THREE ESSAYS IN APPLIED FINANCIAL ECONOMETRICS

by

Horatio Morgan

M.A., York University, 2005
M.Sc., University of the West Indies, 2001
B.Sc., University of the West Indies, 1999

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

In the
Department of Economics

© Horatio Morgan 2009
SIMON FRASER UNIVERSITY
Fall 2009

All rights reserved. However, in accordance with the Copyright Act of Canada, this work may be reproduced, without authorization, under the conditions for Fair Dealing. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.
<table>
<thead>
<tr>
<th>Name:</th>
<th>Horatio Morgan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree:</td>
<td>Doctor of Philosophy</td>
</tr>
<tr>
<td>Title of Thesis:</td>
<td>Three Essays in Applied Financial Econometrics</td>
</tr>
<tr>
<td>Examinining Committee:</td>
<td></td>
</tr>
<tr>
<td>Chair:</td>
<td>Christoph Lülfesmann</td>
</tr>
<tr>
<td></td>
<td>Professor</td>
</tr>
<tr>
<td>Robert Jones</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senior Supervisor</td>
</tr>
<tr>
<td></td>
<td>Professor</td>
</tr>
<tr>
<td>Kenneth Kasa</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Supervisor</td>
</tr>
<tr>
<td></td>
<td>Professor</td>
</tr>
<tr>
<td>Brian Krauth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Supervisor</td>
</tr>
<tr>
<td></td>
<td>Associate Professor</td>
</tr>
<tr>
<td>Andrey Pavlov</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Internal Examiner</td>
</tr>
<tr>
<td></td>
<td>Associate Professor</td>
</tr>
<tr>
<td>Date Defended/Approved:</td>
<td>_____ Tuesday, November 24, 2009___</td>
</tr>
</tbody>
</table>
Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada
ABSTRACT

Efficiency is perhaps one of the most important concepts associated with the functioning of markets in modern economies. When markets are efficient, economic theory suggests that the prices we observe reflect the relative scarcity of resources; and hence, effectively channel those resources to their most productive use. The primary objective of this dissertation is to investigate the efficiency property of the U.S. housing market for single-family homes and the stock market. It does so through the application of advanced techniques in financial and time series econometrics.

In relation to the housing market, the empirical evidence is consistent with the version of the efficiency market hypothesis which suggests that asset prices follow a random walk. However, in relation to the stock market, the empirical evidence is inconsistent with the version of the efficient market hypothesis that attributes price changes to the random arrival of new information. For both markets, however, we do not find the empirical evidence to be definitive.

In the context of the crisis that emerged in the subprime mortgage segment of U.S. housing market in 2006, this dissertation also investigates the interdependency structure of the housing market as a secondary objective. The main result suggests that home prices do not comove systematically over time.
DEDICATION

To Sharlene, Kerene and Keleese,

whom I love;

and my mother, Dorothy Morgan,

who bravely carried on after the death of my father.
ACKNOWLEDGEMENTS

First, I would to thank my senior supervisor, Robert Jones, who went beyond the call of duty in being a mentor and a friend. Had it not been for his wisdom, guidance and unwavering belief in my abilities, this dissertation probably would not have materialized.

Next, I also thank my supervisors, Kenneth Kasa and Brian Krauth, whose invaluable contributions to this dissertation exceed their insightful and timely reviews of earlier drafts.

I also take this opportunity to thank Brenda Spotton Visano and Eli Appelbaum for stimulating my interest in economics throughout my graduate studies at York University.

I also thank my mother, Dorothy Morgan; my sisters; Dennis and Grace Brown; Bonitto & Netilee Gayle; Alister & Charmaine Hylton; Majorie Brown, Leslie Hibbert and his family for all their love, support and encouragement during difficult times.

Last, but by no means least, I thank my wife, Sharlene, and my beautiful daughters, Kerene and Keleese, for their patience, sacrifices, unconditional love and support at every step of the way.
TABLE OF CONTENTS

Approval .............................................................................................................................. ii
Abstract .............................................................................................................................. iii
Dedication ........................................................................................................................... iv
Acknowledgements .......................................................................................................... v
Table of Contents ........................................................................................................... vi
List of Figures ................................................................................................................... viii
List of Tables .................................................................................................................. ix

INTRODUCTION .............................................................................................................. 1

CHAPTER 1 The Efficiency of the Market for Single-Family Homes: A wavelet approach ................................................................. 7
  Returns to Housing........................................................................................................... 8
  Data and Preliminary Analysis ...................................................................................... 10
  Wavelet Methodology .................................................................................................. 12
  Continuous Wavelet Transform .................................................................................. 13
  Discrete Wavelet Transform ........................................................................................ 15
  Maximal Overlap Wavelet Transform ....................................................................... 16
  Empirical Results .......................................................................................................... 17
    Market Efficiency Hypothesis .................................................................................... 17
    Decomposition of Returns to Housing ..................................................................... 20
  Conclusion .................................................................................................................... 21

CHAPTER 2 U.S. Housing Market and Interdependencies ........................................... 24
  Data and Preliminary Analysis ...................................................................................... 25
  Empirical Results .......................................................................................................... 25
  Securitization-based Interdependencies .................................................................. 27
  Summary ....................................................................................................................... 30

CHAPTER 3 Stock Market Efficiency: A behavioral perspective .................................... 31
  Data and Preliminary Analysis ...................................................................................... 32
  Feedback Trading Model .............................................................................................. 35
  Empirical Model Specifications .................................................................................... 38
  Empirical Estimation Procedure ................................................................................ 39
  Empirical Results .......................................................................................................... 40
  Conclusion .................................................................................................................... 42

Appendix A Tables ........................................................................................................... 43
<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Figures</td>
<td>52</td>
</tr>
<tr>
<td>C</td>
<td>Stationary and Non-stationary Processes</td>
<td>78</td>
</tr>
<tr>
<td>D</td>
<td>Basis Functions</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>84</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1: S&P/Case-Shiller housing price indices ........................................53
Figure 2: Monthly (log) returns to housing ..............................................54
Figure 3: Histogram and Normal QQ-Plot of returns: Boston ................55
Figure 4: Histogram and Normal QQ-Plot of returns: Chicago ...............56
Figure 5: Histogram and Normal QQ-Plot of returns: Denver ...............57
Figure 6: Histogram and Normal QQ-Plot of returns: Las Vegas ..........58
Figure 7: Histogram and Normal QQ-Plot of returns: Los Angeles ......59
Figure 8: Histogram and Normal QQ-Plot of returns: Miami ...............60
Figure 9: Histogram and Normal QQ-Plot of returns: New York ..........61
Figure 10: Histogram and Normal QQ-Plot of returns: San Diego .......62
Figure 11: Histogram and Normal QQ-Plot of returns: San Francisco ..63
Figure 12: Histogram and Normal QQ-Plot of returns: Washington, DC .........................................................................................64
Figure 13: Boston – Haar MODWT coefficient vectors .......................65
Figure 14: Chicago – Haar MODWT coefficient vectors .......................66
Figure 15: Denver – Haar MODWT coefficient vectors .........................67
Figure 16: Las Vegas – Haar MODWT coefficient vectors ....................68
Figure 17: Los Angeles – Haar MODWT coefficient vectors .................69
Figure 18: Miami – Haar MODWT coefficient vectors ..........................70
Figure 19: New York – Haar MODWT coefficient vectors ...................71
Figure 20: San Diego – Haar MODWT coefficient vectors ....................72
Figure 21: San Francisco – Haar MODWT coefficient vectors ...............73
Figure 22: Washington DC – Haar MODWT coefficient vectors ..........74
Figure 23: Standard Normal QQ-plots for SPH04 and SPM04 ..............75
Figure 24: Log returns for SPH04 and SPM04 ..................................76
Figure 25: Estimated Conditional Volatilities for SPH04 and SPM04 ...77
LIST OF TABLES

Table 1: Descriptive statistics of monthly home price appreciation rates ........44
Table 2: Unit root tests ..........................................................................................45
Table 3: $Z_a$ statistic values of the Phillips-Ouliaris cointegration test ............46
Table 4: Descriptive statistics of five-minute interval (log) returns .................47
Table 5: Parameter estimates using SPH04 data .............................................48
Table 6: Model diagnostics using SPH04 data ...............................................49
Table 7: Parameter estimates using SPM04 data .............................................50
Table 8: Model diagnostics using SPM04 data ...............................................51
INTRODUCTION

The 1987 stock market crash\(^1\), and the 2006 crisis that emerged in the subprime segment of U.S. housing market are among some of the most destabilizing recent high-impact events experienced by modern economies.\(^2\) Yet economists continue to rely on these markets to determine value, to exhaust all the possible mutually beneficial gains from trade; and ultimately, to allocate the relatively scare financial, physical and human capital to their most productive use.

If markets are efficient, economic theory suggests that the prices we observe reflect the relative scarcity of resources; and hence, effectively channel those resources to their most productive use. The primary objective of this dissertation is to investigate the efficiency property of the U.S. housing market for single-family homes, as well as the stock market. In the context of the crisis that emerged in the subprime mortgage segment of U.S. housing market in 2006, it also investigates the interdependency structure of the housing market as a secondary objective.

According to the efficient market hypothesis, an (informationally) efficient financial market is one in which market prices reflect all available information about the economic values of traded assets. However, three important versions

---

\(^1\) Within a ten-hour trading period (from the close of trade on Friday Oct. 16 to Tuesday Oct. 20): the S&P 500 index fell by 22 percent; the S&P 500 futures index fell by 36 percent (Edwards, 1988).

\(^2\) In 2006, there was a significant spike in defaults on mortgages and foreclosures that were linked to toxic subprime mortgage loans in one way or another. By the following year, there was a major breakdown in global financial intermediation (“credit crunch”). As of March 2008: the International Monetary Fund (IMF) estimates aggregate potential writedowns and losses arising from the crisis to be approximately $945 billion (IMF, 2008).
of this hypothesis emerge based on what one considers to constitute *all available information*. For instance, when *all available information* is limited to past prices, we have the *weak form* of market efficiency. This implies that profits cannot be generated consistently from technical trading rules. The empirical implication is that prices (plus dividends) follow a random walk model (Fama, 1970).

If *all available information* is broadened to constitute all publicly available information (i.e. past prices + news), then we have a *semi-strong form* of market efficiency. This version of market efficiency implies that prices react to news quickly, which is usually evaluated empirically by event studies (Fama, 1970; Fama et. al. 1969; and Ball and Brown, 1968).

Finally, if *all available information* constitutes all information (i.e. past prices + news + private information), then we have the *strong form* of market efficiency. This version of market efficiency is commonly understood to imply that mutual fund managers cannot consistently outperform the market (Fama, 1970; Treynor, 1965; and Jensen, 1968).

In chapter 1, I follow a large body of empirical literature that evaluates the efficiency of the real estate market based on the *weak form* of market efficiency.\(^3\) However, these studies vary considerably along a number of dimensions. For example, real estate may be classified as residential (i.e. single family homes), commercial (i.e. office or industrial), business (i.e. REITs), or land. In addition,

real estate studies on real estate could focus on urban and/or rural areas at the local, national or international level. Furthermore, in evaluating price changes, different levels of aggregate may be applied (i.e. individual prices or home price indices).

Given the different characteristics of these studies, it is not surprising that they have produced mixed evidence on the efficiency of the real estate market. For example, in their study of commercial real estate in the U.S., McIntosh and Henderson (1989) find evidence in support of weak form market efficiency. In relation to residential real estate in the U.S., Green et al. (1988) provide evidence that also appears to be consistent with the weak form of market efficiency. However, Wang’s (2004) study of residential real estate in U.S. provides evidence that appears to be inconsistent with this version of market efficiency. Thus, the evidence on market efficiency for the real estate market remains inconclusive and difficult to reconcile based on the different characteristics of each study undertaken.

There are a number of competing views on why asset prices sometimes diverge from their fundamental level. At the risk oversimplifying nuanced theories on asset pricing, one can perhaps broadly group these theories as either rational or irrational agent models. For example, conventional theories that attribute asset price inflation to low market liquidity are example of the former.

Pavlov and Wachter (2002, 2005) also provide another explanation of inflated asset prices, which is consistent with a model of rational decision-making. In their asset-pricing model, they characterize the marginal homebuyer’s decision
as represented by the purchase of a portfolio that comprises a mortgage loan and an embedded put option. This put option gives the borrower the right, but not the obligation, to default on her mortgage loan by “selling” the real estate asset to the lender for the outstanding mortgage balance. If this put option is correctly priced to reflect the risk of default, then its impact on real estate will be neutral. However, this is not the case if the put option is mistakenly underpriced. In that case, the underpriced put option induces an inflation in real estate prices in excess of their fundamental level. Moreover, if underpricing is severe, then the realignment of real estate prices with their fundamental values may call for significant downward adjustment in prices (Pavlov, 2006). However, such a phenomenon could engender very costly market crashes.

Shiller (1981) is among the first studies to present a convincing alternative explanation of asset pricing outside of a rational agent framework. Specifically, Shiller suggested that the changes in dividends were relatively too smooth to account for the observed volatile stock prices. This observation implies that one had to look beyond fundamental factors in order to understand why asset prices sometimes deviate from the fundamental level.

The study undertaken in Chapter 1 is most similar to that of Case and Shiller (1989) with a focus on the weak form of market efficiency. It investigates whether the housing market for single-family homes is efficient. In particular, it seeks to determine whether home prices have a unit root that is induced by a stochastic trend. However, it departs from their study primarily in the application of wavelet methods. Specifically, it uses conventional time series and wavelet
methods to test for a unit root in home prices. For all the cases considered in this study, Gençay and Fan’s (forthcoming) wavelet-based unit root test and the conventional Augmented Dickey Fuller (ADF) unit root test provide evidence that is consistent with the weak form of market efficiency. However, this finding is tempered by the exclusion of dividends from returns to housing, as well as by concerns about the statistical power of the tests employed.

Chapter 2 goes beyond the efficiency issue tackled in Chapter 1. In particular, it investigates the interdependency structure of the U.S. housing market for single-family homes within the context of the recent housing market crisis. The empirical evidence from cointegration tests suggests that the interrelationships within the housing market itself are not the principal source behind the widespread collapse of home prices in 2006. However, one should consider the potential link between the securitization process and the housing crisis.

In chapter 3, the focus shifts from the housing market to the stock market. It evaluates efficiency from a behavioural perspective in this case. In particular, it uses a family of quadratic GARCH-in-mean models to determine whether trend-chasing behaviour characterizes the trading strategy of market participants in the stock market. If a sufficiently large number of traders pursue this kind of trading strategy, then stock prices could significantly deviate from their fundamental values over some period. The main empirical result suggests that market participants are employing a positive feedback trading strategy. This finding is robust to model specifications and the futures contracts used. At the same time,
positive feedback trading strategy is not a sufficient condition for market inefficiency.
CHAPTER 1
THE EFFICIENCY OF THE MARKET FOR SINGLE-FAMILY HOMES: A WAVELET APPROACH

The recent policy gridlock on the resolution of U.S. housing market crisis appears to reflect different perspectives on the efficiency of the housing market. For instance, Feldstein (2008) advocates home price stabilization policies, while those with greater faith in the efficiency of the housing market regard such policies as distortionary. But the question of market efficiency is ultimately an empirical one. Therefore, the primary objective of this paper is to examine the empirical evidence for random walk type behaviour in home prices. Specifically, it tests whether the U.S. market for single-family homes is (weak form) efficient.

While this study is most similar to Case and Shiller (1989), it makes a major departure in its application of wavelet methods. The wavelet applications in this study feature the use of Gençay and Fan’s (forthcoming) wavelet-based unit root test, and a scale-based decomposition of the variability of the returns to housing. The wavelet-based unit root is used in this study to corroborate the results of the conventional Augmented Dickey Fuller (ADF) unit root test due to concern that the relatively small sample size of the home prices series used could weaken the statistical power of the test. In addition, a wavelet decomposition of the returns to housing could provide us with useful information.

---

4 Under reasonable empirical sizes, Gençay and Fan demonstrate in simulation studies that their wavelet-based unit root test has superior statistical power relative to the ADF.
on potentially fruitful directions for future research on the dynamics of home prices.

**Returns to Housing**

A home is a unique commodity that provides a flow of housing services and an investment opportunity to a potential investor-consumer agent. Following Clayton (1998), I define the expected rate of return to housing investment as

\[
R_t = \frac{E[P_{t+1}|I_t] - P_t}{P_t} + \frac{s_t - (T_t + d_t)}{P_t}
\]  

(1.1)

where \(E\) is the conditional expectation given information available at time \(t\); \(P\) denotes house price; \(s\) is rental income (which is actual for a renter and imputed for a homeowner); \(T\) and \(d\) denote the property tax and operating (i.e. maintenance and depreciation) costs, and \(I\) denotes the information set available to agents at the time they form expectations about future house prices. To simplify the analysis, I further assume that

\[
s_t - (T_t + d_t) \approx 0
\]  

(1.2)

so that the return to housing can be conveniently approximated by

\[
R_t \approx \frac{E[P_{t+1}|I_t] - P_t}{P_t}
\]  

(1.3)

What justifications do I offer for ignoring dividends in my model of returns to housing? I submit the following three:
1. **Data problem**: To the best of my knowledge, I am not aware of any accessible and reliable publicly available data on dividends (e.g. imputed rental income) to housing;

2. **Absence of "best practice"**: There is no unified theoretical framework that identifies the "best practice" among several alternative ways of computing the costs and by extension the returns to housing. Besides, these studies appear to focus exclusively on the relationship between different measures of housing costs and inflation (e.g., consumer price index) (Dougherty and Van Order, 1982; Poole, Ptacek and Verbrugge, 2005).

3. **Convenience**: The approximation of returns to housing as the home price appreciation rate allows for the use of publicly available home price index data. The family of home price indices provided by S&P/Case-Shiller is among the most reliable set of publicly available home price indices.

The cost of ignoring dividends perhaps translates into weaker conclusions about the efficiency of the U.S. housing market. However, given the scope and the objective of this thesis, I find such cost to be tolerable. At the same time, the approximation of returns to housing by equation (1.3) is most appropriate when homebuyers’ decisions are primarily based on the real estate asset itself rather than the embedded put option - which gives the holder the right but not the obligation to return the home to the bank – or the flow of dividends (i.e. imputed rents).
Data and Preliminary Analysis

Monthly data on the original 10 S&P/Case-Shiller U.S. National Home Price Indices (i.e. Composite of 10) over the period January 1987 to April 2008 are obtained from Standard & Poor's website on June 30, 2008. The Composite of 10 includes the following home price indices (HPIs): Boston (BOXR), Chicago (CHXR), Denver (DNXR), Las Vegas (LVXR), Los Angeles (LXXR), Miami (MIXR), New York (NYXR), San Diego (SDXR), San Francisco (SFXR) and Washington DC (WDXR). The purpose of the HPIs is to measure the average change in single-family home prices in a particular metropolitan area. Figure 1 presents plots of the S&P/Case-Shiller housing price indices over the period of study. For all home price indices shown, we see a run-up in home prices followed by a decline around 2006.

When using an index, one has to be mindful that an index and its individual components do not usually perfectly correlate. Moreover, certain statistical properties of an index may not obtain for the individual components. These observations apply to home price indices as well. Under such conditions, it becomes problematic to use a home price index to hedge against price declines.

---

5 S&P/Case-Shiller later added 10 additional regions - Atlanta, Charlotte, Cleveland, Dallas, Detroit, Minneapolis, Phoenix, Portland (Oregon), Seattle, Tampa - to the original 10 metropolitan areas to form the Composite of 20.

6 For all official purposes in the U.S., estimates of metro area populations are given by Metropolitan Statistical Areas (MSAs), or Metropolitan Divisions as defined by the U.S. Office of Management and Budget. With the exception of the New York region - which includes a number of counties within commuting distance of New York City - the other nine areas which make up the Composite of 10 are based on MSAs.
for a single real estate property. However, this study is more concerned about evaluating the efficiency of the housing market rather than about risk management through hedging. In this context, it must be acknowledged that if we establish market efficiency (inefficiency) based on home price index data, it is still possible that the prices of individual real estate properties could be inefficient (efficient). Thus, we cannot simply assume that the efficiency (inefficiency) characteristic of a home price index obtains for the individual properties that comprise the index.

Table 1 provides a summary of descriptive statistics on the monthly (log) returns to housing. Importantly, returns are characterized by nonzero skewness and excess kurtosis which is inconsistent with a normally distributed process. Figure 2 presents plots of monthly (log) returns over the period of study. The question of whether the returns are generated by a normal distribution is also assessed by through the use of quantile-quantile plots (QQ-plots). QQ-plots are produced by plotting the ordered sample return values against the quantiles of a normal distribution, with a reference line of slope one superimposed. The histograms and QQ-plots for returns are shown in Figures 3 to 12. If the return process is generated by a normal distribution, the sample return values should all lie on the superimposed line; however, this is generally not the case here. Therefore, the return series appear to be generated by some non-normal (fat-tailed) distribution. Despite this informal evidence of non-normality - since the hypothesis of normality is not subjected to a formal test - I assume normality in order to apply unit root tests later on in this study.
Wavelet Methodology

The notion of a wavelet is an important building block for the wavelet methods employed in this paper. Informally, one may think of a wavelet as a small wave that rises and recedes over a finite period.\(^7\) Formally, a wavelet is a continuous real-valued function \(\psi(\cdot)\) on the interval \((-\infty, \infty)\), that satisfies the following two properties:

\[
\int_{-\infty}^{\infty} \psi(t) dt = 0, \tag{1.4}
\]

and

\[
\int_{-\infty}^{\infty} \psi(t)^2 dt = 1. \tag{1.5}
\]

That is, \(\psi(\cdot)\) integrates to zero and has unit energy (variance). In other words, equation (1.5) implies that \(\psi(\cdot)\) must exhibit nonzero activity, and equation (1.4) indicates that excursions above zero cancel out excursions below zero. Thus, equations (1.4) and (1.5) taken together implies that \(\psi(\cdot)\) resembles a wave. In addition, equations (1.4) and (1.5) hold, if what is called the wavelet admissibility condition

\[
0 < C_{\psi} \equiv \int_{0}^{\infty} \frac{|\Psi(f)|}{f} df < \infty \tag{1.6}
\]

holds,\(^8\) where \(\Psi(f) \equiv \int_{-\infty}^{\infty} \psi(t)e^{-i2\pi ft} dt\) is the Fourier transform of the \(\psi(\cdot)\), and \(f\) denotes the frequency. Thus, any function \(\psi(t)\) that satisfies equation (1.6) is a wavelet (filter).

\(^7\) This section closely follows Gençay et. al (2001), but also incorporates material from Crowley (2005), Percival and Walden (2000) and Mallat (1989).

\(^8\) That is, equation (1.6) is a sufficient condition for equations (1.4) and (1.5).
For most practical applications, one typically does not use continuous wavelet filters as defined above; we generally use some finite length discrete wavelet filter. The set of wavelet coefficients, \( h_l = \{h_0, \ldots, h_{L-1}\} \) associated with this finite length discrete wavelet filter must satisfy:

\[
\sum_{l=0}^{L-1} h_l = 0 \quad (1.7)
\]

\[
\sum_{l=0}^{L-1} h_l^2 = 1 \quad (1.8)
\]

\[
\sum_{l=0}^{L-1} h_l h_{l+2n} = 0 \quad (1.9)
\]

Equations (1.7) and (1.8) are analogous to equations (1.4) and (1.5) for the continuous wavelet filter; the coefficients must sum to zero and have unit energy. In addition, equation (1.9) requires that the coefficients be orthogonal to even shifts in the filter.\(^9\)

**Continuous Wavelet Transform (CWT)**

Suppose \( x(t) \) is some time series of interest, and \( \psi(t) \) is some wavelet filter; then the continuous wavelet transform (CWT) is given by

\[
W(u, \lambda) = \int_{-\infty}^{\infty} x(t) \psi_{u,\lambda}(t) dt \quad (1.10)
\]

where \( \psi_{u,\lambda}(t) = \frac{1}{\sqrt{\lambda}} \psi(t - \frac{t-u}{\lambda}) \), with \( \lambda \) and \( u \) denoting the scale factor and translation factor, respectively; and \( \frac{1}{\sqrt{\lambda}} \) is being used to ensure that there is variance.

---

\(^9\) The length of a discrete wavelet filter is the number of non-zero wavelet coefficients it has. For example, the Haar wavelet filter has the set of wavelet coefficients \( h_l = \{h_0 = 1/\sqrt{2}, h_1 = -1/\sqrt{2}\} \), and therefore has length \( L = 2 \).

\(^{10}\) In the wavelet literature, equations (1.8) and (1.9) are called the orthonormality conditions.
normalization across different scales. Equation (1.10) shows how a time series \( x(t) \) is decomposed into projections on a set of basis functions, or specifically, wavelets.\(^{11}\) These wavelets are generated from the so-called mother wavelet \( \psi(t) \), by scaling (i.e. dilating or stretching) and translating (i.e. shifting) the mother wavelet. If, for example, the choice of the wavelet is the Haar wavelet filter

\[
\psi(t) = \begin{cases} 
1 & \text{if } 0 \leq t \leq 1/2 \\
-1 & \text{if } 0 \leq t \leq 1/2 \\
0 & \text{otherwise}
\end{cases}
\]

then

\[
\psi^H_{u,\lambda}(t) = \begin{cases} 
1/\sqrt{2\lambda} & \text{if } u - \lambda \leq t \leq u \\
-1/\sqrt{2\lambda} & \text{if } u \leq t \leq t + \lambda \\
0 & \text{otherwise}
\end{cases}
\]

Thus, for \( \lambda = 1 \), and \( u \in (-\infty, \infty) \),

\[
W^H(u,1) = \int_{-\infty}^{\infty} x(t)\psi^H_{u,1}(t)dt = \int_{u}^{u+1} x(t)dt - \int_{u-1}^{u} x(t)dt \equiv D(u,1)
\]

If \( x(t) \) is sampled daily, then a plot of \( D(u,1) \) tells us how the daily average of \( x(t) \) is changing from one day to the next. More generally, for any \( \lambda > 0 \),

\[
W^H(u,\lambda) = \int_{-\infty}^{\infty} x(t)\psi^H_{u,\lambda}(t)dt \propto D(u,\lambda) \equiv \frac{1}{\lambda} \int_{u}^{u+\lambda} x(t)dt - \frac{1}{\lambda} \int_{u-\lambda}^{u} x(t)dt,
\]

and the collection of variables \( \{W^H(u,\lambda) : \lambda > 0, -\infty < u < \infty\} \) represents the Haar CWT of \( x(t) \).

\(^{11}\) See Appendix D for a brief exposition on basis functions in this context.
The CWT of $x(\cdot)$ is an energy preserving transformation; that is, there is no loss of information in using the CWT of $x(\cdot)$ rather than $x(\cdot)$. Therefore, we can recover $x(\cdot)$ from its CWT via

$$x(t) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^{\infty} W(u, \lambda) \frac{1}{\sqrt{\lambda}} \psi \left( \frac{t-u}{\lambda} \right) dt d\lambda$$

(1.11)

where $C_\psi$ is as defined in equation (1.6).

**Discrete Wavelet Transform (DWT)**

The discrete wavelet transform is a subsampling of the continuous wavelet transform in which we deal with just dyadic scales (i.e. the scale $\lambda$ is of the form $\lambda = 2^{j-1}, j = 1, 2, 3, \ldots$). Let $y = \{y_t\}_{t=1}^T$ be a dyadic length vector of observations (i.e. $T = 2^J$). The length $T$ vector of discrete wavelet coefficients $w$ is obtained by

$$w = W y$$

(1.12)

where $W$ is a $T \times T$ real-valued orthogonal matrix defining the DWT which satisfies $W' W = I_T$. The vector of wavelet coefficients can be organized into $J + 1$ vectors

$$w = [w_1, w_2, \ldots, w_J, v_J]'$$

(1.13)

where $w_j$ is a vector of wavelet coefficients associated with changes on a scale of length $\lambda_j = 2^{j-1}$, and $v_j$ is a vector of scaling coefficients associated with averages on a scale of length $2^j = 2\lambda_M$. 
Since the matrix W is orthogonal, the DWT is a variance preserving transformation.\(^{12}\) That is,

\[
\|y\|^2 = \sum_{j=1}^{M} \|w_j\|^2 + \|v_j\|^2
\]  \hspace{1cm} (1.14)

**Maximal Overlap Discrete Wavelet Transform (MODWT)**

Let \(y\) be an arbitrary length \(N\) vector of observations.\(^{13}\) The length \((J + 1)N\) vector of MODWT coefficients \(\tilde{w}\) is obtained via

\[
\tilde{w} = \tilde{W}y
\]  \hspace{1cm} (1.15)

where \(\tilde{W}\) is a \((J + 1)N \times N\) (non-orthogonal) matrix defining the MODWT. The vector of MODWT coefficients may be organized into \(J + 1\) vectors

\[
\tilde{w} = [\tilde{w}_1 \; \tilde{w}_2 \ldots \; \tilde{w}_J \; \tilde{v}_J]',
\]  \hspace{1cm} (1.16)

where \(\tilde{w}_j\) is a vector of wavelet coefficients associated with changes on a scale of length \(\lambda_j = 2^{j-1}\), and \(\tilde{v}_j\) is a vector of scaling coefficients associated with averages on a scale of length \(2^J = 2\lambda_J\).

By using the rescaled wavelet and scaling filters, given by \(\tilde{h}_j = h_j / 2^j\) and \(\tilde{g}_j = g_j / 2^j\), the MODWT is also an energy preserving transformation. That is,

\[
\|	ilde{y}\|^2 = \sum_{j=1}^{M} \|\tilde{w}_j\|^2 + \|\tilde{v}_j\|^2
\]  \hspace{1cm} (1.17)

\(^{12}\) For a non-stationary (unit root) process, the variance may be large but finite provided the time horizon is finite; therefore, the decomposition of the variance is still possible for such a process (see appendix C for a more detailed treatment of unit root processes).

\(^{13}\) This allows for a non-dyadic length vector of observations.
In comparing DWT and MODWT, a number of differences stand out. First, MODWT is a nonorthogonal wavelet transformation that can handle nondyadic sample sizes. Second, unlike DWT, MODWT is invariant to circular shifts of the original time series. Third, MODWT yields an asymptotically more efficient wavelet variance estimator. And fourth, a MODWT multiresolution analysis allows us to align the wavelet details and wavelet smooths with the original time series in the time domain. This is so because the wavelet detail and smooth coefficients of a MODWT multiresolution analysis are associated with zero phase filters; this is not the case for a DWT multiresolution analysis. For all these advantages, MODWT tend to feature more than DWT in empirical work. However, the ultimate cost is the loss of the variance preserving property that DWT possesses. Still, by using appropriately rescaled wavelet and scaling coefficients, the variance preserving property is restored.\(^{14}\)

**Empirical Results**

**Market Efficiency Hypothesis**

Market efficiency implies a unit root in the level of price, which in turn implies that price changes are uncorrelated (Nelson and Plosser, 1982).\(^{15}\)

Suppose the univariate home price (level) series \(\{P_{i,t}\}_{t=1}^{T}\), under consideration, is generated by the process

\[ p_t = p_{t-1} + \epsilon_t, \text{ and } \epsilon_t \sim iid(0, \sigma^2_\epsilon) \]

then the first difference \(\Delta p_t = p_t - p_{t-1} = \epsilon_t\) is now a stationary series.

\(^{14}\) See Percival and Walden (2000) for an exposition on the pyramid algorithm that is used to generate wavelet and scaling filters; and ultimately, the wavelet details and smooth.

\(^{15}\) That is, if \( p_t = p_{t-1} + \epsilon_t, \text{ and } \epsilon_t \sim iid(0, \sigma^2_\epsilon) \); then the first difference \(\Delta p_t = p_t - p_{t-1} = \epsilon_t\) is now a stationary series.
\[ \Delta p_{i,t} = \mu + \gamma p_{i,t-1} + \beta t + \epsilon_t \]  

(1.18)

where \( \Delta p_{i,t} = \Delta p_t = p_t - p_{t-1} = \ln(P_{i,t}) - \ln(P_{i,t-1}) \); \( \epsilon_t \sim \text{iid}(0, \sigma^2_\epsilon) \), and \( i \in \{ \text{Boston, Chicago, Denver, Las Vegas, Los Angeles, Miami, New York, San Diego, San Francisco, Washington DC} \} \). Importantly, equation (1.18) uses the natural logarithm of home price levels, \( \{P_{i,t}\}_{t=1}^T \); hence, we may interpret the right-hand variable as the relative change in prices or log returns. A logarithmic transformation of time series data may be necessary, or meaningful, for a number of reasons. First, if the standard deviation of a time series is proportional to its level, then the data expressed in terms of natural logarithm will stabilize the variance, an assumption required for the validity of the statistical tests to follow. That is, the transformation of prices in natural logarithm makes the assumption of independent and identically distributed (iid) error terms across time more plausible. If they are not iid, then the model is mis-specified and your unit root tests invalid.

Second, one is generally more interested in the relative change of a variable of interest rather than its absolute change. And third, it usually makes sense to take the natural logarithm of data that are positive in empirical work.

In testing whether home prices have a unit root, I add a deterministic trend to an AR(1) with no drift to control for any variation in home prices which may be

\[ p_t = \rho p_{t-1} + \epsilon_t, \]

to get

\[ p_t = \rho p_{t-1} + \beta t + \epsilon_t; \]

which can then be expressed as

\[ \Delta p_t = \mu + \gamma p_{t-1} + \beta t + \epsilon_t \]

with an added drift \( \mu \), with \( \rho - 1 = \gamma \). Thus, \( \rho = 1 \Leftrightarrow \gamma = 0 \).

\[ \text{To obtain equation (15), we simply add a time trend component to the AR(1) model,} \]

\[ p_t = \rho p_{t-1} + \epsilon_t, \]

to get

\[ p_t = \rho p_{t-1} + \beta t + \epsilon_t; \]

which can then be expressed as

\[ \Delta p_t = \mu + \gamma p_{t-1} + \beta t + \epsilon_t \]

with an added drift \( \mu \), with \( \rho - 1 = \gamma \). Thus, \( \rho = 1 \Leftrightarrow \gamma = 0 \).

\[ \text{Time series data tend to exhibit variation that increases in both mean and dispersion, where the dispersion is often proportion to the absolute level of the series.} \]
due to a deterministic trend component\textsuperscript{18}. I then add a drift term $\mu$ to capture the tendency of home prices to move upward (or downward) on average over time. This leads to the following hypotheses:

\[
H_0 : \gamma = 0 \text{ (i.e. } \{p_{i,t}\} \text{ is a unit root process)}
\]

\[
H_1 : \gamma < 0 \text{ (i.e. } \{p_{i,t}\} \text{ is a stationary process)}
\]

Table 2 reports Gençay and Fan’s wavelet unit root test statistic, $\hat{S}^{Ld}_{T,1}$\textsuperscript{19}, and the p-values for the ADF unit root test using Case-Shiller/S&P home prices and the AR(1) model specified in equation (1.18). These tests are conducted under the null hypothesis that home prices contain a unit (i.e. $\gamma = 0$), against the alternative that home prices are trend-stationary (i.e. $\gamma < 0$). For all the metropolitan areas considered, $\hat{S}^{Ld}_{T,1} > \text{critical value}$; hence, we fail to reject the null hypothesis that home prices contain a unit root (induced by a stochastic trend). Similarly, the p-values reported for the ADF test all exceed the standard 5% level of significance; hence, we also fail to reject the null of hypothesis of a unit root in home prices. Therefore, these results suggest that home prices

\begin{footnotesize}
\textsuperscript{18} In this case, we have to first difference and detrend the time series before it becomes a stationary series.

\textsuperscript{19} Gençay and Fan provide Splus routines for computing this statistic and its critical values. These routines are available at http://www.sfu.ca/~rgencay/wavunit/. The null is rejected if $\hat{S}^{Ld}_{T,1} < \text{critical value}$.
\end{footnotesize}
behave similar to securities in financial markets whose prices contain a unit root induced by a stochastic trend.

**Decomposition of Returns to Housing**

One may use wavelets to decompose the returns to housing over different monthly scales. A key reason for doing so is to ascertain which scales account for the variability in the rate of home price appreciation. In other words, we wish to ascertain which scales have information content in relation to the variability of the returns to housing. One of the important decisions in any wavelet analysis is the choice of the wavelet filter. While short width filters may introduce certain undesirable artifacts in the data analysis, wide width filters may unduly expose a larger number of wavelet coefficients to boundary effects.\(^{20}\)

The wavelet filter used is the Haar wavelet filter, which is based on two non-zero coefficients (i.e. \(L = 2\)). In addition to the concerns about boundary effects, this filter is chosen over wider width filters also because there were no discernible benefits in using the latter. The level of decomposition chosen is \(J = 4\), which yields wavelet filter and scaling coefficients \(\tilde{w}_1, \tilde{w}_2, \tilde{w}_3, \tilde{w}_4\) and \(\tilde{v}_4\). The Haar MODWT coefficient vectors for monthly (log) returns to housing are shown in Figures 13 to 22.\(^{21}\) The monthly (log) returns to housing always appear at the top, with the wavelet and scaling coefficients below. Since the data were sampled

---

\(^{20}\) Boundary effects generate spurious wavelet coefficients in relation to the beginning and the ending data.

\(^{21}\) The computation of wavelet and scaling coefficients is performed in R using the wavelet package.
at monthly intervals, we interpret the wavelet coefficients at scales 1, 2, 4 and 8 as corresponding to 2- to 4-month, 4- to 8-month, 8- to 16-month and 16- to 32-month period dynamics, respectively. Importantly, the variability in home price appreciation rates appears to be partly accounted for by fluctuations over 16- to 32-month periods. This is so because the fluctuations over the other (higher frequency) scales appear to be white noise.

**Conclusion**

This study investigates whether the U.S. housing market for single-family homes is (weak form) efficient. Gençay and Fan’s wavelet-based unit root test provides evidence that appears to be consistent with (weak form) market efficiency. The conventional ADF unit root test also provides evidence that is consistent with (weak form) market efficiency. As the same time, these findings though indicative of weak form efficiency must be treated with caution for all least two reasons. First, although the wavelet-based unit root test has been shown to have better statistical power to the conventional ADF unit root test under certain conditions, this attribute may be compromised in relatively small samples. This concern is relevant since the time horizon of the housing price data used in this study is relatively short.

Second, since I have ignored dividends - in the form of imputed rents - in modelling returns to housing, this study does not claim to make a definitive statement about the efficiency of the U.S. housing market for single-family
homes. Thus, one should view the evidence provided as being merely consistent with the efficient market hypothesis, rather than a confirmation of market efficiency.

In an earlier study, Case and Shiller’s (1989) provide evidence that is inconsistent with the (weak form) efficiency market hypothesis in relation to U.S. housing market for single-family homes. Importantly, though, the authors acknowledge that their findings do not convey a definitive statement on market efficiency. However, they attribute inefficiency to the irrational behaviour of uninformed homebuyers who increase their demand for homes based upon expectations that past price increases will persist in the future. But higher housing demand may induce a rapid short-term run-up in house prices in excess of their fundamental values. However, over the longer term, a reversal in the upward trend in prices must occur as prices realign with their fundamental values.

However, one does not need to appeal to irrational investor behaviour to account for evidence of inefficiency in the housing market. For instance, the limited scope for short-selling in the real estate market may tighten liquidity conditions and induce a bubble in real estate prices. Although tradable housing derivatives have been available to the public since 2006 - for example, the Chicago Merchantile Exchange (CME) lists futures and options on the ten S&P/Case-Shiller Home Price Indices - the housing derivatives market is small relative to the underlying residential and commercial housing market. Therefore, even with the development of housing derivatives, the short-selling of home price indices is not likely to contain the potential run-up in real estate prices.
In addition, I also find that the variability in monthly (log) returns to housing over the period of study may be attributed to fluctuations over 16- to 32-month periods. This suggests that cyclical fluctuations over this scale may be an important component of home price changes. It may be worthwhile to explore this finding in future research.
CHAPTER 2
U.S. HOUSING MARKET AND INTERDEPENDENCIES

The nationwide collapse of home prices in the U.S. in 2006 was a costly event, which probably caught a number of people off-guard.\footnote{As of March 2008, the IMF, in collaboration with other financial institutions, estimates aggregate potential writedowns and losses arising from the housing crisis to be approximately $945 billion (IMF, 2008).} Indeed, the likelihood of such an outcome may have been underestimated in an environment in which Federal housing policies\footnote{See the Millennial Housing Commission (Congress, 2002).}, and financial innovations were not only geared towards lowering the barriers faced by low-income ethnic groups in pursuing homeownership; but also possibly making homeownership appear to be less risky than previously thought.

The risk factors which matter for the stability of the housing market are not understood as well as they should be. Therefore, this study makes an important step in this direction by investigating the interdependency structure of the U.S. housing market for single family homes. It follows Simpson (2008) and Liow (2008) in using cointegration tests to empirically investigate pair-wise interdependence in home prices. It also explores how the web of interrelationships which emerge from the securitization process may be linked to the widespread collapse of home prices in 2006.
Data and Preliminary Analysis

Monthly data on the original 10 S&P/Case-Shiller U.S. National Home Price Indices (i.e. Composite of 10) over the period January 1987 to April 2008 are obtained from Standard & Poor’s website on June 30, 2008. The Composite of 10 includes the following home price indices (HPIs): Boston (BOXR), Chicago (CHXR), Denver (DNXR), Las Vegas (LVXR), Los Angeles (LXXR), Miami (MIXR), New York (NYXR), San Diego (SDXR), San Francisco (SFXR) and Washington DC (WDXR). The purpose of the HPIs is to measure the average change in single-family home prices in a particular metropolitan area. Figure 1 presents plots of the S&P/Case-Shiller housing price indices over the period of study. For all home price indices shown, we see a run-up in home prices followed by a decline around 2006.

Empirical Results

The primary empirical exercise is to investigate the interdependency structure within U.S. housing market. One could simply undertake this task by performing a pairwise correlation analysis of the HPIs. However, this approach is problematic when dealing with (weakly) non-stationary series. \(^{24}\) This is so because a simple correlation analysis may yield spurious results for unit root

\(^{24}\) A time series \(\{y_t\}\) is (weakly) stationary if its mean and variance are constant over time, and the covariance between two adjacent time periods is time invariant. Thus, if a time series violates one or more of the conditions for stationarity, we conclude that it is non-stationary. Generally, the trend component is a major feature of a non-stationary process. This trend may be deterministic, and thereby predictable; or stochastic, and hence very hard to predict.
processes (Phillips, 1986; Granger and Newbond, 1977). We overcome this spurious regression problem by using a cointegration test instead. If two variables are cointegrated, then they tend to comove systematically over time. Alternatively, if two home price series are cointegrated, they are said to have a stable long-term equilibrium. Both interpretations suggest an interdependent relationship for the pair of home price series under consideration.

The implementation of a cointegration test requires two key steps. First, we need to determine whether the home price series are non-stationary. But this step has already been completed in Chapter 1. In particular, we found evidence that home prices contain a unit root. (These empirical results are summarized in Table 2). Thus, we simply invoke the non-stationary results obtained in Chapter 1. This leads us to the second step.

I employ the Phillips-Ouliaris cointegration test to determine whether non-stationary pairs of home price series are cointegrated. Specifically, I am testing for a unit root in the residuals of the cointegration regression expressed as a zero mean autoregressive model of order one, without a constant and a linear trend. I then compute the Phillips-Perron $\hat{Z}_\alpha$ test statistic. I continue to use a natural logarithmic transformation of S&P/Case-Shiller home prices in carrying out the cointegration tests under the null hypothesis of no cointegration.

Table 3 presents the Phillips-Perron $\hat{Z}_\alpha$ test statistic for the cointegration test under the null hypothesis of no cointegration. At the 5% level of significance,
the critical value is evaluated at $\hat{Z}_\alpha = -15.6377$. The null hypothesis is rejected if $\hat{Z}_\alpha$ is below the critical value at the 5% level of significance. For all the non-stationary pairs of S&P/Case-Shiller home prices (in natural logarithm) considered, I fail to reject the null hypothesis of no cointegration. That is, the pairs of home prices evaluated do not appear to move together systematically over time.

**Securitization-based Interdependencies**

Over the last decade, the securitization process is probably the most notable financial innovation in mortgage-finance (Rosen, 2007). This process allows banks to remove mortgages - which they originate - off their balance sheets by selling them to third parties (e.g., government-sponsored enterprises (GSEs) such as the Federal Home Loan Mortgage Corporation (Freddie Mac) and the Federal National Association (Fannie Mae) ). These third parties then place these mortgages in a trust (or special-purpose vehicle). The trust issues bonds (e.g., mortgage-backed securities (MBSs)) to investors backed by a pool of these mortgages. In 2006, MBSs accounted for over 50% of the total outstanding mortgage debt (Inside Mortgage Finance Publications Inc., 2007).

---

25 The critical values for the $\hat{Z}_\alpha$ statistic can be found in Tables Ia - Ic of Phillips and Ouliaris (1990). I consider the standard case.

26 Alternatively, we reject the null hypothesis if the absolute value of the $\hat{Z}_\alpha$ if is larger than the absolute value of the critical value at the 5% level of significance.
Much of the policy discussion on the potential problems associated with securitization process tends to be in line with the agency-related, and the non-transparency type of information problems discussed by Swan (2009). For instance, when a bank simply originates subprime mortgage loans with the goal of repackaging and selling them later on, a potential agency problem emerges. That is, the bank lacks the incentive to identify and highlight the risks associated with these mortgage loans because it is not the ultimate bearer of potential losses on future defaults. And by extension, imperfect information precludes investors from effectively assessing and monitoring the risks of highly complex MBSs. However, the web of interrelations which support the securitization process should also raise concerns about the stability of the housing market.

Allen and Gale (2000) suggest that a sudden contraction in the supply of credit or a spike in interest rate has adverse implications for the real estate market during a housing boom. That is, the demand for homes may decline as potential homebuyers begin to lose access to credit. Furthermore, home prices may collapse if a credit contraction persists. One of the major concerns which preoccupied policymakers during the housing crisis was the breakdown of financial intermediation which induced a credit-crunch. At one stage, even individuals with stellar credit histories were experiencing difficulties in obtaining mortgage-loans. Is there any connection between the credit-crunch which emerged during the crisis and possible inherent risks in the securitization process itself? Let us consider how this potential link may be established.
The ownership of residential houses is made possible when home-buyers successfully access mortgage-loans for a home which they desire. But ultimately, these mortgage-loans are included in the portfolios of banks, pension funds, government-sponsored enterprises and private investors. In exchange for a share of the cashflows which accrue from the mortgage-loan payments made by mortgage-holders, these investors bear a portion of the default risk. Importantly, the securitization process spreads risks only by increasing the number of counter-parties who are willing and able to accept a particular allocation of default risk in exchange for a share of the returns from the pool of residential mortgages. Thus, on the one hand, securitization reduces an individual investor's exposure to idiosyncratic risks due to small shocks; but on the other hand, it also introduces greater interdependencies within the economy through the extensive network of counter-party relationships which it creates in order to share such risks.

Thus, securitization provides numerous channels through which a large negative shock can be simultaneously transmitted to multiple interrelated parties. The unfolding of such an adverse event is potentially destabilizing since it may cause the financial constraints of the inter-connected parties to simultaneously bind. For example, if significant default losses on mortgage loans induce a major downgrade of the MBSs which a pension fund holds, the pension fund may be forced to liquidate its MBS portfolio in order to comply with its operating charter. But such fire sales could cause the price of MBSs to dramatically fall. And if significant losses are also transmitted to banks which hold MBSs, then their
capital base could be eroded by such losses, which in turn could trigger a violation of their mandatory capital requirements. In order to comply with regulations, these banks may be forced curtail the extension of credit. Thus, when various financial constraints simultaneously bind for a network of connected counter-parties, the accumulated effect may trigger a breakdown in financial intermediation.

**Summary**

This paper investigates the interdependency structure of the U.S. housing for single-family homes within the context of the housing market crisis in 2006. The empirical evidence from cointegration tests suggests that the interrelationships within the housing market itself are not the principal source behind the widespread collapse of home prices in 2006. However, one should consider the potential link between the securitization process and the housing crisis. While securitization allows individual investors to obtain a share of the returns from a pool of mortgages in exchange for a small share of risk relative to the risk faced by a sole (investor) mortgage-lender, it raises the risk of systemic failure due to the potentially large number counter-parties in the network which it creates.
CHAPTER 3
STOCK MARKET EFFICIENCY: A BEHAVIORAL PERSPECTIVE

Behind several nuanced characterizations of market efficiency in financial economics lies an important proposition: if the prices of financial assets are correct, then financial capital will be directed to the most productive investment activities. Therefore, it is not surprising that over the last three decades, market efficiency has been heavily scrutinized since its emergence as the dominant paradigm in the 1970s.

A key challenge to the market efficiency hypothesis comes from the noise trader literature (e.g., Shiller, 1984), which departs from the assumption that economic agents are rational. This paper evaluates the empirical implications of Sentana and Wadhwani’s (1992) feedback trading model. But the feedback trading model is a special case Shiller’s (1984) heterogeneous agent ‘noise trader’ model. It describes an environment in which market participants may buy, or sell, stocks when prices are rising, or falling, respectively. When this trend-chasing behaviour characterizes the trading strategy of market participants, their strategy is described as a positive feedback trading strategy.

---

As discussed in the introduction section of this thesis, there are three key characterizations of market efficiency: (a) the weak form of market efficiency which claims that the current prices fully reflect all the information implicit in the history of prices; (b) the semi-strong form of market efficiency which claims that prices reflect all publicly available relevant information; and (c) the strong form of market efficiency which asserts that private information is reflected in market prices.
This study is most similar to that of McMillan and Speight (2003) which uses high frequency stock index futures to study investor behavior. Specifically, they use FTSE-100 futures prices\(^{28}\) in their study, while I use S&P 500 futures prices. One of the main appeals of high frequency or tick-by-tick data is the potential insight that they provide into what type of information actually moves prices. At the same time, high frequency data is more susceptible to noise due to mistakes in the recording of transactions. Hence, one must exercise a great deal of care in cleaning the data (MacGregor, 1999).

**Data and Preliminary Analysis**

The data analyzed in this study are five-minute interval S&P 500 futures prices for the SPH04\(^{29}\) and SPM04\(^{30}\) futures contracts over the periods January 2, 2004 to March 11, 2004, and March 11, 2004 to June 10, 2004, respectively.\(^ {31}\) These S&P 500 future prices are used to compute (log) returns.\(^ {32}\) The return series consist of 3,498 and 4,910 observations for SPH04 and SPM04,

---

\(^{28}\) The FTSE-100 index is the principal index for the London Stock Exchange with 100 of UK’s most highly capitalised companies. FTSE-100 futures contracts are traded on the London International Financial Futures and Options Exchange (LIFFE).

\(^{29}\) The SPH04 futures contract specifies that this futures contract is based on the S&P 500 (SP), and final cash settlement is due in March (H) 2004.

\(^{30}\) The SPM04 futures contract specifies that this futures contract is based on the S&P 500 (SP), and final cash settlement is due in June (M) 2004.

\(^{31}\) These data are generated from the daily tick-by-tick data recorded by Tick Data Inc. S&P 500 futures contracts are traded on the Chicago Merchantile Exchange (CME). The size of one contract is $250 times the S&P 500 index; deliveries are made through cash settlements in the contract months of March (H), June (M), September (U) and December (Z). S&P 500 futures prices are quoted in terms of the S&P 500 index points, where one index point equals $250.

\(^{32}\) The period \( t \) log return is defined by \( r_t = \ln(p_t/p_{t-1}) \).
respectively, after excluding overnight trading, and restricting my attention to trading activities during the floor trading time of 9:30 a.m. to 4:15 p.m. In addition, I follow Andersen and Bollerslev’s (1997b) in adjusting for the effects of intraday seasonality. This adjustment further reduces the return series to 3,497 and 4,909 observations for SPH04 and SPM04, respectively.\textsuperscript{33}

As we will see later in this paper, a test of the feedback trading model appears to suggest that we need to actually use the S&P 500 prices rather than the S&P 500 futures prices as I have done. However, the spot and futures market move in tandem.\textsuperscript{34}

Cornell and French’s (1983) cost-of-carry futures pricing model provides a clear link between the spot and futures market. In general, there are two main differences between stocks and futures (Strong, 2004): (a) stocks pay dividends while futures do not; and (b) futures accrue interest on posted margins, while stocks do not. And we account for these differences in the following way.

Suppose the annual T-bill rate, and annualized dividend yield are given by \( R \) and \( D \), respectively; the relationship between the S&P 500 futures price \( F \) and the cash index price \( S_0 \) is described by

\textsuperscript{33} In order to obtain time series data on S&P500 index futures over long periods, one can join successive futures contracts that are approaching their delivery dates. However, this approach introduces an artificial discontinuity in the time series, which calls for some kind of adjustment before computing returns. I proceed, instead, the relatively short time series data for SPH04 and SPM04 to keep the analysis simple. Furthermore, there is no compelling reason why the current time series length of the futures contracts used may undermine the empirical results.

\textsuperscript{34} Particularly in the case of commodities, however, a phenomenon commonly referred to as backwardation sometimes occur. Backwardation is evidenced by relatively higher spot prices. If the prices for near maturities are relatively higher, then backwardation is also evident. However, in either case backwardation occurs if the convenience yield exceeds the storage cost.
\[ F_T = S_0e^{(R-D)T} \] (3.1)

where \( T \) denotes the number of days (in terms of years) until settlement.

Equilibrium in the spot market is achieved when the demand for shares equals the supply of shares; while equilibrium in the forward market is achieved when the number of short positions equals the number of long positions. In both cases, equilibrium is achieved when net demand is zero.

Table 4 provides a summary of descriptive statistics on the (log) returns of the S&P 500 futures contracts, SPH04 and SPM04, over the period January 2, 2004 to March 11, 2004, and March 11, 2004 to June 10, 2004, respectively. The stylized facts of non-normality and volatility clusters of financial time series appear to hold. The non-normality of the returns is apparent from the excess kurtosis and positive skewness reported in Table 4. QQ-plots in Figure 23 also shows fatter tails than would be suggested by a normal distribution. The evidence of volatility clusters is also apparent from the plot of the return series in Figure 24.
Feedback Trading Model

Sentana and Wadhwani's (1992) feedback trading model represents a special case of Shiller's (1984) heterogeneous agent 'noise trader' model. They describe an environment in which the interaction between 'smart' money and feedback traders induces serial correlation in returns. We distinguish between these group of traders based on the trading strategies employed which is inferred from their demand functions. ‘Smart’ money traders’ demand function for stocks is given by

\[ Q_t = \frac{[E_{t-1}(r_t) - \delta]}{\mu_t} \]  

(3.2)

where \( Q_t \) is the fraction of stocks being held, \( r_t \) is the ex-post return in period \( t \), \( E_{t-1} \) is the expectation operator based on the information available as of time \( t - 1 \), \( \delta \) is the return at which the demand for stocks is zero, \( \mu_t \) denotes the risk premium needed to entice smart money traders to hold all the shares. It is assumed that

\[ \mu_t = \mu(\sigma_t) \]  

(3.3)

with \( \mu'(\cdot) > 0 \), and \( \sigma_t \) represents the square root of the conditional variance of returns (i.e. standard deviation of returns) in period \( t \) given the available information in period \( t - 1 \). Equation (3.1) indicates that 'smart' money traders

---

35 This section closely follows the representation of the feedback-trading model by Bohl and Siklos (2004).

36 If smart money traders hold all the shares (i.e. \( Q_t = 1 \)) and we set \( \delta \) to the risk-free rate of return; then we obtain \( \mu_t = E_{t-1}(r_t) - \delta \), which in the language of portfolio theory may be interpreted as the excess of the stock’s expected rate of return over the risk-free rate.
will increase their demand for stocks as the expected excess return, $E_{t-1}(r_t) - \delta$, increases, or as the risk of holding stocks, $\sigma_t$, decreases.

Feedback traders’ demand function for stocks is given by

$$Y_t = \varphi r_{t-1}$$  \hspace{1cm} (3.4)

where $Y_t$ is the fraction of stocks held, and $\varphi$ a real constant. In equilibrium, all stocks are held; thus

$$Q_t + Y_t = 1$$  \hspace{1cm} (3.5)

We can now characterize the interaction between ‘smart’ money and feedback traders in the stock market by substituting (3.2) and (3.4) into (3.5) to get

$$E_{t-1}(r_t) = \delta + \mu(\sigma_t) - \varphi \mu(\sigma_t)r_{t-1}$$  \hspace{1cm} (3.6)

which becomes

$$r_t = \delta + \mu(\sigma_t) - \varphi \mu(\sigma_t)r_{t-1} + \epsilon_t$$  \hspace{1cm} (3.7)

where $\epsilon_t \equiv r_t - E_{t-1}(r_t)$ defines the error term. Equation (3.7) indicates that stock returns follow an autoregressive process of order one, with a time-varying risk premium. Whether there is negative or positive autocorrelation in stock returns depend on the trading strategies pursued by feedback traders. Positive feedback traders react to a price increase and a price decrease by buying and selling shares, respectively; that is, $\varphi > 0$ which induces negatively autocorrelated stock returns. On the other hand, negative feedback traders sell and buy stocks after stock prices rise and fall, respectively; that is, $\varphi < 0$ which induces positively autocorrelated stock returns.
Suppose the relationship between risk premium and risk is characterized by the function

\[ \mu(\sigma_t) = \eta_0 + \eta_1 (\sigma_t - \bar{\sigma}) \]  (3.8)

where \( \eta_0, \eta_1 \in \mathbb{R} \), and \( \bar{\sigma} = (1/T) \sum_{t=1}^{T} \sigma_t \). To ensure that the risk premium function describes a theoretically plausible relationship between risk premium and risk, some restrictions must be imposed on \( \eta_0 \) and \( \eta_1 \). There are two desirable properties which this risk premium function should exhibit:

1. \( \mu(0) = 0 \); and
2. \( \mu'(\sigma_t) > 0 \).

The second property says that risk premium should be increasing in risk. This implies that \( \mu'(\sigma_t) = \eta_1 > 0 \). The first property says that when there is no risk, no risk premium should be awarded. But \( \mu(0) = 0 \Leftrightarrow \eta_0 = \eta_1 \bar{\sigma} \), which leads to \( \eta_0 > 0 \) since \( \eta_1 > 0 \) and \( \bar{\sigma} > 0 \). Therefore, in order to generate a theoretically plausible relationship between risk premium and risk, \( \eta_0 \) and \( \eta_1 \) must be a positive constants.

Substitute equation (3.8) into (3.7) to get

\[ r_t = \delta + \eta_0 - \eta_1 \bar{\sigma} + \eta_1 \sigma_t - \phi \eta_0 r_{t-1} - \phi \eta_1 (\sigma_t - \bar{\sigma}) r_{t-1} + \epsilon_t \]  (3.9)

which may be expressed as

\[ r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \epsilon_t \]  (3.10)

where \( \omega \equiv \delta + \eta_0 - \eta_1 \sigma, \lambda \equiv \eta_1, \theta_1 \equiv -\phi \eta_0 \) and \( \theta_2 \equiv -\phi \eta_1 \). The positive risk premium property imposed on the risk premium function requires that \( \lambda \equiv \eta_1 > 0 \)
with \( \eta_0 > 0 \) as shown above. Thus, evidence for positive feedback trading is given by \( \theta_1, \theta_2 < 0 \) (i.e. a sufficient condition \( \varphi > 0 \)).

**Empirical Model Specifications**

The conditional mean equation

\[
 r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \epsilon_t
\]

\[
 v_t = \sigma_t \epsilon_t
\]

is jointly estimated with one of the following four specifications of conditional volatility:

\[
 \sigma_t = \alpha_0 + \alpha_1 |v_{t-1}| + \beta_1 \sigma_{t-1}
\]

\[
 \sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1}
\]

\[
 \sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \alpha_2 (|v_{t-2}| + \gamma_2 v_{t-2}) + \beta_1 \sigma_{t-1}
\]

\[
 \sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1} + \beta_2 \sigma_{t-2}
\]

By setting \( \gamma_i = 0 \ \forall i \), we obtain the generalized GARCH \( (p,q) \) model under the assumption that return shocks are symmetric (i.e. negative and positive return shocks of the same magnitude have the same impact on the conditional volatility).

---

37 These model specifications belong to the family of quadratic GARCH-in-mean models (Engle and Ng, 1993; Ding, Granger, and Engle, 1993; Sentana, 1995).
volatility). But if we impose no such restriction on $\gamma$, we then allow for asymmetric return shocks which may better characterize the conditional variance.

**Model Estimation Procedure**

A model specification is characterized by equations (3.11) and (3.12) plus one of the four specifications of conditional volatility in equations (3.13) to (3.16). I carry out the following four-step procedure to estimate a given model specification:

- **Step 1:** Jointly estimate the conditional mean equation without the interaction variable $(\sigma_t - \bar{\sigma})r_{t-1}$ (i.e. set $\theta_2 = 0$), along with equation (3.12) and the chosen conditional volatility equation;
- **Step 2:** Generate the estimated conditional volatilities from the jointly estimated model in step 1. For example, Figure 25 shows the estimated conditional volatilities for the stock index futures contracts SPH04 and SPM04 based on the joint estimation of equations (3.11), (3.12) and (3.14).

---

38 A sufficient condition for $\sigma_t^2 > 0$ is that $\alpha_i, \beta_j > 0 \forall i, j$. For many applications, the simple $GARCH(1,1)$ specification is an excellent choice based on its parsimony and its numerical stability property. For instance, higher order $GARCH(p,q)$ models may have multiple local maxima and minima which may lead to numerical instability in the estimation procedure. At the same time, Ma, Nelson and Startz (2006) note that when the $GARCH(1,1)$ model is weakly identified (i.e. $\alpha_1$ is close to zero, and $\beta_1$ is close to unity), inferences based on it may be spurious.

39 When a negative return shock of the same magnitude as a positive return shock has a bigger impact on the conditional variance, a leverage effect is said to exist (Black, 1976; Christie, 1982; Schwert, 1990, and Duffee, 1995). The leverage effect is present if that $\gamma < 0$. 
Step 3: Use the estimated conditional volatilities and lagged returns to compute the interaction variable \((\sigma_t - \bar{\sigma})r_{t-1}\).

Step 4: Return to step 1, but now jointly estimate the conditional mean equation with the interaction variable computed in step 3, along with equation (3.12) and the conditional volatility equation.

One may use the Akaike Information Criterion (AIC) to make model selection from a set of nested econometric models. The model with the smallest AIC is the preferred one.\(^{40}\)

**Empirical Results**

Tables 5 and 7 report parameter estimates for the four models used to test the feedback trading model for positive feedback trading. In all the cases considered, estimates of \(\alpha_0\) and \(\omega\) are supressed. For both S&P 500 futures contracts, estimates of the parameters \(\theta_1\) and \(\theta_2\) are negative and statistically significant; therefore, there appears to be evidence of positive feedback trading. McMillan and Speight (2003) also find evidence of positive feedback trading in their study.

\(^{40}\) Based on Akaike (1973), \(AIC = -lnL + p\), where \(lnL\) is the maximized log-likelihood for an estimated model with \(p\) parameters. AIC imposes a penalty on models that include more parameters without improving the statistical goodness of fit relative to the most parsimonious model; therefore, the preferred model has the smallest AIC.
The statistically significant negative estimates of $\gamma$ provides evidence of the stylized leverage effect (i.e. a negative return shock of the same magnitude as a positive return shock has a bigger impact on the conditional volatility). In addition, the positive risk premium property imposed on the risk premium function is supported by the statistically significant positive estimates of $\lambda$.

Table 6 and 8 report model diagnostics for the four model specifications used to test the feedback trading model for positive feedback trading. I use the Jacque-Bera (JB) statistic under the null hypothesis of normality to determine whether the residuals are normally distributed. For both futures contracts, the normality null hypothesis is rejected at the 5% level of significance for all four model specifications. I also check for evidence of any remaining nonlinear structure in the residuals using the Ljung-Box (Q) statistic under the null hypothesis of no nonlinear effects. For both futures contracts, the null hypothesis of no nonlinear effects is rejected at the 5% level of significance for all four models specified. This finding indicates that the class of models considered here do not fully capture nonlinear dependencies that may characterize stock return dynamics.

As suggested by Engle (1982), I also report the Lagrange multiplier test (LM) under the null hypothesis of no (remaining) ARCH effects (i.e. no serial correlation in the squared residuals). At the 5% level of significance, I reject the null of no ARCH effects for all four models specified in relation to futures contract SPH04. However, this null hypothesis is not rejected at the 5% level of significance in relation to Model I for futures contract SPM04. Finally, for futures
contracts SPH04 and SPM04, Model III and Model II, respectively, have the lowest AIC value. Thus, these model specifications appear to achieve a reasonable balance between parsimony and goodness of fit.

**Conclusion**

This paper uses a family of quadratic GARCH-in-mean models to determine whether trend-chasing behavior characterizes the trading strategy of market participants in the stock market. In particular, it uses high frequency data stock index futures prices to test Sentana and Wadhwani’s (1992) classic feedback trading model for evidence of positive feedback trading which describes such trend-chasing behavior. There appears to be evidence of positive feedback trading. This finding is robust to model specifications and the stock index futures contracts used.

At the same time, evidence of positive feedback trading still does not allow us to make a definitive statement on the efficiency of the stock market. For instance, even if low margin requirements attract positive feedback traders to the market, they need not exert a significant impact on prices. Specifically, their impact on market prices is likely to be small and transitory if they represent a relatively small constituent of the market, and if they are active for relatively short trading periods.
Table 1: Descriptive statistics of monthly (log) returns (%)

<table>
<thead>
<tr>
<th>HPIs</th>
<th>N</th>
<th>Mean</th>
<th>Std.</th>
<th>Ex. Ku.</th>
<th>Skewness</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston</td>
<td>255</td>
<td>0.32</td>
<td>0.83</td>
<td>-0.01</td>
<td>-0.14</td>
<td>0.31</td>
<td>-2.00</td>
<td>2.53</td>
<td>-1.12</td>
<td>1.58</td>
</tr>
<tr>
<td>Chicago</td>
<td>255</td>
<td>0.41</td>
<td>0.69</td>
<td>2.18</td>
<td>-0.21</td>
<td>0.40</td>
<td>-2.28</td>
<td>2.63</td>
<td>-0.75</td>
<td>1.48</td>
</tr>
<tr>
<td>Denver</td>
<td>255</td>
<td>0.37</td>
<td>0.63</td>
<td>1.44</td>
<td>-0.60</td>
<td>0.39</td>
<td>-2.02</td>
<td>2.10</td>
<td>-0.74</td>
<td>1.29</td>
</tr>
<tr>
<td>Las Vegas</td>
<td>255</td>
<td>0.36</td>
<td>1.23</td>
<td>7.79</td>
<td>0.11</td>
<td>0.35</td>
<td>-5.24</td>
<td>5.87</td>
<td>-1.34</td>
<td>1.80</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>255</td>
<td>0.48</td>
<td>1.22</td>
<td>1.79</td>
<td>-0.55</td>
<td>0.54</td>
<td>-4.36</td>
<td>3.79</td>
<td>-1.31</td>
<td>2.15</td>
</tr>
<tr>
<td>Miami</td>
<td>255</td>
<td>0.42</td>
<td>1.00</td>
<td>5.00</td>
<td>-1.25</td>
<td>0.40</td>
<td>-4.61</td>
<td>2.74</td>
<td>-1.30</td>
<td>1.99</td>
</tr>
<tr>
<td>New York</td>
<td>255</td>
<td>0.38</td>
<td>0.72</td>
<td>-0.45</td>
<td>0.07</td>
<td>0.36</td>
<td>-1.35</td>
<td>2.33</td>
<td>-0.73</td>
<td>1.48</td>
</tr>
<tr>
<td>San Diego</td>
<td>255</td>
<td>0.47</td>
<td>1.15</td>
<td>2.31</td>
<td>-0.20</td>
<td>0.49</td>
<td>-3.67</td>
<td>5.18</td>
<td>-1.25</td>
<td>2.15</td>
</tr>
<tr>
<td>San Francisco</td>
<td>255</td>
<td>0.49</td>
<td>1.20</td>
<td>2.66</td>
<td>-0.45</td>
<td>0.50</td>
<td>-4.36</td>
<td>4.06</td>
<td>-1.11</td>
<td>2.30</td>
</tr>
<tr>
<td>Wash. DC</td>
<td>255</td>
<td>0.45</td>
<td>0.90</td>
<td>1.02</td>
<td>-0.01</td>
<td>0.33</td>
<td>-2.51</td>
<td>3.15</td>
<td>-0.75</td>
<td>1.84</td>
</tr>
</tbody>
</table>

N: sample size; Mean: sample mean; Std.: standard deviation; Ex. Ku.: excess kurtosis; Low: 5th percentile; High: 95th percentile.
Table 2: Unit root tests

<table>
<thead>
<tr>
<th>City</th>
<th>$S_{T,1}^{ld}$</th>
<th>p-value_{ADF}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston</td>
<td>-3.88</td>
<td>0.61</td>
</tr>
<tr>
<td>Chicago</td>
<td>-16.30</td>
<td>0.78</td>
</tr>
<tr>
<td>Denver</td>
<td>-4.09</td>
<td>0.88</td>
</tr>
<tr>
<td>Las Vegas</td>
<td>-10.74</td>
<td>0.50</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>-7.71</td>
<td>0.25</td>
</tr>
<tr>
<td>Miami</td>
<td>-7.65</td>
<td>0.18</td>
</tr>
<tr>
<td>New York</td>
<td>-4.32</td>
<td>0.34</td>
</tr>
<tr>
<td>San Diego</td>
<td>-8.02</td>
<td>0.75</td>
</tr>
<tr>
<td>San Francisco</td>
<td>-10.40</td>
<td>0.78</td>
</tr>
<tr>
<td>Washington DC</td>
<td>-7.64</td>
<td>0.74</td>
</tr>
</tbody>
</table>

*Critical value at the 5% level of significance: -27.38

Table 2 reports the wavelet unit root test statistic, $S_{T,1}^{ld}$, of Gençay and Fan (forthcoming), and the p-values corresponding to the Augmented Dickey Fuller (ADF) unit root test based on Case-Shiller/S&P home prices, {$p_{i,t}$}, and the estimated AR(1) model: $\Delta p_{i,t} = \mu + \gamma p_{i,t-1} + \beta t + \epsilon_t$, where $\Delta p_{i,t} = p_{i,t} - p_{i,t-1} = \ln(p_{i,t}) - \ln(p_{i,t-1})$; $\epsilon_t \sim iid(0, \sigma^2_\epsilon)$, and $i \in \{Boston, Chicago, Denver, Las Vegas, Los Angeles, Miami, New York, San Diego, San Francisco, Washington DC\}$. The null hypothesis of a unit root is given by $\gamma = 0$. A lag length of 4 for the ADF test is determined by the software application (R-project).
Table 3: The $\hat{Z}_\alpha$ statistic values of the Phillips-Ouliaris cointegration test*

<table>
<thead>
<tr>
<th></th>
<th>Boston</th>
<th>Chicago</th>
<th>Denver</th>
<th>Las Vegas</th>
<th>Los Angeles</th>
<th>Miami</th>
<th>New York</th>
<th>San Diego</th>
<th>San Francisco</th>
<th>Washington DC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston</td>
<td>-5.12</td>
<td>-0.70</td>
<td>-1.71</td>
<td>-1.74</td>
<td>-1.65</td>
<td>-0.47</td>
<td>-2.72</td>
<td>-5.23</td>
<td>-2.11</td>
<td></td>
</tr>
<tr>
<td>Chicago</td>
<td>-5.12</td>
<td>-3.18</td>
<td>-5.36</td>
<td>-1.21</td>
<td>-4.71</td>
<td>-3.57</td>
<td>-2.28</td>
<td>-3.40</td>
<td>-1.36</td>
<td></td>
</tr>
<tr>
<td>Denver</td>
<td>-0.70</td>
<td>-3.18</td>
<td>-1.05</td>
<td>-0.55</td>
<td>-0.80</td>
<td>-0.64</td>
<td>-0.63</td>
<td>-1.33</td>
<td>-0.54</td>
<td></td>
</tr>
<tr>
<td>Las Vegas</td>
<td>-1.71</td>
<td>-5.36</td>
<td>-1.05</td>
<td>-2.00</td>
<td>-5.52</td>
<td>-1.52</td>
<td>-2.52</td>
<td>-3.80</td>
<td>-4.21</td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>-1.74</td>
<td>-1.21</td>
<td>-0.55</td>
<td>-2.00</td>
<td>-2.40</td>
<td>-3.57</td>
<td>-0.71</td>
<td>-1.47</td>
<td>-2.09</td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>-1.65</td>
<td>-4.71</td>
<td>-0.80</td>
<td>-5.52</td>
<td>-2.40</td>
<td>-2.55</td>
<td>-2.14</td>
<td>-2.75</td>
<td>-8.34</td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>-0.47</td>
<td>-3.57</td>
<td>-0.64</td>
<td>-1.52</td>
<td>-3.57</td>
<td>-2.55</td>
<td>-2.95</td>
<td>-4.47</td>
<td>-5.27</td>
<td></td>
</tr>
<tr>
<td>San Diego</td>
<td>-2.72</td>
<td>-2.28</td>
<td>-0.63</td>
<td>-2.52</td>
<td>-0.71</td>
<td>-2.14</td>
<td>-2.95</td>
<td>-3.89</td>
<td>-0.38</td>
<td></td>
</tr>
<tr>
<td>San Francisco</td>
<td>-5.23</td>
<td>-3.40</td>
<td>-1.33</td>
<td>-3.80</td>
<td>-1.47</td>
<td>-2.75</td>
<td>-4.47</td>
<td>-3.89</td>
<td>-1.12</td>
<td></td>
</tr>
<tr>
<td>Washington DC</td>
<td>-2.11</td>
<td>-1.36</td>
<td>-0.54</td>
<td>-4.21</td>
<td>-2.09</td>
<td>-8.34</td>
<td>-5.27</td>
<td>-0.38</td>
<td>-1.12</td>
<td></td>
</tr>
</tbody>
</table>

*Critical value: -15.6377

Table 3 reports the Phillips-Perron $\hat{Z}_\alpha$ test statistic under the null hypothesis on no cointegration. The critical value for $\hat{Z}_\alpha$ is determined at the 5% level of significance. The Phillips-Ouliaris cointegration test constitutes a test for a unit root in the residuals of the cointegration regression expressed as a zero mean autoregressive model of order one, without a constant and a linear trend.
Table 4: Descriptive statistics of five-minute interval (log) returns (%)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Excess Kurtosis</th>
<th>Skewness</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPH04</td>
<td>3497</td>
<td>-0.0002</td>
<td>13.32</td>
<td>0.20</td>
<td>0.0</td>
<td>-0.70</td>
<td>0.82</td>
<td>-0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>SPM04</td>
<td>4909</td>
<td>0.0005</td>
<td>28.24</td>
<td>0.37</td>
<td>0.0</td>
<td>-1.26</td>
<td>1.20</td>
<td>-0.13</td>
<td>0.13</td>
</tr>
</tbody>
</table>

N: Sample size; Mean: Sample mean; Low: 5th percentile; High: 95th percentile.
Table 5: Parameter estimates using SPH04 data

<table>
<thead>
<tr>
<th>Models</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\lambda$</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model I</td>
<td>0.56</td>
<td>0.03</td>
<td>0.08</td>
<td></td>
<td>-0.15</td>
<td>-0.42</td>
<td>38.21</td>
<td>2.52</td>
<td>3.76</td>
</tr>
<tr>
<td>Model II</td>
<td>0.58</td>
<td>-0.18</td>
<td>0.03</td>
<td>0.0</td>
<td>3.17</td>
<td>0.20</td>
<td>-0.19</td>
<td>-0.43</td>
<td>37.74</td>
</tr>
<tr>
<td>Model III</td>
<td>0.52</td>
<td>-0.48</td>
<td>-0.12</td>
<td>0.07</td>
<td>0.94</td>
<td>0.06</td>
<td>-0.14</td>
<td>-0.43</td>
<td>38.10</td>
</tr>
<tr>
<td>Model IV</td>
<td>0.54</td>
<td>-0.24</td>
<td>0.02</td>
<td>0.21</td>
<td>0.31</td>
<td>0.17</td>
<td>-0.43</td>
<td></td>
<td>39.82</td>
</tr>
</tbody>
</table>

Model I: $r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t$, $\nu_t = \sigma_t \epsilon_t$ where $\epsilon_t \sim N(0,1)$

$\sigma_t = \alpha_0 + \alpha_1 |v_{t-1}| + \beta_1 \sigma_{t-1}$

Model II: $r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t$, $\nu_t = \sigma_t \epsilon_t$ where $\epsilon_t \sim N(0,1)$

$\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1}$

Model III: $r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t$, $\nu_t = \sigma_t \epsilon_t$ where $\epsilon_t \sim N(0,1)$

$\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \alpha_2 (|v_{t-2}| + \gamma_2 v_{t-2}) + \beta_1 \sigma_{t-1}$

Model IV: $r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t$, $\nu_t = \sigma_t \epsilon_t$ where $\epsilon_t \sim N(0,1)$

$\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1} + \beta_2 \sigma_{t-2}$

Estimates of $\alpha_0$ and $\omega$ are suppressed;
The $t$-statistics are in italics.
Table 6: Model diagnostics using SPH04 data

<table>
<thead>
<tr>
<th>Models</th>
<th>JB</th>
<th>Q (12)</th>
<th>LM (12)</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model I</td>
<td>15086</td>
<td>437</td>
<td>29</td>
<td>-9976</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>Model II</td>
<td>13728</td>
<td>463</td>
<td>26</td>
<td>-10002</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.05)</td>
<td></td>
</tr>
<tr>
<td>Model III</td>
<td>9916</td>
<td>477</td>
<td>31</td>
<td>-10130</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>Model IV</td>
<td>12739</td>
<td>466</td>
<td>36</td>
<td>-9982</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.02)</td>
<td></td>
</tr>
</tbody>
</table>

Model I: \( r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + v_t, \quad v_t = \sigma_t \epsilon_t \) where \( \epsilon_t \sim N(0,1) \)
\[
\sigma_t = \alpha_0 + \alpha_1 |v_{t-1}| + \beta_1 \sigma_{t-1}
\]

Model II: \( r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + v_t, \quad v_t = \sigma_t \epsilon_t \) where \( \epsilon_t \sim N(0,1) \)
\[
\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1}
\]

Model III: \( r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + v_t, \quad v_t = \sigma_t \epsilon_t \) where \( \epsilon_t \sim N(0,1) \)
\[
\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \alpha_2 (|v_{t-2}| + \gamma_2 v_{t-2}) + \beta_1 \sigma_{t-1}
\]

Model IV: \( r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + v_t, \quad v_t = \sigma_t \epsilon_t \) where \( \epsilon_t \sim N(0,1) \)
\[
\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1} + \beta_2 \sigma_{t-2}
\]

JB: Jarque-Bera statistic - under the null hypothesis of normality;
Q (12): Ljung-Box statistic - under the null hypothesis of no nonlinear effects;
LM (12): Langrange Multiplier statistic - under the null hypothesis of no ARCH effects in the squared residuals.
p-values are in parentheses.
Table 7: Parameter estimates using SPM04 data

<table>
<thead>
<tr>
<th>Models</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\lambda$</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model I</td>
<td>0.19</td>
<td></td>
<td>0.45</td>
<td></td>
<td>2.03</td>
<td></td>
<td>-0.35</td>
<td>-0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17.97</td>
<td>36.10</td>
<td>11.87</td>
</tr>
<tr>
<td>Model II</td>
<td>0.07</td>
<td>-1.00</td>
<td>0.33</td>
<td></td>
<td>3.33</td>
<td></td>
<td>-0.35</td>
<td>-0.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.90</td>
<td>-8.76</td>
<td>14.91</td>
</tr>
<tr>
<td>Model III</td>
<td>0.03</td>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.82</td>
<td></td>
<td>2.65</td>
<td>-0.24</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.34</td>
<td>0.57</td>
<td>0.68</td>
</tr>
<tr>
<td>Model IV</td>
<td>0.02</td>
<td>-1.00</td>
<td>0.18</td>
<td>0.35</td>
<td>9.16</td>
<td></td>
<td>-0.36</td>
<td>-0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.39</td>
<td>-4.33</td>
<td>5.76</td>
</tr>
</tbody>
</table>

Model I: $r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t$, $\nu_t = \sigma_t \epsilon_t$ where $\epsilon_t \sim N(0,1)$

$\sigma_t = \alpha_0 + \alpha_1 |v_{t-1}| + \beta_1 \sigma_{t-1}$

Model II: $r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t$, $\nu_t = \sigma_t \epsilon_t$ where $\epsilon_t \sim N(0,1)$

$\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1}$

Model III: $r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t$, $\nu_t = \sigma_t \epsilon_t$ where $\epsilon_t \sim N(0,1)$

$\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \alpha_2 (|v_{t-2}| + \gamma_2 v_{t-2}) + \beta_1 \sigma_{t-1}$

Model IV: $r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t$, $\nu_t = \sigma_t \epsilon_t$ where $\epsilon_t \sim N(0,1)$

$\sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1} + \beta_2 \sigma_{t-2}$

Estimates of $\alpha_0$ and $\omega$ are suppressed;

The $t$-statistics are in italics.
### Table 8: Model diagnostics using SPM04 data

<table>
<thead>
<tr>
<th>Models</th>
<th>JB</th>
<th>Q (12)</th>
<th>LM (12)</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model I</td>
<td>15567</td>
<td>168</td>
<td>21</td>
<td>6182</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td>Model II</td>
<td>15113</td>
<td>100</td>
<td>39</td>
<td>6103</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>Model III</td>
<td>11711</td>
<td>297</td>
<td>44</td>
<td>6383</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>Model IV</td>
<td>12569</td>
<td>205</td>
<td>65</td>
<td>6257</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
</tr>
</tbody>
</table>

Model I: \( r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t, \) \( v_t = \sigma_t \epsilon_t \) where \( \epsilon_t \sim N(0,1) \)

\[ \sigma_t = \alpha_0 + \alpha_1 |v_{t-1}| + \beta_1 \sigma_{t-1} \]

Model II: \( r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t, \) \( v_t = \sigma_t \epsilon_t \) where \( \epsilon_t \sim N(0,1) \)

\[ \sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1} \]

Model III: \( r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t, \) \( v_t = \sigma_t \epsilon_t \) where \( \epsilon_t \sim N(0,1) \)

\[ \sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \alpha_2 (|v_{t-2}| + \gamma_2 v_{t-2}) + \beta_1 \sigma_{t-1} \]

Model IV: \( r_t = \omega + \lambda \sigma_t + \theta_1 r_{t-1} + \theta_2 (\sigma_t - \bar{\sigma}) r_{t-1} + \nu_t, \) \( v_t = \sigma_t \epsilon_t \) where \( \epsilon_t \sim N(0,1) \)

\[ \sigma_t = \alpha_0 + \alpha_1 (|v_{t-1}| + \gamma_1 v_{t-1}) + \beta_1 \sigma_{t-1} + \beta_2 \sigma_{t-2} \]

JB: Jarque-Bera statistic - under the null hypothesis of normality;
Q (12): Ljung-Box statistic - under the null hypothesis of no nonlinear effects;
LM (12): Langrange Multiplier statistic - under the null hypothesis of no ARCH effects in the squared residuals.
p-values are in parentheses.
Figure 1: S&P/Case Shiller housing price indices
Figure 2: Monthly (log) returns to housing
Figure 3: Histogram and Normal QQ-Plot of returns: Boston
Figure 4: Histogram and Normal QQ-Plot of returns: Chicago
Figure 5: Histogram and Normal QQ-Plot of returns: Denver
Figure 6: Histogram and Normal QQ-Plot of returns: Las Vegas
Figure 7: Histogram and Normal QQ-Plot of returns: Los Angeles
Figure 8: Histogram and Normal QQ-Plot of returns: Miami
Figure 9: Histogram and Normal QQ-Plot of returns: New York
Figure 10: Histogram and Normal QQ-Plot of returns: San Diego
Figure 11: Histogram and Normal QQ-Plot of returns: San Francisco
Figure 12: Histogram and Normal QQ-Plot of returns: Washington, D.C.
Figure 13: Boston – Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 14: Chicago – Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 15: Denver – Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 16: Las Vegas – Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 17: Los Angeles – Haar MODWT coefficient vectors.
The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 18: Miami - Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 19: New York – Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 20: San Diego – Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 21: San Francisco – Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 22: Washington DC – Haar MODWT coefficient vectors.

The monthly (log) returns to housing is the first graph at the top, with the wavelet and scaling coefficients below.
Figure 23: Standard Normal QQ-plots for SPH04 and SPM04.
Figure 24: Log returns for SPH04 and SPM04.
Figure 25: Estimated Conditional Volatilities for SPH04 and SPM04.
APPENDIX C: PROPERTIES OF STATIONARY AND NON-STATIONARY PROCESSES

Stationary Process

A time series \( \{y_t\} \) is (weakly) stationary if its mean and variance are constant over time, and the covariance between two adjacent time periods is time invariant. That is, a stationary time series \( \{y_t\} \) satisfies the following three conditions:

\[
E(y_t) = E(y_{t-1}) = \cdots = E(y_{t-k}) = \mu \\
var(y_t) = var(y_{t-1}) = \cdots = var(y_{t-k}) = \sigma^2 \\
cov(y_t, y_{t-k}) = cov(y_{t-j}, y_{t-j-k}) = \gamma_k
\]

Suppose \( \{y_t\} \) is governed by the following autoregressive process of order one (i.e. an \( AR(1) \) process):

\[
y_t = \alpha + \rho y_{t-1} + \epsilon_t
\]

where \( \epsilon_t \sim iid(0, \sigma^2) \). By recognizing that this \( AR(1) \) process is also a (first-order) stochastic difference equation, we can obtain a general solution by iteration; that is

\[
y_1 = \alpha + \rho y_0 + \epsilon_1 \\
y_2 = \alpha + \rho \alpha + \rho^2 y_0 + \rho \epsilon_1 + \epsilon_2 \\
y_3 = \alpha + \rho \alpha + \rho^2 \alpha + \rho^3 y_0 + \rho^2 \epsilon_1 + \rho \epsilon_2 + \epsilon_3
\]

\[41\] For a more detailed exposition, see Hamilton (1994) and Box, Jenkins and Reinsel (2008).
continuing this process leads to the following general solution:

\[ y_t = \alpha + \sum_{i=0}^{t-1} \rho^i y_0 + \sum_{i=0}^{t-1} \rho^i \epsilon_{t-i} \]

Suppose \(|\rho| < 1\), then as \(t \to \infty\), \(\rho^t \to 0\) so that

\[ y_t = \alpha + \sum_{i=0}^{t-1} \rho^i + \sum_{i=0}^{t-1} \rho^i \epsilon_{t-i} \]

The mean of \(\{y_t\}\) is given by

\[ y_t = \alpha + \sum_{i=0}^{t-1} \rho^i = \frac{\alpha}{1 - \rho} \]

which is constant. The covariance at lag \(k \geq 0\) is given by

\[
\text{cov}(y_t, y_{t-k}) = E[\{(y_t - E(y_t))(y_{t-k} - E(y_{t-k}))\}]
\]

\[ = E\left[\left(\sum_{i=0}^{\infty} \rho^i \epsilon_{t-i}\right)\left(\sum_{j=0}^{\infty} \rho^{i+j} \epsilon_{t-(j+k)}\right)\right] \]

\[ = E[\rho^k \sum_{i=0}^{\infty} (\rho^2)^i \epsilon_{t-i}^2] \]

\[ = \frac{\rho^k \sigma^2}{1 - \rho^2} \]

which is time invariant. The variance of \(\{y_t\}\) is obtained by setting \(k = 0\); that is

\[ \text{var}(y_t) = \frac{\sigma^2}{1 - \rho^2} \]
which is also constant. Thus, we can describe an AR(1) process as stationary if our sample size is sufficiently large and the restriction that $|\rho| < 1$ is being satisfied.

**Non-stationary Process**

If a time series violates one or more of the conditions for stationarity, we conclude that it is non-stationary. Generally, the trend component is a major feature of a non-stationary process. This trend may be deterministic, and thereby predictable; or stochastic, and hence very hard to predict.

A (pure) unit root process is a typical example of a non-stationary process.\(^{42}\) It can be characterized as a special case of the AR(1) model described above for $\rho = 1$ and $\alpha = 0$ (i.e. no drift); that is

$$y_t = y_{t-1} + \epsilon_t$$

By the method of iteration, we can show that

$$y_t = y_0 + \sum_{i=0}^{t} \epsilon_i$$

The mean of this unit root process is given by

$$E(y_t) = y_0$$

which is a constant; therefore, the first condition of stationarity is satisfied. For $k \geq 0$, the covariance at lag $k$ is given by

\(^{42}\) A unit root process with a stochastic trend is often referred to as a random walk model in financial economics.
\[
\text{cov}(y_t, y_{t-k}) = E[(y_t - E(y_t)) [y_{t-k} - E(y_{t-k})]]
\]
\[
= E[(\sum_{i=0}^{t} \epsilon_i)(\sum_{j=0}^{t-k} \epsilon_j)]
\]
\[
= E[\sum_{i=0}^{t} \epsilon_i^2 + \sum_i \sum_{j \neq i} \epsilon_i \epsilon_j]
\]
\[
\text{cov}(y_t, y_{t-k}) = (t - k)\sigma^2
\]

which is not time invariant. By setting \( k = 0 \), we also see that the variance is given by

\[
\text{var}(y_t) = t\sigma^2
\]

and also depends on time. Thus, a unit root process in this case is not stationary because it violates the second and third conditions for stationarity.

We can also verify that for \( \alpha \neq 0 \) (i.e. with drift), the unit root process is also non-stationary. In this case, it now violates all three conditions for stationarity.\(^{43}\)

Consider the AR(1) model

\[
y_t = \alpha + y_{t-1} + \epsilon_t
\]

By the method of iteration, we can show that

\[
y_t = y_0 + \alpha t + \sum_{i=1}^{t} \epsilon_i
\]

The time series now contains a deterministic trend component \( y_0 + \alpha t \);\(^{44}\) and a stochastic trend component, \( \sum_{i=1}^{t} \epsilon_i \). The mean of \( \{y_t\} \) is now given by

\(^{43}\) A unit root process with a nonzero constant \( \alpha \) is commonly referred to a random walk model with drift. The presence of drift allows for positive or negative trends in the time series which ex ante cannot be predicted.
\[ E(y_t) = y_0 + at \]

which now depends on time. The addition of the constant \( \alpha \) does not change the covariance and variance of \( \{y_t\} \); therefore, we also have

\[ \text{cov}(y_t, y_{t-k}) = (t - k)\sigma^2 \]

and

\[ \text{var}(y_t) = t\sigma^2 \]

which as noted above are time-dependent.

\[ \text{For } \alpha > 0 \text{ and } \alpha < 0, \text{ we have a positive and negative (deterministic) trend, respectively.} \]
APPENDIX D: BASIS FUNCTIONS

The basic idea that underscores basis functions is actually quite similar to that of basis vectors (Strang, 1992). Basis vectors belong to the set of linearly independent or orthogonal vectors (i.e. the basis) that generates all vectors in a (real-valued) vector space (Hoy et. al., 2001). For example, every two-dimensional vector \((u, v) \in \mathbb{R}^2\) is a linear combination of the basis vectors \(u = [0\ 1]'\), and \(v = [1\ 0]'\) with inner product \(u \cdot v = \sum_{j=1}^{2} u_j v_j = 0\) (i.e. \(u\) and \(v\) are orthogonal). Examples of basis functions include the sine and cosine waves, which may be used in Fourier synthesis to produce an alternative representation of a (weakly) stationary time series \(x(t)\). The orthogonality condition is satisfied by choosing the combination of sine and cosine terms whose inner product sums to zero.

Wavelets are more accurately described as scale-varying basis functions (Graps, 1995). By increasing the scale factor \(\lambda\), the time domain becomes more heavily partitioned. Suppose the original time series \(x(t)\) has time support \([0, 1]\). We can use wavelets to decompose \(x(t)\) into two step functions on time supports \([0, 1/2]\) and \([1/2, 1]\), respectively. But by further reducing \(\lambda\), we can decompose \(x(t)\) into four step functions over time supports \([0, 1/4]\), \([1/4, 1/2]\), \([1/2, 3/4]\) and \([3/4, 1]\), respectively. Thus, wavelets or basis functions are generated for different time-resolutions.


