153
	lti1	(2,3)	**6.556**	(1,0.82,1)	6.642	(1,1)	6.683
Jyväskylä	jkla	(3,2)	**5.760**	(3,0.15,2)	6.029	(3,3)	5.779
	jkla1	(3,1)	5.795	(1,-0.05,2)	**5.685**	(1,1)	5.910
	jkla2	(3,3)	6.781	(1,0.37,2)	**6.706**	(1,1)	6.904
Pori	pori	(2,3)	6.746	(1,-0.19,2)	**6.621**	(2,1)	6.898
	pori1	(1,2)	**6.840**	(2,0.13,1)	7.091	(2,1)	7.164
Kuopio	kuo	(3,1)	**5.496**	(2,0.34,1)	5.713	(2,1)	5.726
	kuo1	(2,1)	6.329	(2,0.30,1)	**6.297**	(2,3)	6.310
	kuo2	(3,3)	**6.321**	(2,0.58,3)	6.593	(1,2)	6.659
Joensuu	jnsu1	(2,2)	**6.002**	(1,-0.09,3)	6.065	(1,1)	6.188
Kouvola	kou	(1,3)	**6.551**	(2,0.05,1)	6.605	(1,2)	6.627
Lappeenranta	lrta	(2,2)	6.045	(2,0.42,1)	**5.989**	(1,1)	6.032
	lrta1	(3,3)	6.616	(3,0.42,3)	**6.538**	(1,2)	6.672
Hämeenlinna	hnlina	(3,2)	**6.146**	(1,0.10,1)	6.222	(1,1)	6.264
Kotka	kotka	(2,1)	6.239	(3,0.28,2)	**6.223**	(1,1)	6.303

		Two rooms flats					
		EGARCH		FIGARCH		CGARCH	
		Order (q,p)	AIC	Order (q,d,p)	AIC	Order (q,p)	AIC
Helsinki	hki	(2,3)	4.579	(1,0.37,1)	**4.576**	(1,1)	4.601
	hki1	(2,3)	**5.536**	(1,0.27,1)	5.695	(1,1)	5.738
	hki2	(2,3)	**4.719**	(1,0.73,1)	4.747	(1,1)	4.768
	hki3	(1,3)	5.207	(2,0.08,1)	**5.162**	(2,3)	5.193
	hki4	(1,3)	**5.026**	(2,0.01,1)	5.132	(1,1)	5.085
Tampere	tre1	(1,3)	**5.011**	(1,0.34,2)	5.183	(1,1)	5.255
	tre2	(3,3)	**5.633**	(1,0.27,2)	5.702	(1,1)	5.825
Turku	tku	(3,1)	5.133	(1,0.19,1)	5.102	(1,3)	**5.086**
	tku2	(2,3)	**5.854**	(1,0.11,1)	5.871	(1,1)	5.890
Lahti	lti	(2,2)	**5.056**	(2,0.20,2)	5.120	(2,1)	5.176
Jyväskylä	jkla	(2,2)	**4.956**	(1,0.35,1)	5.085	(1,1)	5.070
	jkla1	(2,2)	**5.233**	(2,0.42,3)	5.308	(1,1)	5.394
	jkla2	(1,3)	**5.745**	(1,0.09,1)	5.811	(1,1)	5.793
Pori	pori	(2,3)	**5.891**	(1,0.23,1)	5.923	(1,2)	5.912
	pori1	(3,3)	**6.211**	(2,0.04,1)	6.316	(2,1)	6.334
	pori2	(1,1)	**6.251**	(1,0.17,1)	6.328	(1,1)	6.414
Kuopio	kuo	(2,1)	5.146	(1,0.26,2)	**5.087**	(1,1)	5.176

(*continued*)

Table A3.
In-sample fit – volatility models

IJHMA

		Two rooms flats					
		EGARCH		FIGARCH		CGARCH	
		Order (q,p)	AIC	Order (q,d,p)	AIC	Order (q,p)	AIC
	kuo1	(3,1)	**5.708**	(3,0.37,1)	5.875	(2,1)	5.896
Joensuu	jnsu1	(2,3)	6.053	(1,-0.08,3)	**6.047**	(1,1)	6.176
Seinäjoki	seoki	(1,1)	6.341	(2,0.44,1)	**6.339**	(1,1)	6.370
Vaasa	vaasa	(3,1)	5.418	(2,0.36,2)	5.329	(2,1)	**5.323**
Kouvola	kou	(3,1)	**5.948**	(1,0.40,2)	6.034	(1,2)	6.129
Lappeenranta	lrta	(3,1)	**5.455**	(3,0.15,1)	5.511	(2,1)	5.566
	lrta1	(1,2)	6.011	(2,-0.32,1)	**5.912**	(1,1)	6.094
Hämeenlinna	hnlina	(2,3)	**5.769**	(1,0.01,1)	5.832	(1,1)	5.818
	hnlina1	(2,2)	**5.943**	(3,0.40,3)	5.964	(1,2)	6.059
Kotka	kotka1	(2,3)	**6.269**	(2,0.42,2)	6.408	(1,2)	6.404
		Three rooms flats					
Regions	Cities/sub-areas	EGARCH		FIGARCH		CGARCH	
		Order (q,p)	AIC	Order (q,d,p)	AIC	Order (q,p)	AIC
Helsinki	hki	(2,2)	**4.908**	(1,0.45,2)	5.011	(1,1)	5.010
	hki1	(3,1)	**5.826**	(1,0.70,1)	5.962	(1,1)	5.968
	hki3	(2,1)	**5.350**	(1,0.44,1)	5.373	(1,1)	5.404
	hki4	(2,2)	**5.193**	(1,0.09,1)	5.335	(1,1)	5.313
Tampere	tre	(3,2)	**5.134**	(2,0.37,1)	5.190	(1,1)	5.185
	tre1	(1,2)	5.759	(3,0.36,1)	**5.743**	(1,2)	5.828
	tre2	(3,1)	**6.035**	(1,0.27,1)	6.109	(1,1)	6.230
	tre3	(1,2)	5.176	(1,0.32,2)	**5.087**	(1,2)	5.199
Turku	tku	(3,2)	**5.419**	(1,0.36,1)	5.442	(1,1)	5.435
	tku1	(2,3)	**6.064**	(1,0.42,1)	6.068	(1,1)	6.074
	tku2	(1,3)	**5.798**	(3,0.54,1)	5.867	(1,1)	5.900
	tku3	(2,3)	**5.547**	(1,0.62,1)	5.679	(1,2)	5.700
Oulu	oulu	(2,3)	**5.275**	(3,0.37,2)	5.369	(1,1)	5.395
	oulu1	(1,2)	**5.680**	(1,0.41,1)	5.828	(1,1)	5.837
Lahti	lti	(1,1)	**5.579**	(1,0.07,1)	5.675	(1,1)	5.687
	lti1	(3,1)	**6.064**	(2,0.11,1)	6.138	(1,1)	6.179
Jyväskylä	jkla	(3,3)	5.681	(3,0.29,1)	5.649	(1,2)	**5.628**
	jkla1	(1,1)	5.965	(2,0.38,2)	**5.935**	(1,2)	5.965
	jkla2	(3,2)	6.271	(3,0.33,1)	**6.243**	(1,1)	6.394
Pori	pori1	(3,1)	6.504	(1,0.27,3)	**6.455**	(1,2)	6.618
Kuopio	kuo	(3,3)	**5.528**	(3,0.24,1)	5.656	(1,2)	5.709
	kuo1	(1,1)	6.501	(2,0.33,2)	**6.381**	(1,1)	6.503
	kuo2	(2,2)	**5.601**	(2,0.15,1)	5.872	(1,1)	5.873
Seinäjoki	seoki	(1,2)	6.651	(1,0.29,1)	**6.522**	(1,1)	6.688
Vaasa	vaasa	(2,1)	**5.776**	(1,0.21,1)	5.820	(1,1)	5.883
	vaasa1	(2,2)	**6.050**	(2,0.16,1)	6.207	(1,1)	6.252
	vaasa2	(1,1)	**6.769**	(2,1.00,2)	6.955	(1,1)	6.781
Lappeenranta	lrta	(2,2)	**5.977**	(2,0.21,1)	6.153	(1,1)	6.209
	lrta2	(3,1)	**6.326**	(1,0.82,3)	6.465	(1,2)	6.583
Hämeenlinna	hnlina1	(3,3)	**6.445**	(2,0.58,3)	6.637	(1,1)	6.685
Kotka	kotka	(1,2)	**6.275**	(3,0.69,1)	6.367	(1,1)	6.344

Notes: This table records, for every city and sub-area, the estimated Akaike information criteria (AICs) for model comparison. The favourable model is the one with the minimum AIC value. The best model's values are marked in bold

Table A3.

Comparison of time series models

Table A4. Results of RMSE and MAE – volatility models

Regions	Cities/sub-areas	One room flats						
		EGARCH		FIGARCH		CGARCH		
		RMSE	MAE	RMSE	MAE	RMSE	MAE	The best model
Helsinki	hki	*0.0112*	*0.0100*	0.0123	0.0113	0.0121	0.0111	EGARCH
	hki1	0.0236	0.0199	0.0193	0.0170	*0.0174*	*0.0156*	CGARCH
	hki2	*0.0118*	*0.0102*	0.0151	0.0134	0.0158	0.0142	EGARCH
	hki4	0.0205	0.0161	0.0188	0.0156	*0.0174*	*0.0142*	CGARCH
Tampere	tre1	0.0398	0.0306	*0.0369*	*0.0319*	0.0374	0.0325	FIGARCH
	tre3	*0.0607*	*0.0431*	0.0635	0.0435	0.0624	0.0436	EGARCH
Turku	tku	0.0217	0.0165	*0.0163*	*0.0138*	0.0387	0.0358	FIGARCH
	tku1	0.0404	0.0296	0.0381	0.0314	*0.0381*	*0.0295*	CGARCH
	tku2	*0.0347*	*0.0297*	0.0445	0.0341	0.0497	0.0390	EGARCH
	tku3	0.0403	0.0335	0.0406	0.0345	*0.0391*	*0.0335*	CGARCH
Oulu	oulu	0.0358	0.0241	0.0354	0.0249	*0.0345*	*0.0236*	CGARCH
	oulu1	0.0569	0.0404	0.0536	0.0383	*0.0515*	*0.0365*	CGARCH
Lahti	lti	0.0553	0.0388	0.0562	0.0393	*0.0541*	*0.0389*	CGARCH
	lti1	0.1681	0.1311	*0.1572*	*0.1212*	0.1586	0.1224	FIGARCH
Jyväskylä	jkla	0.0353	0.0291	0.0431	0.0347	*0.0342*	*0.0286*	CGARCH
	jkla1	0.0388	0.0306	*0.0364*	*0.0319*	0.0366	0.0313	FIGARCH
	jkla2	0.0861	0.0677	0.0792	0.0589	*0.0741*	*0.0584*	CGARCH
Pori	pori	0.0617	0.0522	*0.0612*	*0.0518*	0.0614	0.0524	FIGARCH
	pori1	0.0601	0.0453	*0.0473*	*0.0378*	0.0473	0.0385	FIGARCH
Kuopio	kuo	0.0289	0.0206	*0.0276*	*0.0208*	0.0310	0.0268	FIGARCH
	kuo1	0.0723	0.0463	*0.0623*	*0.0403*	0.0645	0.0386	FIGARCH
	kuo2	0.0959	0.0785	0.0959	0.0774	*0.0928*	*0.0747*	CGARCH
Joensuu	jnsu1	0.0656	0.0404	0.0638	0.0388	*0.0624*	*0.0373*	CGARCH
Kouvola	kou	0.0591	0.0433	0.0567	0.0405	*0.0560*	*0.0404*	CGARCH
Lappeenranta	lrta	0.0384	0.0311	*0.0383*	*0.0320*	0.0388	0.0326	FIGARCH
	lrta1	0.0574	0.0471	0.0466	0.0409	*0.0443*	*0.0381*	CGARCH
Hämeenlinna	hnlina	0.0491	0.0358	*0.0424*	*0.0310*	0.0427	0.0312	FIGARCH
Kotka	kotka	0.0283	0.0230	0.0293	0.0236	*0.0277*	*0.0229*	CGARCH

(*continued*)

Table A4.

		Two room flats						
		EGARCH		FIGARCH		CGARCH		
		RMSE	MAE	RMSE	MAE	RMSE	MAE	The best model
Helsinki	hki	0.0103	0.0090	*0.0097*	*0.0087*	0.0100	0.0089	FIGARCH
	hki1	0.0492	0.0393	*0.0132*	*0.0111*	0.0137	0.0111	FIGARCH
	hki2	*0.0087*	*0.0070*	0.0087	0.0074	0.0088	0.0076	EGARCH
	hki3	0.0234	0.0198	0.0194	0.0169	*0.0180*	*0.0159*	CGARCH
	hki4	0.0220	0.0198	0.0241	0.0213	*0.0211*	*0.0189*	CGARCH
Tampere	tre1	0.0213	0.0184	*0.0171*	*0.0149*	0.0216	0.0201	FIGARCH
	tre2	0.0235	0.0201	*0.0215*	*0.0169*	0.0222	0.0184	FIGARCH
Turku	tku	*0.0133*	*0.0115*	0.0150	0.0131	0.0171	0.0152	EGARCH
	tku2	0.0325	0.0254	0.0373	0.0339	*0.0323*	*0.0295*	CGARCH
Lahti	lti	0.0205	0.0177	0.0178	0.0154	*0.0178*	*0.0152*	CGARCH
Jyväskylä	jkla	0.0226	0.0167	0.0244	0.0172	*0.0209*	*0.0133*	CGARCH
	jkla1	*0.0210*	*0.0143*	0.0212	0.0140	0.0213	0.0153	EGARCH
	jkla2	0.0652	0.0419	*0.0650*	*0.0395*	0.0650	0.0396	FIGARCH
Pori	pori	*0.0428*	*0.0336*	0.0498	0.0372	0.0447	0.0341	EGARCH
	pori1	0.0589	0.0442	0.0589	0.0441	*0.0570*	*0.0428*	CGARCH
	pori2	0.0379	0.0342	*0.0342*	*0.0294*	0.0382	0.0345	FIGARCH
Kuopio	kuo	0.0174	0.0148	0.0177	0.0146	*0.0172*	*0.0148*	CGARCH
	kuo1	0.0216	0.0189	*0.0191*	*0.0177*	0.0195	0.0182	FIGARCH
Joensuu	jnsu1	*0.0209*	*0.0177*	0.0213	0.0177	0.0218	0.0186	EGARCH
Seinäjoki	seoki	*0.0373*	*0.0315*	0.0374	0.0324	0.0381	0.0337	EGARCH
Vaasa	vaasa	0.0236	0.0206	0.0185	0.0148	*0.0174*	*0.0140*	CGARCH
Kouvola	kou	0.0826	0.0473	*0.0821*	*0.0459*	0.0826	0.0480	FIGARCH
Lappeenranta	lrta	0.0270	0.0214	0.0271	0.0236	*0.0249*	*0.0219*	CGARCH
	lrta1	*0.0331*	*0.0302*	0.0383	0.0351	0.0334	0.0309	EGARCH
Hämeenlinna	hnlina	0.0266	0.0221	*0.0261*	*0.0214*	0.0261	0.0216	FIGARCH
	hnlina1	0.0329	0.0246	0.0326	0.0251	*0.0319*	*0.0245*	CGARCH
Kotka	kotka1	0.0792	0.0631	0.0746	0.0604	*0.0745*	*0.0604*	CGARCH

(*continued*)

Comparison of time series models

Table A4.

Regions	Cities/sub-areas	Three rooms flats EGARCH RMSE	EGARCH MAE	FIGARCH RMSE	FIGARCH MAE	CGARCH RMSE	CGARCH MAE	The best model
Helsinki	hki	0.0162	0.0143	0.0158	0.0143	*0.0136*	*0.0123*	CGARCH
	hki1	*0.0203*	*0.0171*	0.0207	0.0173	0.0219	0.0180	EGARCH
	hki3	0.0179	0.0135	0.0174	0.0146	*0.0174*	*0.0139*	CGARCH
	hki4	*0.0184*	*0.0154*	0.0229	0.0199	0.0186	0.0156	EGARCH
Tampere	tre	*0.0131*	*0.0114*	0.0161	0.0140	0.0177	0.0157	EGARCH
	tre1	0.0211	0.0171	*0.0172*	*0.0139*	0.0213	0.0184	FIGARCH
	tre2	0.0531	0.0307	0.0496	0.0391	*0.0495*	*0.0381*	CGARCH
	tre3	0.0177	0.0148	0.0180	0.0144	*0.0177*	*0.0142*	CGARCH
Turku	tku	0.0203	0.0162	0.0225	0.0185	*0.0200*	*0.0160*	CGARCH
	tku1	0.0360	0.0279	*0.0299*	*0.0256*	0.0318	0.0252	FIGARCH
	tku2	0.0318	0.0259	0.0317	0.0261	*0.0308*	*0.0256*	CGARCH
	tku3	0.0373	0.0265	0.0361	0.0259	*0.0357*	*0.0281*	CGARCH
Oulu	oulu	0.0145	0.0128	0.0138	0.0119	*0.0138*	*0.0119*	CGARCH
	oulu1	*0.0208*	*0.0175*	0.0229	0.0201	0.0248	0.0221	EGARCH
Lahti	lti	0.0271	0.0227	0.0269	0.0226	*0.0268*	*0.0226*	CGARCH
	lti1	0.0382	0.0328	*0.0284*	*0.0236*	0.0286	0.0236	FIGARCH
Jyväskylä	jkla	0.0191	0.0157	*0.0188*	*0.0144*	0.0214	0.0183	FIGARCH
	jkla1	*0.0222*	*0.0194*	0.0230	0.0201	0.0222	0.0194	EGARCH
	jkla2	*0.0444*	*0.0362*	0.0489	0.0372	0.0559	0.0414	EGARCH
Pori	pori1	0.0877	0.0601	*0.0788*	*0.0542*	0.0815	0.0549	FIGARCH
Kuopio	kuo	0.0293	0.0246	*0.0228*	*0.0186*	0.0255	0.0218	FIGARCH
	kuo1	0.0351	0.0294	0.0385	0.0296	*0.0349*	*0.0297*	CGARCH
	kuo2	*0.0514*	*0.0409*	0.0532	0.0413	0.0533	0.0412	EGARCH
Seinäjoki	seoki	0.0467	0.0407	0.0529	0.0463	*0.0440*	*0.0367*	CGARCH
Vaasa	vaasa	0.0347	0.0297	0.0342	0.0259	*0.0339*	*0.0259*	CGARCH
	vaasa1	*0.0403*	*0.0304*	0.0418	0.0310	0.0416	0.0309	EGARCH
	vaasa2	0.0319	0.0278	0.0350	0.0296	*0.0304*	*0.0277*	CGARCH
Lappeenranta	lrta	0.0423	0.0336	*0.0356*	*0.0299*	0.0362	0.0305	FIGARCH
	lrta2	0.0139	0.0130	*0.0070*	*0.0068*	0.0080	0.0072	FIGARCH
Hämeenlinna	hnlina1	0.0449	0.0363	*0.0403*	*0.0333*	0.0418	0.0364	FIGARCH
Kotka	kotka	0.0618	0.0425	*0.0561*	*0.0379*	0.0602	0.0410	FIGARCH

Notes: This table records the root mean squared error (RMSE) and the mean absolute error (MAE) values of the three competing models in forecasting the house price volatility. The estimation sample is 1988:*Q*1–2013:*Q*4, whereas the forecasting sample is 2014:*Q*1–2018:*Q*4. The best model's values are marked in bold