Modelling the Dublin Housing Market

A Time Series Analysis

Modelling the Dublin Housing Market

A Time Series Analysis

by

Loretta O'Sullivan

Submitted to the Department of Economics
in fulfilment of the requirements for the degree of

Doctor of Philosophy

at the

UNIVERSITY OF DUBLIN

May 2007

Declaration

I declare that this thesis, submitted to Trinity College Dublin for the degree of Doctor of Philosophy (Ph.D.), has not been submitted as an exercise for a degree at any other university. All research contained herein is entirely my own. I authorise the Library of Trinity College Dublin to lend or copy this thesis upon request.

Date: 8/5/07

Signature: Loretta O'Sullivan

Summary

Applying a range of time series econometric techniques, this thesis comprehensively models the Dublin market for new, private sector housing over the period 1980-2000.

Chapter 1 details the rationale for the study and outlines the structure of the remaining chapters. An overview of the theoretical functioning of the market for housing, along with the baseline model that underlies the empirical sections of the thesis, is provided in Chapter 2.

To facilitate the identification of variables that could potentially influence demand, supply and price inflation in the Dublin marketplace, Chapter 3 presents the empirical findings of selected Irish and international studies that adopt a time series, macroeconomic approach to modelling housing. Overall, the literature appears to be in agreement as to the key factors at play in the housing market. These include price, income, interest rates, building costs, the existing housing stock, demographics, land and government policies.

Chapter 4 provides a detailed description of the data series that feed into the empirical sections, while Chapter 5 presents a synopsis of the modelling techniques applied in the study.

Chapter 6 undertakes an empirical analysis of the Dublin market for new, private sector housing with the aim of uncovering the nature of the long run relations at play in this marketplace. Econometric modelling adopts the Johansen cointegration approach and in line with expectations, identifies the presence of two long run relationships. The demand relation,

which is normalised on the housing stock variable, shows that the stock of housing was below its desired level – that is the level consistent with the prevailing level of house prices, income, mortgage interest rates and demographics – over the time periods 1986-1989, 1993-1997 and 1999. The supply vector is normalised on house prices and suggests that from the end of 1987 to 1989, and from 1994 onwards, prices were relatively high given the prevailing level of mortgage interest rates, building costs and land. However, an analysis of the adjustment properties of the above relationships reveals a failure on the part of both the stock and price of housing to adjust to steady state deviations and restore the market to equilibrium.

Accordingly, Chapter 7 considers adjustment behaviour in the Dublin market for new private sector housing in greater detail. In doing so, short run house price determination is modelled within a linear and non-linear error correction framework. The resultant findings indicate that a range of explanatory variables including income, demographics, the housing stock, government interventionist policies and mortgage interest rates are significant determinants of price inflation. Moreover, house prices are found not to adjust in either of the extreme regimes of the smooth transition regression model. As such, it may be concluded that the long run demand and supply relations identified in Chapter 6 do not define equilibrium for the Dublin market.

Finally, as the existence of a bubble could potentially account for the part of house price determination that is left unexplained by the models estimated in Chapter 7, Chapter 8 examines the Dublin new private sector housing market for evidence of speculative behaviour. The regime-switching approach applied suggests the presence of a fad. Fads give rise to a situation whereby prices deviate temporarily from, but will eventually revert to, underlying fundamentals. On this basis, a 'soft landing' may be anticipated for the future path of the Dublin housing market rather than the collapse type scenario associated with a bubble.

Chapter 9 concludes.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Professor P.J. Drudy for his kindness, support and supervision over the past years.

The financial support of a Government of Ireland Research Scholarship in the Humanities and Social Sciences, an Urban Institute Ireland Scholarship, and a Trinity College Postgraduate Award is gratefully acknowledged.

During my time in the Department of Economics, I have benefited from the help and encouragement of many staff members and postgraduate students, to whom I extend my gratitude. In particular, I would like to thank Professor Alan Matthews, Mr. Michael Harrison and Dr. Jacco Thijssen for their generous advice; and Colette Ding, Orla Doyle and Liam Delaney for their ongoing friendship. I would also like to thank the examiners of this thesis, Dr. Andrew Somerville and Professor Heino Bohn Nielsen.

For further stimulating my interest in economics and econometrics, I thank the lecturers and participants at the LSE, University of Oxford and University of Copenhagen Summer and Easter Schools in Econometrics. The support I received whilst a visiting Marie Curie student at the University of Cambridge is much appreciated.

Above all, I thank my family for the encouragement and unerring love that they have given me throughout this endeavour.

Table of Contents

List of Tables and Figures

Tables

Figures

List of Variables

nph	=	real new house prices
inc	=	real personal disposable income per capita
MR	=	nominal mortgage interest rates
bc	=	real building costs
phs	=	the stock of private housing
hf	=	household formation
ls	=	the stock of zoned housing land
CGT	=	capital gains tax rates
Dumrpt	=	a dummy = 1 for the years in which property taxes applied and 0 otherwise
Dumib	=	a dummy = 1 for the years in which interest on borrowings was tax deductible and 0 otherwise
Dums2327	=	a dummy = 1 for the years in which 'Section 23/27' relief covered all rental income and 0 for the years in which it either did not apply or related solely to rental income from the residential property in question
CPI	=	consumer price index
Dum_1bc	=	a transitory shock dummy = 1 for 2000:4 and 0 otherwise
Dum_1ls	=	a permanent shock dummy = 1 for 1982:2 and 0 otherwise
Dum_2ls	=	a permanent shock dummy = 1 for 1993:2 and 0 otherwise
Dum_3ls	=	a permanent shock dummy = 1 for 1998:3 and 0 otherwise
Dum_4ls	=	a permanent shock dummy = 1 for 1981:3 and 0 otherwise

Dum_2bc = a transitory shock dummy = 1 for 1981:2, 0 for 1981:3, -1 for 1981:4 and 0 otherwise

Dummr = a transitory shock dummy = 1 for 1992:4, -1 for 1993:1 and 0 otherwise

Dum_3bc = a permanent shock dummy = 1 for 1985:2 and 0 otherwise

ecm^D = an error correction term capturing deviations from the long run demand relation

ecm^S = an error correction term capturing deviations from the long run supply relation

Dq = seasonal dummies

R = the return from investing in housing

b = the non-fundamental / bubble component of real new house prices

Chapter 1:
Introduction

1.1. Study Rationale

Developments in national and regional housing markets have a number of implications for the wider economy. From a macroeconomic perspective, investment in housing directly contributes to economic growth, while the wealth effects associated with rising prices may impact on consumer behaviour. At a micro level, the purchase of a house is typically the most important transaction made by a household. Rising prices can therefore lead to affordability concerns which limit access to the marketplace or alternatively, distort household location choice with knock-on effects for physical and social infrastructure.

In an Irish context, the Dublin market has periodically experienced incidents of excess demand, supply constraints and rising house prices; the latest of which has persisted since the mid-nineties. Given the dominance of this marketplace, such incidents can be expected to have had significant economy wide effects. However, before considering the broader impact of housing market disequilibria, a clear understanding of the empirical conditions that characterise the Dublin marketplace is first needed.

Notwithstanding the importance of such knowledge, little research into the factors that lie behind developments in the Dublin housing market has been carried out to date. In order to address this deficit, further empirical modelling is necessary. Accordingly, the aim of this thesis is to comprehensively model the Dublin market for new, private sector housing over the period 1980-2000. Specifically, the empirical work presented in later chapters seeks to identify the nature of the long run demand and supply relations at play in the Dublin marketplace, model short run house price determination and test for the possibility of speculative behaviour. In doing so, a range of time series econometric techniques are adopted.

1.2. Study Structure

An overview of the structure of the study, along with a brief outline of the content of each chapter and its contribution to the literature where relevant, is set out below.

Chapter 2: The Theoretical Framework

As any empirical analysis must be conducted within the framework of economic theory, Chapter 2 begins by considering the theoretical functioning of the housing market. In particular, the special characteristics associated with housing are discussed, as is the manner in which these characteristics are modified so as to enable standard time series modelling of the housing market. The baseline model that underlies the empirical sections of the study is then presented, along with the key theoretical relationships that apply in a housing context.

Chapter 3: Literature Review

Chapter 3 draws together the empirical findings of various Irish and international studies that adopt a time series, macroeconomic approach to modelling housing. The intention in detailing these findings is to facilitate the identification of appropriate factors for inclusion when modelling the Dublin housing market. To this end, attention is paid to the nature of the explanatory variables put forward by the selected studies and to the validity of the modelling techniques adopted.

Chapter 4: Data Description

Chapter 4 provides a detailed description of the data series that feed into the empirical sections of the study. Given the limitations of the available information, it proved necessary to refine much of the data and to generate a measure of the stock of zoned housing land. The methodology and refinement techniques applied when doing so are set out in this chapter. An overview of the data sources and time series properties of the constructed variables is also presented.

Chapter 5: Methodology

Chapter 5 presents a technical synopsis of the modelling techniques adopted in the course of the study's empirical analyses of the Dublin market; namely, the Johansen cointegration methodology, smooth transition regression modelling and regime-switching tests for stochastic bubbles. These techniques draw on the time series properties of the data.

Chapter 6: A Long Run Analysis of the Dublin Housing Market

Chapter 6 undertakes an empirical analysis of the Dublin market for new, private sector housing within a long run framework. Econometric modelling adopts the Johansen cointegration approach and applies it to the time-series dataset described in Chapter 4. A priori, the presence of two equilibrium relations is postulated, one for the demand-side of the market and one representing supply.

In terms of its contribution to the literature, this chapter represents the first attempt to include a land availability variable in a long run model of the housing market. The potentially important role of land is often referred to in international studies, but due to a lack of data on costs and availability, is typically not modelled. While recent work on the national market has incorporated somewhat crude measures of land prices, no effort has been made to consider land in empirical studies of the Dublin marketplace. Moreover, the shortage of zoned and serviced housing land in the Dublin area implies that availability rather than price is the appropriate land measure in this context.

Chapter 7: An Analysis of Adjustment in the Dublin Housing Market

Building on the preceding analysis, Chapter 7 seeks to examine adjustment behaviour in the Dublin market for new private sector housing. In particular, it seems appropriate to examine the response of prices to deviations from the long run relations. As such, this chapter models short run house price determination within an adjustment type framework. In doing so, error correction models of both a linear (ordinary least squares estimation) and non-linear (smooth transition regression modelling) nature are developed and applied to the dataset detailed in Chapter 4.

The estimation of error correction models goes considerably beyond the scope of current empirical work on the Dublin housing market, and in the case of the non-linear model, also beyond that at national level. Indeed, the study's use of the smooth transition regression modelling approach is only the second time in which this technique has been applied in a housing framework. Furthermore, by including a range of variables designed to capture the effectiveness of government interventionist policies, this chapter allows for a more complete modelling of house price determination than has been undertaken to date.

Chapter 8: Testing for the Presence of a Bubble in the Dublin Housing Market

As the existence of a bubble could potentially account for the part of house price determination that is left unexplained by the models estimated in Chapter 7, Chapter 8 examines the Dublin new private sector housing market for evidence of speculative behaviour. A regime-switching approach which enables testing of the hypothesis that house prices are being driven solely by fundamental market factors, against the alternative of a bubble, is adopted for this purpose.

The contribution of Chapter 8 to the wider literature centres on the attention paid to modelling the fundamental component of house prices. Whereas existing empirical research has modelled underlying fundamentals in the Dublin marketplace on a partial basis, this study includes both demand and supply-side variables. Most notably, the fundamentals model presented here is the first to incorporate a measure of land availability.

Chapter 9: Conclusion

Chapter 9 concludes with a summary of the main findings arising from the study's comprehensive modelling of the Dublin housing market.

Overall, this thesis should add to the current body of literature by facilitating an improved understanding of the empirical conditions that characterise the Dublin market for new, private sector housing.

Chapter 2:

The Theoretical Framework

2.1. Introduction

From a theoretical perspective, housing can be viewed as a commodity that is, in a similar manner to other goods and services, subject to the rigours of the market process. However, in contrast to most commodities, it is a good that embodies a number of distinct attributes, the interaction of which serves to complicate economic analyses. While the heterogeneity and durability properties of housing are foremost amongst its distinguishing characteristics, features such as spatial fixity and government involvement also identify it as a unique commodity (Smith, Rosen and Fallis, 1988). Given that an appreciation of both the theoretical and empirical functioning of the housing market requires an understanding of these attributes, the main aim of this chapter is to discuss the special characteristics of housing. Building on this discussion, the modelling implications of housing market peculiarities are then considered, particularly the need for simplifying assumptions. The latter arise from the difficulties associated with any attempt to combine the above diverse range of concepts into a single model. Thus, a further objective of the chapter is to describe the simplified housing model that is typically applied in time series studies of the type undertaken in this thesis.

The remainder of the chapter is organised as follows: Section 2.2 begins by defining the concepts of heterogeneity, durability, spatial fixity and government involvement. The manner in which these characteristics are then modified so as to enable standard modelling of the housing market is detailed in Section 2.3, along with the baseline theoretical model that underlies the empirical estimations presented in Chapters 6 and 7. The theoretical relationships applicable in a housing context are set out in Section 2.4. Section 2.5 concludes.

2.2. Housing Attributes

Housing may be distinguished from other goods and services on the basis of its heterogeneity, durability, spatial fixity and the extensive involvement of government in its marketplace. As noted overleaf, a necessary prerequisite for appreciating the complexity of housing market economics is an understanding of the various attributes of housing, and the means by which the interaction of these features complicates theoretical and empirical analyses of the market. Fortunately, such an understanding has been aided by the emergence of a large volume of literature - the different strands of which 'may be categorised according to which of these characteristics is emphasized and formally modelled in the analysis' (Smith, Rosen and Fallis, 1988). The principal concerns of the literature, as they relate to each characteristic of housing, are briefly summarised below. In addition, the following sub-sections define these attributes.

2.2.1. *Heterogeneity*

Heterogeneity implies that no two houses are the same in every respect. This assertion is borne out by the fact that it is possible to differentiate between dwelling units on many grounds, including those of size, age and design. With regard to the explicit modelling of heterogeneity, housing economics tends to adopt one of two methods; either a characteristics based approach or an approach that focuses on the quantity and quality aspects of housing (Smith, Rosen and Fallis, 1988) [1]. The former encapsulates the idea of hedonic pricing in that each housing unit is described in terms of a bundle of characteristics that commands a price in the marketplace. This price is subsequently regressed on the individual characteristics of the

---

[1] The characteristics based approached was formally proposed by Rosen (1974) within a framework that allowed for product differentiation in a competitive market.

bundle so as to reveal their implicit or hedonic prices. In contrast, the second approach forms part of the literature concerned with the maintenance and renovation decisions of house owners.

2.2.2. *Durability*

A second feature contributing to the distinctiveness of housing is that of durability. This refers to the long lasting nature of housing and implies that 'the existing stock is quite important relative to the flow of newly constructed stock' (Smith, Rosen and Fallis, 1988)[2]. A partial treatment of durability allows for stock-flow adjustment in the housing market, with the intersection of demand and a perfectly inelastic short run supply curve for housing stock / services yielding the unit price of the latter. A capital value per unit of housing stock is then determined by the interaction of the unit price of housing services and factors such as risk premia and the yields on various other assets. Ultimately, the flow of new construction depends on the supply price of new stock relative to the capital value derived above. However, Smith, Rosen and Fallis (1988) argue that a full treatment of durability requires more than an inelastic short run supply curve, and that the markets for housing stock and housing services need to be thought of as distinct concepts. Furthermore, they suggest that these two markets can again be differentiated by means of tenure choice; a topic that in itself has given rise to a huge body of literature.

2.2.3. *Spatial Fixity*

The uniqueness of housing is also reinforced by the importance of the location of the physical stock. This characteristic, known as spatial fixity, is related to the heterogeneity of housing and suggests that housing units, physically similar but located in different neighbourhoods,

---

[2] Note that a further implication of this attribute of housing is the importance of housing finance (Muth and Goodman, 1989).

may sell for different prices. Location in this context comprises three aspects, namely the distance from important areas such as the employment centre, shopping facilitates and transport routes, the nature of land use in the relevant neighbourhood, and local government services and tax levels. Modelling the first aspect of spatial fixity can be undertaken within a general equilibrium framework as proposed by the new urban economics school or alternatively, via computer simulation. In examining the second aspect, the hedonic pricing technique mentioned earlier may be adopted, whereas the third is often discussed as part of the literature arising from Tiebout's (1956) paper on local government.

2.2.4. *Government Involvement*

The final attribute associated with housing is the extent of government involvement in this marketplace. While the intervention mechanisms pursued across countries differ considerably - ranging from taxation and regulatory control to non-market provision - 'the level of involvement compared to other product markets is uniformly high' (Smith, Rosen and Fallis, 1988). The volume of literature in this area is extensive, although it is possible to broadly categorise it in terms of studies that take a normative perspective and those that adopt a positive economics approach. Typically, the latter aim to analyse the end affects of various policy measures and housing programmes whereas, the more normative studies are largely concerned with issues such as what the role of the government vis-à-vis housing ought to be.

2.3. The Theoretical Model

Given the impossibility of incorporating all of the above characteristics into a single model, empirical analyses of housing markets based on time series data and focussing on the relationship between the price or quantity of housing and an array of other variables, necessarily impose a number of simplifying assumptions. The twofold purpose of this section is to firstly set out these assumptions, and secondly, to present the theoretical housing model that they give rise to.

Principal amongst the employed assumptions is the existence of a theoretical and unobservable homogeneous commodity called housing service. Allowing for each dwelling unit 'to yield some quantity of this good during each time period', and assuming that it is 'the only thing in a dwelling unit to which consumers attach value', enables the issue of heterogeneity to be successfully overcome (Olsen, 1969). However, as housing service is an unobservable entity, it is not possible to directly test this proposition, though arguably, a relationship between housing service and certain observable phenomena may be defined and then tested (Muth, 1960 and Olsen, 1969).

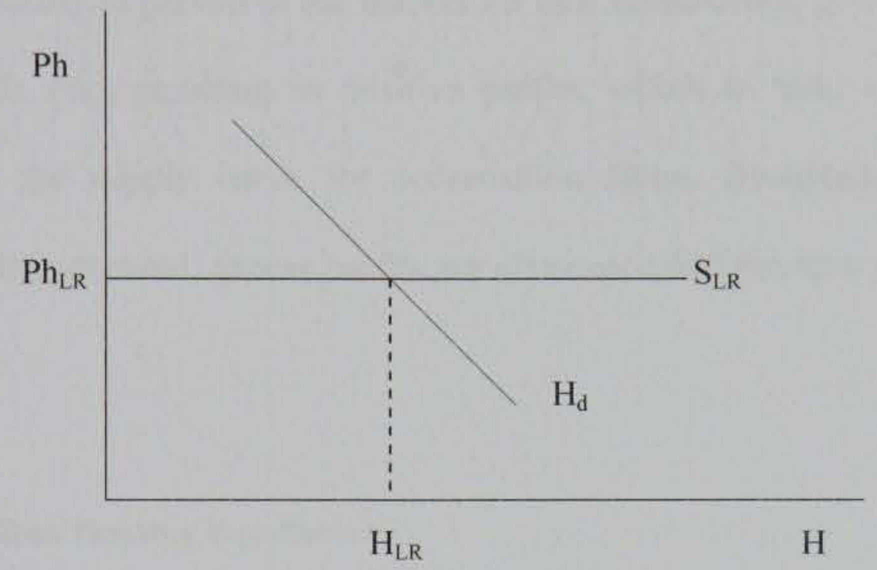
In terms of accounting for the durability characteristic of housing, empirical analyses usually consider some variant of the adjustment process discussed in Section 2.2.2. The key assumption made in this regard concerns the flow of services provided by housing. This flow is assumed to be proportional to the stock of housing whereas a full treatment of durability requires that the markets for housing stock and housing services be further distinguished.

Lastly, the spatial aspects of housing are generally ignored.

The theoretical model derived as a result of combining the above assumptions is discussed in Hendry (1984). This model is commonly adopted as the starting point of time series studies of the housing market, including those undertaken in Chapters 6 and 7 of the present research. Such analyses begin with an equilibrium model for the demand and supply of housing, where the term equilibrium refers to 'no inherent tendency to change' (Hendry, 1984). Within this framework, adjustment in the market occurs in a stock-flow manner; both the markets for existing housing stock and the flow of new construction must be balanced to eliminate disequilibria. The long run static situation is characterised by a perfectly elastic supply curve.

Depreciation of the existing stock is exactly offset by new construction such that $C = \delta H$, where C is new construction, δ represents depreciation and H is the physical stock of housing. Demand is downward sloping and the intersection of demand and supply determine the equilibrium price Ph_{LR} and quantity H_{LR}. At this price, perfect competition implies that the construction market is earning a normal profit. As a result, the stock and flow markets are in long run equilibrium and demand is said to determine quantity but not price (Hendry, 1984).

Figure 2.1: Long Run Housing Equilibrium

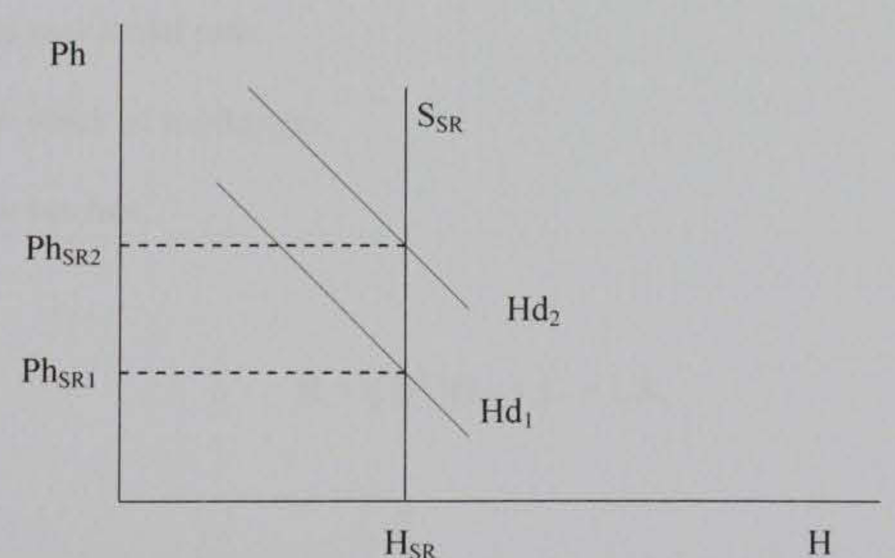


In practice, it is unlikely that supply is perfectly elastic even in the long run. Miles (1994) and Poterba (1984) consider an upward sloping long run housing supply schedule, suggesting that such a schedule can be derived under the assumptions of either diminishing or constant returns to scale – a binding land availability constraint can give rise to the second outcome - and in this case house prices should exhibit some response to changing demand conditions. In an Irish context, the notion of such a constraint is readily applicable given the shortage of zoned and serviced land, especially in the Dublin area (IMF, 2000).

Extending the above analysis, Poterba (1984) employs a type of Tobin q-theory of investment, suggesting that the production decisions of firms are based on a comparison of current house prices and the current costs of production. As this may be considered a little naïve, uncertainty can be incorporated into the model (see for example, Salo (1994)) and the impact on the supply of housing of exogenous variables such as land availability and the proportion of production pre-sold, can also be assessed.

The short run situation differs from that of the long run in the sense that supply is assumed to be perfectly inelastic or fixed initially. Therefore, an increase in the demand for housing necessitates an increase in price in order to restore equilibrium to the market. In this framework changing demand conditions drive house price determination. However, disequilibrium is likely to prevail in the market for new construction given that the short run price Ph_{SR} exceeds Ph_{LR} resulting in positive profits, which in turn, entice entry and an outward shift of the supply curve for construction flows. Eventually supply increases sufficiently to satisfy demand, excess profits are eliminated and the flows market balances at the long run price.

Figure 2.2: Short Run Housing Equilibrium



Figures 2.1 and 2.2 respectively illustrate the above verbal descriptions of the long and short run theoretical situations.

2.4. Theoretical Relationships

Drawing on the above, this section sets out the key theoretical relations that apply in a housing context. The intention in doing so is to provide some indication as to the nature of relationships that might be expected when it comes to the empirical analysis of the Dublin housing market. While comprehensive, note that the list of variables presented below is not exhaustive[3].

$$H_d = f\ \{ph, Y, MR, D, p, M, T) \quad (2.1)$$

where:

H_d	=	housing demand;
ph	=	real house prices;
Y	=	permanent income;
MR	=	nominal mortgage interest rates;
D	=	demographics;
p	=	the real rental rate;
M	=	the stock of mortgages;
T	=	the tax rate;

$$H_s = (1-\delta)H_{t-1} + C_t + LA_t \quad (2.2)$$

where:

H_s	=	stock supply of housing;
δ	=	the rate of depreciation;
H_{t-1}	=	housing stock in the previous period;
C_t	=	completions of new private sector houses at time t;
LA_t	=	sale / transfer of Local Authority houses at time t;

[3] See also Hendry (1984) and Giussani and Hadjimatheou (1990) for further details.

$$C = f\{MR, credit, bc, ls, lp\} \quad (2.3)$$

where:

C	=	completions of new private sector houses;
credit	=	availability of credit;
bc	=	real building costs;
ls	=	the stock of zoned housing land;
lp	=	real land prices;

$$Ph = f\{Y, MR, D, p, M, T, credit, bc, ls, lp\} \quad (2.4)$$

where:

Ph	=	house prices.

2.5. Conclusion

As any empirical analysis must be conducted within the framework of economic theory, this chapter sought to provide a clear outline of the theoretical issues that arise in a housing context. Against this background, the special characteristics associated with housing were discussed, as was the standard time series approach to the modelling of this marketplace. Finally, some key theoretical relationships were presented.

Chapter 3:

Literature Review

3.1. Introduction

A testimony to the growth of academic and policy interest in housing market economics over the past decades is the extensive body of literature that currently exists in this field. To date, housing research covers a broad spectrum of topics ranging from analyses of price determination and tenure and location choice to issues such as urban regeneration, neighbourhood externalities, densities and the housing needs of the under-privileged. While each of these strands is of interest, the focus of the present study is such that it is considered appropriate to limit the following discussion to a review of literature that is of a time series econometric and macroeconomic orientation. Accordingly, this chapter presents the empirical findings of relevant Irish and international studies that adopt a time series, macro approach to modelling housing. In presenting these findings, the intention is to facilitate the identification of variables that could potentially influence demand, supply and price inflation in the Dublin marketplace, bearing in mind the theoretical relationships set out in Chapter 2.

The chapter is structured as follows: to begin with, Section 3.2 provides a brief overview of a priori theoretical expectations. Respectively, Sections 3.3 and 3.4 then consider the findings of recent empirical research on the Irish housing market and the results of selected international studies. Section 3.5 concludes with a summary of the variables, which in terms of precedents set by the literature, may act as determinants of Dublin housing demand, supply and price.

3.2. A Priori Expectations

To lend to the interpretability of the empirical findings discussed below in Sections 3.3 and 3.4, Table 3.1 lists several factors that are generally cited as playing a key role in the housing market. The variables presented are those that, on a priori economic grounds, are assumed to affect the demand, supply and price of housing in the long run. The perceived effect, that is whether an increase in the factor in question impacts in a positive or negative manner on housing demand, supply or price, is signified by either a (+) or (-) sign. As with the variables listed in Table 3.1, these signs follow from theoretical predictions.

Table 3.1: Key Factors affecting Housing Demand, Supply and Prices

Factors	*Demand*	*Supply*	*Price*
House prices	-	+	
Income	+		+
Interest rates	-	-	?
Demographics	+		+
Building costs (labour and material)		-	+
Land		+	-

Notes:
1. Assuming perfect capital markets and a binding land constraint

3.3. Empirical Research on Ireland

Taking as the starting point the above expectations and the validity of the chosen modelling techniques, this section considers the findings of recent empirical work on the Irish market. Thus far, housing research in Ireland has primarily focussed on developments at a national level, though some limited work has been carried out in respect of the Dublin marketplace as detailed below.

3.3.1. *National Housing Market Studies*

The most prominent study on the national housing market to date is a series of three reports known as the 'Bacon Reports'. Of these, the 1998 and 2000 papers (Bacon 1 and 3 respectively) undertake an econometric analysis of house price determinants while the 1999 report provides a descriptive update of the post Bacon 1 situation.

Bacon 1 (Bacon et al., 1998) adopts an econometric approach to the modelling of the national housing market that is comprised of two parts, namely an inverse demand and a house completions equation. Estimation of both equations uses the ordinary least squares methodology and is applied to a dataset covering the period 1974-1996. As regards the inverted demand model, the log of real second-hand house prices is taken as the dependent variable while the set of explanatory variables incorporates the change in the log of real personal disposable income per capita, the log of real personal disposable income per capita lagged, the log of per capita housing stock, user costs and the population aged 25-34 as a percentage of the total population. Excepting the first independent variable, all regressors are correctly signed and statistically significant, with the overall model explaining 86% of the variation in existing house prices.

Turning next to the supply equation, the empirical estimation - which relates the log of private house completions to the current and lagged log of real new house prices and to lags of logged completions - gives rise to an adjusted R squared statistic of 0.821. In line with initial expectations, the above explanatory variables are found to be significant and correctly signed. On the other hand, when potential determinants such as real building costs and real interest rates were tested for possible explanatory power, they proved to be either incorrectly signed or statistically insignificant. The authors also point out that the failure to include the real cost of land as a regressor in the model is due to the fact that a time series of residential land prices is not available in an Irish context. However, it is noted that even if such a series did exist, it would be necessary to allow for high short run correlation between residential land and house prices given that 'the demand for residential land is a derived demand' (Bacon et al, 1998).

Overall, Bacon 1 (Bacon et al., 1998) presents a standard textbook model of the housing market, the results of which appear plausible. However, important limitations of the analysis are the incomplete modelling of the supply-side of the market and the failure to consider the time series properties of the data. As the model is expressed in levels, the presence of non-stationary variables could give rise to spurious regression. In this case, inference using the usual OLS procedures may be misleading. As such, it would have been desirable to test for the possibility of unit roots amongst the data series.

Murphy and Brereton (2001) seeks to update and expand the empirical aspect of the Bacon 1 study. The data range is increased to 1999 and the econometric analysis presents a model of the second-hand housing market using the same technique and variables as in the previous report. Unsurprisingly, the findings are close to those of Bacon 1; though the adjusted R squared statistics appear to be higher for both the inverted demand and house completions equations over the extended sample period. In light of the above discussion, these results should be treated with caution.

Similar to its predecessor, Bacon 3 (Bacon and MacCabe, 2000) estimates an inverse demand and a completions equation. Data for the former is annual and covers the years 1972-1996 and 1972-1999. For both time periods, the inverted demand equation seeks to explain logged second-hand house prices in terms of the current and lagged log of the stock of housing, logged personal disposable income, the base mortgage rate, population aged 25-34 and the log of lagged second-hand prices. Estimation again applies the ordinary least squares technique. Subsequent empirical results indicate adjusted R squared statistics of 0.755 and 0.949 for the respective sample periods. For the years 1972-1996, the income and demographic variables are found to be significant and signed as expected whereas for the extended sample, statistically significant and correctly signed parameter estimates include those of income, mortgage rates and lagged house prices. With respect to the latter, the authors consider 'gathering momentum in demand, based on price expectations' to be 'consistent with the emergence of a significant speculative or transitory demand factor in the Irish housing market' (Bacon and MacCabe, 2000).

In terms of the supply-side of the market, the private house completions equation incorporates as explanatory variables the log of current and lagged house prices, the log of current and lagged building costs (where such costs include material and labour expenses but not land costs) and the log of lagged completions. While the resultant empirical findings reveal an adjusted R squared value of 0.879, only current prices and lagged completions emerge as statistically significant from zero and correctly signed.

In commenting on these results, the starting point is again the failure to examine the statistical properties of the underlying data series. If these are non-stationary in nature, the OLS inference procedures applied in the analysis are likely to be misleading. This calls into question the reliability of the resultant findings. In addition, the incompleteness of the supply-side equation is notable.

In contrast to Bacon 1 and 3, Kenny (1998) does consider the statistical properties of the data and finding evidence of non-stationary behaviour, models the interaction between house prices and a range of macroeconomic factors within a cointegration framework. For the sample period 1975 (first quarter) to 1997 (last quarter), the empirical results identify the presence of two long run / equilibrium relations, one for the demand-side of the market and one representing supply. The demand vector relates the log of the housing stock, the log of real house prices (where house prices are constructed as a weighted average of the price of new and second-hand homes), real gross national product logged and the building society mortgage interest rate. The second vector incorporates supply-side elements, namely the log of real house prices, the building society mortgage interest rate and the log of a composite index of costs that includes material, labour and development land costs. The end results are plausible with each variable bearing the anticipated sign.

In order to capture the dynamic nature of the housing market, the paper also examines the matrix of alpha coefficients in the vector error correction model and undertakes an impulse response analysis. The former reveals that house prices adjust positively (albeit reasonably slowly) to situations of excess demand whereas the latter indicates that in response to a shock to income, 'the stock of housing increased by only $1/10^{th}$ of the proportional increase in the demand for housing' (Kenny, 1998). The author suggests that severe supply-side constraints, such as a lack of available land or services infrastructure, offer a possible explanation for this occurrence.

While the choice of cointegration as a modelling tool is appropriate in this case, one aspect of the above analysis is questionable, namely the derivation of the land cost measure from house prices.

More recently, the Irish housing market has been re-modelled by McQuinn (2004) who presents an econometric analysis based on quarterly data covering the period 1980:1 to 2002:4. As in Kenny (1998), the Johansen technique is adopted to test for possible cointegration. The empirical findings identify two forms of a long run inverted demand / house price equation and a long run supply relation. The first cointegrating house price vector relates the log of average house prices (new and second-hand prices combined), the real rental price, logged housing stock per capita, the level of net migration, logged income per capita and the log of the average mortgage approved. All of the variables are signed correctly with the exception of the rental rate. Given this, the second cointegrating relation for house prices replaces the rental measure with the real interest rate. Regarding the supply-side of the market, the uncovered equilibrium vector is comprised of the log of total house completions (private and Local Authority), average house prices logged, the log of house prices divided by an index of builders' costs and land costs. These variables enter the supply relation signed as expected a priori.

Building on the above cointegration analysis, the paper then estimates error correction models that incorporate the identified long run relationships. As in the static framework, two short run specifications are estimated for the change in the log of house prices while one such equation is estimated for the change in log of total completions. In terms of the impact of deviations from the equilibrium house price and supply relations on the respective dependent variables, the empirical findings indicate that the house price error correction term enters both house price equations significantly and negatively signed, while the supply-side term similarly enters the completions equation significantly and bearing the expected negative sign.

Lastly, the paper plots the fitted values of the second house price error correction model against actual prices. On the basis of this graph, it is suggested that as actual prices were in excess of the fitted 'fundamental' house price for 1998 and 1999, some degree of over-valuation may have existed. However, from late 2000 onwards, the series are 'practically the same' implying that recent movements in the housing market have been motivated by fundamentals (McQuinn, 2004).

Overall, McQuinn (2004) presents a comprehensive model of the Irish housing market, the results of which appear plausible. While the chosen statistical technique is also appropriate, the issue of potential endogeneity between land and house prices is again relevant and merits attention going forward.

In contrast to the above, the main focus of Roche's (2003) analysis of the Irish housing market is on explaining price inflation in the context of fads, bubbles and fundamentals. In this case, the observed house price is decomposed into a 'fundamental' and 'non-fundamental' component and a regime-switching approach is adopted to test the market for evidence of speculative behaviour. The reduced form 'fundamentals' model (which is found to be a long run cointegrating relationship) takes real new house prices as the dependent variable and is estimated using ordinary least squares over the sample period 1979:1 to 2003:1. Explanatory variables include a combination of demand and supply factors, namely an index of builders' costs, land costs, net immigrants, the loan amount, user costs, disposable income per capita and a trend to proxy influences such as household formation[1]. The resultant coefficients are correctly signed and significant.

The difference between the fit of the estimated 'fundamentals' model and actual prices, - the non-fundamental price - acts as a measure of over or under-valuation. In order to examine the claims of the IMF (2003) and 'The Economist' (2003) that Irish house prices were over-valued in 2002, the paper then proceeds to estimate two other models of 'fundamental' house prices based on the methodologies employed by these bodies. Firstly, following the IMF (2003), second-hand house prices are related to real disposable income, real mortgage interest rates and the young house buying population as a percentage of the total population. The second additional specification seeks to explain real second-hand house prices using real personal disposable income per capita and is close to the non-econometric approach adopted by 'The Economist' (2003). Subsequent empirical findings reveal that when Irish house prices in 2002 are modelled solely in terms of demand-side factors (as in the analyses undertaken by the IMF and 'The Economist'); they are over-valued by some 11% to 75%. In contrast, the use of both supply and demand factors (as in Roche's model) confines the extent of over-valuation to between 0.2% and 4.6%.

[1] Using three different methods, the paper tests for the possibility that land costs are endogenous. Such a situation implies that changes in new house prices could cause changes in land prices and vice versa. However, in all cases, it was found that land costs could be treated as exogenous.

With regard to testing for the existence of a bubble in the new housing market (namely a sharp rise in price, with the initial rise generating expectations of further increases - the probability of continued growth falls as the bubble expands and the rise is usually followed by a partial collapse), the fact that the estimated demand and supply 'fundamentals' model constitutes a cointegrating relation suggests that a bubble is not present[2]. Further support for the lack of a bubble component is provided by the empirical outcome of the regime-switching model which finds evidence of a 'fad' in the market (whereby prices deviate temporarily from underlying fundamentals but eventually mean revert) as opposed to a bubble.

On the contrary, Roche (1999) does find evidence of a speculative bubble in the Irish second-hand housing market for the sample period 1976:4 to 1998:4. Moreover, the paper suggests that the probability of a crash had increased to about 2 per cent by the late 1990s. Interestingly, the measures of the 'fundamental' price in this paper assume a relatively inelastic supply curve and therefore, estimate inverted demand type models that relate real second-hand house prices to real permanent income per capita, expected real mortgage rates, demographics and the real rental price.

3.3.2. *Dublin Housing Market Studies*

The variables adopted when analysing the Dublin housing market should in principle be the same as those applied in the above assessments of the national market. However, Bacon et al. (1998) found that gaps in the data prevented them from undertaking a carbon copy of their national analysis for the Dublin market. As such, they derive a model for logged Dublin house prices relative to national house prices using an average of new and second-hand prices. The estimated OLS inverted demand equation for the sample 1974-1996 exhibits some evidence of instability, and hence, the empirical results are deemed 'suggestive but appear very

---

[2] However, as acknowledged by Roche, the use of cointegration as a test for speculative behaviour is not very powerful if a partially collapsing bubble is present (Evans, 1991).

plausible' (Bacon et al., 1998). In particular, lagged real per capita personal disposable income and the population share aged 25-34 were found to favour Dublin houses prices relative to the rest of the country.

In terms of modelling the supply-side of the market, the log of private house completions in Dublin is taken as the dependent variable with the explanatory variables including lagged completions, lagged real Dublin house prices (comprised of average new and second-hand house prices) and the change in the log of real Dublin house prices. With the exception of the latter, the regressors are statistically significant and all are correctly signed. The adjusted R squared statistic indicates that the estimated model explains 57% of the variation in completions. Furthermore, the long run price elasticity of supply in the Dublin market is found to be roughly half the national figure. The authors note that 'this is consistent with the view that non-economic constraints reduce the elasticity of supply in the Dublin area' (Bacon et al., 1998).

Murphy and Brereton (2001) expand the timeframe of the above empirical house completions model from 1996 to 1999. As expected, the findings are much the same, though the estimated equation now explains 73% of the variation in the dependent variable.

However, to the extent that the models presented in these studies are expressed in levels and include non-stationary variables; OLS is not an appropriate estimation technique and its use may undermine the reliability of the reported results. Furthermore, as in the national models, the supply-side equations are incomplete - no attempt made is made to include cost variables or a land measure. The latter is particularly striking given the above reference to the impact of non-economic constraints on housing supply in the Dublin area.

Roche (2001) seeks to explain whether bubbles, fads or fundamentals are driving new house prices in Dublin over the period 1976:1 to 1999:1. As in his corresponding papers for the national market (1999 and 2003), the observed house price is decomposed into a 'fundamental' and 'non-fundamental' component and a regime-switching approach is adopted to test the market for evidence of speculative behaviour. With respect to the calculation of the 'non-fundamental' price, four possible proxies are estimated. The first adopts an inverted demand type approach and relates real new house prices to expected real disposable income, expected real mortgage interest rates and net migrants. The resultant OLS coefficient estimates are statistically different from zero and are correctly signed. The residual from this 'fundamentals' equation is then taken as a measure of the 'non-fundamental' price. In contrast, the second proxy is based on the standard asset pricing model and defines the 'non-fundamental' price as the real house price minus the mean price-rent ratio multiplied by the real rent. The third method assumes that the 'non-fundamental' house price is correlated with the real growth rate of new house mortgages, while the last measure assumes a correlation between the 'non-fundamental' house price and the new house price-building cost ratio (where development land prices are excluded).

These various measures of the 'non-fundamental' price are subsequently fed into a general regime-switching model in order to test their predictive power. The overall empirical findings indicate that there is some evidence of a speculative bubble in Dublin house prices.

While Roche's (2001) analysis is thorough in the sense that alternative measures of the fundamental price are considered, when modelling underlying fundamentals in the Dublin market, demand and supply variables should be considered in tandem. Moreover, the use of OLS as an estimation technique in this case is also questionable.

3.4. International Empirical Research

In addition to the above findings, further insight into the range of variables that may potentially influence demand, supply and price inflation in the Dublin housing market can be gained by considering the empirical results of relevant international research. Reflecting this, the following aims to provide an overview of the broader literature that informed the present study.

When considering international housing literature of a time series, macroeconomic nature, research undertaken in respect of the UK market represents a natural starting point given the structural similarities between the markets for housing in Ireland and the United Kingdom. For this same reason, empirical studies on the UK market also proved the most informative and consequently, account for the bulk of the international literature reviewed below.

Of the many studies carried out on housing in the United Kingdom, Hendry (1984) presents what is perhaps the best known econometric model of UK house prices[3]. Using quarterly data covering the period 1959:1 to 1982:2, the paper adopts an error correction approach that enables the estimation of a dynamic model of second-hand house prices, which in turn, is solved to yield a static equation. With regard to the latter, the empirical results find that the loan to income ratio, real income per household, the interest rate and the rate of inflation constitute the long run determinants of the average house price to household income ratio.

In terms of the dynamics, the estimated specification relates the change in the dependant variable to the lagged change in second-hand house prices, the change in real personal disposable income, an error correction term defined as the ratio of borrowing to own equity, the ratio of real personal disposable income to the stock of housing, retail prices, real lending and levels and differences of the interest rate. With the exception of the interest rate, all the variables are expressed as logs. In addition, the coefficients are correctly signed, significant and explain some 78% of the change in second-hand house prices.

[3] For a detailed overview of UK housing market studies, see Meen (1993).

A particularly interesting feature of the above specification is the inclusion of a cubic price difference term, which allows for a non-linear response on the part of house prices to excess demand. While the inclusion of this term is valid in so far as its coefficient is significant, Hendry (1984) asserts that 'in a sense, the cubic transformation is too extreme: for $|\Delta_1 ph| < 0.1$, the effect of the cubic term is trivial, whereas it dominates for $|\Delta_1 ph| > 0.3$'.

Teräsvirta (1998) re-estimates the dynamic equation for house prices set out in Hendry (1984) using a smooth transition regression (STR) model. The objective in doing so is to determine if an improved specification can be achieved using this specific non-linear technique. The author concludes that this is indeed the case, stating that the estimated STR model 'contributes to the general understanding of the mechanism generating UK house price expectations during 1960-1981' (Teräsvirta , 1998).

Similar to Hendry (1984), Giussani and Hadjimatheou (1990) model the static and dynamic determinants of second-hand house prices in the United Kingdom. For the sample period 1961:1 to 1988:2, the empirical results suggest a long run relationship between the log of second-hand house prices and the log of the number of households, the log of real personal disposable income per capita, the log of new construction costs, the rental rate and the log of the housing stock. The potentially important role of land is also referred to, but due to a lack of data on costs and availability, is not modelled. The resultant coefficients all bear the expected sign and are statistically significant. Moreover, on the basis of the Dickey Fuller and Cointegrating Regression Durbin-Watson (CRDW) tests, the residuals of the estimated long run equation are found to be stationary, implying that the above variables are in fact cointegrated. While valid, one drawback of the Engle-Granger approach adopted in this paper is the difficultly in testing for rank. This implies that the number of cointegrating relations must be known a priori, which is not always the case.

The associated short run model seeks to relate the change in logged second-hand house prices to changes in, inter alia, the log of second-hand house prices squared and cubed, the log of real personal disposable income per capita, the log of the stock of mortgages, the log of retail prices, the log of new construction costs, the log of net advances of mortgages and the rental rate. The lagged residuals of the estimated long run equation are also included as an error correction term. This term is negatively signed implying the gradual elimination of any deviations from the equilibrium house price relation. Overall, the coefficients of the dynamic model are significant and signed as expected a priori.

Lastly, note that the presence of the squared and cubic terms in the short run specification gives rise to a non-linear relationship between current and lagged house price changes. The nature of this relationship is illustrated by Giussani and Hadjimatheou (1990) who, taking the range of values for Δph_t from their sample as the starting point, show that 'for values in the range between 0 and 0.15, the relationship is positive, ie. increases in house prices will trigger expectations of even higher prices. For values outside this range, that is negative and larger than 0.15, the relationship is negative'.

In contrast to the Engle-Granger methodology adopted by Giussani and Hadjimatheou (1990), Drake (1993) models UK house prices using the Johansen cointegration technique. Over the period 1981:1 to 1990:1, the findings identify a long run equilibrium relation between logged house prices and the log of real personal disposable income, the building society mortgage interest rate, the log of private sector housing starts and a dummy variable. The latter is designed to capture the impact on house prices of the 1988 Budget announcement that dual tax relief on mortgage borrowing would end later that year. The coefficients of the estimated relation are all correctly signed.

Interestingly, changes in income and the interest rate variable do not enter the preferred dynamic specification. Conversely, the error correction term, which is constructed as the lagged residuals of the cointegrating vector, enters significantly and bearing a negative sign as anticipated.

However, the use of the Johansen technique may not be appropriate in this case given the very small size of the sample. As Juselius (forthcoming) notes 'many simulation studies have demonstrated that the asymptotic distribution [of the Trace statistic] can be a poor approximation to the true distribution when the sample size is small resulting in substantial size and power distortions'[4].

Finally, Salo (1994) presents two models of the Finnish housing market over the period 1960 to 1988. The first is a standard demand model based on slow adjustment, while the second takes the form of a simultaneous model of demand and supply. Particularly noteworthy are the inclusion of an explicit land variable and the adoption of an inter-temporal focus when developing the supply-side equation. In this respect, the paper assumes that the firm's production decision depends not only on current variables such as the real price of the existing stock, construction costs and an index of the price of building lots, but also on expectations concerning the future market situation[5]. Reflecting the fact that all dwellings may not necessarily be sold in the period in which they are produced, a variable measuring the expected cost of unsold production is incorporated into the model. In addition, technological development is captured by the inclusion of a trend.

The empirical results indicate that the coefficients of the estimated supply equation are for the most part significant and signed as expected a priori. Specifically, current house prices and technology play a positive role in influencing production, whereas the price of building lots is found to negatively impact on the decision to supply housing. On the other hand, the t value

---

4 [ ] parenthesis inserted.

5 Note that in the presence of a binding constraint for building lots, the production decision would depend on the availability rather than on the price of such lots (Salo, 1994).

for the variable capturing future market conditions is somewhat on the low side and the index of construction costs proves incorrectly signed throughout.

3.5. Conclusion

In keeping with the focus of the present research, this chapter sought to provide an overview of the findings of Irish and international studies that have adopted a time series, macroeconomic approach to modelling housing. The aim in doing so was to facilitate the identification of appropriate factors for inclusion in the empirical models of the Dublin marketplace presented in later chapters. To this end, particular attention was paid to the nature of the explanatory variables put forward by the selected studies and to the validity of the modelling techniques used.

Overall, the literature appears to be in agreement as to the key factors at play in the housing market. Allowing for definitional variations, these broadly encompass price, income, interest rates, credit, building costs, the existing housing stock, demographics, land and government interventionist policies. As anticipated, the empirical findings in this respect also fit well with a priori theoretical expectations as to the determinants of housing demand, supply and price inflation.

As such, the above variables can be viewed as representing a sound starting point when modelling the Dublin housing market. However, it should be noted that of these potential determinants, credit is typically only relevant if explicit rationing is in place, while the shortage of zoned and serviced land in the Dublin area implies that availability rather than price is the appropriate land measure.

Chapter 4:

Data Description

4.1. Introduction

The aim of this chapter is to provide a detailed description of the data series that underlie the empirical sections of the study. Such an aim is motivated by the fact that it was necessary to refine much of the available data, the consequent result being a need to explain how the individual series are constructed.

The chapter is organised as follows: Section 4.2 defines the variables, data sources and based on a graphical analysis, discusses their time series properties[1]. Section 4.3 briefly concludes.

4.2. Data Description

The data needed to construct an econometric model of the housing market is reasonably extensive in nature. In identifying potential data series, Chapter 3's analysis of existing housing literature proved useful as it facilitated the identification of a range of variables, which in terms of precedents, may act as key determinants of demand, supply and price inflation in the Dublin housing market. A priori expectations in this respect posit that variables such as income, mortgage interest rates, building costs, the existing housing stock, demographics, land availability and government interventionist policies all play an important role in the market, and therefore, should be considered when modelling housing. The precise

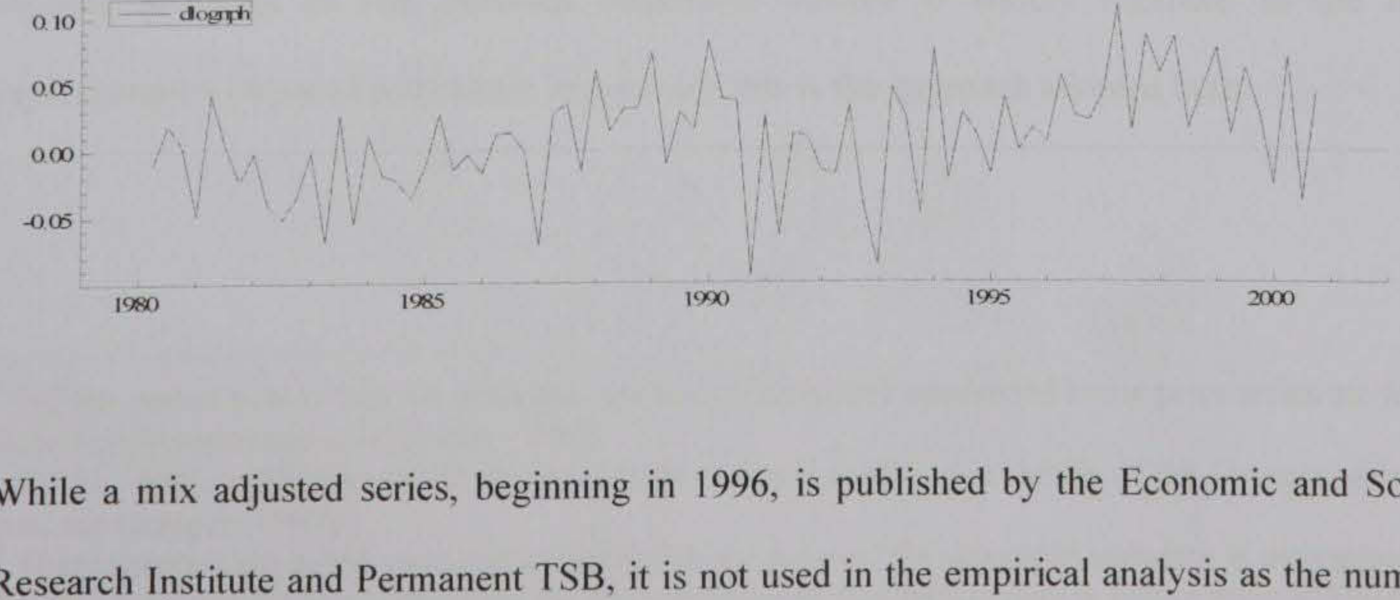
[1] These properties are formally tested in Chapter 6.

definition of the variables used in this study, along with the time series characteristics of the data, are outlined below for the sample period 1980 quarter one to 2000 quarter four. Note that most variables are transformed into logs. This allows for the coefficient estimates presented in later chapters to be interpreted as elasticities. For variables expressed in levels, the corresponding coefficients are interpretable as semi-elasticities.

4.2.1. *New House Prices (nph)*

The variable *nph* measures the average price of new houses in Dublin in real terms. Nominal house price data is available on a quarterly basis from the 'Housing Statistics Bulletin' published by the Department of the Environment, Heritage and Local Government. These figures are in turn derived from information on loan approvals supplied by the main mortgage lenders. The data is not mix adjusted and therefore does not account for heterogeneity across new completions.

Figure 4.1: Log of New House Prices expressed in Levels and First Differences



While a mix adjusted series, beginning in 1996, is published by the Economic and Social Research Institute and Permanent TSB, it is not used in the empirical analysis as the number of observations available is considered too small for making sensible inference in a long run framework[2]. In converting house prices from nominal to real terms, the consumer price index (sourced from the Central Statistics Office) is used as the deflator.

With respect to the time series properties of the new house price variable, it can be seen from Figure 4.1 that prices are non-stationary in levels but are potentially just stationary in first differences.

Lastly, the interaction between new and second-hand house prices over the sample period was examined. The results of a standard Granger causality test indicate that movements in new house prices caused movements in second-hand house prices, but not vice versa[3]. As such, second-hand prices can be disregarded from the present analysis of the Dublin housing market.

4.2.2. *Income (inc)*

Ideally, permanent income should be modelled given that a household / individual's house purchase decision making process tends to be related to lifetime earnings along the lines of Friedman's Permanent Income and Modigliani and Ando's Life Cycle Hypotheses[4]. However, as this variable is unobservable in practice, a second best outcome is needed. Measuring income in terms of real personal disposable income is widely regarded as the most representative proxy of permanent income and this is the approach adopted here.

---

[2] For the period where both are available, the mix adjusted and unadjusted house price series are found to be highly correlated (coefficient = 0.99).

[3] Roche (2003) reports a similar finding for the national market. For details of the Granger causality test, see Granger (1969).

[4] These theories are based on the proposition that the demand for consumer spending is determined by average long run income rather than by current earnings. For further details please refer to Friedman (1957) and Modigliani and Ando (1963).

In defining the income variable as above, it was necessary to overcome a number of data constraints. Firstly, nominal personal disposable income per capita figures are only available on a regional basis for the years 1991-1999 in the CSO's 'Households, Incomes, Regions and Counties 1991-1999' report. As such, there is a need for approximation. Annual total personal disposable income figures are given for the State as a whole in the 'National Income and Expenditure Accounts' published by the Central Statistics Office for the entire sample period 1980-2000. In addition, national and Dublin population data is available from the CSO's 'Labour Force Survey' and 'Population and Migration Estimates' for 1983-2000. National figures for the preceding three years are obtainable from the statistical databank at www.cso.ie but unfortunately, population figures for Dublin are not available. Combining these various data series, it is possible to construct a measure of real personal disposable income per capita at a national level by taking total personal disposable income, dividing by population and then deflating by the consumer price index.

When constructing a Dublin income measure, real personal disposable income per capita figures are assumed to be 115.4% of the national figure for each year. This assumption is based on an analysis of the State and Dublin figures for the 1990s as presented in the 'Households, Incomes, Regions and Counties 1991-1999' report, and is in line with the view of the Economic and Social Research Institute[5]. Figures 4.2 and 4.3 provide a graphical overview of this analysis. Expressing Dublin figures as a percentage of national figures over the period 1991-1999, it is clear that they are moving together at a stable rate and that the growth rate is relatively constant. There is no reason to expect that the relationship between the series would have differed during the 1980s.

[5] It would of course be preferable to have a data series measuring income in the Dublin region - be it personal disposable income, household income or Gross National Product - for the entire sample period. Unfortunately, no such series is available.

Figure 4.2: Real Personal Disposable Income per Capita

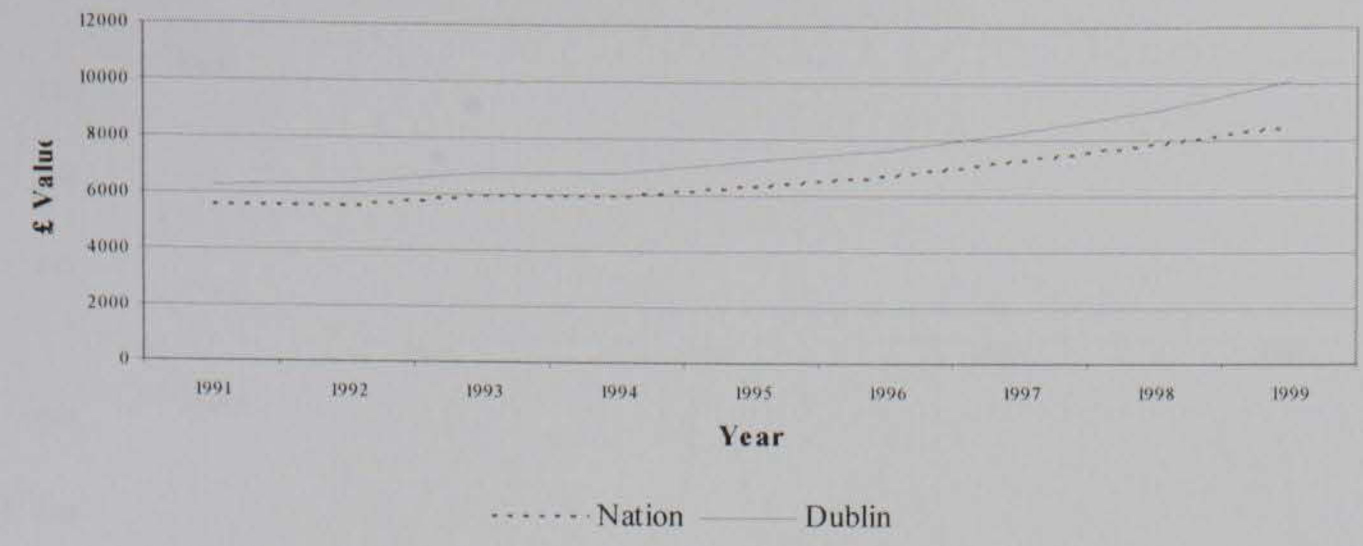
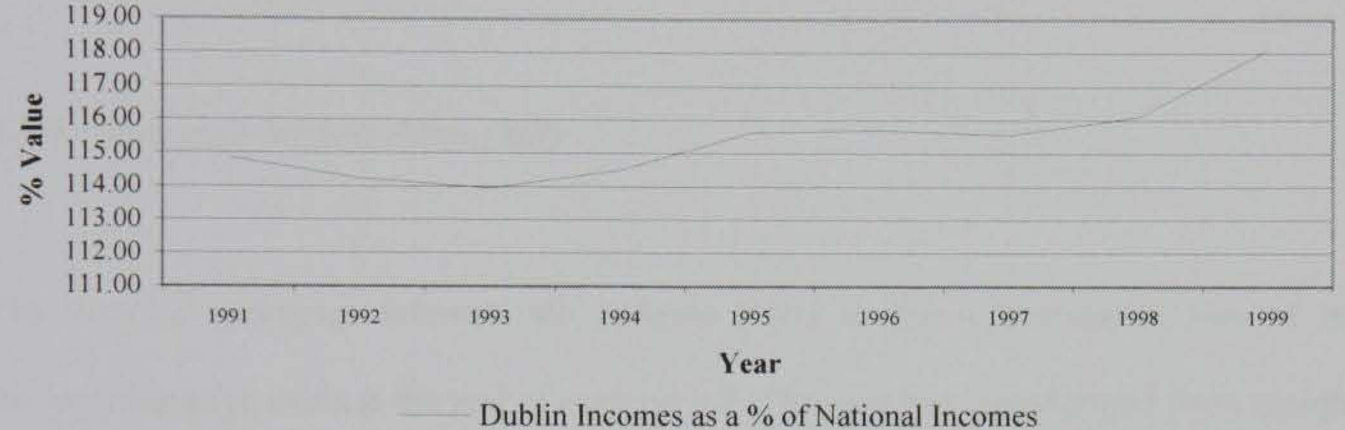


Figure 4.3: Real Personal Disposable Income per Capita Relationship

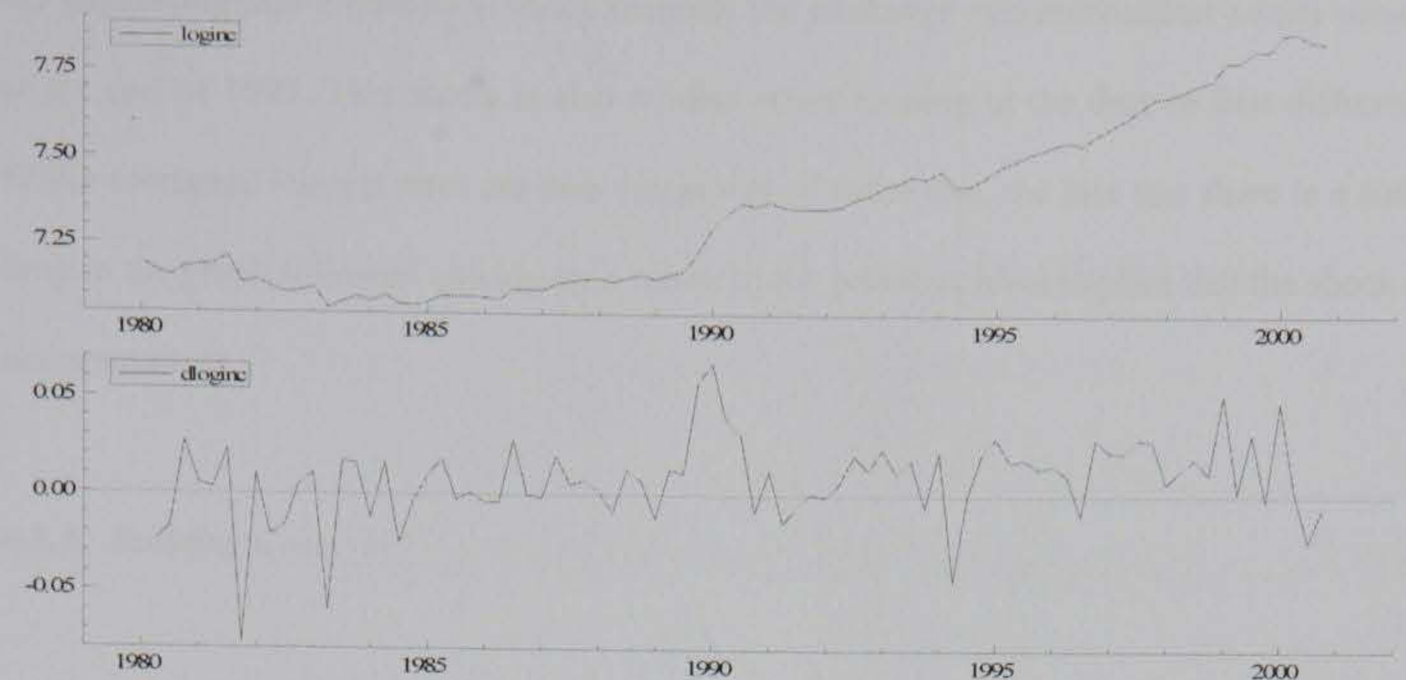


A second issue that arises when working with income data relates to its frequency. The empirical models of later chapters are expressed in quarterly terms, but as income data is unavailable at this frequency, annual figures were interpolated using the Denton method[6]. The seasonally adjusted volume of retail sales index, published by the Central Statistics Office, acted as the interpolation series used to obtain quarterly values.

From Figure 4.4, it can be seen that the final income variable is non-stationary in levels and contains both a trending component and possibly a level shift in the data about 1990. The variable is however stationary when differences are taken.

[6] See http://www.imf.org/external/pubs/ft/qna/2000/Textbook/index.htm for further details.

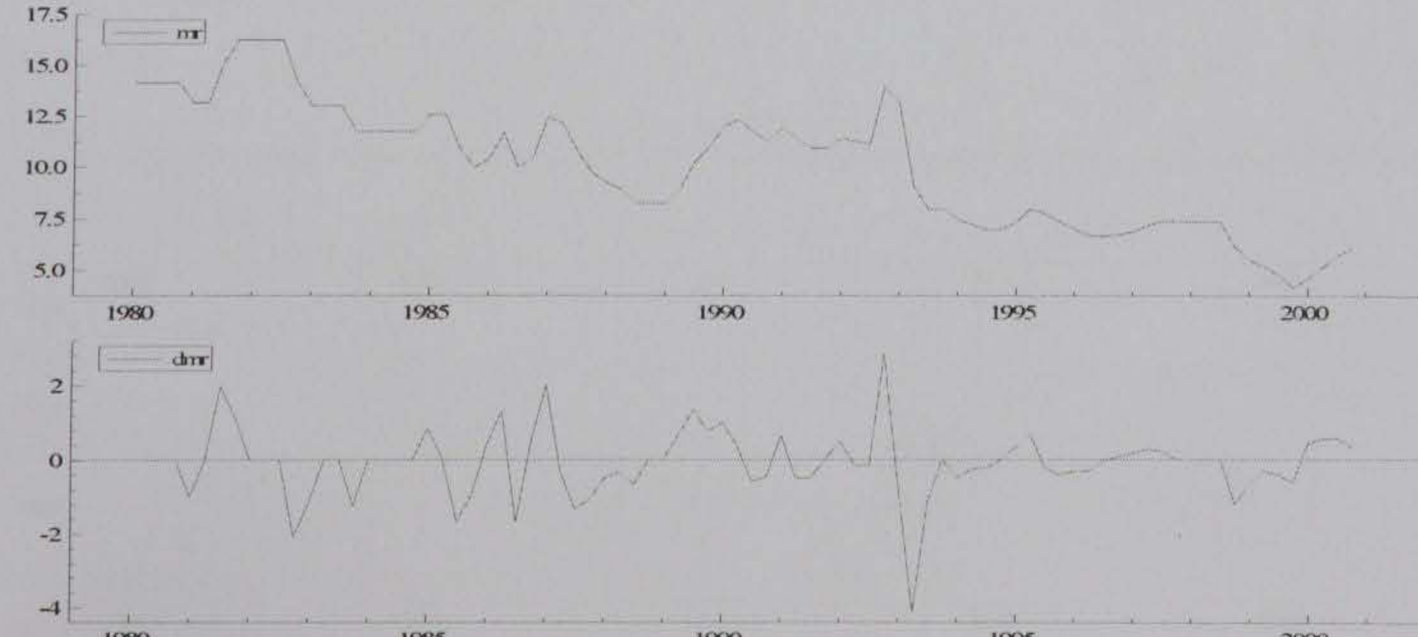
Figure 4.4: Log of Income expressed in Levels and First Differences



4.2.3. *Mortgage Interest Rates (MR)*

The nominal mortgage interest rate variable refers to the representative rate of building society mortgage loans at the end of each period. The data was transformed from monthly into quarterly values by simple averaging and is available in its raw format from the Central Bank's 'Quarterly Bulletin' and the 'Economic Series' published by the CSO.

Figure 4.5: Mortgage Interest Rates expressed in Levels and First Differences

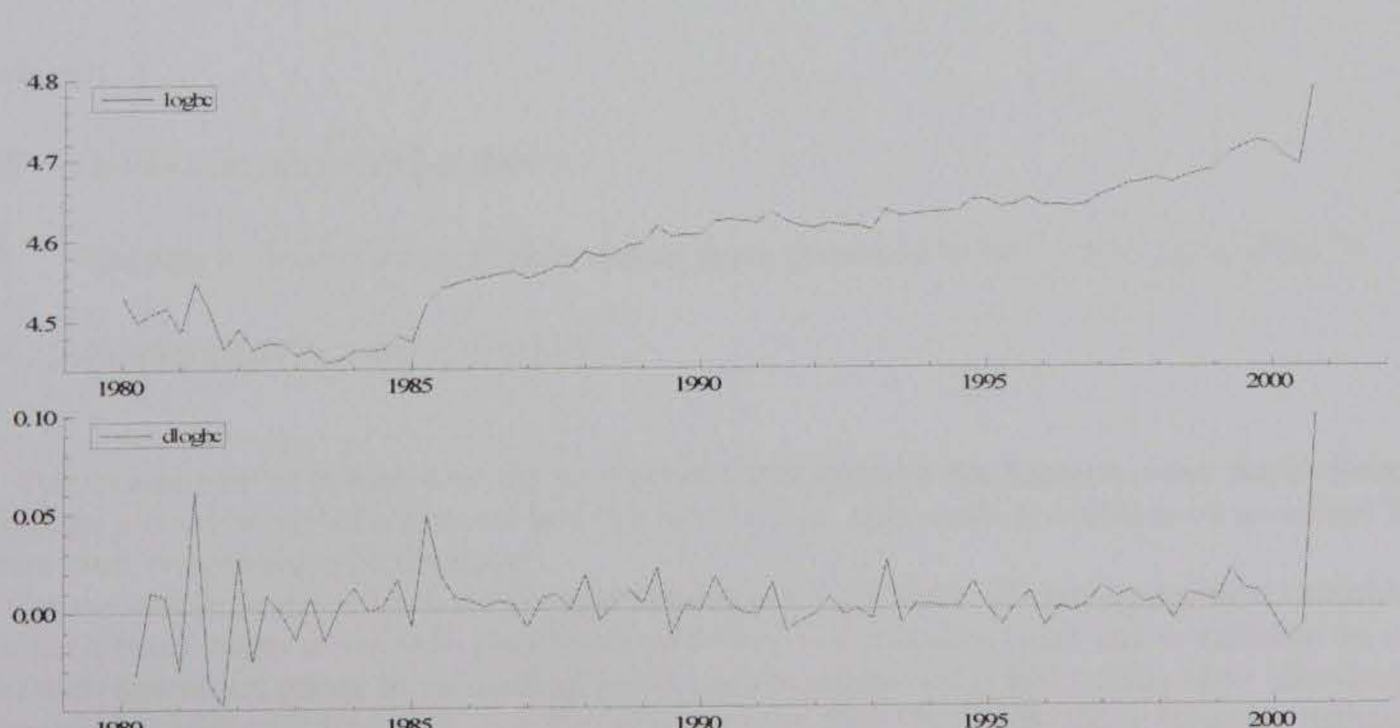


Regarding the time series properties of the variable, in levels it is non-stationary with Figure 4.5 suggesting that a transitory shock (namely the exchange rate mechanism crisis) occurred at the end of 1992. This shock is also evident when looking at the data in first differences. While mortgage interest rates are now integrated of order one, the fact that there is a sudden drop in the graph followed quickly by a return to the previous level implies that the shock was not permanent.

4.2.4. *Building Costs (bc)*

The variable *bc* is defined as the nationwide index of building costs taken from the 'Housing Statistics Bulletin' published by the Department of the Environment, Heritage and Local Government. A separate index for the Dublin region is not available. Monthly figures are averaged to obtain quarterly values and the series is then deflated by the consumer price index to convert it into real terms[7]. This index relates solely to material and labour costs – these normally do not exceed 65% of the total price of a house. Other costs facing suppliers in the construction industry, such as land, are not included.

Figure 4.6: Log of Building Costs expressed in Levels and First Differences



[7] As the base quarter for both indices is set at 1990:1 = 100, it is possible to divide one by the other.

From the graphical analysis presented in Figure 4.6, it can be seen that building costs are non-stationary in levels but stationary when differenced. There is also evidence of a trend and a level shift in the data. This shift is reflected in the levels graph by the movement to a higher path in 1985 and in the corresponding blip in the graph of the first differences. In addition, a transitory shock occurs in 1981 and an outlier is clearly evident for the final observation of the *bc* variable. The latter may be explainable by the fact that the index, since October 2000, includes the first phase of a review of rates of pay and grading structures for the construction industry, and, the first phase increase under the 'Programme for Prosperity and Fairness' ('Housing Statistics Bulletin', DELG).

4.2.5. *Private Housing Stock (phs)*

The private housing stock variable is compiled using data on private house completions and the sale of Local Authority houses. As the starting point, the stock of private housing in Dublin in 1971 is assumed to be 5% greater than the number of households in the region as given by the 1971 'Census of Population'[8]. The stock is then cumulated applying the perpetual inventory methodology which is formalised as follows[9]:

$$H_t = (1-\delta)H_{t-1} + C_t + LA_t \qquad (4.1)$$

where:

H_t = private housing stock at time t;

δ = the rate of depreciation of the existing stock (assumed to be 0.125% per quarter[10]);

H_{t-1} = private housing stock at time t-1;

[8] This assumption is in line with the assumptions that underlie the Economic and Social Research Institute's construction of a national housing stock series. This series is available on an annual basis from their 'Macroeconomic Database'.

[9] Census 2006 provides a figure for the total housing stock in Dublin. If such figures were available for earlier Census years, it would be possible to take them as a reference point, and to calculate the stock for each intercensal period by subtracting newly constructed dwellings and making some allowance for depreciation of the existing stock over the relevant period (See Fitzgerald, 2005). However, data on the stock of housing in Dublin is not available prior to 2006.

[10] Again, this assumption is consistent with the approach adopted by the Economic and Social Research Institute, which assumes a depreciation rate of 0.5% per annum for the above national housing stock series.

C_t = completions of new private sector houses at time t;

LA_t = sale / transfer of Local Authority houses at time t.

Data on the number of private completions and sales / transfers of Local Authority houses in the Dublin region is sourced from the Department of the Environment, Heritage and Local Government's 'Housing Statistics Bulletin'. While the former series is compiled on a quarterly basis for the entire sample period, data for the latter is only available in this format from 1993 onwards. As such, annual figures prior to 1993 are simply divided by four to obtain quarterly values.

Figure 4.7: Log of Private Housing Stock expressed in Levels and First Differences

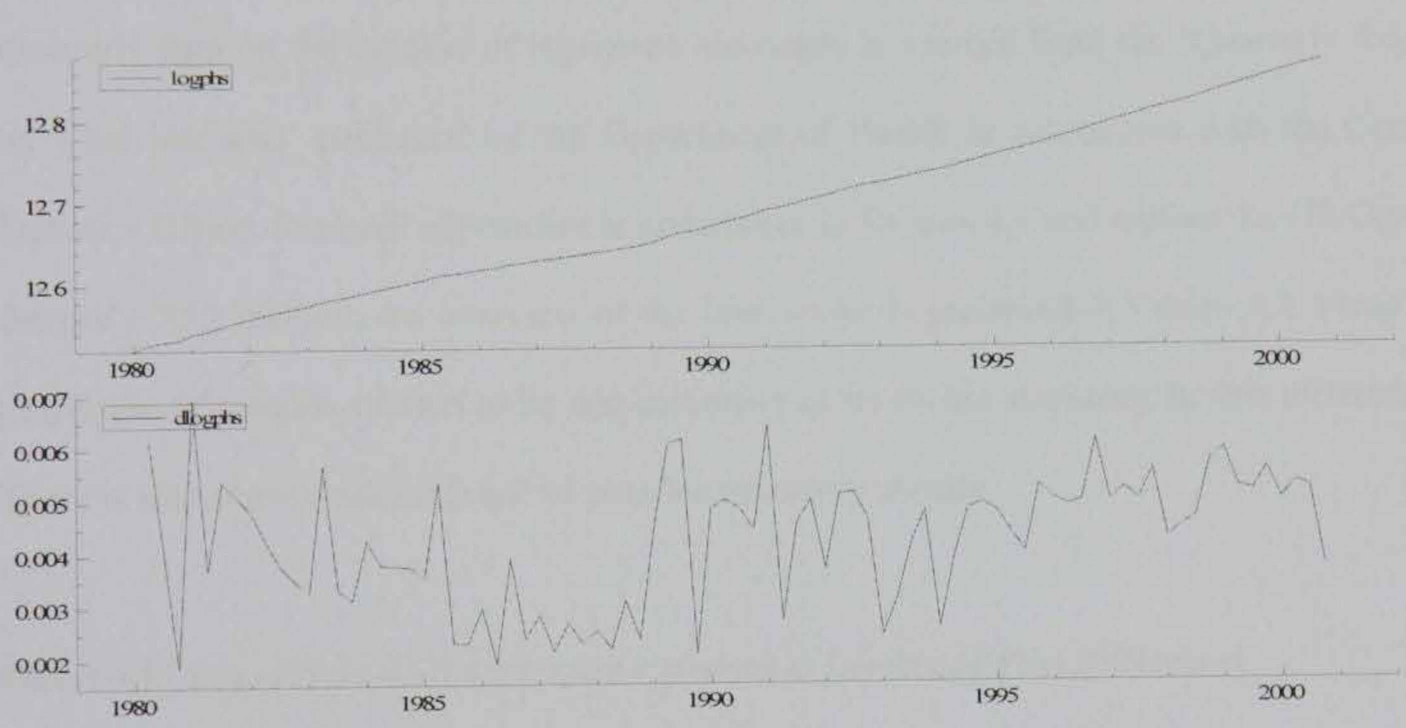


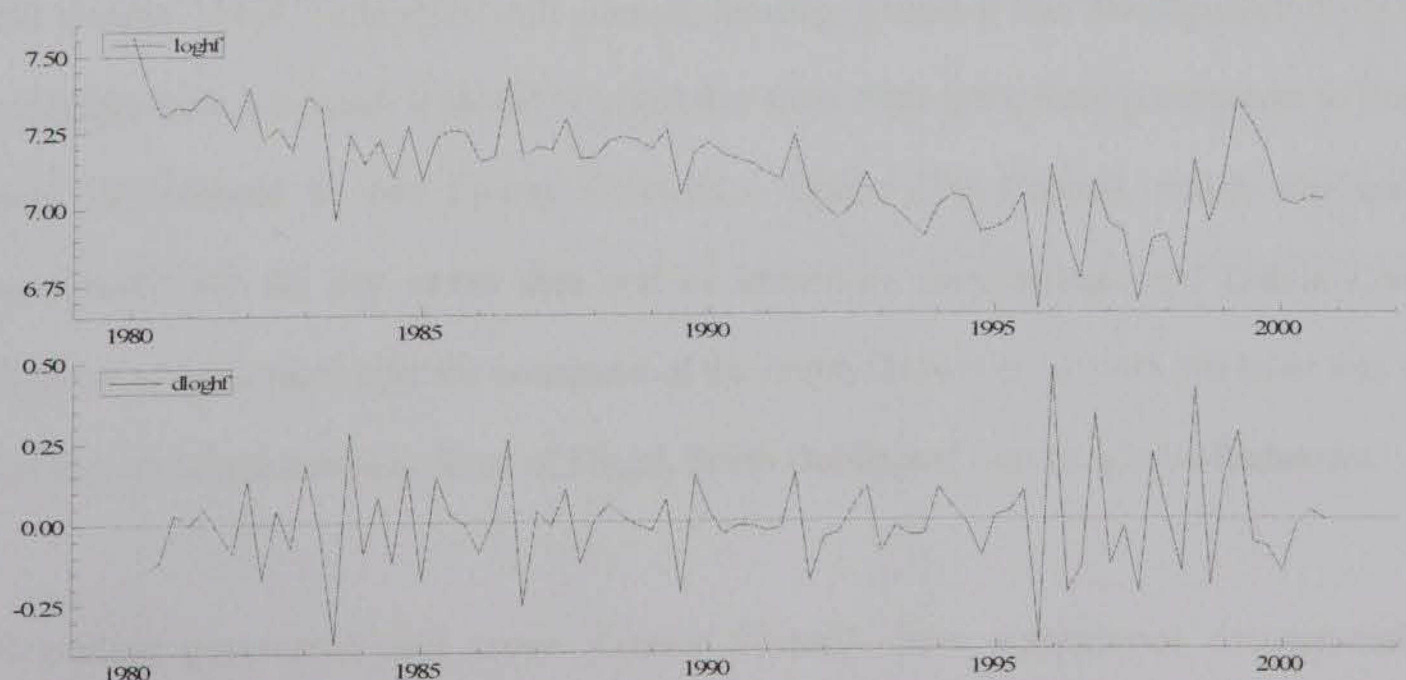
Figure 4.7 provides a graphical illustration of the constructed private housing stock variable. In levels, the data exhibits a strong trending component and is obviously non-stationary. It is less clear if this finding of non-stationarity carries over to the first difference of the variable.

4.2.6. *Household Formation (hf)*

The variable *hf* measures new household formation. Ideally, this series should be constructed as the change in the number of households but while the Central Statistics Office's 'Labour Force Survey' details household numbers on an annual basis, data for Dublin was not gathered prior to 1988. It is therefore necessary to proxy household formation and following Hadjimatheou (1976); the seasonally adjusted number of marriages is used. In justifying this choice of proxy, note that marriages undoubtedly constituted the largest component of household formation throughout the 1980s and early 1990s. This implies that the proxy is reasonable for most of the sample period.

Quarterly data on the number of registered marriages is sourced from the 'Quarterly Report on Vital Statistics' published by the Department of Health in association with the Central Statistics Office. Seasonal adjustment is undertaken in Eviews 4.1 and applies the US Census Bureau's X12 method. An overview of the final series is presented in Figure 4.8. From the graph, the *hf* variable appears to be non-stationary in levels but stationary in first differences. There is also some weak evidence of possible transitory shocks.

Figure 4.8: Log of Household Formation expressed in Levels and First Differences



Finally, with respect to the construction of a demographic variable, alternative measures to household formation that were considered in the course of this study included the number of individuals in Dublin of house buying age in absolute terms, along with these individuals as a percentage of total Dublin population. Annual figures for these potential series are available from the CSO's 'Labour Force Survey' and 'Population and Migration Estimates'. However, as discussed in Section 4.2.2, population data is not available for Dublin for the years 1980-1982. Moreover, while the 1981 'Census of Population' does contain such data for Dublin, the age classifications differ from those adopted by other data sources, and thus, a consistent series cannot be constructed. Overall, the range of difficulties encountered meant that it proved impossible to capture demographic effects on housing using any series other than that of the proxy for new household formation.

4.2.7. *Land Stock (ls)*

The *ls* variable represents an attempt to quantify the stock of land available for housing in Dublin over the sample period 1980-2000. The term 'available for housing' refers to the fact that the land is both zoned residential and is provided with services such as water and sewerage. As this information has not been gathered by any state or private organisation prior to 1997, it was necessary to generate it for inclusion in the study. In doing so, the minutes of the various Local Authorities' full council, housing, planning and development committee meetings were consulted. It should be noted that from 1980-1993, local government in Dublin was administered by two County Councils – Dublin City Council, which was mainly responsible for the city centre area and its immediate surroundings, and Dublin County Council which looked after the remainder of the county. However, in 1993, the latter was sub-divided into three councils, those of Fingal, South Dublin and Dun Laoghaire-Rathdown.

Regarding governance and scope, County Councils have autonomous decision-making powers with respect to issues of land use zoning and the granting of planning permissions. In addition, each is required by legislation to prepare a Development Plan for its administrative area on a five yearly basis. The aim of these plans is to set out the proposed nature of development for the given area for the five-year period in question. Plans must also be accompanied by a series of maps indicating the zoning objective(s) applicable to individual areas within the broad administrative area of the Local Authority. Unfortunately, Development Plans do not give a breakdown of the amount of land zoned for a specific purpose, and as the associated maps are not to scale; it is not possible to obtain a quantifiable figure for zoned housing land from this source. However, the making of the above Plans requires a public viewing process of draft versions, the tabling of motions, discussions amongst elected Council members and finally, the carrying or otherwise of the proposed motions by the Council and the passing of the Plan. The outcome of the voting process is recorded at each stage in the minutes of the relevant meeting and thus it is possible to glean information on land use decisions related to housing from this source.

There are however two drawbacks to the above approach. The first relates to the fact that occasionally the number of hectares involved in a zoning decision is simply not stated. In this case, the construction of a land availability variable using data recorded in council minutes may be somewhat biased. Secondly, it is not possible to reconcile data concerning a decision to zone a particular plot of land for housing with a later decision to service it. This disparity arises due to the manner in which the details of initiated serviced schemes are reported – they tend to be referred to under broad title headings and cover large, loosely defined locations. As such, it is impossible to isolate out a particular plot of housing land to determine whether it has been included in a specific service scheme or not.

While the bulk of large-scale zoning / re-zoning decisions are undertaken as part of the preparation process for the making of the next Development Plan[11], land use changes also

[11] This process usually includes a review by Council members of the existing Development Plan, along with the preparation of a draft version of the forthcoming Plan.

come about as a result of amendments to the existing Plan, Section 4 motions and material contraventions. The latter two refer to situations whereby planning permission is sought but the proposed land use in the application is in conflict with the zoning objective(s) for that land as outlined in the relevant Development Plan. For example, a developer may apply to a Local Authority for permission to build a residential estate on land that is zoned industrial in that County Council's Development Plan. If planning permission is granted in these circumstances, a material contravention is said to have occurred.

The final decision to grant a planning application that contravenes the existing Development Plan rests with the elected members of the Council, and is decided by a vote following discussion and consideration of the views of relevant County Council Departments such as Planning, Development and Services. These decisions, along with amendments, are recorded as part of the minutes of Council meetings and therefore, this source can again be exploited to obtain data needed for the construction of a land availability variable. Note though that the drawbacks discussed above also apply here.

Overall, it was decided to proceed and measure the availability of housing land in Dublin on the basis of data sourced from the minutes of the meetings of the various Local Authorities. Despite the limitations mentioned previously, working with this data represents the sole means of compiling a sensible land availability series. There is simply no other information source that quantifies the changing patterns of land usage in Dublin for the sample period 1980-2000[12].

[12] The Department of the Environment, Heritage and Local Government publishes a series (beginning in 1999) in its 'Housing Statistics Bulletin' that details the number of units of zoned serviced land available on an annual and county basis. Some ad hoc sources that provide one off estimates include the Bacon 3 Report (Bacon and MacCabe, 2000), which gives an annual figure for available units of zoned serviced land in Dublin in 1998, the Bacon 1 Report and a joint publication on the part of the Dublin Local Authorities (Bacon et al, 1998 and Dublin Corporation et al, 1999). The latter two both provide a figure for 1997.

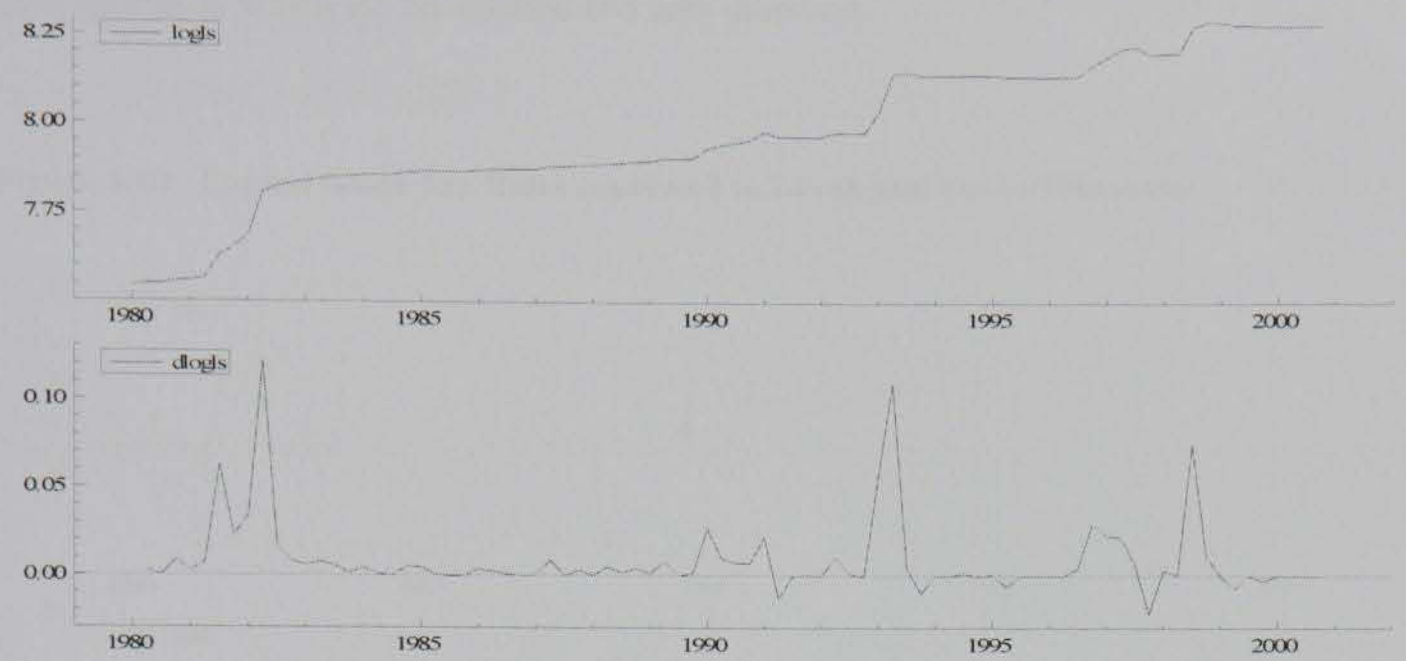
The *ls* series used in the empirical analyses of later chapters is precisely defined as the stock of zoned housing land. Using the perpetual inventory method[13], the chosen base stock is augmented in each quarter by the net number of hectares newly zoned (mainly in Development Plan years) and re-zoned (mainly material contraventions and amendments) for residential purposes. The initial value of the stock is taken from a survey carried out by Dublin County Council's Planning Department in 1977; that quantifies the amount of privately owned land zoned for housing within the Council's administrative area at that time. Unfortunately, Dublin City Council did not carry out a similar survey, though minutes do refer to the fact that most privately owned housing land within the City's boundaries had been fully developed by 1977. Given this, the stock identified by Dublin County Council is assumed to be representative and as such, it is taken as the starting point of the analysis. With respect to the practical construction of the flows aspect of the *ls* variable, the methodology applied involved noting the size of the plot of land concerned and the month in which a housing related zoning motion was passed, and subsequently recorded in the various Local Authorities' minutes. These figures were then combined to obtain quarterly flow values for land availability in Dublin as a whole.

Figure 4.9 illustrates the time series properties of the land stock variable. While clearly non-stationary in levels, the series appears to be stationary when differenced once and to contain a number of shift effects. These shifts are reflected in the levels graph by the movement to a higher path about 1981, 1982, 1993 and 1998 and also in the corresponding blips in the graph of the first differences. In fact, both shifts and blips correspond perfectly to the years in which the different Dublin Local Authorities authorised massive residential zoning / re-zoning as part of the making of their respective Development Plans. The 1981-1982 shift captures the changing land usage that materialised in the run up to Dublin County Council's 1983 Plan, with the 1993 shift measuring change under its' 1993 Plan. The activity around 1998-1999

[13] Please refer back to Section 4.2.5 for a formal definition of this methodology.

corresponds to the review processes and the making of Development Plans on the part of the new Local Authorities of Fingal, South Dublin and Dun Laoghaire-Rathdown.

Figure 4.9: Log of Land Stock expressed in Levels and First Differences



4.2.8. *Government Policies (CGT, dumrpt, dumib and dums2327)*

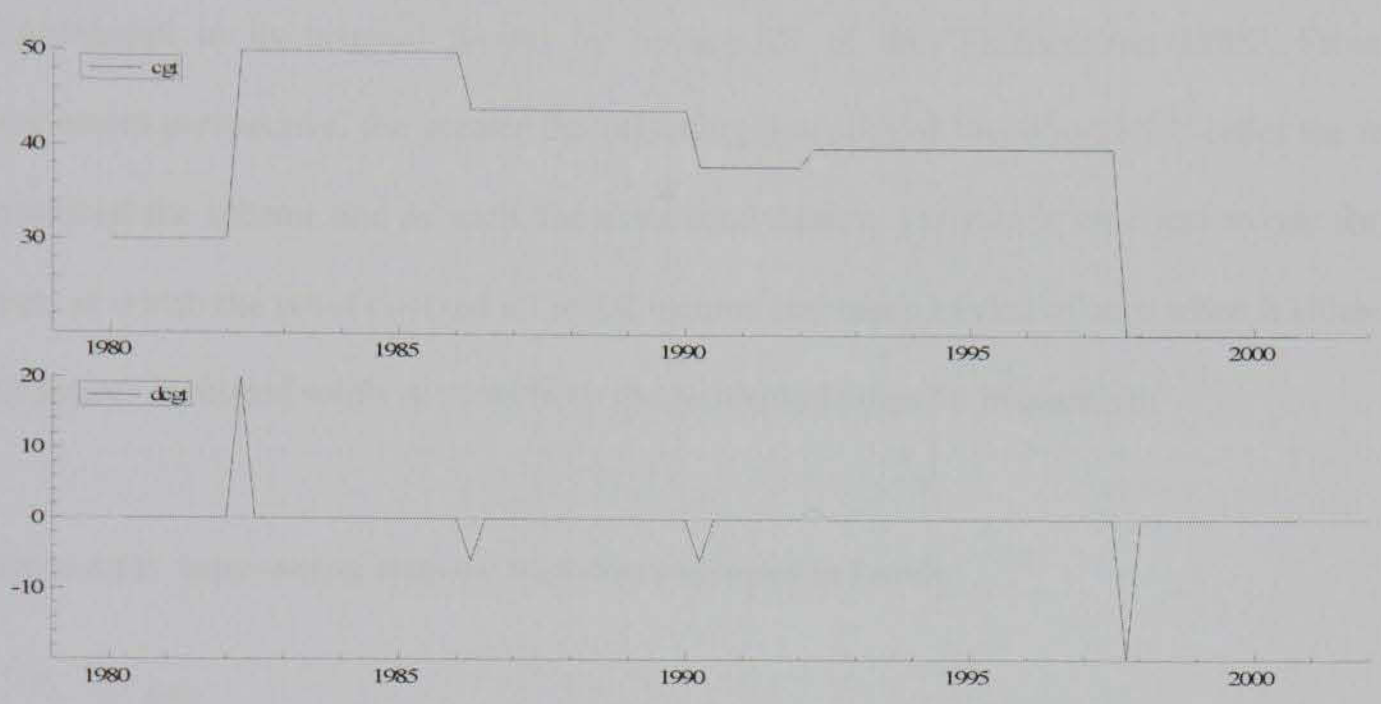
Government intervention in the residential property market may take many forms. In an Irish context, intervention in the new housing market has ranged over the years from property and capital gains taxation to investment incentives such as 'Section 23/27' relief and the tax deductibility of interest on borrowings[14]. The following discussion considers each of these policies in turn while Figures 4.10 and 4.11 present a graphical overview of the constructed series.

In April 1983, a property tax chargeable on the market value of residential property owned and occupied on April 5 each year, came into being. The tax was charged at a rate of 1.5% on

[14] Additionally, on foot of recommendations made in the Bacon 3 Report (Bacon and MacCabe, 2000), an anti-speculative tax was proposed as part of the measures in the 'Finance (No.2) Bill 2000'. This tax was aimed at reducing speculation in the marketplace and was due to take a self-assessment format lasting three years and to apply to residential property in the State that was not the principal private residence of the new owner. However, changes to the Bill announced in February 2001 reversed the above decision and the proposed anti-speculative tax did not go ahead. Therefore, it is not included as part of the study's empirical analysis.

the excess of the market value of the property over an exemption limit, provided that the income of the household also exceeded an income exemption limit. The abolition of the tax was announced in December 1996 with effect from April 1997. With regard to the empirical analysis, residential property taxation is modelled as a dummy variable set equal to one for each quarter in which the tax applied and zero otherwise.

Figure 4.10: Capital Gains Tax Rates expressed in Levels and First Differences

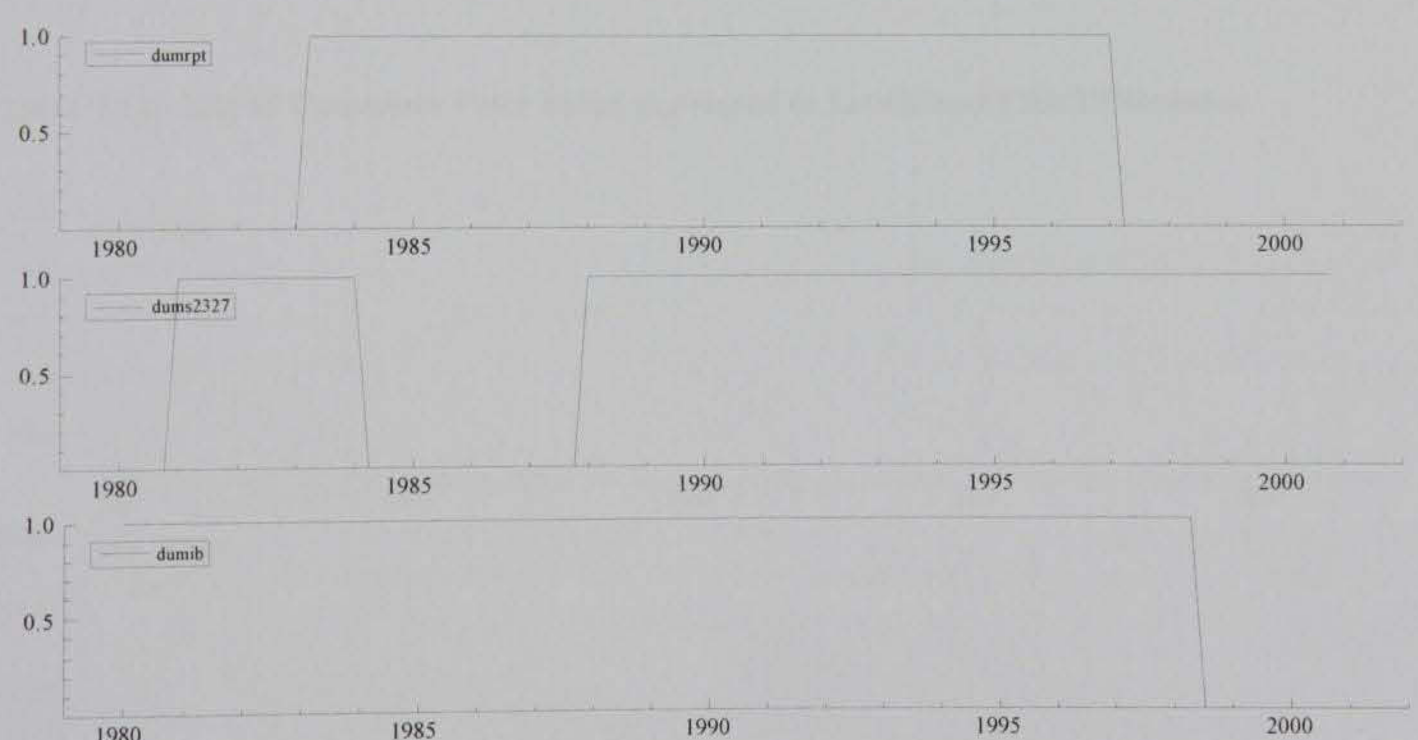


Capital gains taxes on the other hand are not assigned dummy values. Rather, this variable measures the rate of tax payable on the disposal of non-principal private residences[15]. Capital gains taxation was levied at a single rate of 30% over the period 1980 to March 1982 whereas a rate of 40% applied from April 1992 up to the beginning of December 1997, when it was replaced by a lower one of 20%. For the intervening period (April 1982 to March 1992), the tax was levied at various rates ranging from 30% to 60% depending on the length of time the asset was owned. Given this, the empirical analysis takes an average of the applicable rates as an approximation of the typical rate payable at any point during these years.

[15] For the period covered by this study, capital gains were indexed to the consumer price index.

With respect to investment incentives, Section 23 of the 'Finance Act 1981' enacted the first tax relief scheme linked to the provision of rented residential accommodation. This scheme allowed for expenditure incurred in the course of constructing, converting or refurbishing rented residential accommodation to be offset in full against all rental income. The initial scheme, which was due to end in March 1984, was extended for a further three years but modified so that costs could only be offset against rents from the property in respect of which the expenditure was incurred. The scheme subsequently lapsed for a short while before being reintroduced in its original format by Section 27 of the 'Finance Act 1988'. From an investment perspective, the greater the offsetting potential of 'Section 23/27' relief the more beneficial the scheme and as such, the associated dummy variable is set equal to one for the years in which the relief covered all rental income and takes a value of zero when it either did not apply, or related solely to rents from the residential property in question.

Figure 4.11: Intervention Dummy Variables expressed in Levels



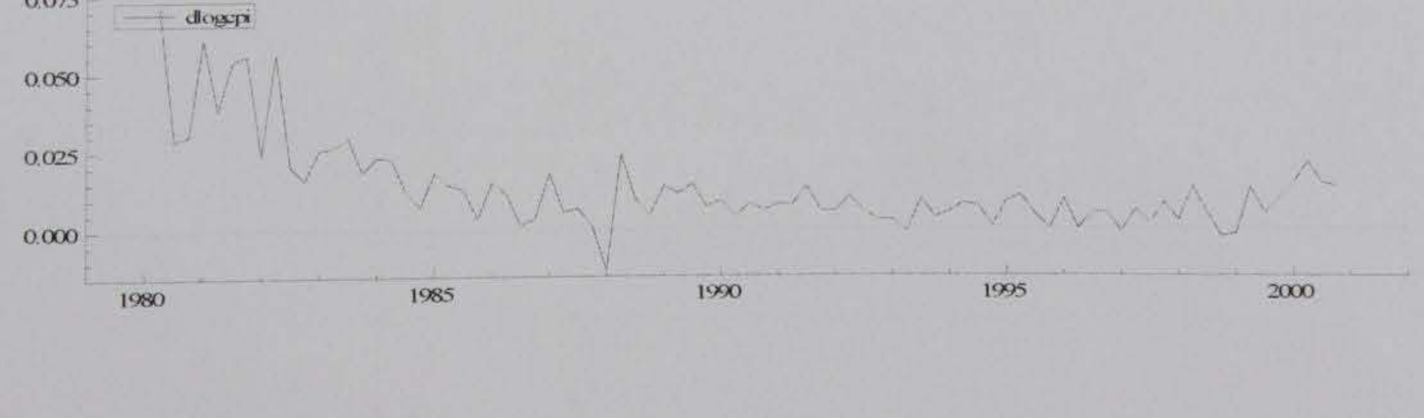
Finally, a dummy is constructed to capture the effect of interest relief on borrowings. This incentive allowed for the interest payable on loans used to purchase residential accommodation for rental purposes, to be written off in full when calculating tax liabilities. On the basis of recommendations outlined in the first Bacon Report (Bacon et al, 1998), the incentive was abolished in April 1998. The dummy variable is thus assigned a value of zero for sample observations post quarter two of 1998 and one for the earlier periods in which tax deductibility was in place.

Data relating to the various taxes and incentives discussed in this section is available from www.irishstatutebook.ie and www.revenue.ie.

4.2.9. *Consumer Price Index (CPI)*

In the context of this study, the consumer price index acts as the deflating mechanism used to convert variables from nominal into real terms. The index, which represents the value of a basket of typical consumer goods and services, is available from the Central Statistics Office at State level[16] . Averages of monthly figures are taken to obtain quarterly values and the base quarter is set at 1990:1 = 100.

Figure 4.12: Log of Consumer Price Index expressed in Levels and First Differences



[16] A 'Dublin and Outside of Dublin' breakdown of prices is not available prior to 2004. Of the 73 items included in the most recent analysis (November 2006), average prices in Dublin were higher for 39 items and lower for 34.

Regarding time series properties, Figure 4.12 shows that the data is trending upwards over the sample period 1980-2000 and appears to be potentially stationary in differences. Although it is not entirely clear from the graphs, there is some weak evidence that a level shift occurs in the index about 1988.

4.3. Conclusion

This chapter sought to provide a detailed description of the data series that underlie the empirical sections of the study. In doing so, source materials and refinement techniques were discussed, along with the time series properties of the constructed variables. Overall findings indicate considerable limitations with respect to the available data. In terms of scope, there are a number of gaps and a general failure to adequately disaggregate data on a Dublin regional basis. In addition, the annual frequency of many series results in the extensive use of interpolation procedures. However, as is evident from the discussion in Section 4.2, the study has largely overcome the above limitations to compile a comprehensive dataset of housing market variables for the period 1980-2000.

Chapter 5:
Methodology

5.1. Introduction

This chapter provides a detailed overview of the modelling techniques applied in the course of the study's empirical analyses of the Dublin housing market. As the estimation methods adopted are reasonably technical, a thorough discussion of the underlying econometric, statistical and mathematical theory is warranted. Accordingly, the following sections are concerned with describing the key features of the chosen methodologies from a theoretical perspective. Intuitive explanations and issues associated with implementation from a practical viewpoint are typically reserved for discussion, as they arise, in later chapters.

The chapter is organised as follows: Section 5.2 describes the cointegrating technique that underlies the long run analysis of the Dublin housing market while Appendix 5A and Section 5.3 respectively, consider the ordinary least squares and smooth transition regression models estimated as part of the adjustment analysis. The regime switching approach adopted when testing for the presence of a bubble component in the marketplace is outlined in Section 5.4. Section 5.5 concludes.

5.2. Cointegration Methodology

The concept of cointegration dates back to the 1980s and a number of influential papers published at that time including Granger (1981), Granger and Weiss (1983), Engle and

Granger (1987), Stock (1987), Johansen and Juselius (1990) and Johansen (1988 and 1991). Originating from the important finding that macroeconomic time series may contain unit roots, which in turn spurred the theoretical development of non-stationary time series analysis; the main idea underlying cointegration theory is the realisation that a set of non-stationary variables may be linearly related in a stationary manner. If such a linear combination exists, the non-stationary series are cointegrated. Another way of expressing this is to say that if two or several variables have common stochastic (or deterministic) trends, they will exhibit a tendency to move together. Moreover, from an economic viewpoint, a cointegrating vector can be interpreted as an equilibrium relationship.

With respect to the empirical specification of a housing model for Dublin, the decision to test for possible cointegration amongst the chosen variables followed naturally from the non-stationary nature of the series, and from reasonable a priori expectations that long run housing demand, supply and price relationships should be identifiable from the data[1]. As regards actual estimation of the model, the Johansen cointegration technique is applied in preference to the more traditional Engle-Granger methodology. The reasoning that underlies this choice is presented below. The remainder of the section discusses representation theory and inference in Johansen's cointegrated vector autoregressive model.

5.2.1. *Johansen versus Engle-Granger Cointegration Analysis*

Table 5.1 overleaf outlines the most popular methodologies employed in testing for possible cointegration amongst a set of non-stationary variables, namely those of Engle and Granger (1987) and Johansen (1995). The advantages and disadvantages of these contrasting techniques are also noted. Overall, the Johansen VAR analysis allows for greater flexibility

[1] Table 6.1 reports the results of Dickey-Fuller and Augmented-Dickey Fuller tests for unit roots in the housing dataset. These tests fail to reject the null hypothesis of a unit root when the series are expressed in levels. However, once the data is differenced, the tests reject the null for each of the variables. Chapter 4's graphical analysis of the time series properties of the data also serves to highlight the non-stationary nature of these series.

and contains a number of desirable properties that are lacking in the Engle-Granger approach. As such, the empirical model outlined in Chapter 6 of this study is estimated on the basis of the former.

Table 5.1: Cointegration Testing – A Comparison of Methods

Let (y_t, x_{1t}, x_{2t}) be the dataset to be investigated for cointegrating properties:

Engle-Granger Regression	*Johansen VAR Analysis*[2]
(1) Regress y_t on x_{1t} and x_{2t} and define the residuals: $ecm_t = y_t - \hat{\beta}_1 x_{1t} - \hat{\beta}_2 x_{2t} - \hat{\beta}_0$	(1) Regress $X_t = (y_t, x_{1t}, x_{2t})'$ on its lags to obtain white noise errors: VAR: $X_t = \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \varepsilon_t$ VECM: $\Delta X_t = \Pi X_{t-1} + \Gamma_1 \Delta X_{t-1} + \varepsilon_t$
(2) Using some parametric or non-parametric method, test if the ecm_t process is stationary.	(2) Using reduced rank regression, test the rank of Π by estimating the model: Hr: $\Delta X_t = \alpha\beta' X_{t-1} + \Gamma_1 \Delta X_{t-1} + \varepsilon_t$ Next, determine the cointegrating rank by applying the likelihood ratio test.
(3) Take $\hat{\beta}_i$ as given and estimate the model: $\Delta y_t = \alpha ecm_{t-1} + \gamma_1 \Delta x_{1t} + \gamma_2 \Delta x_{2t} +$ lagged $\Delta + \varepsilon_t$ The usual chi squared inference can be made on the remaining parameters.	(3) For fixed $\hat{\beta}$, estimate the model: $\Delta X_t = \alpha \hat{\beta}' X_{t-1} + \Gamma_1 \Delta X_{t-1} + \varepsilon_t$ Chi squared inference can be made on α and Γ.
Characteristics: ▪ Simple regression analysis; ▪ $\hat{\beta}$ is super consistent for β; ▪ Inference on β_1 and β_2 is complicated by the presence of nuisance parameters; ▪ In a multivariate setting, the models corresponding to r = 1, 2, 3 are not nested in each other and it is difficult to test for rank; ▪ x_{1t} and x_{2t} should not cointegrate.	*Characteristics:* ▪ Usual inference on β; ▪ Simultaneous modelling of short and long run effects; ▪ Models are nested and it is possible to test for rank; ▪ A combination of I(1) and I(0) variables can be included; ▪ The model needs to describe variation in the data which may prove problematic.

[2] It is important to note that the Johansen approach allows for long run effects and short term adjustment patterns to be modelled simultaneously. The sequential manner in which the technique is presented here is merely for convenience.

5.2.2. *Johansen Cointegration Analysis* [3]

The Johansen approach to the estimation of full system models containing cointegrated variables is based on estimating a vector autoregressive model using maximum likelihood. The use of full information maximum likelihood (FIML) to estimate the linear space spanned by the cointegrating vectors is advantageous for a number of reasons. Firstly, it overcomes the need to impose arbitrary normalisations to uniquely define cointegrating vectors, and secondly, it is possible to test for rank as opposed to forming ex-ante assumptions regarding the number of cointegrating relations (Hamilton, 1994).

As the starting point, consider an unrestricted vector autoregressive model (VAR) expressed in levels:

$$X_t = \Pi_1 X_{t-1} + \ldots. + \Pi_k X_{t-k} + \Phi D_t + \epsilon_t \qquad (t = 1, \ldots., T) \qquad (5.1)$$

where the errors ϵ_t are independent Gaussian with mean zero and variance Ω. The deterministic term D_t is a vector of non-stochastic variables and can contain a constant, a linear trend, seasonal dummies, intervention dummies and so on, as desired. The initial values $X_{-k+1}, \ldots., X_0$ are fixed while the parameters $\Pi_1, \ldots., \Pi_k$, Φ and Ω are unrestricted. The individual variables in X_t are assumed to be integrated of order one at most, but could be stationary.

Now define a stacked vector Z_t of dimension pk+m as $Z'_t = (X'_{t-1}, \ldots., X'_{t-k}, D'_t)$ and a matrix B' of dimension p x (pk+m) as $B' = (\Pi_1, \ldots., \Pi_k, \Phi)$ for the corresponding parameters. On the basis of this stacking, model (5.1) may be rewritten as:

$$X_t = B'Z_t + \epsilon_t \qquad (t = 1, \ldots., T) \qquad (5.2)$$

In estimating the unrestricted VAR, the ordinary least squares estimators are derived as

[3] See Johansen (1995) for a more thorough treatment of the models and concepts presented in this section of the chapter.

maximum likelihood estimators, with the Gaussian errors facilitating an analysis of the log likelihood function as follows:

$$\log L(B,\Omega) = - \tfrac{1}{2}\, T\log(2\Pi) - \tfrac{1}{2}\, T\log|\Omega| - \tfrac{1}{2} \sum_{t=1}^{T} (X_t - B'Z_t)'\, \Omega^{-1} (X_t - B'Z_t) \tag{5.3}$$

In turn, this analysis lends itself to the formulation of equations for estimating B and the regression estimators $\hat{B}$ and $\hat{\Omega}$. These are presented below in expressions (5.4), (5.5) and (5.6) respectively. The maximal value (excluding the constant term) is given by (5.7).

$$\sum_{t=1}^{T} X_t Z'_t = \hat{B} \sum_{t=1}^{T} Z_t Z'_t \tag{5.4}$$

$$\hat{B} = \left(\sum_{t=1}^{T} Z_t Z'_t \right)^{-1} \left(\sum_{t=1}^{T} Z_t X'_t \right) = S_{zz}^{-1} S_{zx} \tag{5.5}$$

$$\hat{\Omega} = T^{-1} \sum_{t=1}^{T} (X_t - \hat{B} Z_t)(X_t - \hat{B} Z_t)' = S_{xx.z} = S_{xx} - S_{xz} S_{zz}^{-1} S_{zx} \tag{5.6}$$

$$L_{\max}^{-2/T} = |\hat{\Omega}| \tag{5.7}$$

As inference is based on asymptotic results, the asymptotic properties of the estimators are detailed in the appendix to this chapter for the case of a stationary process.

While the above unrestricted vector autoregressive model with k lags is easily fitted, the validity of the derived procedures requires the satisfaction of a number of assumptions; namely that the linear conditional mean is explained by past observations and the deterministic trends, the conditional variance is constant and the errors are independent and normal. Though not all of these assumptions are crucial, it is important to test for model misspecification. As such, the statistical formulae underlying the tests applied for this purpose in Chapter 6, along with those relating to model selection, are presented in Table 5.2[4].

---

[4] Misspecification testing was carried out using CATS 1.0 and Givewin 2.20. For further details as to how these tests are implemented in the above software programmes, refer respectively to Hansen and

Table 5.2: VAR Model Selection and Misspecification Tests – Statistical Formulae

Residual Correlations:

Correlation Coefficient

$$\hat{\rho}_{ij} = \frac{\hat{\sigma}_{ij}}{\sqrt{\hat{\sigma}_{ii}\hat{\sigma}_{jj}}} \quad (i = 1, \ldots, p)$$

Information Criteria:

Likelihood Ratio Test for Model Reduction

$$-2\ln Q(H_k/H_{k+1}) = T(\ln|\hat{\Omega}_k| - \ln|\hat{\Omega}_{k+1}|)$$

H_k is the null hypothesis that k lags are appropriate;
H_{k+1} is the alternative that the VAR needs k+1lags;
-2lnQ is asymptotically distributed as chi squared with p squared degrees of freedom.

Schwartz Information Criteria (SIC)

$$\ln|\hat{\Omega}| + (p^2 k)\frac{\ln T}{T}$$

Hannan-Quinn (HQ)

$$\ln|\hat{\Omega}| + (p^2 k)$$

The above tests are based on the maximal value of the likelihood function. They each impose a different penalty factor associated with the increase in the model's parameters as a result of adding more lags. The decision criteria involves calculating the tests for various lag lengths k and choosing the value of k that corresponds to the smallest test statistic.

Residual Autocorrelation:

LM Test

$$LM(j) = -(T-p(k+1)-½)\ln\left(\frac{|\tilde{\Omega}(j)|}{|\hat{\Omega}|}\right)$$

$$(j = 1,4)$$

The LM test for first and fourth order autocorrelation is calculated using an auxiliary regression as outlined in Godfrey (1988). Wilks' ratio test with a small sample correction is then applied (Anderson, 1984 and Rao, 1973);
The test is asymptotically distributed as chi squared with p degrees of freedom.

AR Test

See Table 5A.2 for a description of this test.

Juselius (1995) and Doornik and Hendry (2001). Table 5.2 reports additional references when appropriate.

Table 5.2 (contd): VAR Model Selection and Misspecification Tests – Statistical Formulae

Residual Heteroscedasticity:

ARCH Test

$$(T-k)\, R^2$$

where R^2 is taken from the auxiliary regression:

$$\hat{\varepsilon}^2_{it} = \gamma_0 + \sum_{j=1}^{k} \gamma_j\, \hat{\varepsilon}^2_{i,t-j} + \text{error}$$

Normality Tests:

Multivariate Test

$$\sum_{i=1}^{p} (\tau^2_{1i} + \tau^2_{2i})$$

The test is approximately chi squared distributed with 2p degrees of freedom and includes a small sample correction;
See Doornik and Hansen (1994) and Shenton and Bowman (1977) for further details.

Univariate Jarque-Bera Test

$$T(\text{skewness})^2 / 6 + T(\text{kurtosis}-3)^2 / 24$$

This test is asymptotically chi squared distributed with 2 degrees of freedom.

where:

$$\text{Skewness}\,(\hat{\varepsilon}_{i,t}) = T^{-1} \sum_{t=1}^{T} \left(\frac{\hat{\varepsilon}_{i,t}}{\hat{\sigma}_i} \right)^3$$

$$\text{Kurtosis}\,(\hat{\varepsilon}_{i,t}) = T^{-1} \sum_{t=1}^{T} \left(\frac{\hat{\varepsilon}_{i,t}}{\hat{\sigma}_i} \right)^4$$

Parameter Constancy:

Trace Test

$$-2\log Q(H(r)|H(p)) = -T \sum_{i=r+1}^{p} \log(1-\hat{\lambda}_i)$$

Refer to equation (5.26) for further details.

Test of Constancy of Beta

$$-2\ln(Q(H_{\beta\tau} | \hat{\beta}_{(t_1)})) =$$

$$t_1 \sum_{i=1}^{r} \left(\ln(1-\hat{\rho}_{i,(t_1)}) - \ln(1-\hat{\lambda}_{i,(t_1)}) \right)$$

$$t_1 = (T_0, \ldots, T)$$

The hypothesis tested is $H\beta\tau$: $\hat{\beta} \in sp(\beta_{(t_1)})$, $t_1 = (T_0, \ldots, T)$, in which $\hat{\beta}$ is a known matrix;

where $\hat{\rho}_{i,(t_1)}$ are the solutions of:

Table 5.2 (contd): VAR Model Selection and Misspecification Tests – Statistical Formulae

The test statistic is asymptotically distributed as chi squared with (p1-r)r degrees of freedom (Hansen and Johansen, 1993).	$\lvert \rho \hat{\beta}' S_{11,(t1)} \hat{\beta} - \hat{\beta}' S_{10,(t1)} S^{-1}_{00,(t1)} S_{01,(t1)} \hat{\beta} \rvert = 0$
	and $\hat{\lambda}_{i(\tau)}$ are the r largest eigenvalues in the unrestricted eigenvalue problem:
	$\lvert \lambda S_{11,(t1)} - S_{10,(t1)} S^{-1}_{00,(t1)} S_{01,(t1)} \rvert = 0$
Log Likelihood Value	$-2/t_1 \ln(\ell(r)) =$
	$\left(\ln \lvert S_{00,}(t_1) \rvert + \sum_{i=1}^{r} \ln(1 - \hat{\lambda}_{i,}(t_1)) \right)$
	$t_1 = (T_{0,} \ldots, T)$
	with the 95% confidence bound calculated as follows:
	$\pm 2 \sqrt{2p / t_1}$

The next step in the analysis is to move from consideration of the unrestricted VAR model to consideration of the cointegrated version of such a model. However, before doing so, some basic definitions and concepts are needed[5].

Definition 5.1: A stochastic process X_t is said to be integrated of order d, I(d), d = 1,2, …. if $\Delta^d(X_t - E(X_t))$ is I(0).

Definition 5.2: Let X_t be an I(1) process. Then X_t is cointegrated with the cointegrating vector $\beta_i \neq 0$ if $\beta' X_t$ can be made stationary by a suitable choice of its initial distribution. The cointegrating rank refers to the number of linearly independent cointegrating relations while the space spanned by these relations is known as the cointegrating space.

[5] See Johansen (1995) for examples and a more detailed discussion of these definitions.

Definition 5.3: The I(d) process X_t is said to be cointegrated CI(d,b) with cointegrating vector $\beta_i \neq 0$ if $\beta' X_t$ is I(d-b) where b = 1,...., d and d = 1,....

Using the above definitions as building blocks, consider again the p dimensional process X_t set out in equation (5.1). Reparametrising the model in so-called error correction form gives the following expression:

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \epsilon_t \qquad (t = 1, ..., T) \qquad (5.8)$$

where $\Pi = \sum_{i=1}^{k} \Pi_i - I$, with the short-term dynamics Γ_i and Γ defined as $\Gamma_i = -\sum_{j=i+1}^{k} \Pi_j$ and $\Gamma = I - \sum_{i=1}^{k-1} \Gamma_I$ (Johansen, 1995). Furthermore, the characteristic polynominal for the process X_t in (5.8) may be written as[6]:

$$A(z) = (1-z)I - \Pi z - \sum_{i=1}^{k-1} \Gamma_i (1-z)z^i \qquad (5.9)$$

In describing a non-stationary I(1) process, that is allowing for unit roots, $A(1) = -\Pi$ must be singular. Such a singular matrix of rank r has the representation $\Pi = \alpha\beta'$ for some p x r matrices α and β. For illustrative purposes, Johansen (1995) suggests letting β' denote r linearly independent rows of Π, with α denoting the coefficients that are necessary in order to express the rows of Π as linear combinations of the rows β'. The rank condition $\Pi = \alpha\beta'$ then becomes clear. Decomposing the Π matrix within the autoregressive framework, β determines the common long run relations while α reflects the loadings. More formally, the 'Granger Representation Theorem' states that:

[6] The reciprocal values of the roots of the characteristic polynominal are referred to as the eigenvalues of the companion matrix. The cointegrated VAR model discussed here requires that these eigenvalues lie inside or on the unit circle. Values outside the unit circle correspond to explosive processes and in such cases, the I(1) model provides a poor description of the data.

Theorem 5.1: (Engle and Granger, 1987) If $|A(z)| = 0$ implies that $|z| > 1$ or $z = 1$, and Π has rank $r < p$, then there exist $p \times r$ matrices α and β of rank r such that:

$$\Pi = \alpha\beta' \tag{5.10}$$

A necessary and sufficient condition for ΔX_t -E(ΔX_t) and $\beta' X_t$ -E($\beta' X_t$) to be given initial distributions so as to induce stationarity is:

$$|\alpha'_{\perp}\Gamma\beta_{\perp}| \neq 0 \tag{5.11}$$

or, alternatively, that there are exactly p–r unit roots[7]. In either case, the solution X_t to equation (5.8) has the following representation:

$$X_t = C\sum_{i=1}^{t}(\in_i + \Phi D_i) + C_1(L)(\in_t + \Phi D_t) + A \tag{5.12}$$

where A depends on initial conditions such that $\beta' A = 0$ and $C = \beta_{\perp}(\alpha'_{\perp}\Gamma\beta_{\perp})^{-1}\alpha'_{\perp}$. Thus, it follows that X_t is an I(1) process with cointegrating vectors β. In addition, note that $C_1(z)$ is convergent for $|z| < 1+\delta$ for some $\delta > 0$.

The importance of Theorem 5.1 lies in the fact that it shows the equivalence between the notions of cointegration and error correction models. From (5.12), it can be seen that $\beta' X_t$-E($\beta' X_t$) is stationary as $\beta' C = 0$, while $\beta' C_1(L)(\in_t + \Phi D_t)$ may be viewed as a representation of the equilibrium error $\beta' X_t$. The C matrix also makes a valuable contribution towards the understanding of I(1) models and is interpretable as being indicative of the manner in which the common trends $\alpha'_{\perp}\Sigma_{i=1}^{t}\varepsilon_i$ influence the various variables through the matrix $\beta_{\perp}$.

In analysing the structure of these common trends, the vector moving average representation, as opposed to the vector autoregressive model, is used. There is, however, complete duality between the two. For example, the decomposition of the matrix C (as presented in Theorem

[7] If β is any $p \times r$ matrix of full rank, $\beta_{\perp}$ can be defined as a $p \times (p\text{-}r)$ matrix of full rank, such that $\beta'\beta_{\perp} = 0$.

5.1) is similar to that of the Π matrix. If the former is expressed as $C = \tilde{\beta}_\perp \alpha'_\perp$ with $\tilde{\beta}_\perp = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}$, $\alpha'_\perp$ determines the common stochastic trends and $\tilde{\beta}_\perp$ denotes the loadings to these trends, whereas for the reduced rank condition $\Pi = \alpha\beta'$ in the AR representation, β determines the common long run relations while α reflects the loadings. Thus, taken together the above definitions of the C matrix show that 'non-stationarity in the process X_t originates from the cumulative sum of the p-r combinations $\alpha'_\perp \Sigma^{t}_{i=1}$' (Juselius, forthcoming).

For a more intuitive understanding of the above discussion, briefly consider a simple reduced form error correction model[8]:

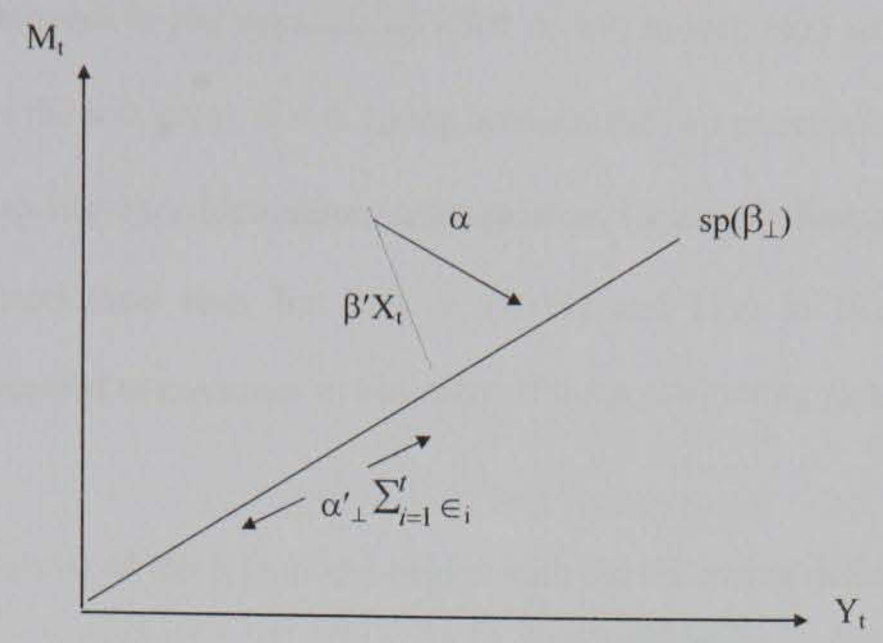
$$\Delta X_t = \alpha\beta' X_{t-1} + \mu + \epsilon_t \qquad (t = 1, \ldots, T) \tag{5.13}$$

where the initial value is X_0 and as before, α and β are p x r matrices. Note also that in this model short run dynamics have been excluded. Defining the underlying economic relations as $\beta' X_t - E(\beta' X_t) = 0$, equation (5.13) illustrates how the process is pulled towards the steady state with the α coefficient capturing the reaction of agents to any disequilibrium error $\beta' X_t$.

The common trends representation on the other hand, shows how the cumulated disturbances $\alpha'_\perp \Sigma^{t}_{i=1}$ push the economic variables around in the space spanned by $\beta_\perp$, the attractor set, which is in turn defined by the long run relations. The two concepts of error correction and common trends are of course complementary and respectively represent the pulling and pushing forces of the system. Figure 5.1 provides a graphical illustration of these forces at work.

[8] This example draws primarily on Juselius (forthcoming). A similar discussion is also presented in Johansen (1995).

Figure 5.1: Pulling and Pushing Forces



In the diagram, a bivariate system is depicted with the process X_t comprising M_t and Y_t Assume that the steady state position is given by M-Y = μ. In this case, $\beta' = [1,-1]$ and the attractor set may be written as $\beta_\perp = [1,1]$. The 45° line captures the notion of such a set and shows that $M_t = Y_t$ in the steady state. If there is any deviation from this long run equilibrium relationship, the adjustment coefficients α will become active pulling the process back towards the attractor set, with the speed of adjustment depending on the size of both α and the disequilibrium error $\beta' X_t$. Movement up and down the attractor set is dictated by the nature of the cumulated shocks or common trends. Positive shocks will push the process higher along $\beta_\perp$ whereas negative shocks will move it downwards.

The above example thus concludes a relatively detailed introduction to the concepts of integration, cointegration and the cointegrated VAR. The remainder of this section presents the method of reduced rank regression that solves the estimation problem for the cointegrating vectors and facilitates the derivation of a test statistic for determining the rank.

As the first step, recall the reduced rank condition $\Pi = \alpha\beta'$ for some p x r matrices α and β. Secondly, define the I(1) model H(r) as the sub-model of the VAR obtained under the above condition. Note that the I(1) model forms part of a nested sequence of models:

$$H(0) \subset \ldots. H(r) \subset \ldots.H(p) \quad (5.14)$$

where H(p) corresponds to the unrestricted VAR or I(0) model, H(r) sets the rank of $\Pi \leq r$ and H(0) imposes the restriction $\Pi = 0$. Lying between the two extremes of (5.14) are a range of models, H(1) to H(p-1), which ensure cointegration. Given this formulation, it is possible to derive likelihood ratio tests for H(r) in H(r+1) and H(r) in H(p). These tests can subsequently be applied to construct an estimator of the cointegrating rank.

The statistical analysis of the I(1) model begins with the following definition of the reduced form error correction model:

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-1} + \Phi D_t + \epsilon_t \qquad (t = 1, \ldots., T) \quad (5.15)$$

where ϵ_t are independent $N_p(0, \Omega)$ and the parameters α, β, $\Gamma_1, \ldots., \Gamma_{k-1}$, Φ and Ω vary freely. As before, the β vectors are the cointegrating relations while α is the adjustment coefficient. Now define $Z_{0t} = \Delta X_t$, $Z_{1t} = X_{t-1}$, Z_{2t} = the stacked variables $\Delta X_{t-1}, \ldots., \Delta X_{t-k+1}$ and D_t, and let Ψ be the matrix of parameters $\Gamma_1, \ldots., \Gamma_{k-1}$ and Φ that correspond to Z_{2t} Expressing the model in terms of these variables gives the reduced rank regression:

$$Z_{0t} = \alpha\beta' Z_{1t} + \Psi Z_{2t} + \epsilon_t \qquad (t = 1, \ldots., T) \quad (5.16)$$

and the Gaussian log likelihood function (apart from a constant):

$$\log L(\Psi,\alpha,\beta,\Omega) = -\tfrac{1}{2} T\log|\Omega| - \tfrac{1}{2} \sum_{t=1}^{T} (Z_{0t}-\alpha\beta' Z_{1t}-\Psi Z_{2t})' \Omega^{-1} (Z_{0t}-\alpha\beta' Z_{1t}-\Psi Z_{2t}) \quad (5.17)$$

Concentrating equation (5.17) by regressing Z_{0t} and Z_{1t} on Z_{2t} leads to a definition of the residuals R_{0t} and R_{1t} and the cross moment matrices:

$$S_{ij} = T^{-1} \sum_{t=1}^{T} R_{it} R'_{jt} \qquad (i, j = 0,1) \quad (5.18)$$

The associated concentrated likelihood function also takes the form of a reduced rank regression (the parameters Ψ having been eliminated in the course of the regression) and may be written as[9]:

$$R_{0t} = \alpha\beta' R_{it} + \hat{\epsilon}_t \qquad (5.19)$$

Fixing β, it is now possible to estimate α and Ω by regressing R_{0t} on $\beta' R_{it}$. The subsequent estimates are presented below in expressions (5.20) and (5.21) respectively, with the maximal value given by (5.22).

$$\hat{\alpha}(\beta) = S_{01}\beta(\beta' S_{11}\beta)^{-1} \qquad (5.20)$$

$$\hat{\Omega}(\beta) = S_{00} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' S_{10} \qquad (5.21)$$

$$L_{\max}^{-2/T}(\beta) = |S_{00}||\beta'(S_{11} - S_{10}S_{00}^{-1}S_{01})\beta| \,/\, |\beta' S_{11}\beta)| \qquad (5.22)$$

To determine β, equation (5.22) is minimised by means of solving the eigenvalue problem:

$$|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}| = 0 \qquad (\lambda = 1\text{-p}) \qquad (5.23)$$

for the eigenvalues $1 > \hat{\lambda}_1 > \ldots. > \hat{\lambda}_p > 0$ and the corresponding eigenvectors $\hat{V} = (\hat{v}_1, \ldots., \hat{v}_p)$, where the latter are normalised by $\hat{V}' S_{11} \hat{V} = I$. Taking the eigenvectors for the r largest eigenvalues, the cointegrating relations are thus estimated by:

$$\hat{\beta} = (\hat{v}_1, \ldots., \hat{v}_r) \qquad (5.24)$$

Normalising $\hat{\beta}$ on S_{11} (so that $\hat{\beta}' S_{11} \hat{\beta} = I$ and $\hat{\beta}'(S_{11} - S_{10}S_{00}^{-1}S_{01})\hat{\beta} = \text{diag}(\hat{\lambda}_1, \ldots., \hat{\lambda}_r)$), the maximised likelihood function is subsequently found to be:

$$L_{\max}^{-2/T}(H(r)) = |S_{00}| \prod_{i=1}^{r} (1 - \hat{\lambda}_i) \qquad (5.25)$$

Finally, a comparison of two expressions like (5.25) leads to a definition of the likelihood ratio test statistic Q(H(r)|H(p)) for H(r) in H(p):

[9] Refer to Anderson (1971) for a further discussion.

$$-2\log Q(H(r)|H(p)) = -T\sum_{i=r+1}^{p}\log(1-\hat{\lambda}_i) \qquad (5.26)$$

This LR test statistic (also known as the Trace statistic) determines the cointegrating rank. Under I(1) assumptions and setting the rank equal to r, the asymptotic distribution of (5.26) is a function of a Brownian motion of dimension p-r. The precise distribution depends on the type of deterministic trends present[10].

Referring back to equation (5.15) for a moment, note that the parameters α and β are not uniquely identified in the sense that 'given any choice of α and β and any non-singular matrix ξ(r x r), the choice $\alpha\xi$ and $\beta(\xi')^{-1}$ will give the same matrix Π and hence determine the same probability distributions for the variable' (Johansen, 1995). Given this outcome, it is necessary to identify β before its individual coefficients can be estimated. Typical methods of doing so include the imposition of zero restrictions and normalisations or alternatively, different restrictions can be imposed on the individual relations, as can cross equation restrictions. However, attention is also drawn to the fact that some hypotheses on the parameters α and β can be tested without necessarily identifying the system. Table 5.3 concludes this section by providing an overview of the latter[11].

[10] Johansen (1988, 1991 and 1994) mathematically derives asymptotic distributions for models with different types of deterministic components. Also, see Johansen and Nielsen (1993).

[11] These model specific data properties are empirically tested in Chapter 6 of the study using the CATS 1.0 software programme. Further details of the precise formulation of the hypotheses can be found in the accompanying manual (Hansen and Juselius, 1995).

Table 5.3: Model Specific Data Properties – Hypothesis Testing

A general linear hypothesis for β may be expressed as:

$$H_\beta:\ \beta = (H_1\varphi_1, H_2\varphi_2, \ldots, H_r\varphi_r)$$

where H_i is p x s_i and known, the parameter to be estimated φ_i is s_i x 1 and $1 \leq s_i \leq p$. All hypotheses concerning β are nested in this formulation with H_β specifying structural restrictions on each of the cointegrating vectors.

Now consider some special cases of the above:

(1) The same p-s restrictions can be imposed on all r cointegrating vectors by applying the hypothesis:

$$H_1:\ \beta = H\varphi$$

where H is p x s and φ is s x r. This condition may also be written geometrically as sp(β) ⊂ sp(H), thus restricting the subspace to lie in the given subspace sp(H) of R^p.

Example 1: *Testing for long run exclusion*

Assume that $X_t = (y_t, x_{1t}, x_{2t})'$. The hypothesis that say x_{2t} can be omitted from the cointegrating space can be formulated as H_1 with:

$$H = \begin{bmatrix} 1\,0 \\ 0\,1 \\ 0\,0 \end{bmatrix}$$

(2) The following hypothesis completely specifies the coefficients of the first cointegration relationship while allowing the remaining r_2 vectors to vary freely:

$$H_2:\ \beta = (H, \varphi)$$

where H is p x r_1, φ is p x r_2 and $r_1 + r_2 = r$.

Example 2: *Testing for stationarity*

Assume again that $X_t = (y_t, x_{1t}, x_{2t})'$. H_2 can be combined with H (as given below) in order to investigate if the variable x_{2t} is stationary:

$$H = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

As with restrictions on β, hypotheses can also be formed with respect to the adjustment coefficient α; the most common being a hypothesis of long run weak exogeneity which is formulated as:

$$\alpha = A\phi$$

where A is p x m and known and ϕ is the m x r parameter to be estimated. Geometrically, sp(α) =sp(A).

Example 3: *Testing for long run weak exogeneity*

As in the previous examples, let $X_t = (y_t, x_{1t}, x_{2t})'$. In addition assume a cointegrating rank of one. By imposing restrictions on α it is possible to test if say x_{1t} and x_{2t} are weakly exogeneous; that is, while these variables may influence the long run stochastic path of other variables in the system, they are not influenced by them. This hypothesis is expressed as:

$$A = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

5.3. Smooth Transition Regression Methodology

While linear regression models often constitute a valid approximation to underlying non-linear behaviour, in certain cases it may hold that an explicitly non-linear specification is capable of better describing such phenomena. Reflecting this, an array of non-linear time series models, loosely grouped as those that do not nest a linear model and those that do, have been proposed in recent years[12]. Popular examples of the latter include the switching regression model originally introduced by Quandt (1958), and the smooth transition regression model developed by Bacon and Watts (1971) as a generalisation of a particular switching regression specification. Both of the above approaches attempt to characterise the theoretical notion 'that the economy behaves differently if some variable lies in one region rather than in another'; the main distinction being that the smooth transition regression model allows for a smooth transition between regimes rather than a discrete, sharp switch (Granger and Teräsvirta , 1993).

With regard to the Dublin housing market, a prompt and uniform response on the part of all economic agents to news requiring action is unlikely given the range of individuals and firms at play in this marketplace. From an aggregate viewpoint, this implies that a smooth model is more appropriate in a housing context than the discrete switching alternative. As such, the non-linear dynamic model estimated in Chapter 7 applies the smooth transition regression technique. Drawing on Granger and Teräsvirta (1993) and Teräsvirta (1998), the following discussion provides an overview of this methodology.

As the starting point, consider the smooth transition regression model (STR):

$$y_t = x_t' \varphi + (x_t' \theta) G(\gamma, c; s_t) + u_t \qquad (t = 1, \ldots, T) \qquad (5.27)$$

[12] See Granger and Teräsvirta (1993) for a detailed treatment of non-linear modelling techniques.

where $x_t = (1, x_{1t}, ..., x_{pt})' = (1, y_{t-1}, ..., y_{t-k}; z_{1t}, ..., z_{mt})'$, $p = k+m$, represents a vector of explanatory variables, $\varphi = (\varphi_0, \varphi_1, ..., \varphi_p)'$ and $\theta = (\theta_0, \theta_1, ..., \theta_p)'$ are parameter vectors and u_t is a sequence of error terms which are independent and identically distributed. A priori, some of the φ_i and θ_j parameters can be set equal to zero and the restriction that $\varphi_i = -\theta_j$ may also be imposed. The transition function $G(\gamma, c; s_t)$, which is continuous and typically bounded between zero and one, comprises the continuous transition variable s_t (a stochastic variable that is often an element of x_t, though it could be a linear combination of several variables or a linear time trend), the slope parameter γ (indicating the smoothness of the regime change) and the location parameter c (identifying the point at which transition occurs).

In terms of defining G, a transition function that takes the logistic form and monotonically increases as a function of s_t is labelled an LSTR1 model:

$$G_1(\gamma, c; s_t) = (1 + \exp\{-\gamma(s_t - c)\})^{-1} \qquad (\gamma > 0) \qquad (5.28)$$

where $\gamma > 0$ is an identifying restriction. A simple non-monotonic alternative is the LSTR2 model with the restrictions on γ, c_1 and c_2 acting as identifying restrictions in this case:

$$G_2(\gamma, c; s_t) = (1 + \exp\{-\gamma(s_t - c_1)(s_t - c_2)\})^{-1} \quad (\gamma > 0, c_1 \leq c_2) \qquad (5.29)$$

In addition, an exponential ESTR model may be defined as follows:

$$G(\gamma, c; s_t) = 1 - \exp\{-\gamma(s_t - c)^2\} \qquad (\gamma > 0) \qquad (5.30)$$

Note that when $\gamma = 0$ and $G(\gamma, c; s_t) \equiv 1/2$, the STR model presented in equation (5.27) nests the linear model whereas when $\gamma \to \infty$, the LSTR1 model approaches a switching regression model with two regimes while the LSTR2 model becomes a three regime switching regression model in which the outer two regimes are equal.

With respect to inference in the smooth transition regression model, hypothesis testing requires that additional assumptions be made concerning equation (5.27). Firstly, the individual stochastic variables among $z_{1t}, ..., z_{kt}$ are assumed to be stationary, with the non-stochastic ones representing dummy variables. It may also hold that some of the variables in

z_{jt} are stationary linear combinations of I(1) variables[13]. Secondly, it is assumed that all cross moments exist and that the error terms are uncorrelated with x_t and s_t.

Taking the above conditions as given, inference in the STR model necessarily begins with linearity testing. Following Teräsvirta (1998), rewrite equation (5.27) as:

$$y_t = x_t'\varphi + (x_t'\theta)G^*(\gamma, c; s_t) + u_t \qquad (\gamma > 0) \qquad (5.31)$$

where $G^* = G-1/2$, with G referring to any of the transition functions defined in equations (5.28) to (5.30), and $u_t \sim nid(0,\sigma^2)$. The conditional log likelihood function of the model is given by:

$$\sum_{t=1}^{T} \ell(\varphi,\theta,\gamma,c; y_t \mid x_t, s_t) = \alpha - \frac{T}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T} u_t^2 \qquad (5.32)$$

The null hypothesis of linearity, expressed as H_0: $\gamma = 0$ in equation (5.31), is then tested against the alternative of STR type non-linearity, H_A: $\gamma > 0$. However, the fact that the null could equally be formulated as H_1: $\theta = 0$ implies an identification issue as the model is identified under the alternative, but not under the null hypothesis. Consequently, applying a likelihood ratio test in order to test H_0 is problematic given that θ and c are now nuisance parameters whose values do not affect the likelihood value. As such, the parameters of equation (5.31) cannot be consistently estimated under H_0 and the likelihood ratio and Lagrange multiplier statistics no longer have their standard asymptotic chi squared distributions.

The property of nuisance parameters is not exclusive to smooth transition regression modelling; many other non-linear models nesting a linear model share this same property[14]. However in the STR case, the identification problem can be circumvented by approximating the transition function $G^* = G-1/2$ in equation (5.31) by a Taylor expansion around the null

---

[13] The super consistency of the least squares estimates of these parameters means that they can be treated as known (Teräsvirta , 1998).

[14] See Hansen (1996) for a general discussion.

hypothesis $\gamma = 0$[15]. Taking s_t to be an element of x_t and applying a third order approximation yields the auxiliary regression:

$$y_t = x_t' \beta_0 + (\tilde{x}_t s_t)'\beta_1 + (\tilde{x}_t s_t^2)'\beta_2 + (\tilde{x}_t s_t^3)'\beta_3 + u_t^* \quad (5.33)$$

where $u_t^* = u_t + (x_t' \theta) R_3(\gamma, c; s_t)$ with remainder $R_3(\gamma, c; s_t)$, and $\beta_j = \gamma \tilde{\beta}_j$, j = 1, 2, 3. The null hypothesis H_0: $\gamma = 0$ in equation (5.31) now becomes H_0': $\beta_j = 0$, j = 1, 2, 3, in equation (5.33) with H_A' : at least one $\beta_j \neq 0$. Moreover, given that equation (5.33) is linear in parameters and that $u_t^* = u_t$ under the null, an LM-type test can be applied to test H_0':

$$\chi_{LM}^2 = \hat{\sigma}^{-2}\left(\sum_{t=1}^{T} \hat{u}_t w_t\right)' (\hat{M}_{11} - \hat{M}_{10}\hat{M}_{00}^{-1}\hat{M}_{01})^{-1}\left(\sum_{t=1}^{T} w_t \hat{u}_t\right) \quad (5.34)$$

where $\hat{M}_{00} = \Sigma_{t=1}^{T} z_t z_t'$, $\hat{M}_{01} = \hat{M}_{10}' = \Sigma_{t=1}^{T} z_t w_t'$, $\hat{M}_{11} = \Sigma_{t=1}^{T} w_t w_t'$, $\hat{\sigma}_2 = (1/T)\Sigma_{t=1}^{T} \hat{u}_t^2$, $\hat{u}_t$ is the error term estimated under the null hypothesis, $z_t = x_t$ and $w_t = (\tilde{x}_t' s_t, \tilde{x}_t' s_t^2, \tilde{x}_t' s_t^3)'$. When the moments implied by equation (5.34) and w_t exist, the LM statistic has an asymptotic chi squared distribution with 3p degrees of freedom.

Unfortunately, the above chi squared statistic is likely to be severely size distorted in small samples. Given this, Granger and Teräsvirta (1993) recommend an F approximation:

$$F = \frac{(SSR_0 - SSR_1)/3p}{SSR_1/(T - 4p - 1)} \quad (5.35)$$

where the residual sum of squares $SSR_0 = (1/T)\Sigma_{t=1}^{T} \hat{u}^2$ and $SSR_1 = (1/T)\Sigma_{t=1}^{T} \hat{v}_t^2$ are respectively obtained from a regression of y_t on x_t, and a regression of either $\hat{u}_t$ or y_t on x_t, $\tilde{x}_t s_t$, $\tilde{x}_t s_t^2$ and $\tilde{x}_t s_t^3$. Under H_0': $\beta_1 = \beta_2 = \beta_3 = 0$, the corresponding F statistic has an approximate F distribution with 3p and T-4p-1 degrees of freedom.

---

[15] This method of circumvention originates from the work of Saikkonen and Luukkonen (1988) and Luukkonen, Saikkonen and Teräsvirta (1988).

In terms of selecting an appropriate variable to act as the transition variable under the alternative, it may arise that the choice is obvious from economic theory. If not, Luukkonen, Saikkonen and Teräsvirta (1988) suggest choosing from all the elements of $\tilde{x}_t$ as follows. Firstly, define the linear combination $a'\tilde{x}_t$ such that a is a px1 vector = (0, …., 1, 0, 0, …., 0)′, with the unit element corresponding to the true but unknown transition variable. Secondly, substituting this for s_t in equation (5.28) leads to the auxiliary regression:

$$y_t = x_t'\beta_0 + \sum_{i=1}^{p}\sum_{j=i}^{p}\beta_{1ij}x_{it}x_{jt} + \sum_{i=1}^{p}\sum_{j=1}^{p}\beta_{2ij}\tilde{x}_{it}\tilde{x}_{jt}^2 + \sum_{i=1}^{p}\sum_{j=1}^{p}\beta_{3ij}\tilde{x}_{it}\tilde{x}_{jt}^3 + u_t^* \qquad (5.36)$$

and a re-formulated null of linearity, namely H_0': $\beta_{1ij} = 0$ where i = 1, …., p, j = i, i+1, …., p and $\beta_{2ij} = \beta_{3ij} = 0$, i, j = 1, …., p. As before, an LM-type test statistic, which in this case is asymptotically distributed under H_0' as chi squared with $p(p+1)/2+2p^2$ degrees of freedom, can then be derived[16].

Thus, the statistical theory discussed so far results in standard asymptotic distributions about the null and testing procedures that can be undertaken using the ordinary least squares method. If these tests subsequently reject the hypothesis of linearity in favour of a STR model with transition variable x_{jt}, the next step in the specification process is the selection of a particular type of smooth transition regression model. In making this selection, the following test sequence is applicable:

$$H_{04}: \beta_3 = 0$$

$$H_{03}: \beta_2 = 0 \mid \beta_3 = 0$$

$$H_{02}: \beta_1 = 0 \mid \beta_2 = \beta_3 = 0$$

where H_{04}, H_{03} and H_{02} represent a series of null hypotheses within equation (5.33). The associated decision rule is based on the fact that the coefficient vectors β_j, j = 1, 2, 3, are functions of the parameters of the original STR specification and depend on the type of

[16] Note that as p increases, the test statistic also tends to grow rapidly. As such, Luukkonen, Saikkonen and Teräsvirta (1988) propose an 'economy version' of equation (5.36). Please refer to this paper for further details.

model. Accordingly, if the test of H_{03} yields the strongest rejection when measured in p values, either an ESTR or LSTR2 model should be selected; otherwise, an LSTR1 model is chosen [17].

Having decided on a model type, estimation of the parameters of the STR model is carried out using conditional maximum likelihood. The algorithm used for the purpose of numerically maximising the log likelihood is dependent on good starting values, which can be obtained by fixing γ and c in the transition function so that the initial STR model, as presented in equation (5.27), is linear in parameters[18]. This in turn, suggests the construction of a grid and estimation of the remaining parameters φ and θ conditional on (γ, c) in the LSTR1 case, and on (γ, c_1, c_2) for the LSTR2 model. The residual sum of squares is then computed for N combinations of these parameters, with the parameter values that minimise this sum being taken as the starting values. However, note that as γ is not a scale-free parameter, effective grid construction requires the standardisation of the exponent of the transition function, by dividing it by the K^{th} power of the sample standard deviation of the transition variable.

Finally, as with linear modelling, it is important to evaluate the quality of the estimated STR model and to check the residuals for evidence of misspecification. In this respect, the goodness of fit measures and information criteria outlined in Table 5.4 are applicable. Furthermore, Teräsvirta (2004) discusses generalisations of standard misspecification tests to the STR framework, including those of no error autocorrelation, parameter constancy and no additive non-linearity. Given that these tests dominate the diagnostic checks implemented in Chapter 7, the remainder of this section is concerned with detailing their statistical formulae[19].

---

[17] See Teräsvirta (1994) for a simulation of this testing procedure.

[18] The software package JMulTi (which is used for the empirical STR estimations in Chapter 7) applies the iterative BFGS algorithm with numerical derivatives. Further details are provided in Hendry (1995).

[19] Misspecification testing was carried out using JMulTi 3.11. For further details as to how these tests are implemented in this software programme, refer to the package 'Help' system.

Table 5.4: STR Model Misspecification Tests – Statistical Formulae

Residual Autocorrelation:

AR Test

This test, which is a variant of Godfrey's LM test, is approximately distributed as F(q, T-n-q);
As the LM asymptotic chi squared statistic can be severely size distorted in small samples, the empirical analysis in Chapter 7 reports the F-form of the test;
Refer to Teräsvirta (1998) for a detailed discussion.

$$\frac{(SSR_0 - SSR_1)/q}{SSR_1/(T-n-q)}$$

where:

SSR_0 = the sum of squared residuals from the STR model

And

SSR_1 = the sum of squared residuals from the auxiliary regression of

$\tilde{u}_t$ on $\tilde{u}_{t-1,\ldots,}\tilde{u}_{t-q}$ and the

partial derivatives of the log likelihood function with respect to the parameters of the model

---

Residual Heteroscedasticity:

ARCH Test

The ARCH LM test is approximately distributed as chi squared with q degrees of freedom;
As above, the empirical analysis reports an F version of this test so as to counteract small sample distortionary effects.

TR^2

where R^2 is taken from the auxiliary regression:

$$\hat{u}_t^2 = \beta_0 + \beta_1\hat{u}_{t-1}^2 + \ldots + \beta_q\hat{u}_{t-q}^2 + \text{error}_t$$

Normality Test:

The associated test statistic is that of the Jarque-Bera test described in Table 5.2.

Parameter Constancy:

Again, the F version of the test for parameter constancy is reported in Chapter 7. The associated degrees of freedom are $3(p_0+p_1)$ and $T-4(p_0+p_1)$;
Further details are provided in Teräsvirta (1998).

Consider the STR model:

$$y_t = (x_t^0)'\varphi^0(t) + (x_t^1)'\theta^0(t)G(\gamma, c; s_t) + u_t$$

where:

$$\varphi^0(t) = \varphi^0 + \lambda_1 H(t;\gamma_1, c_1)$$

and

$$\theta^0(t) = \theta^0 + \lambda_2 H(\gamma_1, c_1; s_t);$$

Next define the transition functions H_j, j = 1, 2, 3:

$$H(\gamma_1, c_1; s_t) = (1+\exp\{-\gamma_1(t-c_1)\})^{-1} - 1/2$$

Table 5.4 (contd): STR Model Misspecification Tests – Statistical Formulae

$$H_2(\gamma_1, c_1; s_t) = (1+\exp\{-\gamma_1(t-c_{11})(t-c_{12})\})^{-1} - 1/2$$

$$H_3(\gamma_1, c_1; s_t) = (1+\exp\{-\gamma_1(t-c_{11})(t-c_{12})(t-c_{13})\})^{-1} - 1/2$$

where $\gamma_1 > 0$ and $c_{11} \le c_{12} \le c_{13}$;

Assuming that the parameters in the exponent of G are constant, the null hypothesis is then H_0: $\gamma_1 = 0$.

In testing parameter constancy against $\varphi^0(t)$ and $\theta^0(t)$, let $H = H_3$ (H_1 and H_2 are special cases of H_3) and apply a Taylor approximation to overcome the identification problem. This yields the auxiliary regression:

$$y_t = (x_t^0)'\beta_0 + (x_t^0 t)'\beta_1 + (x_t^0 t^2)'\beta_2 + (x_t^0 t^3)'\beta_3 + \{(x_t^1)'\beta_4 + (x_t^1 t)'\beta_5 + (x_t^1 t^2)'\beta_6 + (x_t^1 t^3)'\beta_7\}G(\gamma, c; s_t) + u_t^*$$

where:

$\beta_j = \gamma_1 \tilde{\beta}_j$, j = 1, 2, 3, 4, 5, 6, 7

and

$$u_t^* = u_t + \{(x_t^0)'\varphi^0 + (x_t^1)'\theta^0 G(\gamma_1, c; s_t)\}R_1(\gamma, c_1; t)$$

The null hypothesis of parameter constancy now becomes:

$H_0^{'}: \beta_1 = \beta_2 = \beta_3 = 0, \beta_5 = \beta_6 = \beta_7 = 0$

and the test statistic given in equation (5.34) has an asymptotic chi squared distribution under $H_0^{'}$

No Remaining Non-Linearity:

The empirical analysis in Chapter 7 reports the F form of the test for no remaining non-linearity. The test statistic is distributed with 3p and T-4p-1 degrees of freedom;
Refer to Teräsvirta (1998) for a detailed discussion.

Define an additive STR model as:

$$y_t = x_t'\varphi + (x_t'\theta)G(\gamma_1, c_1; s_t) + (x_t'\psi)H(\gamma_2, c_2; s_{2t}) + u_t$$

(t = 1, …., T)

where H is another transition function and s_{2t} is assumed to be an element of x_t;

Table 5.4 (contd): STR Model Misspecification Tests – Statistical Formulae

The corresponding null hypothesis is H_0: $\gamma_2 = 0$.

As the additive STR model is not identified under this null, a Taylor approximation is again applied. The resultant auxiliary regression is as follows:

$$y_t = x_t'\beta_0 + (x_t'\theta)G(\gamma_1, c_1; s_t)$$
$$+(\tilde{x}_t s_{2t})'\beta_1 + (\tilde{x}_t s_{2t}^2)'\beta_2 + (\tilde{x}_t s_{2t}^3)'\beta_3 + u_t^*$$

where:

$$\beta_j = \gamma_2 \tilde{\beta}_j \text{, j = 1, 2, 3}$$

and

$$u_t^* = u_t + (x_t'\psi)R_3(\gamma_2, c_2; s_{2t})$$

In this case, the null hypothesis of no remaining non-linearity is given by:

$$H_0' : \beta_1 = \beta_2 = \beta_3 = 0$$

and the associated test statistic is asymptotically chi squared distributed under H_0'

Notes:
1. In order for the asymptotic statistical theory underlying the tests of no error autocorrelation, constant parameters and no remaining non-linearity to be valid, the maximum likelihood estimators must be consistent and asymptotically normal. Further details are provided in Wooldridge (1994) and Escribano and Mira (1995)

5.4. Regime-Switching Methodology and Bubble Testing

A relatively new approach to testing for behaviour indicative of a bubble is to estimate a regime-switching model along the lines proposed by van Norden (1993) and van Norden and Schaller (1993a, 1993b). The move towards such models, away from the more traditional unit root and cointegration tests, has been largely motivated by the findings of Evans (1991) who used Monte Carlo simulations to show that the latter tests frequently reject the presence of a bubble, even when a bubble is present by construction. Developing this idea further, van Norden and Vigfusson (1996a) find that the finite sample properties of regime-switching tests for bubbles are also better than those of the alternatives. As such, Chapter 8 adopts a regime-switching approach when empirically testing for the presence of a bubble component in the Dublin housing market. An overview of this technique is presented below[20].

To begin, consider the following simplified asset-pricing model:

$$p_t = X_t + aE_t(p_{t+1}) \qquad (5.37)$$

where p_t is the price of the asset, X_t is a vector of explanatory variables, E_t is the expectations operator conditional on information available at time t and $0<a<1$. Solving equation (5.37) forward gives rise to the general result:

$$p_t = \left(\sum_{j=0}^{T} a^j . E_t(X_{t+j}) \right) + a^{T+1} E_t(p_{T+1}) \qquad (5.38)$$

to which there is more than one solution. For example, when it holds that:

$$\lim_{T \to \infty} a^{T+1} . E_t(p_{T+1}) = 0 \qquad (5.39)$$

the standard fundamentals solution applies:

$$p_t^F = \left(\sum_{j=0}^{\infty} a^j . E_t(X_{t+j}) \right) \qquad (5.40)$$

whereby the asset price is treated solely as a function of the current and expected behaviour of the other variables in the model. However, a bubbles solution may also exist. Defining the latter as any alternative set of actual and expected asset prices that satisfy equation (5.37), but where $p_t \neq p_t^F$, leads to an expression for the size of the bubble:

$$b_t \equiv p_t - p_t^F \qquad (5.41)$$

Furthermore, as p_t^F satisfies equation (5.37), it follows from equations (5.37) and (5.41) that the bubbles solution is of the form:

$$b_t = aE_t(b_{t+1}) \qquad (5.42)$$

where $a<1$ indicates the anticipated growth of the bubble over time.

---

[20] Further details are provided in van Norden (1996).

With respect to the specification of a process that satisfies equation (5.42), Blanchard (1979) and Blanchard and Watson (1982) propose a stochastic bubbles model in which two states of nature are considered - state S in which the bubble survives and continues to grow, and state C in which it crashes. The expected value of the bubble over both these states of nature may then be written as:

$$E_t(b_{t+1}) = (1-q)\, E_t(b_{t+1} \mid C) + q\, E_t(b_{t+1} \mid S) \tag{5.43}$$

where state S occurs with a constant probability q. In addition, the process assumes that:

$$E_t(b_{t+1} \mid C) = 0 \tag{5.44}$$

which implies:

$$E_t(b_{t+1} \mid S) = \frac{b_t}{a.q} \tag{5.45}$$

when this condition is combined with that of equation (5.42).

While the above process offers a solution to equation (5.42), van Norden (1996) and Schaller and van Norden (1997) argue that the assumptions underlying the model are overly restrictive. Instead, they suggest a process that allows for a partial collapse in state C and for the probability of the bubble's continued growth to fall as the bubble expands[21]. Formally re-defining equation (5.44), so that the expected size of the collapse becomes a function of the relative size of the bubble, gives rise to equation (5.46):

$$E_t(b_{t+1} \mid C) = u(b_t) \tag{5.46}$$

where u( ) is a continuous and everywhere differentiable function such that u(0) = 0 and $1 \geq u' \geq 0$. Similarly, a revised counterpart to equation (5.45) can be derived as:

$$E_t(b_{t+1} \mid S) = \frac{b_t}{a.q(b_t)} - \left(\frac{1-q(b_t)}{q(b_t)} u(b_t) \right) \tag{5.47}$$

[21] The underlying argument against an instantaneous and complete collapse of the bubble relates to the fact that financial institutions or the government would be expected to intervene to prevent such an occurrence. For a detailed description of the typical life cycle of bubbles, see Kindleberger (1989).

by drawing on equations (5.42), (5.46) and replacing q in expression (5.45) with:

$$q = q(b_t), \quad \frac{d}{d|b_t|} q(b_t) < 0 \tag{5.48}$$

where $q(b_t)$ allows for the probability of survival to decrease as the bubble grows. In turn, equations (5.46) and (5.47) may be used to formulate a series for excess returns in each state:

$$E_t(R_{t+1} \mid C) = u(b_t) - \frac{b_t}{a} \tag{5.49}$$

$$E_t(R_{t+1} \mid S) = \frac{1-q(b_t)}{a.q(b_t)}.[b_t - a.u(b_t)] \tag{5.50}$$

With regard to the actual estimation of equations (5.49) and (5.50), the particular state generating a given observation of R_{t+1} is not necessarily known a priori. As such, the use of ordinary least squares results in inconsistent and biased estimates[22]. Fortunately, 'consistent, efficient, asymptotically normal parameter estimates of such systems can be obtained by maximum likelihood methods' as illustrated in van Norden (1996)[23]. As the starting point, suppose that in regime C:

$$R_{t+1} = R^c_{t+1} = h_c(b_t) + e^c_{t+1} \tag{5.51}$$

whereas in regime S:

$$R_{t+1} = R^s_{t+1} = h_s(b_t) + e^s_{t+1} \tag{5.52}$$

On the basis of these expressions, the probability density function of an observation, conditional on the fact that it has been generated by a given regime, may be written as:

$$\phi_c(e^c_{t+1}) = \phi_c(R_{t+1} - h_c(b_t)) \tag{5.53}$$

and

$$\phi_s(e^s_{t+1}) = \phi_s(R_{t+1} - h_s(b_t)) \tag{5.54}$$

[22] Refer to Lee and Porter (1984) for a detailed discussion.

[23] Note that inference requires the series under consideration to be stationary. As the bubble process described above is clearly not, it is necessary to formulate the null hypothesis as a hypothesis of no bubbles. Under this null, the data is assumed to be stationary and as such, standard inference procedures apply. Simulation results are provided in van Norden and Vigfusson (1996a).

Next, assuming a set of variables M_t and imperfect information as to which regime generates each observation, the probability that $R_{t+1}=R^s_{t+1}$ is denoted by $q(M_t)$. The unconditional probability density function of each observation is thus:

$$q(M_t)\,.\,\phi_s(R_{t+1}-h_s(b_t)) + [1-q(M_t)].\phi_c(R_{t+1}-h_c(b_t)) \tag{5.55}$$

while the likelihood function for a set of T observations is given by:

$$\prod_{t=1}^{T}\{q(M_t)\,.\,\phi_s(R_{t+1}-h_s(b_t)) + [1-q(M_t)].\phi_c(R_{t+1}-h_c(b_t))\} \tag{5.56}$$

Subsequent maximisation of this likelihood function facilitates the simultaneous estimation of both equations (5.51) and (5.52) and a set of parameters for $q(M_t)$. Moreover, the resultant estimates are not only consistent and efficient, but are also obtained without prior knowledge as to which observations correspond to a given regime.

However, additional information is required before the bubble process described in equations (5.49) and (5.50) can be estimated; namely functional forms for $h_c(b_t)$, $h_s(b_t)$ and $q(M_t)$, along with distributional forms for e^c_{t+1} and e^s_{t+1}. In acquiring this information, the starting point is an approximation of $E_t(R_{t+1}|C)$ and $E_t(R_{t+1}|S)$ by a first order Taylor expansion about some arbitrary value $\overline{b}$. This yields the following expressions for $h_c(b_t)$ and $h_s(b_t)$:

$$h_c(b_t) = \beta_{c0} + \beta_{cb}b_t \tag{5.57}$$

$$h_s(b_t) = \beta_{s0} + \beta_{sb}b_t \tag{5.58}$$

where $\beta_c < 0$ and $\beta_{sb} > 0$. Secondly, equation (5.45) implies that M_t can be set equal to b_t. Furthermore, as the logit function satisfies equation (5.48) and ensures that $q(M_t) = q(b_t)$ is bounded between zero and one, it constitutes an appropriate functional form for the latter:

$$q(b_t) = \Phi(\beta_{q0} + \beta_{qb}\,b_t^2) \tag{5.59}$$

where $\beta_{qb} > 0$ and Φ is the logistic cumulative distribution function. Thirdly, e^c_{t+1} and e^s_{t+1} are assumed independently and identically distributed with mean zero and standard deviations σ_c and σ_s.

Combining the above information gives rise to the general switching regression model:

$$E_t(R_{t+1} \mid C) = \beta_{c0} + \beta_{cb} b_t \tag{5.60}$$

$$E_t(R_{t+1} \mid S) = \beta_{s0} + \beta_{sb} b_t \tag{5.61}$$

$$\text{Prob}(\text{State}_{t+1}=S) = q(b_t) = \Phi(\beta_{q0} + \beta_{qb} b_t^2) \tag{5.62}$$

which, as previously discussed, can be estimated by maximising the associated log likelihood function:

$$\log L(\beta_{c0}, \beta_{cb}, \beta_{s0}, \beta_{sb}, \beta_{q0}, \beta_{qb}, \sigma_c, \sigma_s) = \sum_{t=1}^{T} \ln \left| \begin{matrix} \left([1-\Phi(\beta_{q0} + \beta_{qb} b_t^2)].\phi\left(\dfrac{R_{t+1} - \beta_{c0} - \beta_{cb} b_t}{\sigma_c} \right) / \sigma_c \right) \\ + \left(\Phi(\beta_{q0} + \beta_{qb} b_t^2)].\phi\left(\dfrac{R_{t+1} - \beta_{s0} - \beta_{sb} b_t}{\sigma_s} \right) / \sigma_s \right) \end{matrix} \right| \tag{5.63}$$

where ϕ ( ) is the standard normal probability density function. Note, however, that a property of switching regressions is that they allow for the model to be identified only up to a particular renaming of the parameters, that is, the names of the C and S regimes could be swapped around. The implication for the bubbles model of such equivalence is the need to impose the following restrictions on equations (5.60) to (5.62):

$$\beta_{c0} \neq \beta_{s0} \tag{5.64}$$

and either

$$(\beta_{cb} < 0, \beta_{sb} > 0, \beta_{qb} > 0) \text{ or } (\beta_{cb} > 0, \beta_{sb} < 0, \beta_{qb} < 0) \tag{5.65}$$

Using Wald and t tests, these parameter restrictions may then be tested and if found to hold, they provide evidence in favour of the stochastic bubbles model proposed by van Norden (1996) and Schaller and van Norden (1997). To lend further support to the presence of this type of bubble, the general switching regressing model should also be able to reject a range of simpler models nested within it. One such model is that of the mixture of normal distributions which implies $\beta_{cb} = \beta_{sb} = \beta_{qb} = 0$ in equations (5.60) through to (5.62). Applying the likelihood ratio test, this model can be tested against the switching regression alternative, with the failed rejection of the null having the interpretation that fundamentals are driving excess returns.

Finally, after estimating the model and carrying out the above hypothesis tests, diagnostic checks should be implemented so as to eliminate the possibility of inconsistent estimates and invalid inferences that may arise in the case of model misspecification. Table 5.5 details the tests implemented for this purpose in Chapter 8 of the study[24]. In addition, an overview of the Wald and likelihood ratio test principles is provided.

Table 5.5: Switching Regression Hypothesis and Misspecification Tests – Statistical Formulae

Hypothesis Testing:	Consider a K dimensional parameter vector θ estimated by maximum likelihood. Restrictions on θ may be written as: H_0: Rθ = q where: q = a fixed J dimensional vector And R = a J x K matrix. Subsequent testing of these restrictions can be achieved using the Wald and likelihood ratio test statistics defined below.
Wald Test The Wald test statistic reported in the empirical analysis is obtained by running a Gauss procedure written by van Norden and Vigfusson (1996b). This procedure calculates the Wald statistic using the inverse of the hessian; The test statistic has a chi squared distribution with J degrees of freedom; See Verbeek (2000) for a general discussion.	Generally: $T(R\hat{\theta}-q)'[R\hat{V}R']^{-1}(R\hat{\theta}-q)$ where: $\sqrt{T}(R\hat{\theta}-R\theta)\rightarrow N(0, RVR')$
Likelihood Ratio Test The LR test also follows a chi squared distribution with J degrees of freedom; Further details are provided in Verbeek (2000).	$2[\log L(\hat{\theta})-\log L(\tilde{\theta})]$

[24] The misspecification tests presented in Table 5.5 are included as part of the Gauss procedures for estimating regime-switching models kindly provided by Simon van Norden. For further details on the nature of these tests, see van Norden and Vigfusson (1996b), White (1987) and Hamiliton (1990 and 1996).

Table 5.5 (contd): Switching Regression Hypothesis and Misspecification Tests – Formulae

Misspecification Testing: Residual Autocorrelation and Heteroscedasticity	Let $L(y_t \mid x_t, \theta)$ denote the likelihood function of a given observation y_t conditional on x_t and a set of parameters θ. The gradient of $L(y_t \mid x_t, \theta)$ with respect to θ defines the score $h_t(\theta)$. Subject to certain conditions, it holds that at the true parameter estimates θ_0, $h_t(\theta_0)$ cannot be forecasted from any information available at time t-1, including that of $h_{t-1}(\theta_0)$. As the vector $h_t(\theta)$ has the same dimensions as θ, it follows that the absence of first order serial correlation in $h_t(\theta_0)$ gives rise to m x m testable restrictions. A general test of such restrictions may then be constructed. In doing so, the *l* elements of the m x m matrix $h_t(\theta).h_t(\theta)'$ to be tested as part of the *l* x1 vector $c_t(\theta)$ are listed. The resultant matrix product has a chi squared asymptotic distribution with *l* degrees of freedom: $T^{-1}\left[\sum_{t=1}^{T} c_t(\hat{\theta})\right]' .\hat{A}. \left[\sum_{t=1}^{T} c_t(\hat{\theta})\right]$ In turn, this matrix product can be used to define a range of misspecification test statistics. Note that the Gauss procedures written by van Norden and Vigfusson (1996b) require: (1) the T x m matrix of gradients evaluated at the parameter estimates; (2) a column vector that lists the numbers of the columns in the gradient vector to be included in the test. On the basis of the above, test statistics and associated chi squared significance levels are then produced.
AR Test This LM test is approximately distributed as chi squared with 1 degree of freedom; Testing the gradients is carried out under the Assumption of a constant in each of the level equations.	The gradients ε_{t1} and ε_{t2} are tested for evidence of serial correlation. If present, the implication is that there is persistence / serial correlation in the residuals
ARCH Test The ARCH LM test is also approximately distributed as chi squared with 1 degree of freedom.	The gradients of σ^c and σ^s are tested for persistence effects

Notes:
1. As these tests have a tendency to over-reject the null hypothesis in small samples, interpretation at the 1% level is deemed more appropriate than at the 5% level. Please refer to Hamilton (1990) for further details

Allowing for the rejection of nested models and satisfaction of the coefficient requirements and diagnostic checks implied by the stochastic bubbles model, it would seem logical to conclude that the bubbly component of the asset has a significant influence on returns in the following period. However, it should be noted that a major drawback of the regime switching approach, and indeed of all such bubble tests, is that behaviour indicative of a bubble component is not proof that a bubble exists (Flood and Hodrick, 1990). For example, Flood and Hodrick (1986) point out that bubbles are observationally equivalent to regime switching in fundamentals. Conversely, the failure to find evidence of a bubble may be suggestive of misspecification errors with respect to the chosen measure of non-fundamentals[25]. Alternatively, it could be that the stochastic bubble described in this section simply does not exist.

5.5. Conclusion

This chapter aimed to provide a technical overview of the modelling techniques adopted in the course of the study's empirical analyses of the Dublin housing market; namely those of Johansen cointegration, ordinary least squares estimation, smooth transition regression modelling and regime-switching tests for stochastic bubbles. In achieving this objective, attention was paid to describing these methodologies from the perspectives of estimation, inference and hypothesis testing. In addition, the tabulations of the statistical formulae underlying the diagnostic checks applied in forthcoming chapters should act as a reference point for subsequent model evaluation.

---

[25] In the case of the van Norden bubble test, misspecification is highly unlikely to result in the false finding of a bubble. As both the likelihood ratio tests and the signs on the coefficient restrictions are invariant to linear transformations of the bubble, the test is robust to misspecification in either the level or the scale of the bubble (van Norden, 1996).

5A. Appendix

5A.1. Cointegration Methodology

As inference in the unrestricted vector autoregressive model (VAR) is based on asymptotic results, the asymptotic properties of the estimators are set out below for the case of a stationary process.

Theorem 5A.1: (Johansen, 1995) Assuming that the errors are independent and identically distributed with mean zero and variance Ω, that the process X_t is stationary and that $\Phi D_t = \mu$; then the asymptotic distribution of the maximum likelihood estimator of the parameters $B' = (\Pi_1, \ldots, \Pi_k, \mu)$ is Gaussian and given by:

$$T^{1/2}(\hat{B} - B) \overset{w}{\rightarrow} N(0, \Sigma^{-1} \otimes \Omega) \tag{5A.1}$$

where:

$$\hat{\Sigma} = T^{-1} \sum_{t=1}^{T} Z_t Z'_t \overset{P}{\rightarrow} \Sigma,$$

$$\hat{\Omega} \overset{P}{\rightarrow} \Omega \text{ and } Z'_t = (X'_{t-1}, \ldots, X'_{t-k}, 1)$$

As such, for any vectors ξ and η, it holds that:

$$\frac{T^{1/2} \xi' (\hat{B} - B)\eta}{\sqrt{\xi'\hat{\Sigma}^{-1}\xi\eta'\hat{\Omega}\eta}} \overset{w}{\rightarrow} N(0,1) \tag{5A.2}$$

5A.2. Ordinary Least Squares Methodology

To quote Verbeek (2000), the linear regression model and the ordinary least squares method of estimation (OLS) respectively represent 'one of the cornerstones of' and the 'most important technique in' econometrics. This assertion is borne out by the fact that undergraduate textbooks typically take the standard linear regression model and OLS as the starting point of their econometric analyses[1]. As such, the following discussion assumes some prior knowledge and focuses primarily on re-iterating the main features of this methodology from a time series perspective.

To begin, consider the multiple extension of the simple linear regression model:

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots. + \beta_k x_{tk} + u_t \qquad (t = 1, \ldots., T) \tag{5A.3}$$

where $x_{tj} = (x_{t1}, \ldots., x_{tk})$ and u_t is a sequence of disturbance or error terms. The corresponding estimates $\hat{\beta}_0, \hat{\beta}_1, \ldots., \hat{\beta}_k$ are obtainable by means of the ordinary least squares method which chooses k+1 estimates to minimise the sum of squared residuals:

$$\sum_{t=1}^{T}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{t1} - \ldots. - \hat{\beta}_k x_{tk})^2 \tag{5A.4}$$

Solving the above minimisation problem results in k+1 linear equations in k+1 unknowns referred to as the OLS first order conditions:

$$\sum_{t=1}^{T}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{t1} - \ldots. - \hat{\beta}_k x_{tk}) = 0$$

$$\sum_{t=1}^{T} x_{t1}\,(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{t1} - \ldots. - \hat{\beta}_k x_{tk}) = 0$$

.

.

$$\sum_{t=1}^{T} x_{tk}\,(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{t1} - \ldots. - \hat{\beta}_k x_{tk}) = 0 \tag{5A.5}$$

---

[1] See, for example, Gujarati (1995).

If it holds that none of the independent variables in equation (5A.3) are perfectly collinear, the equations in (5A.5) can in turn be uniquely solved for $\hat{\beta}_j$.

In estimating the parameters of the underlying population model, the statistical properties of the OLS estimators come into play. The large sample properties of OLS, namely those of consistency and asymptotic normality are discussed in Theorems 5A.2 and 5A.3[2]. The assumptions underlying these asymptotic properties are outlined in Table 5A.1.

Table 5A.1: Large Sample Assumptions

(i): *Linear in Parameters and Weak Dependency*

The stochastic process (x_{t1}, …., x_{tk}, y_t) follows the linear model:

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots. + \beta_k x_{tk} + u_t$$

where u_t is the sequence of errors. In addition, ($\mathbf{x}_t$,y_t) where t = 1, …., is weakly dependent. As the β_i do not change over time, a form of stationarity is also assumed.

(ii): *Zero Conditional Mean*

$$\text{For each t: } E(u_t | \mathbf{x}_t) = 0$$

(iii): *No Perfect Collinearity*

In the sample and consequently, in the underlying time series process, no independent variable is either constant or a perfect linear combination of the other explanatory variables.

(iv): *Homoscedasticity*

$$\text{For all t: } Var(u_t | \mathbf{x}_t) = \sigma^2$$

(v): *No Serial Correlation*

$$\text{For all } t \neq s\text{: } E(u_t,u_s | \mathbf{x}_t,\mathbf{x}_s) = 0$$

[2] For a proof of the large sample properties of OLS estimators, refer to Wooldridge (1994).

Theorem 5A.2: Under the large sample assumptions (i) to (iii), the OLS estimators are consistent:

$$\hat{\beta}_j \xrightarrow{P} \beta_j \qquad (j = 0, 1, \ldots, k) \qquad (5A.6)$$

Theorem 5A.3: Assumptions (i) to (v) imply that the OLS estimators are asymptotically normally distributed. Thus, OLS standard errors along with the usual t, F and LM statistics are asymptotically valid.

Having established the asymptotic properties of the OLS estimators, the remainder of this section is concerned with inference. In particular, the statistical formulae that motivate the hypothesis and misspecification tests applied in Chapter 7 are presented in Table 5A.2. Furthermore, the criteria adopted when choosing between various OLS estimations of dynamic price determination in the Dublin housing market are also defined[3].

Table 5A.2: OLS Model Selection, Hypothesis and Misspecification Tests – Statistical Formulae

Goodness of Fit:

R Squared

$$SSE / SST = 1 - \frac{\sum_{i=1}^{T} \hat{u}_i^2}{\sum_{i=1}^{T} (y_i - \bar{y})^2}$$

where:

SSE = the explained sum of squares

and

SST = the total sum of squares

[3]Testing was carried out using Givewin 2.20 and Eviews 3.1. For further details as to how these tests are implemented in the above software programmes, refer respectively to Doornik and Hendry (2001) and the Eviews User's Guide.

Table 5A.2 (contd): OLS Model Selection, Hypothesis and Misspecification Tests – Formulae

Adjusted R Squared	$1 - \dfrac{SSE/(T-k)}{STT/(T-1)}$
Information Criteria:	
Akaike Information Criteria (AIC)	$\log \tilde{\sigma}^2 + 2k / T$
Schwartz Information Criteria (SIC)	$\log \tilde{\sigma}^2 + k (\log T) / T$
Hannan-Quinn (HQ)	$\log \tilde{\sigma}^2 + 2k [\log (\log T)] / T$
As in the VAR specification, these tests make use of the maximum likelihood function. With respect to model selection, smaller values of the criteria are preferred.	where: $\tilde{\sigma}^2 = (T\text{-}k / T)\,\hat{\sigma}^2$
Hypothesis Testing:	
T Values The test approximates a student t distribution with T-k degrees of freedom.	$\dfrac{\hat{\beta}_j}{se(\hat{\beta}_j)}$
Exclusion Test If the null hypothesis that the coefficient of the variable in question equals zero is true, the F statistic is distributed as $F_{(J,T-k)}$.	$F = \dfrac{R^2_{ur} - R^2_r / J}{(1 - R^2_{ur})/(T-k)}$ where: R^2_{ur} = the R squared from the unrestricted model and R^2_r = the R squared from the restricted model
Residual Autocorrelation:	
AR Test This test is approximately distributed as F(s, T-k-s); The empirical analysis in Chapter 7 reports the F-form of the test, as opposed to the LM version, on the grounds that the former is likely to be better behaved in small samples; Refer to Harvey (1990) for a detailed discussion.	$\dfrac{R^2}{1-R^2} . \dfrac{T-k-s}{s}$ where R^2 is taken from the regression of $\hat{u}_t$ on $x_{t1,\ldots,} x_{tk,} \hat{u}_{t-1,\ldots,} \hat{u}_{t-r}$

Table 5A.2 (contd): OLS Model Selection, Hypothesis and Misspecification Tests – Formulae

Residual Heteroscedasticity:	
ARCH Test	$$\frac{R^2}{1-R^2}\cdot\frac{T-k-s}{s}$$
As in the AR case, the ARCH test is approximately distributed as F(s, T-k-s); Again, the empirical analysis reports the F version of the test when testing for the presence of ARCH as it tends to be better behaved in small samples than its LM counterpart.	where R^2 is taken from the regression of $\hat{u}_t^2$ on a constant and $\hat{u}_{t-1}^2$ to $\hat{u}_{t-r}^2$
Normality Test:	
The associated test statistic is that of the multivariate test described in Table 5.2.	

Chapter 6:

A Long Run Analysis of the Dublin Housing Market

6.1. Introduction

Chapter 2's discussion of the theoretical structure of the housing market suggests that similar to other commodities, housing can be modelled within a demand and supply framework. It then follows that it should be possible to identify a long run demand and supply curve for housing in a Dublin context. Moreover, a clear understanding of the empirical conditions that characterise the market is a necessary prerequisite for solving current housing disequilibrium. This chapter seeks to facilitate such an understanding in that it undertakes an empirical analysis of the Dublin market for new, private sector housing with the aim of uncovering the nature of the long run relationships at play in this marketplace. Econometric modelling adopts the Johansen cointegration approach and applies it to the time-series dataset described in Chapter 4[1].

The chapter is organised as follows[2]: Section 6.2 provides an overview of a priori expectations while the time series properties of the data are discussed in Section 6.3. Sections 6.4 and 6.5 present respectively, the statistical model and apply a range of diagnostic checks. Following the determination of the cointegrating rank in Section 6.6, Section 6.7 carries out an examination of model specific data properties. The findings of the latter indicate the need

[1] As noted in Chapter 5, the Johansen approach allows for long run effects and short term adjustment patterns to be modelled simultaneously. In practice, long run identification is often considered first. This is a matter of convenience and does not reflect the actual estimation procedure.

[2] Econometric modelling undertaken in the course of this chapter was carried out in RATS 5 / CATS 1.0, GiveWin 2.20 / PcGive 10.0 and EViews 4.1.

to move to a partial model as discussed in Section 6.8. Long run identification is undertaken in Section 6.9 while Section 6.10 concludes the analysis.

6.2. A Priori Expectations

Before embarking upon the empirical aspect of this study, it is necessary to form initial expectations as to what factors are likely to have a role to play in the housing market. On the basis of economic theory and precedents set by relevant studies, a number of variables can be identified as potentially important when modelling housing in the long run[3]. These include house prices, income, mortgage interest rates, building costs, the stock of housing, household formation and land availability[4]. This chapter begins its analysis of the Dublin new housing market by postulating the existence of two equilibrium relations amongst these variables - one for the demand-side of the market and one representing supply.

With respect to specifying a housing demand vector, four variables - house prices, mortgage interest rates, income and household formation – are predicted to play a prominent part in determining the desired stock of housing in the long run. Of these, prices and interest rates should exhibit a negative relationship with demand whereas income and the demographic measure are expected to enter the equation with a positive sign.

Given the theoretical inter-connection between the housing and construction markets, a priori expectations imply a long run relationship between the supply of housing units, house prices, land availability and construction and interest costs. A positive sign is expected on the house price and land coefficients while cost coefficients should be characterised by negative signs.

---

[3] Please refer back to Chapter 3 for further details.
[4] These variables are defined as in Section 4.2.

6.3. Data Properties

Formal testing of the long run relationships postulated above can be achieved by applying the Johansen cointegration technique[5]. The objective of this method is to uncover stationary relationships amongst a set of non-stationary data. Assuming that all series are integrated of order one or lower, the presence of one or more cointegrating vectors implies that non-stationary variables are linearly related in a stationary manner, and can thus be interpreted as an equilibrium relationship.

Table 6.1: Unit Root Testing and Degree of Integration[6]

Variable	*DF & ADF Test Statistics*		*Degree of Integration*
	Levels	First Differences	
lognph	2.30	-2.67*	I(1)
loginc	1.84	-8.11***	I(1)
MR	-1.67	-4.48***	I(1)
logbc	0.33	-10.39***	I(1)
logphs	3.20	-6.62***	I(1)
loghf	-2.22	-11.48***	I(1)
logls	-1.67	-6.60***	I(1)

Notes:

1. Ho: The process contains a unit root
2. * (***) denotes rejection of the hypothesis at the 10% (1%) level
3. The test equations include a constant in each case. The addition of a linear trend to the housing stock regression results in test statistics of 0.93 and –7.43 in levels and first differences respectively
4. Lag lengths are chosen on the basis of the Schwartz Information Criteria. This automatic selection method chooses P lagged difference terms to minimise the term:

$$-2(l\,/\,T) + kl\,\mathrm{og}(T)\,/\,T$$

where l is the value of the log of the likelihood function and k refers to the number of estimated parameters

Table 6.1 reports the results of Dickey-Fuller and Augmented-Dickey Fuller tests for unit roots in the Dublin housing market data[7]. These tests fail to reject the null hypothesis of a unit root when the series are expressed in levels. However, once the data is differenced, the tests reject the null for each of the variables. Chapter 4's graphical analysis of the time series

[5] For a detailed discussion of this technique refer to Section 5.2.1 and Johansen (1995).

[6] For further practical information on unit root testing, refer to the Eviews User's Guide.

[7] See Davidson and MacKinnon (1993) and Hamilton (1994) for further details of these tests.

properties of the data lends additional weight to these findings. As there is no evidence of I(2) behaviour, it is appropriate to proceed using the Johansen technique.

In addition, note that a nominal to real transformation has been performed in respect of the house price, income and building cost series. The validity of this approach is dependent on the assumed long run homogeneity of consumer prices and the nominal values of these variables. When tested, this hypothesis is found to hold. The mortgage interest rate is not transformed given the explanatory power associated with the decomposition of this series into its component parts – an inflationary and a real effect - in the empirical analysis[8].

6.4. The Statistical Model

The statistical vector autoregressive (VAR) model estimated as part of the empirical analysis of the Dublin housing market is defined as:

$$X_t = \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \Phi_{tr1} Dum_1bc_t + \Phi_{p1} Dum_1ls_t + \Phi_{p2} Dum_2ls_t + \Phi_{p3} Dum_3ls_t$$
$$+ \Phi_{p4} Dum_4ls_t + \Phi_{tr2} Dum_2bc_t + \Phi_{tr3} Dummr_t + \Phi_{p5} Dum_3bc_t$$
$$+ \Phi'_q Dq_t + \mu + \varepsilon_t \qquad (6.1)^9$$

where:

X = a vector comprising the log of real new house prices, the log of real personal disposable income per capita, nominal mortgage interest rates, the log of real building costs, the log of private housing stock, the log of household formation and the log of the stock of zoned housing land;

Dum_1bc = a transitory shock dummy = 1 for 2000:4 and 0 otherwise;

Dum_1ls = a permanent shock dummy = 1 for 1982:2 and 0 otherwise;

[8] Kenny (1998) adopts a similar approach.
[9] While a linear trend was included as part of the initial VAR specification, subsequent testing indicated that this could be excluded from the analysis.

Dum_2ls = a permanent shock dummy = 1 for 1993:2 and 0 otherwise;

Dum_3ls = a permanent shock dummy = 1 for 1998:3 and 0 otherwise;

Dum_4ls = a permanent shock dummy = 1 for 1981:3 and 0 otherwise;

Dum_2bc = a transitory shock dummy = 1 for 1981:2, 0 for 1981:3, -1 for 1981:4 and 0 otherwise;

Dummr = a transitory shock dummy = 1 for 1992:4, -1 for 1993:1 and 0 otherwise;

Dum_3bc = a permanent shock dummy = 1 for 1985:2 and 0 otherwise;

Dq = three centred seasonal dummies Dq_1, Dq_2 and Dq_3, where $Dq_i = 0.75$ in quarter 1 and -0.25 in quarters i+1, i+2 and i+3 [10];

Π & Φ = unrestricted parameters;

μ = a vector of unrestricted constants;

ε = a vector of error terms.

The sample (adjusted for a lag length of two) begins in the third quarter of 1980 and ends in the last quarter of 2000[11].

Note that the above model can be re-parameterised in vector error correction (VECM) format as follows:

$$\Delta X_t = \Pi_1 X_{t-1} + \Gamma_1 \Delta X_{t-1} + \Phi_{tr1} Dum_1bc_t + \Phi_{p1} Dum_1ls_t + \Phi_{p2} Dum_2ls_t + \Phi_{p3} Dum_3ls_t + \Phi_{p4} Dum_4ls_t + \Phi_{tr2} Dum_2bc_t + \Phi_{tr3} Dummr_t + \Phi_{p5} Dum_3bc_t + \Phi'_q Dq_t + \mu + \varepsilon_t \quad (6.2)$$

where:

Γ = a matrix of short term dynamics.

[10] Johansen (1995) suggests using centred rather than standard seasonal dummies. Centred dummy variables shift the mean but do not affect the trend.

[11] Rather than ending the analysis in the third quarter of 2000, a dummy is included for the last observation. While this is somewhat trivial, it allows for consistency between this and later chapters, vis-à-vis the end point of the sample period.

6.5. Model Selection and Misspecification Testing

Estimation of the VAR outlined in equation (6.1) relies on a number of assumptions, namely that the linear conditional mean is explained by past observations and the deterministic trends, that the conditional variance is constant and that errors are normal and independent with mean zero and variance Ω. While not all of these assumptions are crucial, it is important to test for model misspecification. As such, the following tables report the information criteria that motivated the choice of lag length and the outcome of tests for residual normality, autocorrelation and homoscedasticity. In addition, the constancy of the model's parameters is checked using recursive analysis and the findings of the study in this respect are discussed.

6.5.1. *Information Criteria and Diagnostic Checking*

Tables 6.2 to 6.6 illustrate the multivariate properties of the VAR specification for various lag lengths. Initial modelling excludes the permanent and transitory shock dummies and chooses an appropriate lag length on the basis of the standard tests of the Schwartz (SC) and Hannan-Quinn (HQ) information criteria, along with the Likelihood Ratio test for model reduction. As Tables 6.2 and 6.3 indicate, these three tests unanimously select a lag length of one.

Table 6.2: Information Criteria for Choice of VAR Lag Length – No Dummies

Lag Length	*SC*	*HQ*
1	-46.025	-47.398
2	-44.189	-46.436
3	-42.644	-45.765
4	-40.986	-44.982

Notes:
1. To ensure validity, the models are nested

Table 6.3: Likelihood Ratio Test for Model Reduction – No Dummies

Lag Length	*Test Statistic*	*Probability*
Lag 4 to Lag 3	F(49,217) = 0.995	0.49
Lag 4 to Lag 2	F(98,274) = 1.144	0.20
Lag 4 to Lag 1	F(147,291) = 1.131	0.19
Lag 4 to Lag 0	F(196,298) = 31.633	0.00***
Lag 3 to Lag 2	F(49,253) = 1.300	0.10
Lag 3 to Lag 1	F(98,318) = 1.201	0.12
Lag 3 to Lag 0	F(147,337) = 43.484	0.00***
Lag 2 to Lag 1	F(49,288) =1.071	0.36
Lag 2 to Lag 0	F(98,363) = 68.999	0.00***
Lag 1 to Lag 0	F(49,324) = 221.76	0.00***

Notes:
1. To ensure validity, the models are nested
2. Ho: Model reduction is acceptable
3. The F version of this test is used as it corrects for small sample size
4. *** denotes rejection of Ho at the 1% significance level

With respect to diagnostic checking and serial correlation, the LM test for first and fourth order autocorrelation implies that the null hypothesis of no serial correlation is largely acceptable. The vector heteroscedasticity test supports the notion of homoscedastic residuals where applicable but the normality assumption is clearly violated given considerable residual non-normality as evident from Table 6.5.

Table 6.4: Testing for Serial Correlation in the VAR Residuals – No Dummies

Lag Length	*Test Statistic*	*Probability*
1	LM (1) = 60.962	0.12
	LM (4) = 48.719	0.48
2	LM (1) = 78.005	0.01
	LM (4) = 43.391	0.70
3	LM (1) = 63.734	0.08
	LM (4) = 49.372	0.46
4	LM (1) = 57.165	0.20
	LM (4) = 51.338	0.38

Notes:
1. Ho: No serial correlation at lag order h
2. Probabilities are distributed around a chi-squared distribution with 49 degrees of freedom
3. A small sample correction is included

Table 6.5: Testing for Normality in the VAR Residuals – No Dummies

Lag Length	*Test Statistic*	*Probability*
1	111.218	0.00***
2	81.071	0.00***
3	62.713	0.00***
4	52.499	0.00***

Notes:
1. Ho: Residuals are multivariate normal
2. Probabilities are distributed around a chi-squared distribution with 14 degrees of freedom
3. A small sample correction is included
4. *** denotes rejection of Ho at the 1% significance level

Table 6.6: Testing for Heteroscedasticity in the VAR Residuals – No Dummies

Lag Length	*Test Statistic*	*Probability*
1	F(392,388) = 1.060	0.28
2	F(784,253) = 0.569	1.00
3	Chi^2(1176) = 1254	0.06
4	N/A	N/A

Notes:
1. Ho: Residuals are homoscedastic

However, it should be noted that the presence of non-normal residuals does not render the Johansen technique invalid. For example, studies by Cheung and Lai (1993) and Gonzolo (1994) have been successful in demonstrating the robustness of the procedure given non-normal error terms. Nevertheless, Juselius (forthcoming) argues that it is a mistake to treat significant interventions and reforms as mere statistical nuisances. As such, an attempt is made in the following analysis to improve the above baseline VAR(1) specification by the inclusion of a number of dummy variables as defined in equation (6.1).

In terms of identifying the location and type of dummies needed to account for shocks to the variables in the system, the standardised residuals from the estimated VAR(1) model with no dummies were examined to pinpoint outliers. Values in excess of the rule of thumb critical figure of 3.73 (in absolute terms) were isolated out one at a time and were modelled as permanent or transitory dummies as appropriate.

As a further check, attention was paid to the graphs of the individual data series presented in Chapter 4. Figure 4.9's illustration of the land stock variable indicates that the series appears to be stationary when differenced once and to contain a number of shift effects. These shifts are reflected in the levels graph by the movement to a higher path about 1981, 1982, 1993 and 1998 and also in the corresponding blips in the graph of the first differences. As the shocks represent permanent interventions, they are modelled in difference terms by dummies that take a value of one for the relevant observation and zero otherwise. These dummies then cumulate to form a shift in the data in levels and thus the desired effect is obtained. Similarly, Figure 4.6 suggests a level shift in the building cost series at the beginning of 1985 and consequently, a permanent shock dummy is also included to account for this occurrence. Overall, it is expected that the various shifts evident in the data will cancel and as such, it is not necessary for them to separately enter the cointegrating space.

In addition to the above non-reversible shocks, there appears to be some instances of temporary overshooting with respect to the interest rate and building cost variables. Figure 4.5 depicts a blip in the mortgage interest rate data expressed in levels at the end of 1992 (related to the exchange rate mechanism crisis) and a corresponding drop in the difference graph followed quickly by a return to the previous level. Likewise, a temporary shock to the building cost series is depicted in Figure 4.6 for 1981 and an outlier is clearly evident for the final observation of this variable. Given the transitory nature of these shocks, it is appropriate to model them using dummies that assign values of one and minus one to the observations concerned so that the effect is eliminated almost immediately.

Following the inclusion of these dummy variables, the model was then re-tested to choose a correct lag length and the various diagnostic checks reported previously were carried out a second time. Tables 6.7 to 6.11 report the findings.

Table 6.7: Information Criteria for Choice of VAR Lag Length – Dummies

Lag Length	*SC*	*HQ*
1	-47.442	-49.814
2	-46.201	-49.448
3	-44.817	-48.937
4	-43.677	-48.671

Notes:
1. To ensure validity, the models are nested

Table 6.8: Likelihood Ratio Test for Model Reduction – Dummies

Lag Length	*Test Statistic*	*Probability*
Lag 4 to Lag 3	F(49,177) = 1.285	0.12
Lag 4 to Lag 2	F(98,223) = 1.297	0.06
Lag 4 to Lag 1	F(147,237) = 1.458	0.005***
Lag 3 to Lag 2	F(49,212) = 1.265	0.13
Lag 3 to Lag 1	F(98,268) = 1.486	0.007***
Lag 2 to Lag 1	F(49,248) = 1.665	0.007***

Notes:
1. To ensure validity, the models are nested
2. Ho: Model reduction is acceptable
3. The F version of this test is used as it corrects for small sample size
4. *** denotes rejection of Ho at the 1% significance level

There is now greater inconclusiveness amongst the information criteria regarding the choice of lag length. While the SC and HQ methods select a lag of one, the Likelihood Ratio test for model reduction suggests a lag length of two. In both cases, the residuals are reasonably well behaved in that they are neither serially correlated or heteroscedastic.

Table 6.9: Testing for Serial Correlation in the VAR Residuals – Dummies

Lag Length	*Test Statistic*	*Probability*
1	LM (1) = 66.671	0.05
	LM (4) = 62.910	0.09
2	LM (1) = 56.707	0.21
	LM (4) = 52.802	0.33
3	LM (1) = 53.598	0.30
	LM (4) = 50.134	0.43
4	LM (1) = 50.269	0.42
	LM (4) = 46.787	0.56

Notes:
1. Ho: No serial correlation at lag order h
2. Probabilities are distributed around a chi-squared distribution with 49 degrees of freedom
3. A small sample correction is included

Table 6.10: Testing for Normality in the VAR Residuals – Dummies

Lag Length	*Test Statistic*	*Probability*
1	30.339	0.007***
2	32.212	0.004***
3	37.633	0.001***
4	27.821	0.015**

Notes:
1. .Ho: Residuals are multivariate normal
2. Probabilities are distributed around a chi-squared distribution with 14 degrees of freedom
3. A small sample correction is included
4. **(***) denotes rejection of Ho at the 5% (1%) significance level

Table 6.11: Testing for Heteroscedasticity in the VAR Residuals – Dummies

Lag Length	*Test Statistic*	*Probability*
1	F(392,288) = 0.665	0.99
2	F(784,94) = 0.147	1.00
3	Chi^2(1176) = 1131	0.82
4	N/A	N/A

Notes:
1. Ho: Residuals are homoscedastic

However, as is clear from Table 6.10, non-normality is still present though the associated test statistics are vastly improved when compared with the VAR(1) and VAR(2) specifications containing no dummy variables.

After some consideration a decision was made to proceed on the basis of two lags. This choice was motivated in part by the findings of the information selection methods and also by the outcome of the residual tests, which indicate that a lag of two renders serial correlation in the new housing market insignificant with a much higher probability than a lag of one. Moreover, a lag length of two fits better with expectations and is likely to be less restrictive than a model estimated using only one lag.

6.5.2. *Analysis of the Residuals of the Estimated VAR Model*

Bearing in mind that the chosen specification must satisfy a number of assumptions in order for the Johansen cointegration technique to be valid, the following tables report the multivariate and univariate statistics associated with the estimated VAR(2) model.

Findings are generally favourable though the model and the data do contain some negative features. Firstly, the correlation matrix indicates slightly high, but not unexpected, correlations between particular variables - most notably between house prices and income. On the other hand, the multivariate statistics indicate that the residuals are serially uncorrelated and homoscedastic.

Secondly, the univariate analysis implies a lack of normality for the household formation and land stock variables. This appears to arise from the kurtosis problem that these series exhibit and while serving to undermine normality in the multivariate framework, the problem is of an insufficient magnitude to be of any real concern. Moreover, as pointed out in Section 6.5.1 the presence of non-normal residuals does not render the methodology invalid. As such, the empirical model is statistically sound and does not suffer from misspecification.

Table 6.12: Estimated VAR(2) Model – Correlation Matrix

Variable	*dlognph*	*dloginc*	*dMR*	*dlogbc*	*dlogphs*	*dloghf*	*dlogls*
dlognph	1.000						
dloginc	0.437	1.000					
dMR	-0.309	-0.168	1.000				
dlogbc	0.274	0.056	-0.277	1.000			
dlogphs	-0.073	-0.014	0.069	-0.062	1.000		
dloghf	-0.038	0.014	-0.306	0.336	-0.096	1.000	
dlogls	0.068	0.180	0.179	0.050	0.227	0.181	1.000

Table 6.13: Estimated VAR(2) Model – Multivariate Statistics

Test Type	*Test Statistic*	*Probability*
Information Criteria:		
SC	-46.139	
HQ	-49.336	
Serial Correlation:		
Lagrange Multiplier	LM(1) = 46.143	0.59
	LM(4) = 54.909	0.26
Vector AR(1-4) Test	F(196,161) = 0.856	0.85
Normality	34.430	0.00***
Heteroscedasticity	F(784,134) = 0.198	1.00

Notes:

1. Ho: Residuals are respectively serially uncorrelated, multivariate normal and homoscedastic
2. Probabilities are distributed around a chi-squared distribution with 49 degrees of freedom for the LM test and 14 degrees of freedom when testing for normality
3. Where applicable, F versions of the above tests are used to correct for small sample size. The LM test for serial correlation and the normality test also include small sample corrections
4. *** denotes rejection of Ho at the 1% significance level

Table 6.14: Estimated VAR(2) Model – Univariate Statistics

Variable	*Mean*	*Std Dev*	*Skewness*	*Kurtosis*
lognph	-0.000	0.033	-0.063	2.852
loginc	-0.000	0.016	0.025	2.610
MR	0.000	0.511	0.143	3.126
logbc	-0.000	0.008	-0.193	3.424
logphs	-0.000	0.001	0.356	3.214
dloghf	0.000	0.098	-0.082	4.649
logls	0.000	0.006	0.127	5.481

Variable	*Test Type*	*Test Statistic*	*Probability*
lognph	AR(1-4)	F(4,52) = 0.775	0.55
loginc		F(4,52) = 1.239	0.31
MR		F(4,52) = 1.228	0.31
logbc		F(4,52) = 2.148	0.09
logphs		F(4,52) = 0.322	0.86
loghf		F(4,52) = 0.350	0.84
logls		F(4,52) = 2.764	0.04**

Variable	*Test Type*	*Test Statistic*	*Probability*
lognph	Normality	0.122	0.94
loginc		0.111	0.95
MR		0.904	0.64
logbc		2.231	0.33
logphs		2.085	0.35
loghf		12.175	0.00***
logls		21.184	0.00***
lognph	ARCH (1-4)	F(4,48) = 1.264	0.30
loginc		F(4,48) = 0.683	0.61
MR		F(4,48) = 2.023	0.11
logbc		F(4,48) = 0.324	0.86
logphs		F(4,48) = 0.476	0.75
loghf		F(4,48) = 1.697	0.17
logls		F(4,48) = 2.975	0.03**

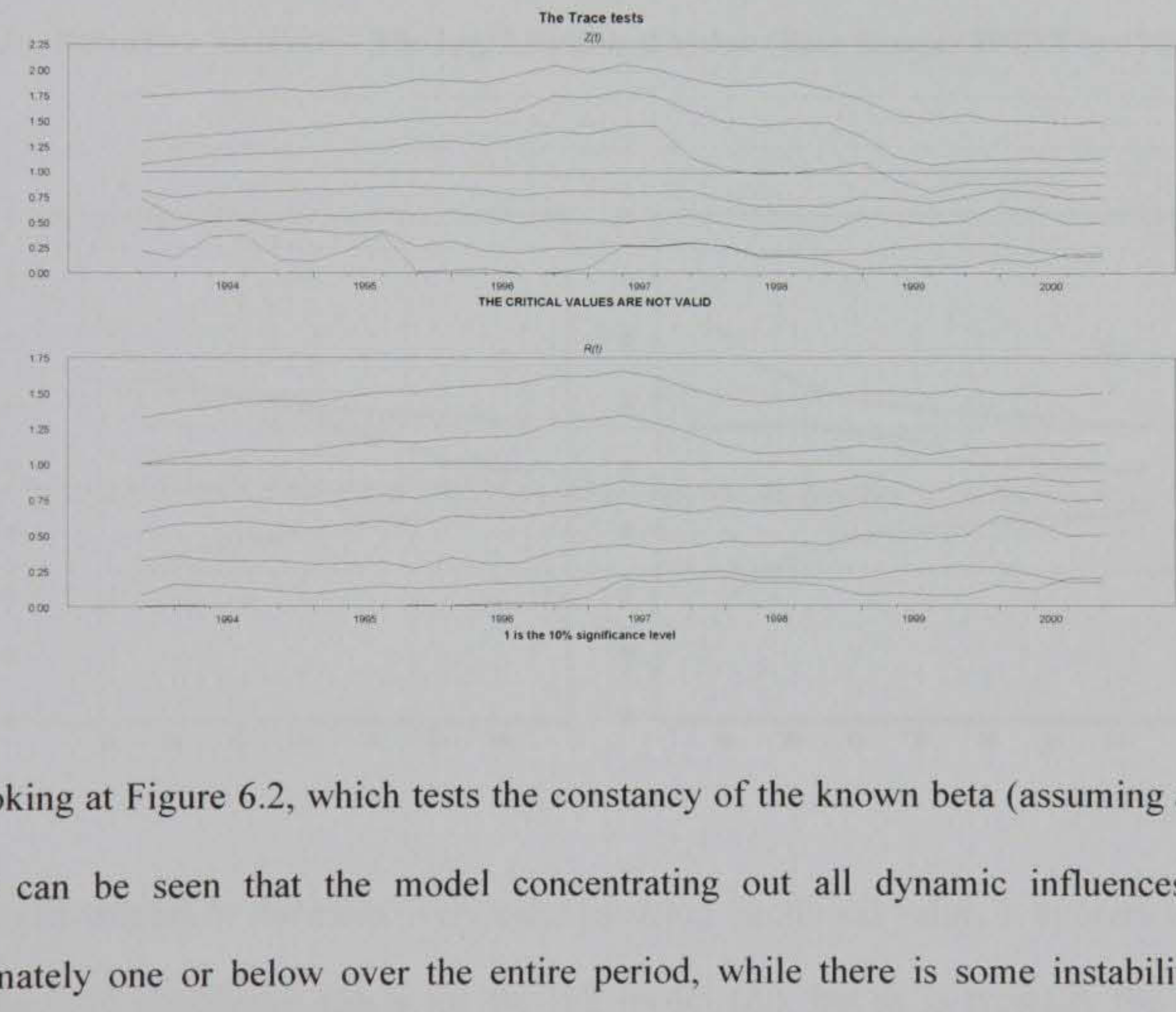
Notes:
1. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
2. Probabilities for the normality test are distributed around a chi-squared distribution with 2 degrees of freedom
3. ** (***) denotes rejection of Ho at the 5% (1%) significance level

6.5.3. *Recursive Analysis*

Parameter constancy is an important assumption underlying the VAR methodology, and while a model may pass the various misspecification tests discussed above, it may be the case that it suffers from non-constant parameters. To test for this possibility, a recursive analysis was undertaken on a base sample covering the period 1980:3 to 1993:4[12].

Beginning with Figure 6.1, the recursively calculated Trace test suggests five unit roots and two stationary relationships amongst the variables. Intuitively, the presence of a unit root can be rejected / accepted for components of the Trace statistic which lie above / below the critical value of one. Note that a critical value in excess of one is deemed more appropriate for VAR models which contain dummy variables, as in this case.

Figure 6.1: Recursive Analysis – Trace Test (Base Sample 1980:3 to 1993:4) [13]



Next looking at Figure 6.2, which tests the constancy of the known beta (assuming a rank of two), it can be seen that the model concentrating out all dynamic influences (R_t) is approximately one or below over the entire period, while there is some instability in the unconcentrated model (Z_t) especially around 1998.

[12] In order to test the sensitivity of the recursive analysis, a variety of time periods were modelled as the base sample. As the results are similar to those presented in Figures 6.1 to 6.3, the associated graphs are not reproduced here.

[13] The critical values presented in this figure are inappropriate as the model contains dummy variables. The presence of such dummies implies that the line at one should be higher.

Figure 6.2: Recursive Analysis – Test of Known Beta, rank = 2 (Base Sample 1980:3 to 1993:4)

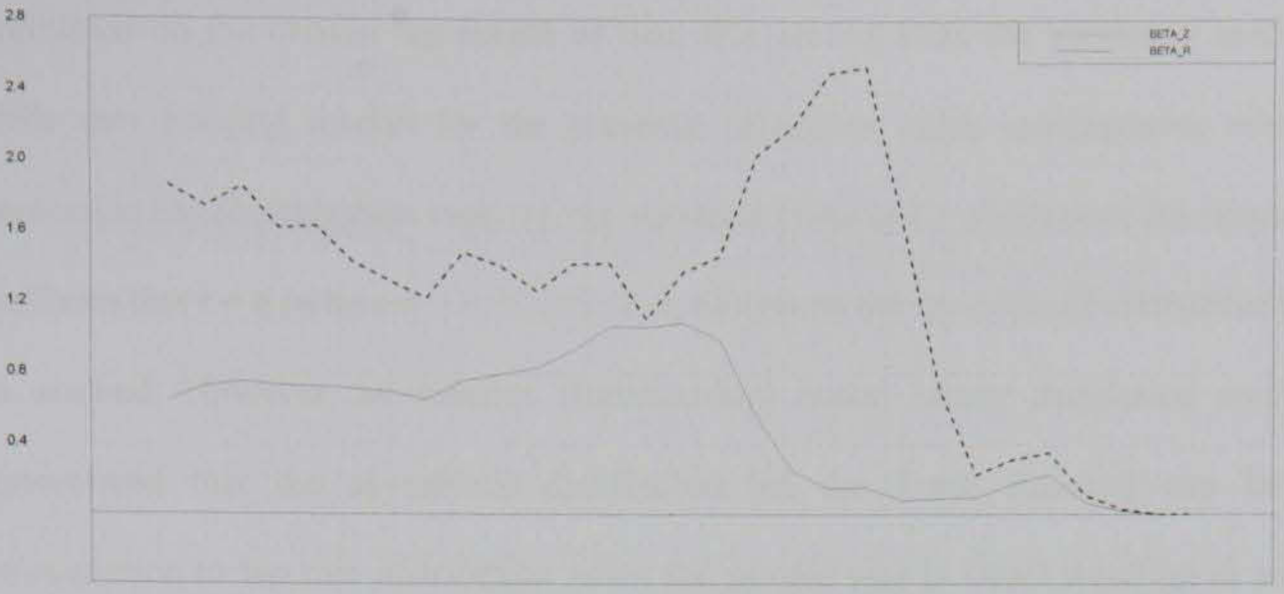
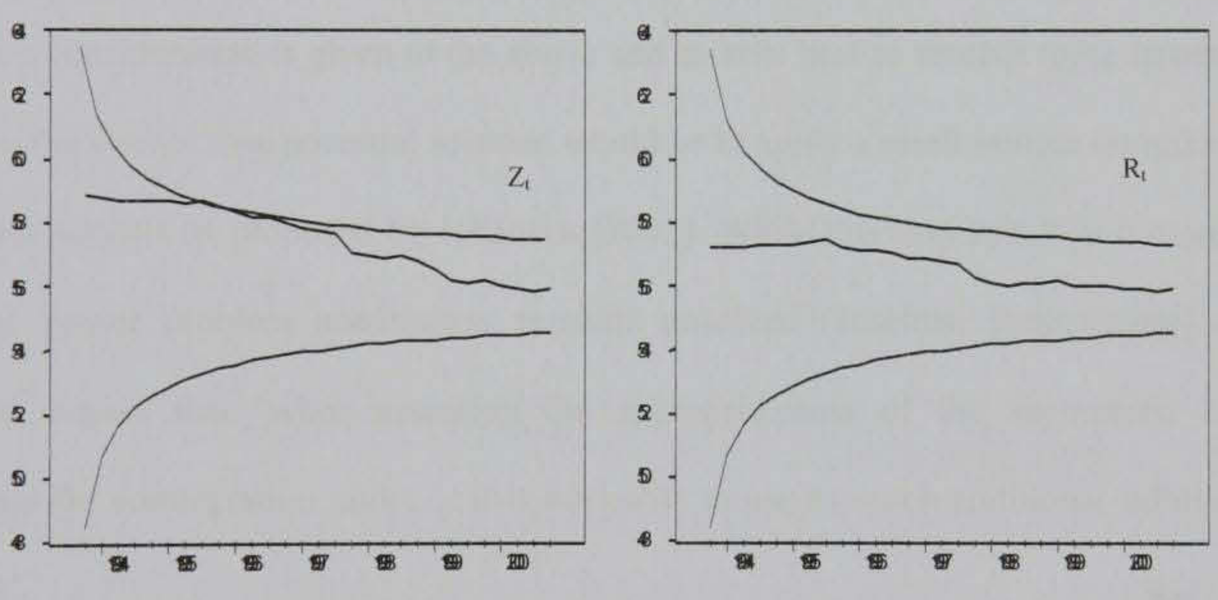


Figure 6.3: Recursive Analysis – The Log Likelihood Value (Base Sample 1980:3 to 1993:4)



Finally, with respect to the recursively calculated log likelihood value, it appears to just lie within the 95% confidence bands for the full model (Z_t), but as is typically the case, the calculated value is more stable for the concentrated (R_t) version. Combined, these tests allow for a tentative acceptance of the null hypothesis of constancy[14].

---

[14] It is also unlikely that the tests for autocorrelation would have been passed if the parameters of the model are not constant.

6.6. Determination of the Cointegrating Rank

Conditional on the chosen lag length of two, this section tests the estimated model of the Dublin new housing market for the presence of one or more cointegrating relations. In determining the cointegration rank (r), the standard Trace test - which tests the restricted null hypothesis that $r = q$ (where $q = 0, 1, 2, 3, 4, 5, 6$) against the unrestricted alternative that $r = 7$ - is applied. However, as Juselius (forthcoming) notes, 'many simulation studies have demonstrated that the asymptotic distribution [of the Trace statistic] can be a poor approximation to the true distribution when the sample size is small resulting in substantial size and power distortions'[15]. The latter refers to the fact that the data itself may not be particularly informative about a hypothetical long run relation.

Given the reasonably small sample of 84 observations employed in this study, it is appropriate that some consideration is given to the above and to how best to resolve these issues[16]. With regard to the former, one potential solution would be to apply a small sample correction to the Trace test statistic as proposed by Johansen (2002). While this may result in a more correct size, the power problem nonetheless remains unsolved (Juselius, forthcoming). Juselius therefore argues that 'when assessing the appropriateness of the asymptotic tables to determine the cointegration rank .... it is advisable to use as much additional information as possible'.

Accordingly, a range of informal criteria, namely the roots of the companion matrix, the t values of the alpha coefficients of the $r^{th}+1$ cointegrating vector, the graphs of the potential relations and economic theory, are used in tandem with the formal Trace test to motivate the choice of rank in the following. Tables 6.15 to 6.17 present the empirical findings.

---

[15] [ ] parenthesis inserted.

[16] Note that Juselius (forthcoming) views a sample of 50-70 observations as moderately sized. Applying this classification to the current sample of 84 observations implies that the latter is far from small.

Beginning with the Trace test, the results presented in Table 6.15 indicate the presence of two cointegrating vectors amongst the data at the 5% significance level. Next, Table 6.16 details the largest modulus roots of the companion matrix for ranks seven, four, three and two. Setting r = 7, four unit roots are evident implying three cointegrating relations. However, imposing a rank of three is clearly incorrect as the fifth root is equal to one indicating the need to move to a lower rank. As a sensitivity analysis, consider setting r = 4 and r = 2. The latter leaves a reasonably small root of 0.87 in the model whereas the former gives rise to an unacceptably large fourth root of one. Two therefore appears to represent an appropriate choice of rank.

Table 6.15: Testing for the Cointegrating Rank – The Trace Test

Eigenvalues	*Ho:r*	*p-r*	*Trace Statistic*	*95% Critical Value*
0.5996	0	7	176.74	123.04
0.4208	1	6	101.68	93.92
0.2559	2	5	56.91	68.68
0.2110	3	4	32.67	47.21
0.1258	4	3	13.23	29.38
0.0204	5	2	2.21	15.34
0.0064	6	1	0.52	3.84

Notes:
1. Ho: r = q (where q = 0, 1, 2, 3, 4, 5, 6)
2. The 95% critical values are taken from Table 15.3 in Johansen (1995)

Table 6.16: Testing for the Cointegrating Rank – Roots of the Companion Matrix

p-r	*r=7*	*r=4*	*r=3*	*r=2*
7	1.00	1.00	1.00	1.00
6	0.99	1.00	1.00	1.00
5	0.93	1.00	1.00	1.00
4	0.93	1.00	1.00	1.00
3	0.84	0.83	1.00	1.00
2	0.84	0.83	0.74	0.87
1	0.41	0.40	0.42	0.67

Notes:
1. p refers to the number of variables in the analysis. p-r = 7 unit roots corresponds to zero cointegration while p-r = 6 unit roots implies the presence of one cointegrating relation amongst the data series

Figure 6.4: The Cointegrating Vectors – Beta 1 and 2

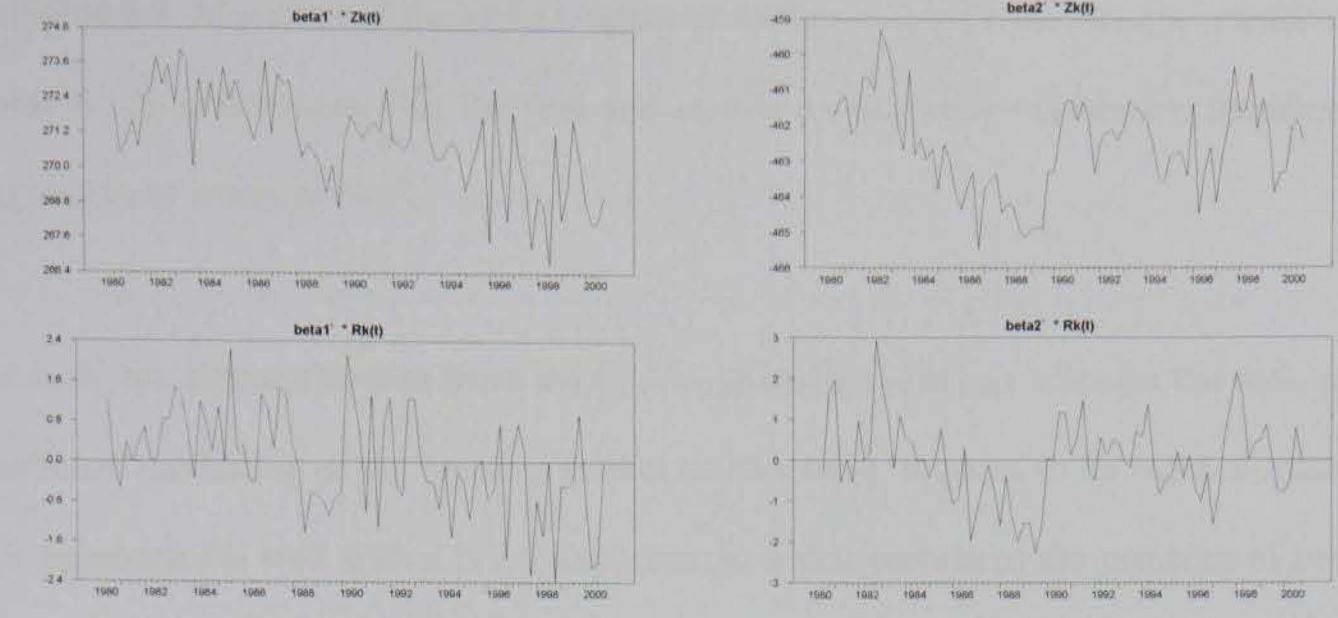


Table 6.17: Testing for the Cointegrating Rank – Unrestricted Estimates of the VAR(2) Model

Beta	*lognph*	*loginc*	*MR*	*logbc*	*logphs*	*loghf*	*logls*
1	-7.752	8.670	0.461	-19.608	22.903	10.672	1.460
2	2.655	8.593	0.437	7.908	-50.939	-5.972	11.656
3	-6.045	14.245	-0.033	-28.774	56.707	6.467	-22.752
4	4.840	-24.353	0.501	-10.878	113.173	-4.064	-24.209
5	1.174	-5.444	0.485	11.626	9.429	7.250	10.875
6	2.495	-8.635	-0.110	-43.470	62.060	1.237	-7.151
7	-9.581	15.743	0.096	-8.174	-9.946	1.951	4.967
Alpha							
dlognph	-0.017	-0.008	0.005	0.001	0.000	0.001	0.002
dloginc	-0.004	-0.001	-0.002	0.004	0.004	-0.000	0.001
dMR	-0.055	-0.299	0.011	0.047	-0.023	-0.013	-0.027
dlogbc	0.002	-0.001	-0.001	-0.000	-0.000	0.001	0.000
dlogphs	-0.000	0.000	0.000	-0.000	0.000	0.000	-0.000
dloghf	-0.029	0.038	-0.027	0.019	-0.019	0.005	0.001
dlogls	0.001	-0.000	0.002	0.003	-0.000	0.000	-0.000
Alpha T Values							
dlognph	-4.487	-2.278	1.415	0.185	0.091	0.165	0.589
dloginc	-2.400	-0.846	-1.177	2.545	2.223	-0.039	0.304
dMR	-0.979	-5.312	0.192	0.833	-0.411	-0.222	-0.482
dlogbc	2.501	-1.497	-1.677	-0.339	-0.401	1.062	0.264
dlogphs	-4.634	1.610	2.256	-0.216	1.283	0.662	-0.327
dloghf	-2.665	3.526	-2.505	1.750	-1.747	0.418	0.075
dlogls	1.716	-0.143	2.496	3.778	-0.430	0.374	-0.057

Notes:
1. Beta corresponds to the cointegrating vectors while alpha reflects the adjustment coefficients

A rank of two is also supported by an examination of the corresponding graphs as presented in Figure 6.4. Moreover, if the alpha t values of the unrestricted estimates are considered (see Table 6.17), it is evident that the first and second cointegrating vectors are significant and suggestive of a rank of two[17].

As such, the assumption that there exist two equilibrium relations amongst the data, and the continued modelling of the housing market on this basis, appears to be valid. Furthermore, this approach fits well with a priori expectations which postulated the presence of two such relations – one for the demand-side of the market and one representing supply.

6.7. Model Specific Data Properties

The following section examines a range of model specific data properties. The variables included in the estimated VAR(2) model are firstly tested for long run exclusion, and then for stationarity. Conditional on a rank of two, the former provides information as to whether the particular variable under consideration can or cannot be omitted from the cointegration space, while the latter determines if any of the series are individually stationary. A test for long run weak exogeneity is also undertaken. This imposes restrictions on alpha and allows for the fact that a variable may influence the long run stochastic path of other variables in the system without being influenced by them, that is, a 'no levels feedback' effect. The results of these three tests are reported below in Table 6.18.

Firstly, with respect to the hypothesis of long run exclusion, the null is rejected at the 5% significance level for most of the variables. The test fails to reject exclusion of the housing and land stock series but it is only the probability of 0.26 associated with the latter that is a

[17] With respect to the significance levels of the alpha coefficients, as their distribution is unknown a t value of roughly 3.5 (in absolute terms) is regarded as more appropriate than the standard rule of thumb figure of 2. In addition, note that the significance of a given beta vector requires a high t value for only one of the related alpha coefficients.

cause of concern. Moreover, caution is advised when implementing this test, as it may be the case that a relevant variable obtains an insignificant test value as a result of multicollinearity with other variables in the system (Juselius, forthcoming). Regressing the land series on the housing stock variable indicates a strong degree of collinearity. As such, a decision was made not to exclude the stock of land from the empirical analysis.

Table 6.18: Estimated VAR(2) Model – Model Specific Data Properties

Exclusion Test (Chi squared (2) = 5.99)						
lognph	*loginc*	*MR*	*logbc*	*logphs*	*loghf*	*logls*
22.78 (0.00)	6.56 (0.04)	23.66 (0.00)	10.63 (0.00)	5.18 (0.07)	31.61 (0.00)	2.72 (0.26)
Stationarity Test (Chi squared (5) = 11.07)						
lognph	*loginc*	*MR*	*logbc*	*logphs*	*loghf*	*logls*
27.04 (0.00)	29.47 (0.00)	21.73 (0.00)	33.49 (0.00)	36.90 (0.00)	38.42 (0.00)	39.38 (0.00)
Weak Exogeneity Test (Chi squared (2) = 5.99)						
lognph	*loginc*	*MR*	*logbc*	*logphs*	*loghf*	*logls*
19.87 (0.00)	4.64 (0.10)	20.56 (0.00)	7.07 (0.03)	16.91 (0.00)	10.77 (0.00)	1.90 (0.39)

Notes:
1. The above tests are conditional on a cointegrating rank of two
2. Respectively, the null hypothesis states that the variable under consideration should be excluded, is individually stationary and is weakly exogeneous
3. p values are given in brackets

Secondly, testing for stationarity reveals that none of the variables are individually stationary. The rejection of the null at the 5% significance level across all series confirms the findings of the earlier unit root tests, and is important as it rules out the possibility that one of the uncovered cointegrating relations can be accounted for by the presence of a stationary variable.

The final section of Table 6.18 reports the findings of a test for long run weak exogeneity, with income and land stock falling into this category. While the evidence is poor with respect to the former, the p value of 0.39 corresponding to the land variable points towards the acceptance of this variable as weakly exogenous. Given this, it proved possible (and indeed advantageous) to re-estimate the VAR(2) model as a partial system as discussed below.

6.8. The Partial Model

The weak exogeneity of the land variable implies that the series is influential in determining the long run path of the other variables in the housing model, but that they in turn do not influence it. Thus, the variable fails to adjust to the cointegrating relations and may be viewed as a common driving trend[18]. The overall implication of this finding is that valid inference on beta can be obtained from a six dimensional system comprised of new house prices, income, mortgage interest rates, the housing stock and household formation conditioned on the land variable (Juselius, forthcoming).

With respect to rank determination and partial systems, the asymptotic distribution of the rank test statistics for the latter differs from that of a full system and can consequently prove problematic (Johansen, 1995). However, as the rank has already been determined within the full system and the model tested for weak exogeneity, there is no further need to test for rank in the partial system (Hansen and Juselius, 1995). The standard approach is to proceed on the basis of the previously chosen rank, that of $r = 2$ in the case of the housing model.

Often the move to a partial system results in an improved statistical specification and can be particularly useful in eliminating insignificant dummies relating to the weakly exogenous variable (Hendry and Juselius, 2000). With respect to the estimated VAR(2) model, four of

---

[18] The refers to the fact that the unexplained variation of $\varepsilon_{logls,t}$ cumulates in the system so that $\Sigma_{i=1}^{t} \varepsilon_{logls,t}$ is one of the common trends. In this sense, the stock of land 'drives' the system.

the included dummy variables relate to the land series. By re-running the model conditioning on this series, it is possible to check if these dummy variables are excludable. Table 6.19 presents the associated short run matrices. While it appears that Dum_2ls is still extremely important, the remaining three land dummies are insignificant[19]. On balance it is felt that Dum_2ls should be retained as part of the analysis, but that Dum_1ls, Dum_3ls and Dum_4ls can be dropped. Thus, the partial system modelled from now on is as defined in equation (6.1), with the exclusion of the above dummies and conditioning on land. As in the full model, two lags are included.

Table 6.19: The Partial Model – Short Run Matrices

Variable	Dum_1bc	Dum_1ls	Dum_2ls	Dum_3ls	Dum_4ls	Dum_2bc	$Dummr$	Dum_3bc
dlognph	-0.001	-0.065	-0.002	-0.043	-0.042	0.005	0.024	0.045
dloginc	-0.028	-0.069	-0.026	-0.036	-0.013	0.055	0.011	0.027
dMR	0.327	-0.871	-5.301	-0.508	2.397	-0.873	2.364	-0.151
dlogbc	0.102	-0.034	0.028	0.006	-0.028	0.060	0.002	0.042
dlogphs	-0.002	-0.003	-0.004	-0.004	-0.001	-0.001	0.002	0.003
dloghf	0.060	-0.348	-0.214	-0.088	-0.119	0.057	0.093	0.087
T Values								
dlognph	-0.023	-0.915	-0.035	-0.750	-0.737	0.185	0.831	1.247
dloginc	-1.604	-2.142	-0.859	-1.393	-0.502	4.182	0.885	1.610
dMR	0.568	-0.817	-5.339	-0.592	2.781	-2.033	5.534	-0.276
dlogbc	11.860	-2.105	1.908	0.444	-2.191	9.351	0.302	5.166
dlogphs	-1.940	-1.592	-2.219	-2.411	-0.515	-1.281	2.139	2.693
dloghf	0.545	-1.701	-1.121	-0.534	-0.715	0.692	1.136	0.831

6.8.1. *Analysis of the Residuals of the Partial Model*

The residuals of the partial system pass the battery of diagnostic checks discussed previously, and in some respects, with greater ease than those of the full model. Tables 6.20 to 6.22 below illustrate this point.

[19] The rule of thumb that dummy variables should have an absolute t value of at least 3.73 in order to be considered significant is applied here. This is consistent with the approach adopted earlier when identifying outliers (see Section 6.5.1).

Despite the correlation matrix remaining largely unchanged, the multivariate statistics indicate an overall absence of serial correlation and again, find no evidence of heteroscedasticity in the model. However, the most noteworthy feature with respect to Table 6.20 is that the residuals now approximate a normal distribution. As such, the end result of the move to a partial system is an improvement in the statistical specification of the housing model and given this, it can be concluded that re-specifying the VAR(2) model to allow for the weak exogeneity of the land stock variable is both appropriate and desirable.

Table 6.20: The Partial Model – Multivariate Statistics

Test Type	*Test Statistic*	*Probability*
Information Criteria:		
SC	-37.526	
HQ	-40.055	
Serial Correlation:		
Lagrange Multiplier	LM(1) = 33.511	0.59
	LM(4) = 42.237	0.22
Vector AR(1-4) Test	F(144,177) = 1.135	0.21
Normality	14.138	0.29
Heteroscedasticity	F(546,268) = 0.469	1.00

Notes:

1. Ho: Residuals are respectively serially uncorrelated, multivariate normal and homoscedastic
2. Probabilities are distributed around a chi-squared distribution with 36 degrees of freedom for the LM test and 12 degrees of freedom when testing for normality
3. Where applicable, F versions of the above tests are used to correct for small sample size. The LM test for serial correlation and the normality test also include small sample corrections

Table 6.21: The Partial Model – Correlation Matrix

Variable	*dlognph*	*Dloginc*	*dMR*	*dlogbc*	*dlogphs*	*dloghf*
dlognph	1.000					
dloginc	0.437	1.000				
dMR	-0.297	-0.121	1.000			
dlogbc	0.273	0.073	-0.293	1.000		
dlogphs	-0.067	-0.004	0.075	-0.086	1.000	
dloghf	-0.035	0.025	-0.284	0.355	-0.115	1.000

Table 6.22: The Partial Model – Univariate Statistics

Variable	*Mean*	*Std Dev*	*Skewness*	*Kurtosis*
lognph	0.000	0.034	-0.026	2.881
loginc	0.000	0.016	0.029	2.618
MR	0.000	0.552	0.257	3.297
logbc	0.000	0.008	0.017	2.876
loghs	0.000	0.001	0.273	3.180
loghf	-0.000	0.098	-0.116	4.501

Variable	*Test Type*	*Test Statistic*	*Probability*
lognph	AR(1-4)	F(4,54) = 0.667	0.62
loginc		F(4,54) = 1.021	0.40
MR		F(4,54) = 1.003	0.41
logbc		F(4,54) = 1.527	0.21
logphs		F(4,54) = 0.441	0.78
loghf		F(4,54) = 0.359	0.84
lognph	Normality	0.118	0.94
loginc		0.104	0.95
MR		1.810	0.40
logbc		0.107	0.95
logphs		1.511	0.47
loghf		10.626	0.00***
lognph	ARCH (1-4)	F(4,50) = 1.243	0.30
loginc		F(4,50) = 0.718	0.58
MR		F(4,50) = 1.528	0.21
logbc		F(4,50) = 0.189	0.94
logphs		F(4,50) = 0.260	0.90
loghf		F(4,50) = 1.601	0.19

Notes:

1. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
2. Probabilities for the normality test are distributed around a chi-squared distribution with 2 degrees of freedom
3. *** denotes rejection of Ho at the 1% significance level

6.9. Long Run Identification[20]

Following rank determination and the re-specification of the model as a partial system, the next step is to ensure that the uncovered cointegrating relationships are unique and to impose some structure on the given vectors (a priori expectations suggest that one of these vectors should represent the demand-side of the housing market and the other supply). To this end, a

[20] For a more technical discussion of this issue, see Section 5.2.2.

just / exactly identified model is derived and compared with a series of restricted models using the likelihood ratio test as described in Johansen and Juselius (1994).

As a general rule, exact identification is achieved by imposing one normalisation and r-1 restrictions on each vector (Hendry and Juselius, 2000). In this case, r = 2 which implies one restriction on both vectors and an appropriate normalisation or re-scaling of each. As the first relation (beta 1) is intended as the demand relationship, it is normalised on the housing stock variable and excludes the measure of residential land availability - theoretically this variable should appear only on the supply-side of the model. The second cointegrating vector (beta 2) represents supply and as such, normalisation is on house prices with the stock of housing eliminated from this relation. Table 6.23 presents the just identified model.

Table 6.23: The Exactly Identified Model

Beta	*lognph*	*loginc*	*MR*	*logbc*	*logphs*	*loghf*	*logls*
1	-0.225	0.095	0.005	-0.537	1.000	0.365	0.000
2	1.000	-1.135	-0.146	2.377	0.000	-1.641	-2.538
Standard Errors of Beta							
1	0.047	0.095	0.003	0.155	0.000	0.049	0.000
2	0.000	0.369	0.025	1.062	0.000	0.305	0.435
Alpha							
1	0.125	-0.043	11.180	0.065	-0.011	-2.430	
2	0.091	0.011	1.763	-0.002	0.001	-0.153	
Standard Errors of Alpha							
1	0.192	0.098	3.153	0.047	0.005	0.593	
2	0.025	0.013	0.411	0.006	0.001	0.077	
Alpha T Values							
1	0.647	-0.440	3.546	1.370	-2.058	-4.098	
2	3.608	0.828	4.286	-0.248	1.096	-1.972	

Notes:

1. Log likelihood = 1137.68127
2. Beta corresponds to the cointegrating vectors while alpha reflects the adjustment coefficients

However, to lend to the interpretability of these vectors as demand and supply relationships, it is necessary to impose further over-identifying restrictions. Economic arguments would suggest that a valid starting point in this respect is the exclusion of the index of building costs from the demand relation, and the demographic and income variables from the supply-side of the model. As a consequence of dropping these variables, the coefficients of the remaining series are altered somewhat – particularly noteworthy are the larger coefficients of -1.145, 0.610 and 0.833 respectively on the house price, income and supply-side interest rate variables. Building on these findings, additional restrictions which impose unit coefficients on the house price and income series in the demand relation, and on the mortgage interest rate and building cost variables in the supply vector, are also tested. The final over-identified model is set out in Table 6.24[21].

Table 6.24: The Over-Identified Model

Beta	*lognph*	*loginc*	*MR*	*logbc*	*logphs*	*loghf*	*logls*
1	1.000	-1.000	0.092	0.000	1.000	-1.722	0.000
2	1.000	0.000	-1.000	-1.000	0.000	0.000	-10.628
Standard Errors of Beta							
1	0.000	0.000	0.013	0.000	0.000	0.192	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000	1.087
Alpha							
1	0.052	0.019	-0.833	-0.010	0.003	0.364	
2	0.018	0.003	0.132	-0.002	0.000	0.020	
Standard Errors of Alpha							
1	0.029	0.014	0.472	0.007	0.001	0.087	
2	0.004	0.002	0.072	0.001	0.000	0.013	
Alpha T Values							
1	1.822	1.340	-1.765	-1.480	3.263	4.200	
2	4.065	1.178	1.844	-1.608	3.847	1.521	

Notes:
1. Log likelihood = 1130.77036
2. Likelihood ratio test: Chi^2 (7) = 13.822 with a probability of 0.05
3. Beta corresponds to the cointegrating vectors while alpha reflects the adjustment coefficients

[21] While a wide range of additional hypotheses were considered and tested, only the chosen model is presented here.

More formally, the two long run relations uncovered for the Dublin new housing market (beta 1 and 2 in Table 6.24) may be written as follows:

Demand relation: logphs + lognph − loginc + 0.092 MR − 1.722 loghf

$$\text{logphs} = -\ \text{lognph} + \text{loginc} - 0.092\ \text{MR} + 1.722\ \text{loghf} + \varepsilon_t \quad (6.3)$$

Supply relation: lognph − MR − logbc − 10.628 logls

$$\text{lognph} = \text{MR} + \text{logbc} + 10.628\ \text{logls} + \varepsilon_t \quad (6.4)$$

where ε_t - I(0).

In analysing the results for the demand-side of the market, the equi-proportional relationship between income and the housing stock is noteworthy, as is the unit price elasticity of housing demand. As housing is a necessary good for most, a small coefficient (as in the unrestricted case) might be anticipated[22].

With respect to the interest rate variable, the negatively signed coefficient accurately captures the depressing effect of higher interest rates on house prices due to their depressing effect on demand. However, attention is drawn to the fact that the study uses the nominal mortgage interest rate as the interest rate variable. As such, it is possible to decompose this variable into a real interest rate component and a component that reflects inflationary expectations. If the real interest rate is assumed to be constant, then the negative sign on the interest rate coefficient may also reflect the depressing effect of inflation on housing demand due to repayment tilt[23]. Finally, as expected a priori, the sign of the coefficient on the demographic

[22] Kenny (1998) comments on a similar finding of a higher than expected long run price elasticity of demand (also equal to one) for the national housing market.
[23] This relates to the ability of inflation to 'tilt' forward the real burden of repaying mortgage debt. For a detailed discussion see Thom (1983).

variable implies a positive relationship between the demand for housing and household formation.

In terms of the supply relation, the unit coefficient on the building costs and interest rate series respectively indicate that increased material and labour costs are fully passed onto consumers, and that the opportunity cost of borrowing is an important factor in affecting the willingness of suppliers to provide housing. The magnitude of the latter effect is particularly striking, if somewhat unexpected[24].

Again, the nominal nature of the interest rate variable is worth mentioning as the positive sign on the associated coefficient can also be explained by the use of housing as a hedging instrument. In times of high uncertainty and inflation, housing acts as a hedge against future inflation and thus, as the nominal mortgage interest rate increases, the demand for housing increases. In this case, the effect of the nominal interest rate on housing demand is captured on the supply-side of the model as opposed to the more traditional demand-side.

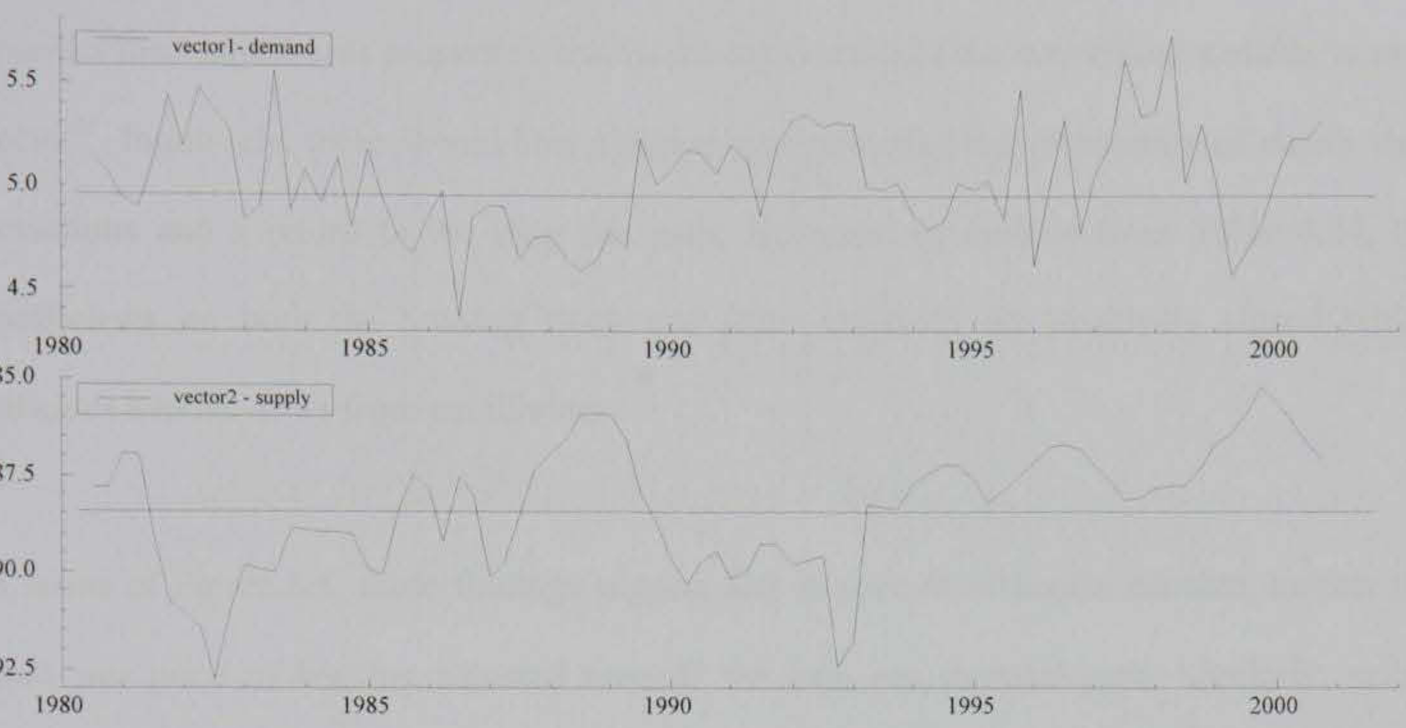
The last variable to play a role in the long run supply relationship is land. Interestingly, this series enters with a positive as opposed to a negatively signed coefficient and as such, the initial expectation that an increase in the stock of zoned housing land would translate into an expansion in supply and a subsequent fall in the price of housing does not materialise. A possible explanation, and one that would seem to fit well with on the ground opinion, is that of hoarding on the part of landowners. The failure to release residentially zoned land onto the market, as a strategy for artificially inflating the price, translates into lower housing supply and higher house prices. However, it should be remembered that while the land concerned may have been zoned for housing purposes, it might not be in receipt of services or planning permission, which would also result in a delay in its coming on stream.

---

[24] Note however that Kenny (1999) also finds a strong negative relationship between the long run supply of housing and interest rates (coefficient of 1.16).

Figure 6.5 depicts the identified long run demand and supply relationships. These seem to paint an accurate picture of the Dublin new housing market over the sample period.

Figure 6.5: Long Run Demand and Supply



Looking first at the demand vector, it appears that the stock of housing was below its desired level - the level consistent with the prevailing level of house prices, income, mortgage interest rates and household formation - from 1986-1989, most of 1993-1997 and during 1999. In periods such as these, increases in the housing stock, house prices or both would be needed to eliminate excess demand and restore the market to equilibrium (the horizontal line in Figure 6.5).

Secondly, the supply relation shows that from the end of 1987 to 1989, and from 1994 onwards, house prices were relatively high given the prevailing level of mortgage interest rates, building costs and land. Of particular interest is the consistency of this occurrence towards the latter end of the decade. In this situation, an upward adjustment in the supply of housing would be required to eliminate disequilibria. Alternatively, house prices could decline.

In addition, note that the supply vector presented in Figure 6.5 is close to an inverse, trend adjusted version of the mortgage interest rate series graphed in Figure 4.5. This reflects the magnitude of the interest rate coefficient in the identified supply relation.

Finally, before accepting the above relations as equilibrium relationships, it is essential to consider their adjustment properties, that is, the alpha value of the normalised variable in each vector[25]. Intuitively, these should bear a minus sign reflecting the elimination of steady state deviations and a return to the long run path. However, as evident from Table 6.24, the coefficients on both the housing stock and price variables are positively signed which indicates a move away from equilibrium.

In terms of Figure 6.5, these findings suggest that in periods of excess demand, neither the stock nor price of housing adjusted towards the long run demand path. Similarly, prices appear to have moved away from the long run supply path when above their equilibrium level, rather than adjusting towards it.

One explanation for this lack of error correction is that the demand and supply relationships identified above do not define equilibrium for the Dublin housing market[26]. Alternatively, it could be the case that adjustment behaviour in this marketplace is non-linear in nature, as opposed to linear. Hence, prior to accepting a hypothesis of no cointegration, this latter possibility should be considered.

---

[25] Again, the sequential manner in which the long run relations and their adjustment properties are discussed in this section is only for convenience.

[26] For example, the household formation variable could be viewed as the equilibrating force in the demand relation.

6.10. Conclusion

This chapter sought to undertake a long run analysis of the market for new, private sector housing in Dublin over the period 1980-2000. In order to achieve this objective econometric modelling adopted the Johansen cointegration technique, applying it to a dataset comprised of new house prices, income, mortgage interest rates, building costs, housing stock, household formation and the stock of land. A priori, the existence of two long run relations amongst these variables was postulated - one for the demand-side of the market and one representing supply.

In line with expectations, the results identify the presence of two cointegrating vectors. The demand relation, which is normalised on the housing stock variable, shows that the stock of housing was below its desired level – that is the level consistent with the prevailing level of house prices, income, mortgage interest rates and demographics – over the time periods 1986-1989, 1993-1997 and 1999. The supply vector is normalised on house prices and suggests that from the end of 1987 to 1989, and from 1994 onwards, the price of housing exceeded its equilibrium relationship with mortgage interest rates, building costs and land.

However, an analysis of the adjustment properties of the above relations reveals a failure on the part of both the stock and price of housing to adjust to steady state deviations and restore the market to equilibrium. This lack of error correction is of concern as it implies that the identified demand and supply relationships may not in fact define equilibrium for the Dublin housing market.

Given these findings, a natural next step is to consider in greater detail the manner in which the Dublin housing market adjusts. In doing so, the possibility of non-linear as well as linear adjustment behaviour should be examined. Chapter 7 takes up this gauntlet.

Chapter 7:

An Analysis of Adjustment in the Dublin Housing Market

7.1. Introduction

Building on the preceding analysis, this chapter seeks to examine adjustment behaviour in the Dublin market for new private sector housing. Knowledge of the adjustment mechanisms inherent in the marketplace acts as a guide for policy-making and given this, it seems appropriate to examine the manner in which, and indeed if, deviations from the identified steady state paths are eliminated so as to facilitate a return to the equilibrium level. In the context of this study, the response of the price variable to such deviations is of particular interest. Thus, the key aim of the present chapter is to model short run house price determination in Dublin within an adjustment type framework. In order to achieve this objective, error correction models of both a linear and non-linear nature are developed and applied to the time series dataset described in Chapter 4. Furthermore, to enable a direct comparison between the two models, the linear analysis moves from the VAR framework to a conditional single equation. This approach has the added advantage of allowing for additional regressors, namely government interventionist policies, to be included as part of the dynamic specification.

The remainder of the chapter is organised as follows: Section 7.2 begins by presenting the baseline statistical model. A priori expectations with respect to the signs and characteristics of the coefficients of the model are outlined in Section 7.3, while Sections 7.4 and 7.5 respectively, estimate a linear and non-linear version of the proposed specification and discuss subsequent empirical findings[1].Sec tion 7.6 concludes.

7.2. The Statistical Model

In terms of dynamic behaviour, housing market theory postulates the existence of an inelastic supply curve. This notion of fixed short run supply arises from the manner in which the stock of housing is assumed to evolve, that is, the current stock of private housing is related to the stock in the last period (adjusted for depreciation) and to net additions (new private sector completions and Local Authority sales)[2]. However, as such additions are typically very small relative to the existing level of the stock and are to a considerable extent 'predetermined by the level of construction in progress' (Hendry, 1984), it follows that the stock at time t-1 may be viewed as the fixed supply of housing in the very short run. In this case, changing demand conditions drive house price determination[3].

From a modelling perspective, such a supply curve implies that the dynamic housing market is best represented by an inverted demand equation whereby house prices are regressed on a range of demand-side variables taking the stock of housing as given. Nesting this approach within a framework that allows for the adjustment of prices to deviations from the long run demand and supply relationships gives rise to the following error correction model:

$$\Delta \text{lognph}_t = \beta_0 + \Sigma_{i=0}^{4}\beta_{1i}\,\Delta \text{lognph}_{t-i} + \Sigma_{i=0}^{4}\beta_{2i}\,\Delta \text{loginc}_{t-i} + \Sigma_{i=0}^{4}\beta_{3i}\,\Delta \text{MR}_{t-i} + \Sigma_{i=0}^{4}\beta_{4i}\,\Delta \text{logphs}_{t-i}$$

$$+ \Sigma_{i=0}^{4}\beta_{5i}\,\Delta \text{loghf}_{t-i} + \Sigma_{i=0}^{4}\beta_{6i}\,\Delta \text{CGT}_{t-i} + \beta_7\,\text{Dumrpt}_t + \beta_8\,\text{Dumib}_t$$

$$+ \beta_9\,\text{Dums2327}_t + \beta_{10}\,\text{ecm}^{D}_{t-1} + \beta_{11}\,\text{ecm}^{S}_{t-1} + \beta'_{q}\text{Dq}_t + \varepsilon_t \qquad (7.1)$$

[1] Econometric modelling undertaken in the course of this chapter was carried out in GiveWin 2.20 / PcGive 10.0, EViews 3.1 and JMulTi 3.11.

[2] Equation (4.1) provides a formal definition of the inter-temporal evolution of the housing stock.

[3] Please refer back to Section 2.3 for a more detailed discussion.

where[4]:

nph = real new house prices;

inc = real personal disposable income per capita;

MR = nominal mortgage interest rates;

phs = the stock of private housing;

hf = household formation;

CGT = capital gains tax rates;

Dumrpt = a dummy = 1 for the years in which property taxes applied and 0 otherwise;

Dumib = a dummy = 1 for the years in which interest on borrowings was tax deductible and 0 otherwise;

Dums2327 = a dummy = 1 for the years in which 'Section 23/27' relief covered all rental income and 0 for the years in which it either did not apply or related solely to rental income from the residential property in question;

ecm^D = an error correction term capturing deviations from the long run demand relation;

ecm^S = an error correction term capturing deviations from the long run supply relation[5];

Dq = a vector of seasonal dummies;

ε = the error term;

Δ = the first difference operator.

The above baseline specification underlies the empirical estimations in Sections 7.4 and 7.5. In each case, the sample (which has been adjusted to allow for four lags of the variables) begins in the second quarter of 1981 and ends in the last quarter of 2000.

---

[4] Empirical studies typically include contemporaneous variables in error correction models of this type. See, for example, McQuinn (2004).

[5] These error correction terms are derived from the demand and supply relations presented in Table 6.24.

7.3. A Priori Expectations

As highlighted in Chapters 3 and 6, economic theory and precedents set by relevant studies are useful when seeking to identify potential explanatory variables. With respect to the modelling of dynamic price determination in the Dublin market for new private sector housing, it should be noted that many of the variables included in equation (7.1) follow directly from the long run analysis, with additional regressors attempting to measure the effectiveness of government interventionist policies in the housing market and to capture the impact of deviations from the long run demand and supply relations on short run price determination.

Ceteris paribus, a priori expectations regarding the signs and characteristics of the coefficients of the model propose that current and lagged changes in income and household formation, along with the dummy variables accounting for 'Section 23/27' relief and the tax deductibility of interest on borrowings, should put upward pressure on Dublin house prices. Similarly, a positive coefficient on the lagged house price series is anticipated as past increases often lead to self-fulfilling expectations of further price rises. On the other hand, changes in current and lagged mortgage interest rates and capital gains taxation, as well as an outward shift of the stock of housing and the property tax dummy, are expected to exhibit a negative relationship with price inflation in the Dublin market.

Lastly, the coefficients on the two error correction terms should bear a minus sign reflecting the elimination of steady state deviations and the return to the long run path. The corresponding size of these coefficients quantifies the speed at which equilibrium in the housing market is restored. Intuitively, a negative demand error correction term implies that if the stock of housing is in excess of its equilibrium level given prevailing house prices, income, mortgage interest rates and demographics, downward price adjustment will occur so as to ensure that the existing dwelling stock is willingly held. Such negative error correction should also be evident in response to deviations on the supply-side of the market. For example, a rise in the price of housing above its equilibrium given the prevailing levels of mortgage interest rates, building costs and land (as occurred in the Dublin market between 1987 and 1989, and from 1994 onwards), necessitates downward price adjustment in order to restore the stationary relationship between house prices and suppliers' costs.

7.4. The Linear Error Correction Model

Equation (7.1) can be readily estimated using the standard ordinary least squares technique[6]. Initial estimation takes into account the quarterly nature of the data by imposing four lags on the model. The resulting empirical estimates are presented in Table 7.1.

The presence of a number of insignificant explanatory variables suggests moving away from this general model and adopting a more specific approach to the determination of the dynamic lag structure. As such, a simple specification search is employed, beginning with the removal of the regressor for which the t statistic of the corresponding parameter estimate is lowest in absolute value (excepting the error correction terms). Subsequent model reduction is achieved through successive re-estimation and the elimination of the most insignificant variable per iteration. At each stage, the removal of the potentially redundant regressor is formally tested. If the null hypothesis of a zero coefficient cannot be rejected, and the diagnostics of the resultant model are sound, the variable is dropped from the analysis. Table 7A in the appendix to this chapter details the step-by-step empirical outcomes associated with the specification search. For convenience, the end model is reproduced below in Table 7.2 and in equation format in equation (7.2).

With regard to overall model selection, attention is paid to the adjusted R squared statistic, to the information provided by the Akaike (AIC), Schwartz (SC) and Hannan-Quinn (HQ)

[6] Note that standard OLS assumes weak exogeneity of the explanatory variables.

criteria, and to the results of a range of misspecification tests[7]. The former indicates that equation (7.2) represents a good fit while the values of the Schwartz and Hannan-Quinn information criteria are also smallest for this specification. Additional support in favour of the model is given by the outcome of the exclusion / redundant variable test, in that, the low elimination probability associated with the regressor $Dums2327_t$ suggests that the general model has been sufficiently reduced in size and complexity.

[7] Section 7.4.1 presents the findings of these tests for the chosen model.

Table 7.1: The Linear Error Correction Model – Initial Estimation

Dependent Variable: Δ $lognph_t$

Independent Variable	*Coefficient*	*Standard Error*	*T Value*	*Probability*
Δ $lognph_{t-1}$	-0.319	0.153	-2.09	0.04**
Δ $lognph_{t-2}$	0.100	0.158	0.63	0.53
Δ $lognph_{t-3}$	0.010	0.151	0.06	0.95
Δ $lognph_{t-4}$	-0.029	0.143	-0.20	0.84
Δ $loginc_t$	0.622	0.215	2.90	0.01***
Δ $loginc_{t-1}$	0.279	0.230	1.21	0.23
Δ $loginc_{t-2}$	-0.233	0.254	-0.92	0.36
Δ $loginc_{t-3}$	-0.038	0.231	-0.17	0.87
Δ $loginc_{t-4}$	-0.091	0.259	-0.35	0.73
Δ MR_t	-0.005	0.006	-0.92	0.36
Δ MR_{t-1}	-0.007	0.006	-1.19	0.24
Δ MR_{t-2}	0.015	0.006	2.33	0.03**
Δ MR_{t-3}	-0.011	0.006	-1.81	0.08*
Δ MR_{t-4}	0.008	0.006	1.22	0.23
Δ $logphs_t$	4.536	4.320	1.05	0.30
Δ $logphs_{t-1}$	-7.699	4.216	-1.83	0.08
Δ $logphs_{t-2}$	-2.214	4.355	-0.51	0.61
Δ $logphs_{t-3}$	-2.972	4.264	0.70	0.49
Δ $logphs_{t-4}$	2.215	4.349	0.51	0.61
Δ $loghf_t$	0.023	0.037	0.62	0.54
Δ $loghf_{t-1}$	0.053	0.081	0.66	0.52
Δ $loghf_{t-2}$	0.046	0.071	0.66	0.52
Δ $loghf_{t-3}$	-0.050	0.059	-0.85	0.40
Δ $loghf_{t-4}$	-0.032	0.041	-0.79	0.44
Δ CGT_t	-0.002	0.002	-1.29	0.21
Δ CGT_{t-1}	-0.002	0.001	-1.64	0.11*
Δ CGT_{t-2}	-0.001	0.002	-0.48	0.64
Δ CGT_{t-3}	-0.001	0.001	-0.80	0.43
Δ CGT_{t-4}	-0.001	0.001	-1.03	0.31
$Dumrpt_t$	-0.024	0.014	-1.77	0.08*
$Dumib_t$	0.032	0.019	1.67	0.10*
$Dums2327_t$	0.015	0.013	1.16	0.25
ecm^D_{t-1}	0.002	0.058	0.03	0.98
ecm^S_{t-1}	0.015	0.008	2.04	0.05
Dq_{1t}	0.008	0.013	0.61	0.55
Dq_{2t}	0.021	0.014	1.54	0.13
Dq_{3t}	0.017	0.012	1.40	0.17
Constant	1.333	0.472	2.82	0.01***

Equation standard error	0.030
R squared	0.744
Adjusted R squared	0.513
Akaike information criteria	-6.725
Schwartz criteria	-5.585
Hannan-Quinn criteria	-6.268

Notes:

1. * (**) (***) denotes significant at the 10% (5%) (1%) level

Table 7.2: The Linear Error Correction Model – Final Model

Dependent Variable: Δ lognph$_t$

Independent Variable	*Coefficient*	*Standard Error*	*T Value*	*Probability*
Δ lognph$_{t-1}$	-0.480	0.102	-4.73	0.00***
Δ loginc$_t$	0.769	0.151	5.09	0.00***
Δ loginc$_{t-1}$	0.476	0.171	2.78	0.01***
Δ MR$_{t-2}$	0.010	0.004	2.57	0.01***
Δ MR$_{t-3}$	-0.008	0.004	-2.08	0.04**
Δ logphs$_{t-1}$	-9.535	3.109	-3.07	0.00***
Δ logphs$_{t-2}$	-7.335	3.121	-2.35	0.02**
Δ loghf$_{t-1}$	0.133	0.039	3.45	0.00***
Δ loghf$_{t-2}$	0.110	0.028	3.94	0.00***
Δ CGT$_{t-1}$	-0.002	0.001	-1.93	0.06**
Dumrpt$_t$	-0.027	0.010	-2.75	0.01***
Dumib$_t$	0.035	0.013	2.64	0.01***
Dums2327$_t$	0.017	0.010	1.78	0.08*
ecm^D_{t-1}	0.075	0.027	2.84	0.01***
ecm^S_{t-1}	0.022	0.004	6.10	0.00***
Dq_{2t}	0.014	0.008	1.83	0.07*
Dq_{3t}	0.014	0.008	1.84	0.07*
Constant	1.600	0.239	6.68	0.00***
Equation standard error		0.027		
R squared		0.690		
Adjusted R squared		0.603		
Akaike information criteria		-7.038		
Schwartz criteria		-6.498		
Hannan-Quinn criteria		-6.823		

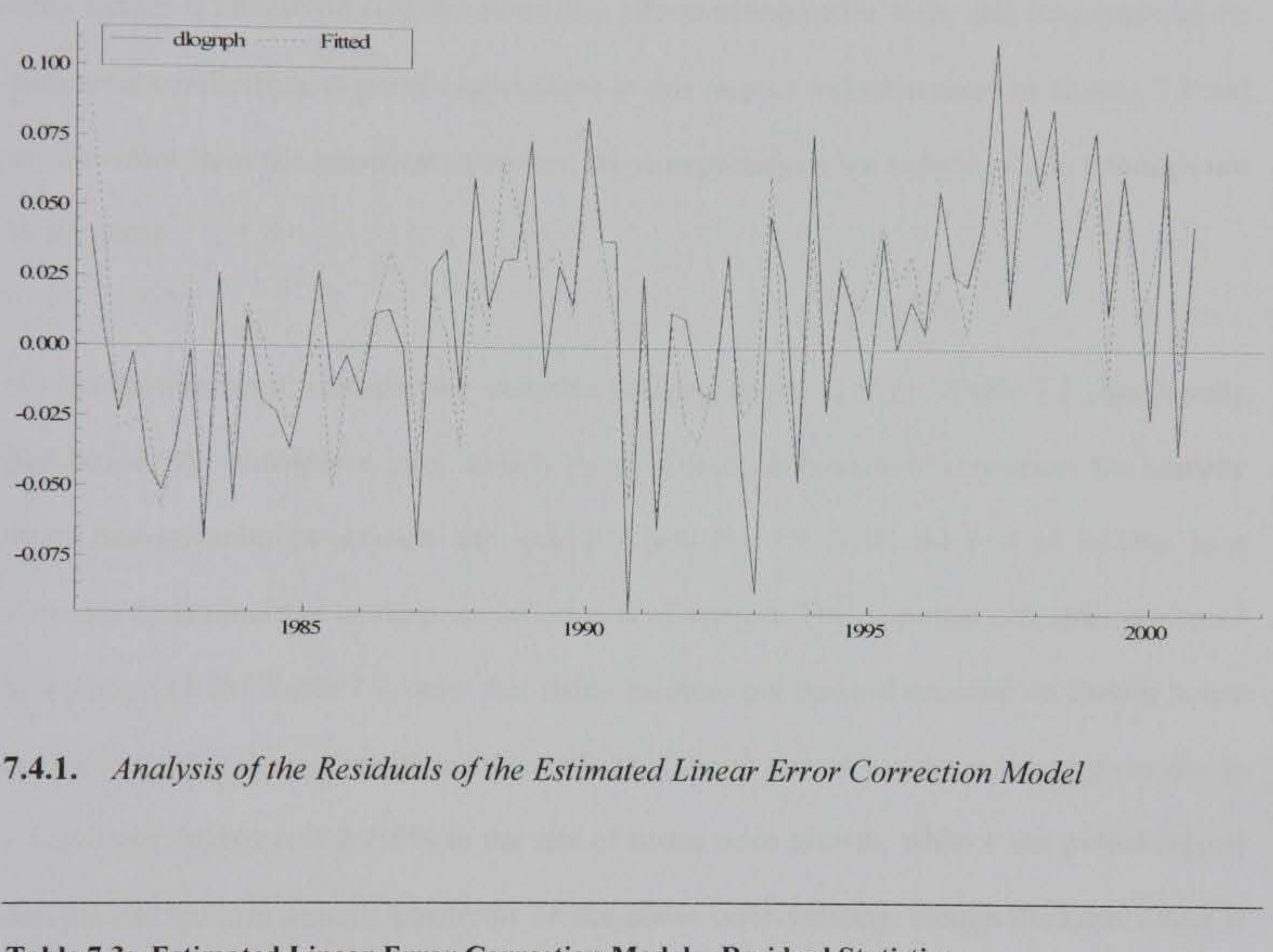
Notes:
1. * (**) (***) denotes significant at the 10% (5%) (1%) level

More formally, the selected linear error correction model is written as:

$$\begin{aligned}\Delta \text{lognph}_t = {} & 1.600 - 0.480\,\Delta \text{lognph}_{t-1} + 0.769\,\Delta \text{loginc}_t + 0.476\,\Delta \text{loginc}_{t-1} + 0.010\,\Delta \text{MR}_{t-2} \\ & - 0.008\,\Delta \text{MR}_{t-3} - 9.535\,\Delta \text{logphs}_{t-1} - 7.335\,\Delta \text{logphs}_{t-2} + 0.133\,\Delta \text{loghf}_{t-1} \\ & + 0.110\,\Delta \text{loghf}_{t-2} - 0.002\,\Delta \text{CGT}_{t-1} - 0.027\,\text{Dumrpt}_t + 0.035\,\text{Dumib}_t \\ & + 0.017\,\text{Dums2327}_t + 0.075\,\text{ecm}^{D}_{t-1} + 0.022\,\text{ecm}^{S}_{t-1} \\ & + 0.014\,\text{Dq}_2 + 0.014\,\text{Dq}_3 + \hat{\varepsilon}_t \end{aligned} \quad (7.2)$$

Figure 7.1 below provides an overview of the fitted and actual values of the dependent variable. As can be seen from the graphical representation, the estimated linear error correction model captures the general trend of the series reasonably well, though its tendency to imprecisely fit turning points is reflected in the model's relatively poor explanatory power of 0.60.

Figure 7.1: The Estimated Linear Error Correction Model – Actual and Fitted Values



7.4.1. *Analysis of the Residuals of the Estimated Linear Error Correction Model*

Table 7.3: Estimated Linear Error Correction Model – Residual Statistics

Test Type	*Test Statistic*	*Probability*
Serial Correlation:		
AR(1-4) Test	F(4,57) = 0.685	0.61
AR(1) Test	F(1,60) = 0.048	0.83
Normality	1.66	0.44
Autoregressive Conditional Heteroscedasticity:		
ARCH (1-4) Test	F(4,53) = 0.870	0.49

Notes:
1. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
2. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test

Table 7.3 reports the outcome of a range of diagnostic tests imposed on the residuals of equation (7.2). The results imply a statistically well-behaved specification in that no evidence of serial correlation, non-normality or autoregressive conditional heteroscedasticity is provided.

7.4.2. *Analysis of the Coefficients of the Estimated Linear Error Correction Model*

This section is concerned with the economic interpretation of the signs and magnitude of the parameter coefficients. A priori expectations in this respect were discussed in Section 7.3 and as is evident from the empirical estimates, these expectations are upheld in most, though not in all, cases.

As the starting point, consider the variables entering equation (7.2) / Table 7.2 significantly and bearing the anticipated sign, namely those of income, household formation, the housing stock and government taxation and incentive policies. Of these, the role of income as a dynamic determinant of house price inflation is of interest. The empirical estimates presented in equation (7.2) / Table 7.2 show that rising incomes put upward pressure on Dublin house prices. For example, a percentage increase in the growth rate of income at time t gives rise to a significant increase of 0.769% in the rate of house price growth, while a one period lagged income change also impacts positively on the house price variable, though the latter effect is smaller. Similarly, a positive relationship holds between price inflation in the Dublin market and household formation. On the contrary, the signs of the housing stock parameter estimates indicate a negative impact on house price growth of an outward shift in the short run supply curve. The magnitude of the estimated coefficients serves to highlight the strength of the negative relationship between the two series.

The estimated linear model also includes a range of variables designed to capture the effectiveness of government taxation and incentive policies vis-à-vis investors in the housing market. As far as changes in the rate of capital gains tax are concerned, the minus sign on the associated coefficient implies that a one unit increase at time t-1 induces a fall of 0.002% in the growth rate of Dublin prices. In addition, the 'Section 23/27' relief, the tax deductibility of interest on borrowings and the property tax dummy variables all enter equation (7.2) / Table 7.2 significantly. The positive coefficient of the 'Section 23/27' relief dummy (where such relief was granted for expenditure incurred in the course of providing rented residential accommodation) suggests that the introduction and extension of this offsetting relief, from rental income arising solely from the property in question to all rental income, has as expected, been effective in stimulating both investor demand and house price inflation in the Dublin market. Equally, the fact that interest payable on loans used to purchase residential accommodation for rental purposes can be written off in full when calculating tax liabilities, is found to positively impact on house price inflation. On the other hand, residential property tax appears to have a dampening effect on the growth rate of house prices. There would therefore appear to be some scope for manipulating these policy tools so as to moderate instances of price inflation in the Dublin housing market.

Of further importance are the contradictory coefficient signs exhibited by lagged changes in the mortgage interest rate variable. At period t-3, the parameter estimate for this variable behaves as postulated - a unit rise in interest rates gives rise to a significant fall of 0.008% in the growth rate of Dublin house prices. However, the significant and positively signed coefficient of the variable at period t-2 is suggestive of the opposite response. In offering an explanation for this occurrence, the nominal nature of the interest rate series comes into play, as does the use of housing as a hedging instrument. In times of high uncertainty and inflation, housing acts as a hedge against future inflation and thus, as the nominal mortgage interest rate rises, the demand for housing increases putting upward pressure on prices. The scale of the parameter estimates for the lagged interest rate changes in the Dublin housing market imply that, in the short run, the latter positive effect slightly outweighs the more traditional negative price response.

In contrast to the above, the parameter estimates for the demand and supply-side error correction terms in equation (7.2) / Table 7.2 do not adhere to a priori expectations of negative signs. Instead, both terms are significant and positive which indicates upward price adjustment and the movement of the market away from the long run path, as opposed to the anticipated elimination of steady state deviations and a return to equilibrium. This lack of error correction has serious implications. It implies that the long run demand and supply relations are not informative on prices or alternatively, that prices move in the wrong direction.

In addition, the negative coefficient on the lagged house price series is somewhat unexpected. As past increases in house prices typically lead to expectations of further price rises, a positive sign had been anticipated a priori. A possible explanation for the opposite occurrence is non-linear behaviour on the part of Dublin house prices. In particular, a situation could be envisaged whereby a rise in the price of housing leads to self-fulfilling expectations of further price increases up to a certain threshold. After that point is reached, the relationship between current and lagged house price changes might be expected to turn negative. Giussani and Hadjimatheou (1990) report such a finding in respect of UK house prices.

Lastly, note that the significance of the constant term in equation (7.2) / Table 7.2 suggests drift. Setting the various explanatory variables to zero, a positive drift parameter implies that the expected value of current house prices is growing over time, while a negative figure indicates that the expected value is falling.

Overall, the linear error correction model estimated and discussed in this section is well specified from a statistical perspective and for the most part, is comprised of correctly signed and significant explanatory variables. However, the reasonably poor fit of equation (7.2), combined with the apparent lack of error correction suggests that there is some room for improvement.

7.5. The Non-Linear Error Correction Model

The failure of house prices to adjust to deviations from the long run demand and supply relations in equation (7.2) suggests that these relationships may not in fact define equilibrium for the Dublin housing market. However, before accepting a hypothesis of no cointegration, it would seem both valid and interesting to test for non-linear adjustment, as an alternative to the linear behaviour assumed to date. Accordingly, this section seeks to model short run price determination in the Dublin market for new private sector housing within a non-linear framework.

With regard to the selection of an appropriate non-linear specification, the smooth transition regression (STR) model is considered appropriate in a housing context. Given the range of individuals and firms active in this market, a prompt and uniform response on the part of all economic agents to news requiring action is unlikely. From an aggregate viewpoint, this implies that a model which allows for a smooth transition between regimes is more suitable than a discrete switching alternative. As such, the following section is concerned with modelling non-linear error correction in the Dublin housing market using the STR model as a tool.

Chapter 5 provides a detailed overview of the STR methodology and its properties; therefore, the basic concept is only briefly outlined below[8]. Intuitively, the STR model allows for two extreme regimes in which the transition from one to the other is a function of a continuous transition variable s_t. At time t, the type of switching behaviour that arises is determined by the definition of the associated transition function $G(s_t; \gamma, c)$, where G is continuous and typically bounded between zero and one, γ is the slope parameter indicating the smoothness of the regime change and c is the location parameter identifying the point at which transition

[8] A thorough treatment of the family of smooth transition regression models is also given in Granger and Teräsvirta (1993) and Teräsvirta (1998).

occurs. A transition function that takes the logistic form and monotonically increases as a function of s_t is labelled an LSTR1 model while a simple non-monotonic alternative is the LSTR2 model. In addition, an exponential ESTR model may be defined.

In terms of modelling house price inflation, the LSTR1 model is of particular interest given that it is capable of characterising asymmetric behaviour in the housing market. In this model, the two regimes are associated with small and large values of the transition variable s_t relative to c and thus, the model can describe 'processes whose dynamic properties are different in expansions from what they are in recessions, and the transition from one extreme regime to the other is smooth' (Teräsvirta , 2004).

Table 7.4: Testing Linearity against STR – P Values

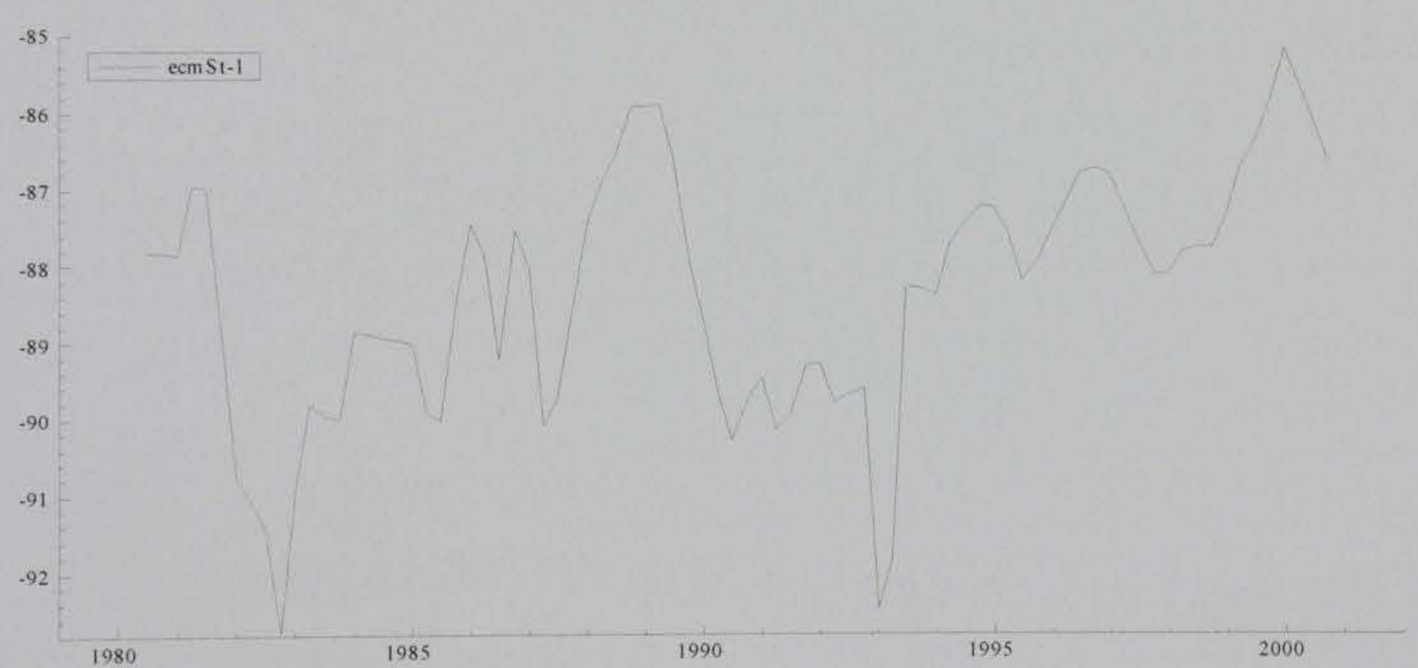
	Linearity Test			
Transition Variable	*F*	F_4	F_3	F_2
Δ lognph$_{t-1}$	0.729	0.648	0.646	0.413
Δ loginc$_t$	0.789	0.943	0.716	0.268
Δ loginc$_{t-1}$	0.125	0.547	0.350	0.310
Δ MR$_{t-2}$	0.552	0.285	0.887	0.338
Δ MR$_{t-3}$	0.380	0.443	0.904	0.860
Δ logphs$_{t-1}$	N/A	N/A	0.229	0.616
Δ logphs$_{t-2}$	N/A	N/A	0.581	0.049
Δ loghf$_{t-1}$	0.564	0.168	0.913	0.584
Δ loghf$_{t-2}$	0.446	0.178	0.493	0.708
Δ CGT$_{t-1}$	N/A	N/A	N/A	0.488
ecm$^D_{t-1}$	0.079	0.130	0.566	0.048
ecm$^S_{t-1}$	0.000	0.000	0.902	0.050

Notes:
1. All variables in equation (7.2) enter the linear part of the STR model but only the error correction terms (ecm^D_{t-1} and ecm^S_{t-1}) and Δ lognph$_{t-1}$are assumed to enter non-linearly
2. Ho: the model is linear
3. N/A indicates that it was not possible to compute p values due to matrix inversion problems
4. F_4, F_3 and F_2 refer to the corresponding F versions of the null hypotheses H_{04}, H_{03} and H_{02} discussed in Section 5.3

From an applied econometric perspective, the first step in smooth transition regression modelling is to test if the estimated linear error correction model - equation (7.2) – contains non-linearity of the STR type[9]. A priori, it is assumed that all variables enter the STR model linearly but that only lagged house price changes and the demand and supply error correction terms enter the non-linear part of the specification[10]. Furthermore, as theory does not offer any guidance concerning the choice of transition variable, the linearity test is repeated allowing each of the variables in turn to act as the transition variable. Table 7.4 presents the outcomes.

In interpreting these results, consider the column labelled *F* in the above table. The corresponding p values indicate that the strongest case for rejecting the null of linearity arises when the supply-side error correction term, ecm^{S}_{t-1}, is taken to be the transition variable. Figure 7.2 presents a graphical illustration of this series.

Figure 7.2: The Transition Variable



[9] Note that in common with many other non-linear models, the STR model is only identified under the alternative and not under the null hypothesis of linearity. Fortunately in the STR case, this identification problem can be circumvented as discussed in Section 5.3, Teräsvirta (1998) and Granger and Teräsvirta (1993).

[10] These restrictions are justified on the grounds that past changes in house prices and the error correction terms are the only variables in equation (7.2) to exhibit any evidence of counterintuitive behaviour. As explicit modelling of non-linearity could potentially explain such behaviour, these variables are included in the non-linear part of the STR model.

Building on this finding of STR type non-linearity and the selection of ecm^S_{t-1} as the transition variable, it is necessary to choose an appropriate model type. In doing so, the remaining columns in Table 7.4 come into play, as does the test sequence detailed in Section 5.3. The latter states that if the F_3 test yields the strongest rejection, then either the ESTR or LSTR2 model should be selected. Otherwise, the LSTR1 model is chosen. Referring back to Table 7.4, the strength of the rejection p value associated with the F_4 test suggests that an LSTR1 model best describes the nature of the non-linearity present in the Dublin housing market.

Estimation of the parameters of the LSTR1 model is carried out using conditional maximum likelihood. The algorithm used for the purpose of numerically maximising the log likelihood is dependent on good starting values which can be obtained by the construction of a two dimensional grid for γ and c, as discussed in Section 5.3. The resulting empirical STR estimates are set out overleaf in Table 7.5.

Table 7.5: The Smooth Transition Regression Model – Initial Estimation

Dependent Variable: Δ lognph$_t$

Independent Variable	*Coefficient*	*Standard Error*	*T Value*	*Probability*
Linear Part of Model				
Δ lognph$_{t-1}$	3.047	61.374	0.05	0.96
Δ loginc$_t$	0.823	0.146	5.63	0.00***
Δ loginc$_{t-1}$	0.531	0.166	3.19	0.00***
Δ MR$_{t-2}$	0.010	0.004	2.43	0.02**
Δ MR$_{t-3}$	-0.009	0.004	-2.47	0.02**
Δ logphs$_{t-1}$	-10.327	3.072	-3.36	0.00***
Δ logphs$_{t-2}$	-7.233	2.992	-2.42	0.02**
Δ loghf$_{t-1}$	0.151	0.038	4.02	0.00***
Δ loghf$_{t-2}$	0.128	0.028	4.59	0.00***
Δ CGT$_{t-1}$	-0.002	0.001	-1.67	0.10*
Dumrpt$_t$	-0.023	0.011	-2.19	0.03**
Dumib$_t$	0.029	0.013	2.18	0.03**
Dums2327$_t$	0.022	0.010	2.26	0.03**
ecm^D_{t-1}	3.102	18.336	0.17	0.87
ecm^S_{t-1}	0.192	1.028	0.19	0.85
Dq_{2t}	0.014	0.008	1.74	0.09*
Dq_{3t}	0.012	0.008	1.53	0.13
Constant	1.342	0.259	5.19	0.00***
Non-Linear Part of Model				
Δ lognph$_{t-1}$	-3.563	61.377	-0.06	0.95
ecm^D_{t-1}	-3.022	18.335	-0.16	0.87
ecm^S_{t-1}	-0.173	1.027	-0.17	0.87
$\hat{\gamma}$		10.669		
$\hat{c}$		-92.028		
Residual standard deviation		0.026		
R squared		0.740		
Adjusted R squared		0.743		
Akaike information criteria		-7.088		

Notes:
1. * (**) (***) denotes significant at the 10% (5%) (1%) level

As with the linear error correction model, the presence of a number of insignificant explanatory variables in Table 7.5 suggests the need to better specify the parameter structure of the STR model. In particular, restricting the combined coefficient of the lagged change in house prices to zero for $G(s_t; \gamma, c) = 0$, so that past house prices do not contribute in this regime, gives rise to an improved specification. Table 7.6 presents the reduced version of the STR model.

Table 7.6: The Smooth Transition Regression Model – Final Model

Dependent Variable: Δ lognph$_t$

Independent Variable	*Coefficient*	*Standard Error*	*T Value*	*Probability*
Linear Part of Model				
Δ loginc$_t$	0.823	0.145	5.67	0.00***
Δ loginc$_{t-1}$	0.531	0.165	3.23	0.00***
Δ MR$_{t-2}$	0.010	0.004	2.45	0.02**
Δ MR$_{t-3}$	-0.009	0.004	-2.49	0.02**
Δ logphs$_{t-1}$	-10.336	3.043	-3.40	0.00***
Δ logphs$_{t-2}$	-7.239	2.966	-2.44	0.02**
Δ loghf$_{t-1}$	0.151	0.037	4.05	0.00***
Δ loghf$_{t-2}$	0.128	0.028	4.63	0.00***
Δ CGT$_{t-1}$	-0.002	0.001	-1.72	0.09*
Dumrpt$_t$	-0.023	0.010	-2.25	0.03**
Dumib$_t$	0.029	0.013	2.21	0.03**
Dums2327$_t$	0.022	0.009	2.28	0.03**
ecm^D_{t-1}	2.301	1.028	2.24	0.03**
ecm^S_{t-1}	0.148	0.060	2.48	0.02**
Dq$_{2t}$	0.013	0.008	1.75	0.09*
Dq$_{3t}$	0.012	0.008	1.55	0.13
Constant	1.340	0.255	5.25	0.00***
Non-Linear Part of Model				
Δ lognph$_{t-1}$	-0.516	0.100	-5.16	0.00***
ecm^D_{t-1}	-2.221	1.023	-2.17	0.03**
ecm^S_{t-1}	-0.128	0.059	-2.18	0.03**
$\hat{\gamma}$		18.657		
$\hat{c}$		-91.757		
Residual standard deviation		0.025		
R squared		0.740		
Adjusted R squared		0.743		
Akaike information criteria		-7.113		

Notes:
1. * (**) (***) denotes significant at the 10% (5%) (1%) level

More formally, the estimated smooth transition regression model is written as:

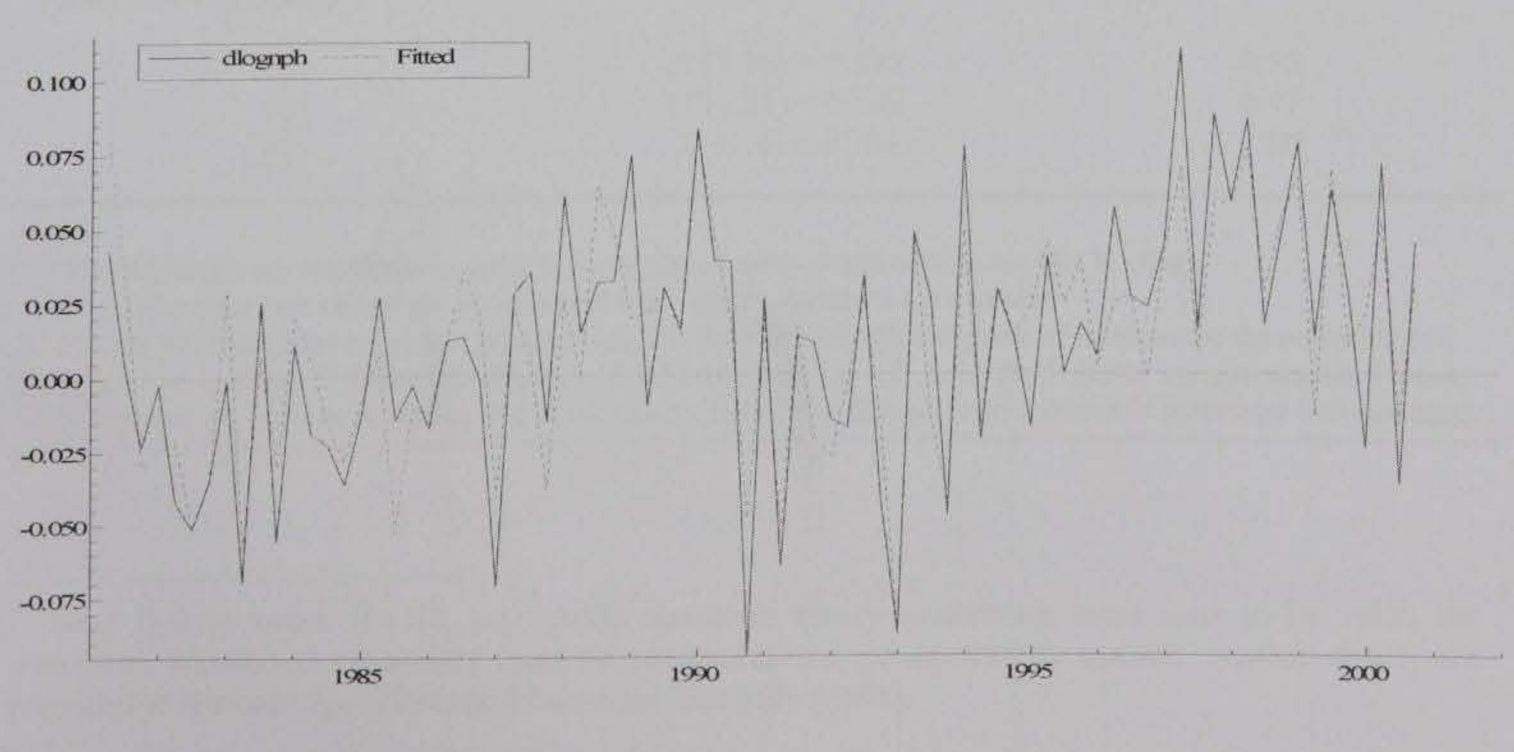
$$\begin{aligned}\Delta \text{lognph}_t = {} & 1.340 + 0.823\,\Delta\,\text{loginc}_t + 0.531\,\Delta\,\text{loginc}_{t-1} + 0.010\,\Delta\,\text{MR}_{t-2} - 0.009\,\Delta\,\text{MR}_{t-3} \\ & - 10.336\,\Delta\,\text{logphs}_{t-1} - 7.239\,\Delta\,\text{logphs}_{t-2} + 0.151\,\Delta\,\text{loghf}_{t-1} + 0.128\,\Delta\,\text{loghf}_{t-2} \\ & - 0.002\,\Delta\text{CGT}_{t-1} - 0.023\,\text{Dumrpt}_t + 0.029\,\text{Dumib}_t + 0.022\,\text{Dums2327}_t \\ & + 2.301\,\text{ecm}^D_{t-1} + 0.148\,\text{ecm}^S_{t-1} + 0.013\,\text{Dq}_2 + 0.012\,\text{Dq}_3 \\ & - \{0.516\,\Delta\,\text{lognph}_{t-1} + 2.221\,\text{ecm}^D_{t-1} + 0.128\,\text{ecm}^S_{t-1}\} \\ & \times [1 + \exp\{-18.657\,(\text{ecm}^S_{t-1} - (-91.757)) / \hat{\sigma}(\text{ecm}^S_{t-1})\}]^{-1} \\ & + \hat{\varepsilon}_t \qquad (7.3)\end{aligned}$$

where:

$\hat{\sigma}(\text{ecm}^S_{t-1})$ = the sample standard deviation of ecm^S_{t-1}.

Figure 7.3 provides a graphical view of the fitted and actual values of the dependent variable. With respect to overall goodness of fit, the adjusted R squared statistic of 0.74 associated with the reduced STR model / equation (7.3) represents a considerable improvement on the linear specification's corresponding figure of 0.60. As is evident from the graph, much of this additional explanatory power relates to the fact that the STR model is more successful than its linear counterpart in fitting the 1993 turning point.

Figure 7.3: The Estimated Smooth Transition Regression Model – Actual and Fitted Values



7.5.1. *Analysis of the Residuals and Evaluation of the Estimated Smooth Transition Regression Model*

As in the linear case, it is important to check the residuals of the non-linear model for evidence of misspecification. Eitrheim and Teräsvirta (1996) derive a range of misspecification tests within a univariate setting that can be readily generalised to the STR framework, including those of an LM test of no error autocorrelation and an LM-type test of parameter constancy[11]. In addition to reporting the outcome of the above tests, Table 7.7 also imposes an LM test of autoregressive conditional heteroscedasticity and the Lomnicki-Jarque-Bera normality test on the residuals of the estimated STR model.

Table 7.7: Estimated Smooth Transition Regression Model – Residual Statistics

Test Type	*Test Statistic*	*Probability*
Serial Correlation[a]:		
AR(1)	F(1,55) = 0.169	0.68
AR(2)	F(2,53) = 0.414	0.66
AR(3)	F(3,51) = 0.517	0.67
AR(4)	F(4,49) = 0.574	0.68
Normality[a]	0.555	0.76
Autoregressive Conditional Heteroscedasticity[a]:		
ARCH (1-4) Test	F = 1.266	0.29
Parameter Constancy[b]:		
F_1	F(17,38) = 0.895	0.58
F_2	F(34,21) = 0.500	0.97
F_3	F(51,4) = 0.101	1.00

Notes:

1. [a] Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
   [b] Ho: All parameters except the government intervention dummies are constant
2. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test
3. F_j is a test against a STR model with transition function H_j, j = 1, 2, 3 where H_1 yields a single structural break, H_2 implies two structural breaks and H_3 allows for non-symmetric and non-monotonic parameter non-constancy

[11] Note that in order for the asymptotic statistical theory underlying these tests to be valid, the maximum likelihood estimators must be consistent and asymptotically normal. Further details are provided in Wooldridge (1994) and Escribano and Mira (1995).

Diagnostically, the STR model appears to be well behaved. Findings of the serial correlation, normality and ARCH tests indicate that the null cannot be rejected, while the tests of parameter constancy suggest selection of a STR model with constant parameters over one with either a single *(F_1)* or two structural breaks *(F_2)*, or with non-symmetric and non-monotonic parameter non-constancy *(F_3)*.

Of further importance in evaluating the estimated STR model is ensuring that it successfully captures all the non-linear features present in the data. In doing so, the null hypothesis of no remaining non-linearity is tested against the alternative of an additive STR model, that is, a model with two additive non-linear components and a second transition function[12]. The practical testing procedure is similar to that of linearity testing - firstly, a set of potential transition variables is defined and secondly, the test is repeated taking each variable in turn as the additional transition variable. Table 7.8 details the findings in the case of equation (7.3).

Table 7.8: Testing for No Remaining Non-Linearity – P Values

	Linearity Test			
Transition Variable	*F*	F_4	F_3	F_2
Δ lognph$_{t-1}$	0.975	0.677	0.967	0.702
Δ loginc$_t$	0.117	0.565	0.228	0.025
Δ loginc$_{t-1}$	0.248	0.295	0.397	0.116
Δ MR$_{t-2}$	0.443	0.207	0.996	0.152
Δ MR$_{t-3}$	0.387	0.286	0.756	0.129
Δ logphs$_{t-1}$	0.308	0.630	0.113	0.262
Δ logphs$_{t-2}$	0.757	0.927	0.707	0.172
Δ loghf$_{t-1}$	0.710	0.407	0.843	0.335
Δ loghf$_{t-2}$	0.444	0.225	0.332	0.582
Δ CGT$_{t-1}$	N/A	N/A	N/A	N/A
ecm^D_{t-1}	0.267	0.330	0.481	0.093
ecm^S_{t-1}	0.321	0.137	0.492	0.349

Notes:
1. Ho: No remaining non-linearity in the estimated STR model
2. N/A indicates that it was not possible to compute p values due to non-positive definite matrices

[12] See Section 5.3 and Teräsvirta (1998) for a detailed discussion of this test.

The p values presented in column F of the above table indicate that regardless of which variable is chosen as the transition variable, the null hypothesis of no additive non-linearity cannot be rejected at the 1% significance level[13]. Indeed, the test results for all variables are in excess of 0.1. Furthermore, when ecm^{S}_{t-1} is taken to be the transition variable, a p value of 0.321 is obtained indicating that the non-linearity initially associated with this variable (see Table 7.4) has been satisfactorily modelled. As such, it may be concluded that the estimated STR model captures much of the non-linearity present in the Dublin housing data.

7.5.2. *Analysis of the Coefficients of the Estimated Smooth Transition Regression Model*

This section is concerned with the economic interpretation of the parameters of the linear and non-linear parts of the estimated STR model. With respect to the former, equation (7.3) / Table 7.6 indicates that there is little change in the magnitude of most coefficients from those of the linear error correction model presented in equation (7.2). However, it seems that explicitly modelling non-linearity provides additional insights into price behaviour in the Dublin market.

As the starting point, consider the empirical estimates of the transition function:

$$G(\hat{\gamma}, \hat{c}; ecm^{S}_{t-1}) = [1 + \exp\{-18.657(\mathrm{ecm}^{S}_{t-1} - (-91.757)) / \hat{\sigma}(\mathrm{ecm}^{S}_{t-1})\}]^{-1} \qquad (7.4)$$

and recall that as the estimated STR model is of the LSTR1 type, it is capable of characterising asymmetric behaviour in the Dublin housing market. As such, it is possible to divide the market into two regimes, those of expansion (upturn) and recession (downturn), where the transition from one to the other is associated with large and small values of the transition variable ecm^{S}_{t-1} relative to the location parameter c. On this basis, a downturn is said to occur if the level of house price disequilibrium (given the prevailing levels of mortgage

---

[13] To protect against the risk of over-fitting, Teräsvirta (2004) suggests that low significance levels should be applied when evaluating the outcomes of the no additive non-linearity test.

interest rates, building costs and land) is lower than the estimated location parameter of -91.757. Conversely, house price disequilibrium in excess of this figure represents an expansionary phase. In terms of the smoothness of regime change, the 18.657 estimate of the parameter γ is fairly small, indicating a gradual switching between up and downturns.

Figure 7.4 presents a graphical overview of equation (7.4). Values of the transition function close to one imply that the entire model is necessary when describing house price inflation while zero values indicate that the linear part is adequate. Given that the function obtains a value of one for most of the sample period, it may be concluded that both the linear and non-linear components of the STR specification are required when modelling house price growth in the Dublin new housing market.

Figure 7.4: Values of the Transition Function of the Estimated STR Model

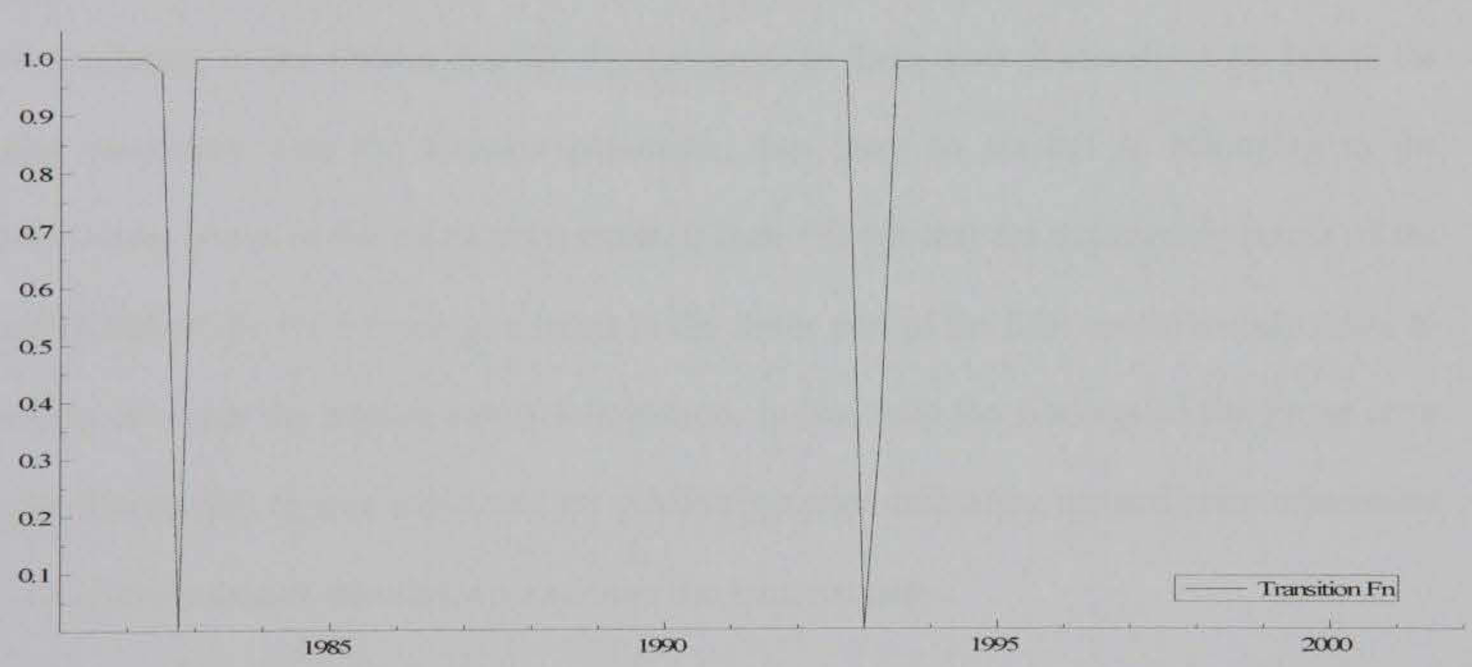
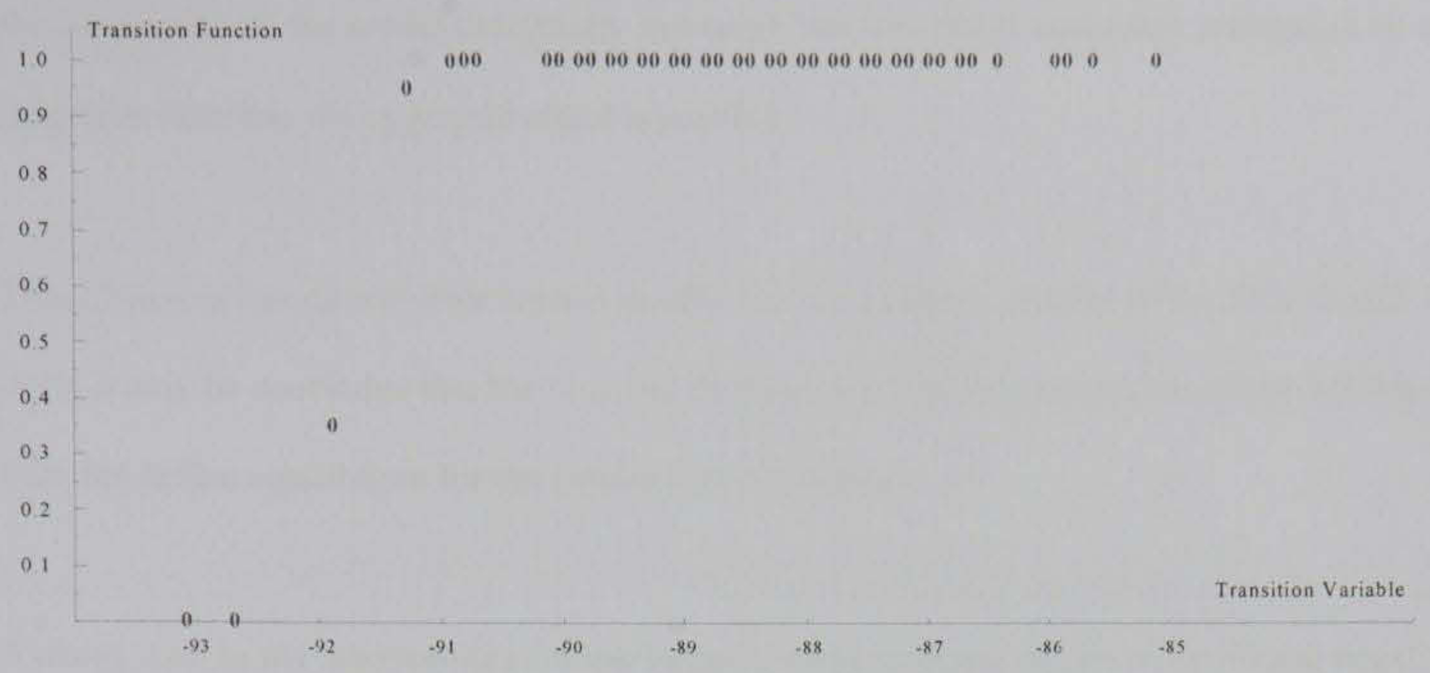


Figure 7.5 plots the transition function as a function of the transition variable. The interpretation of this graph is similar to that of Figure 7.4 in that values of the function close to zero again imply that the linear component of the STR model sufficiently describes the rate of house price growth, whereas at the opposite scale of one, both the linear and non-linear parts of the specification are needed to capture dynamic price determination. Note that each dot corresponds to at least one observation.

Figure 7.5: The Transition Function of the Estimated STR Model as a Function of the Transition Variable



As is evident from this graphical representation, only a small number of observations are equal to zero implying a limited ability on the part of the linear component to solely model price inflation in the Dublin market. Furthermore, as these zero observations lie below the value associated with the location parameter, they may be classed as belonging to the recessionary phase of the house price cycle. It then follows that the explanatory power of the coefficients of the error correction terms in the linear part of the STR model mainly relate to periods in which the market enters a downturn. In line with the findings of the linear error correction model, these coefficients are positively signed indicating upward price adjustment and the movement of the market away from the long run path.

However, it appears that for most of the sample period, the Dublin housing market was in an expansionary phase given that the majority of observations in Figure 7.5 exceed the value of the location parameter. Moreover, nearly all are equal to the upper bound of one and in this context; the complete STR model best explains dynamic price determination.

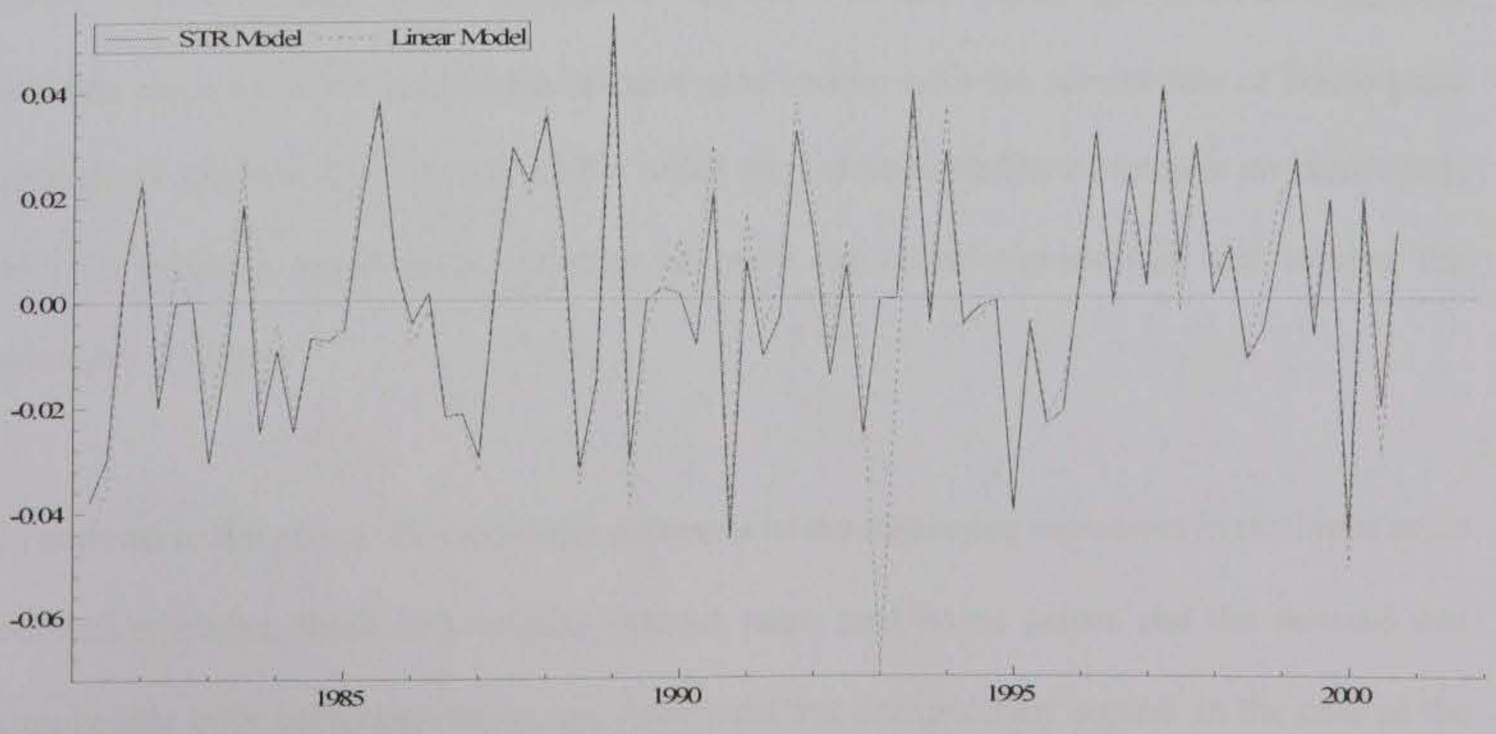
Acceptance of the complete STR model brings into play the non-linear estimates presented in Table 7.6 / equation (7.3). Of particular interest in this respect are the negative signs on the coefficients of the error correction terms. These suggest adjustment to steady state deviations and a return to equilibrium. However, as the magnitudes of the coefficients on these terms in the linear part of the model marginally outweigh the non-linear estimates multiplied by the transition function, the aggregate effect is positive.

Thus, house prices do not error correct in either of the extreme regimes of the STR model. As such, it may be concluded that the long run demand and supply relations identified in Chapter 6 do not define equilibrium for the Dublin housing market.

Turning next to the interpretation of the lagged change in house prices, a significant negative relationship is evident between current and past house price changes in the non-linear part of the STR model. This reflects a tendency for expectations of further price increases in the Dublin market to fall after a certain point, namely once the value of the location parameter has been exceeded.

Finally, the residuals of the estimated linear error correction and smooth transition regression models are plotted against time in Figure 7.6.

Figure 7.6: Residuals of the Linear Error Correction and STR Models



Overall, the smooth transition regression model is well specified from a statistical perspective and in terms of explanatory power, represents a considerable improvement on the fit of the estimated linear error correction model. This primarily reflects its success in modelling outlying episodes. However, as in the linear case, a lack of error correction is apparent.

7.6. Conclusion

This chapter sought to model dynamic price determination in the Dublin new, private sector housing market within an adjustment type framework. In doing so, error correction models of both a linear and non-linear nature were developed and applied to a dataset of explanatory variables comprised of income, mortgage interest rates, the housing stock, household formation - all of which follow directly from the previous long run analysis - and a range of measures that attempt to capture the effectiveness of government interventionist policies in the housing market and the impact of deviations from the static demand and supply relations on short run price inflation.

With respect to the empirical findings, the estimated linear error correction model contains, for the most part, correctly signed and significant explanatory variables. As expected a priori, current and lagged changes in income and household formation are statistically significant and positively related to the dependent variable whereas, capital gains taxation and the housing stock exhibit a significant negative relationship with the current rate of house price growth. In addition, the 'Section 23/27' relief, the tax deductibility of interest on borrowings and the property tax dummy variables all enter the model significantly and bearing the anticipated sign.

In contrast to the above, the parameter estimates of the remaining regressors in the linear error correction model, those of mortgage interest rates, past house prices and the demand and supply-side error correction terms, are significant but unexpectedly signed. In the case of the former, the dominant positive response of current price changes to an increase in interest rates is explainable by the fact that in times of high uncertainty and inflation, housing acts as a hedge against future inflation and thus, as the nominal mortgage interest rate rises, the demand for housing increases exerting upward pressure on prices. With regard to the latter, the positive signs on both error correction terms indicate upward price adjustment and the movement of the market away from equilibrium. Likewise, the coefficient on the lagged house price series exhibits a somewhat counterintuitive sign.

The empirical outcomes associated with the smooth transition regression model indicate little change in the magnitude of most coefficients from those of the linear specification. However, a significant negative relationship between current and past house price changes is evident in the non-linear part of the STR model. This reflects a tendency for expectations of further price increases in the Dublin market to fall after a certain point, namely once the value of the location parameter has been exceeded. Moreover, as house prices fail to adjust in either of the extreme regimes of the STR model, it may be concluded that the long run demand and supply relations identified in Chapter 6 do not define equilibrium for the Dublin housing market.

Overall, the empirical findings presented in this chapter indicate that the estimated smooth transition regression model is reasonably successful in modelling dynamic price determination in the Dublin new, private sector housing market over the period 1981-2000. Nonetheless, there clearly remains some variation in the data that has not been fully explained to date. A possible explanation in this context, and one that fits with the finding of no cointegration, is the presence of a speculative component in the marketplace. Accordingly, Chapter 8 tests for the possibility of a bubble in the Dublin housing market.

7A. Appendix

Table 7A: The Linear Error Correction Model – Specification Testing

Dependent Variable: $\Delta lognph_t$			Model			
Independent Variable	(1)		(2)		(3)	
	Coefficient	*Std Error*	*Coefficient*	*Std Error*	*Coefficient*	*Std Error*
$\Delta lognph_{t-1}$	-0.319	0.153	-0.319	0.151	-0.322	0.148
$\Delta lognph_{t-2}$	0.100	0.158	0.096	0.143	0.096	0.142
$\Delta lognph_{t-3}$	**0.010**	**0.151**				
$\Delta lognph_{t-4}$	-0.029	0.143	-0.031	0.135	**-0.030**	**0.133**
$\Delta loginc_t$	0.622	0.215	0.626	0.206	0.627	0.203
$\Delta loginc_{t-1}$	0.279	0.230	0.280	0.227	0.281	0.224
$\Delta loginc_{t-2}$	-0.233	0.254	-0.231	0.250	-0.225	0.244
$\Delta loginc_{t-3}$	-0.038	0.231	**-0.032**	**0.207**		
$\Delta loginc_{t-4}$	-0.091	0.259	-0.089	0.253	-0.088	0.250
ΔMR_t	-0.005	0.006	-0.005	0.006	-0.005	0.006
ΔMR_{t-1}	-0.007	0.006	-0.007	0.006	-0.007	0.006
ΔMR_{t-2}	0.015	0.006	0.015	0.006	0.015	0.006
ΔMR_{t-3}	-0.011	0.006	-0.011	0.006	-0.011	0.006
ΔMR_{t-4}	0.008	0.006	0.008	0.006	0.007	0.006
$\Delta logphs_t$	4.536	4.320	4.529	4.267	4.488	4.211
$\Delta logphs_{t-1}$	-7.699	4.216	-7.656	4.114	-7.621	4.061
$\Delta logphs_{t-2}$	-2.214	4.355	-2.232	4.294	-2.368	4.156
$\Delta logphs_{t-3}$	2.972	4.264	2.934	4.173	2.876	4.109
$\Delta logphs_{t-4}$	2.215	4.349	2.134	4.111	2.036	4.016
$\Delta loghf_t$	0.023	0.037	0.023	0.037	0.024	0.036
$\Delta loghf_{t-1}$	0.053	0.081	0.054	0.077	0.056	0.076
$\Delta loghf_{t-2}$	0.046	0.071	0.048	0.067	0.048	0.067
$\Delta loghf_{t-3}$	-0.050	0.059	-0.049	0.057	-0.049	0.057
$\Delta loghf_{t-4}$	-0.032	0.041	-0.032	0.040	-0.031	0.039
ΔCGT_t	-0.002	0.002	-0.002	0.002	-0.002	0.002
ΔCGT_{t-1}	-0.002	0.001	-0.002	0.001	-0.002	0.001
ΔCGT_{t-2}	-0.001	0.002	-0.001	0.002	-0.001	0.001
ΔCGT_{t-3}	-0.001	0.001	-0.001	0.001	-0.001	0.001
ΔCGT_{t-4}	-0.001	0.001	-0.001	0.001	-0.001	0.001
$Dumrpt_t$	-0.024	0.014	-0.025	0.013	-0.025	0.013
$Dumib_t$	0.032	0.019	0.032	1.019	0.033	0.018
$Dums2327_t$	0.015	0.013	0.016	0.013	0.015	0.013
ecm^D_{t-1}	0.002	0.058	0.002	0.056	0.004	0.055
ecm^S_{t-1}	0.015	0.008	0.016	0.007	0.016	0.007
Dq_{1t}	0.008	0.013	0.008	0.012	0.008	0.012
Dq_{2t}	0.021	0.014	0.022	0.014	0.022	0.013
Dq_{3t}	0.017	0.012	0.017	0.012	0.017	0.012
Constant	1.333	0.472	1.345	0.432	1.345	0.427
Adjusted R squared		0.51		0.52		0.54
SC information criteria		-5.58		-5.64		-5.69
AR(1-4) test p value		0.21		0.23		0.21
Normality test p value		0.90		0.89		0.88
ARCH (1-4) test p value		1.00		1.00		1.00
Exclusion test p value		0.93		0.83		0.76

Notes: 1. Figures in bold refer to the most insignificant regressor in each model
2. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
3. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test
4. Exclusion test Ho: the coefficient of the bolded variable is zero. The associated test statistic is the log likelihood ratio. If the null cannot be rejected, the variable is dropped from the analysis and the model is re-estimated

Table 7A(contd): The Linear Error Correction Model – Specification Testing

Dependent Variable: Δ lognph$_t$		Model				
Independent Variable	(4)		(5)		(6)	
	Coefficient	*Std Error*	*Coefficient*	*Std Error*	*Coefficient*	*Std Error*
Δ lognph$_{t-1}$	-0.323	0.146	-0.325	0.145	-0.331	0.142
Δ lognph$_{t-2}$	0.091	0.139	0.101	0.136	0.093	0.133
Δ lognph$_{t-3}$						
Δ lognph$_{t-4}$						
Δ loginc$_t$	0.623	0.200	0.632	0.198	0.658	0.180
Δ loginc$_{t-1}$	0.276	0.221	0.282	0.219	0.282	0.216
Δ loginc$_{t-2}$	-0.218	0.239	-0.205	0.236	-0.214	0.232
Δ loginc$_{t-3}$						
Δ loginc$_{t-4}$	-0.114	0.218	**-0.065**	**0.194**		
Δ MR$_t$	-0.005	0.005	-0.005	0.005	-0.005	0.005
Δ MR$_{t-1}$	-0.007	0.006	-0.007	0.006	-0.007	0.005
Δ MR$_{t-2}$	0.015	0.006	0.015	0.006	0.014	0.006
Δ MR$_{t-3}$	-0.011	0.006	-0.011	0.006	-0.011	0.005
Δ MR$_{t-4}$	0.007	0.006	0.007	0.006	0.007	0.006
Δ logphs$_t$	4.554	4.156	4.111	4.029	4.067	3.987
Δ logphs$_{t-1}$	-7.474	3.965	-7.722	3.902	-7.783	3.860
Δ logphs$_{t-2}$	-2.552	4.032	-2.726	3.984	-2.902	3.911
Δ logphs$_{t-3}$	2.946	4.053	2.657	3.979	2.505	3.915
Δ logphs$_{t-4}$	2.085	3.967	1.897	3.917	**1.932**	**3.877**
Δ loghf$_t$	0.025	0.036	0.020	0.034	0.019	0.034
Δ loghf$_{t-1}$	0.054	0.074	0.067	0.069	0.069	0.068
Δ loghf$_{t-2}$	0.047	0.066	0.057	0.062	0.057	0.062
Δ loghf$_{t-3}$	-0.050	0.056	-0.046	0.055	-0.048	0.054
Δ loghf$_{t-4}$	-0.030	0.038	-0.028	0.038	-0.031	0.037
Δ CGT$_t$	-0.002	0.002	-0.002	0.002	-0.002	0.002
Δ CGT$_{t-1}$	-0.002	0.001	-0.002	0.001	-0.002	0.001
Δ CGT$_{t-2}$	**-0.001**	**0.001**				
Δ CGT$_{t-3}$	-0.001	0.001	-0.001	0.001	-0.001	0.001
Δ CGT$_{t-4}$	-0.001	0.001	-0.001	0.001	-0.001	0.001
Dumrpt$_t$	-0.025	0.013	-0.024	0.012	-0.025	0.012
Dumib$_t$	0.033	0.018	0.035	0.018	0.035	0.017
Dums2327$_t$	0.015	0.012	0.013	0.012	0.013	0.012
ecm^D_{t-1}	0.003	0.054	0.014	0.049	0.014	0.048
ecm^S_{t-1}	0.015	0.007	0.017	0.006	0.017	0.006
Dq_{1t}	0.008	0.011	0.008	0.011	0.007	0.011
Dq_{2t}	0.020	0.012	0.022	0.012	0.022	0.012
Dq_{3t}	0.016	0.011	0.016	0.011	0.016	0.011
Constant	1.329	0.416	1.381	0.400	1.388	0.396
Adjusted R squared		0.55		0.55		0.56
SC information criteria		-5.75		-5.80		-5.85
AR(1-4) test p value		0.26		0.24		0.23
Normality test p value		0.86		0.89		0.89
ARCH (1-4) test p value		1.00		0.99		0.99
Exclusion test p value		0.50		0.66		0.51

Notes: 1. Figures in bold refer to the most insignificant regressor in each model
2. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
3. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test
4. Exclusion test Ho: the coefficient of the bolded variable is zero. The associated test statistic is the log likelihood ratio. If the null cannot be rejected, the variable is dropped from the analysis and the model is re-estimated

Table 7A(contd): The Linear Error Correction Model – Specification Testing

Dependent Variable: Δ lognph$_t$	Model					
Independent Variable	(7)		(8)		(9)	
	Coefficient	*Std Error*	*Coefficient*	*Std Error*	*Coefficient*	*Std Error*
Δ lognph$_{t-1}$	-0.335	0.141	-0.338	0.140	-0.357	0.134
Δ lognph$_{t-2}$	0.084	0.131	0.075	0.129	**0.057**	**0.123**
Δ lognph$_{t-3}$						
Δ lognph$_{t-4}$						
Δ loginc$_t$	0.667	0.178	0.674	0.176	0.666	0.174
Δ loginc$_{t-1}$	0.262	0.211	0.277	0.208	0.307	0.198
Δ loginc$_{t-2}$	-0.219	0.230	-0.193	0.223	-0.157	0.210
Δ loginc$_{t-3}$						
Δ loginc$_{t-4}$						
Δ MR$_t$	-0.005	0.005	-0.006	0.005	-0.006	0.005
Δ MR$_{t-1}$	-0.007	0.005	-0.007	0.005	-0.007	0.005
Δ MR$_{t-2}$	0.014	0.005	0.014	0.005	0.014	0.005
Δ MR$_{t-3}$	-0.011	0.005	-0.011	0.005	-0.010	0.005
Δ MR$_{t-4}$	0.007	0.005	0.007	0.005	0.006	0.005
Δ logphs$_t$	4.387	3.904	4.399	3.874	4.569	3.831
Δ logphs$_{t-1}$	-7.420	3.761	-7.726	3.686	-7.638	3.654
Δ logphs$_{t-2}$	-2.659	3.850	-2.929	3.786	-3.268	3.700
Δ logphs$_{t-3}$	2.448	3.882	**1.904**	**3.710**		
Δ logphs$_{t-4}$						
Δ loghf$_t$	**0.017**	**0.033**				
Δ loghf$_{t-1}$	0.070	0.068	0.068	0.067	0.080	0.063
Δ loghf$_{t-2}$	0.057	0.061	0.055	0.060	0.063	0.058
Δ loghf$_{t-3}$	-0.049	0.054	-0.048	0.053	-0.041	0.051
Δ loghf$_{t-4}$	-0.034	0.036	-0.033	0.036	-0.029	0.034
Δ CGT$_t$	-0.002	0.001	-0.002	0.001	-0.002	0.001
Δ CGT$_{t-1}$	-0.002	0.001	-0.002	0.001	-0.002	0.001
Δ CGT$_{t-2}$						
Δ CGT$_{t-3}$	-0.001	0.001	-0.001	0.001	-0.001	0.001
Δ CGT$_{t-4}$	-0.001	0.001	-0.001	0.001	-0.001	0.001
Dumrpt$_t$	-0.025	0.012	-0.025	0.012	-0.026	0.011
Dumib$_t$	0.033	0.017	0.033	0.017	0.033	0.017
Dums2327$_t$	0.014	0.012	0.012	0.011	0.013	0.011
ecm^D_{t-1}	0.016	0.048	0.025	0.044	0.033	0.041
ecm^S_{t-1}	0.017	0.006	0.017	0.006	0.018	0.006
Dq_{1t}	0.008	0.011	0.008	0.011	0.007	0.010
Dq_{2t}	0.023	0.011	0.023	0.011	0.022	0.011
Dq_{3t}	0.017	0.010	0.016	0.010	0.015	0.010
Constant	1.377	0.392	1.394	0.387	1.427	0.379
Adjusted R squared		0.57		0.58		0.58
SC information criteria		-5.90		-5.95		-6.00
AR(1-4) test p value		0.19		0.21		0.25
Normality test p value		0.76		0.68		0.56
ARCH (1-4) test p value		0.97		0.99		0.98
Exclusion test p value		0.50		0.51		0.56

Notes: 1. Figures in bold refer to the most insignificant regressor in each model
2. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
3. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test
4. Exclusion test Ho: the coefficient of the bolded variable is zero. The associated test statistic is the log likelihood ratio. If the null cannot be rejected, the variable is dropped from the analysis and the model is re-estimated

Table 7A(contd): The Linear Error Correction Model – Specification Testing

Dependent Variable: Δ lognph$_t$	Model					
Independent Variable	(10)		(11)		(12)	
	Coefficient	*Std Error*	*Coefficient*	*Std Error*	*Coefficient*	*Std Error*
$\Delta lognph_{t-1}$	-0.383	0.121	-0.391	0.120	-0.401	0.117
$\Delta lognph_{t-2}$						
$\Delta lognph_{t-3}$						
$\Delta lognph_{t-4}$						
$\Delta loginc_{t}$	0.666	0.173	0.668	0.172	0.687	0.166
$\Delta loginc_{t-1}$	0.322	0.194	0.346	0.189	0.351	0.187
$\Delta loginc_{t-2}$	**-0.122**	**0.195**				
$\Delta loginc_{t-3}$						
$\Delta loginc_{t-4}$						
ΔMR_{t}	-0.006	0.005	-0.006	0.005	-0.005	0.005
ΔMR_{t-1}	-0.006	0.005	-0.006	0.005	-0.006	0.005
ΔMR_{t-2}	0.014	0.005	0.014	0.005	0.014	0.005
ΔMR_{t-3}	-0.010	0.005	-0.011	0.005	-0.011	0.005
ΔMR_{t-4}	0.007	0.005	0.006	0.005	0.006	0.005
$\Delta logphs_{t}$	4.589	3.800	4.582	3.778	4.489	3.745
$\Delta logphs_{t-1}$	-8.029	3.526	-8.106	3.503	-8.397	3.427
$\Delta logphs_{t-2}$	-3.527	3.628	-3.824	3.576	-4.007	3.530
$\Delta logphs_{t-3}$						
$\Delta logphs_{t-4}$						
$\Delta loghf_{t}$						
$\Delta loghf_{t-1}$	0.092	0.056	0.092	0.056	0.095	0.055
$\Delta loghf_{t-2}$	0.074	0.053	0.073	0.053	0.076	0.052
$\Delta loghf_{t-3}$	-0.032	0.047	-0.029	0.046	**-0.026**	**0.046**
$\Delta loghf_{t-4}$	-0.024	0.032	-0.021	0.032	-0.018	0.031
ΔCGT_{t}	-0.002	0.001	-0.002	0.001	-0.001	0.001
ΔCGT_{t-1}	-0.002	0.001	-0.002	0.001	-0.002	0.001
ΔCGT_{t-2}						
ΔCGT_{t-3}	-0.001	0.001	-0.001	0.001	-0.001	0.001
ΔCGT_{t-4}	-0.001	0.001	-0.001	0.001	-0.001	0.001
$Dumrpt_t$	-0.026	0.011	-0.027	0.011	-0.027	0.011
$Dumib_t$	0.035	0.016	0.037	0.016	0.038	0.015
$Dums2327_t$	0.013	0.011	0.012	0.011	0.013	0.011
ecm^{D}_{t-1}	0.040	0.038	0.042	0.037	0.044	0.037
ecm^{S}_{t-1}	0.019	0.005	0.019	0.005	0.020	0.005
Dq_{1t}	0.006	0.010	**0.005**	**0.010**		
Dq_{2t}	0.020	0.010	0.018	0.010	0.016	0.008
Dq_{3t}	0.015	0.010	0.013	0.010	0.011	0.008
Constant	1.523	0.315	1.507	0.312	1.518	0.309
Adjusted R squared		0.59		0.59		0.60
SC information criteria		-6.05		-6.10		-6.15
AR(1-4) test p value		0.36		0.36		0.38
Normality test p value		0.46		0.36		0.47
ARCH (1-4) test p value		0.98		0.98		0.95
Exclusion test p value		0.43		0.54		0.48

Notes: 1. Figures in bold refer to the most insignificant regressor in each model
2. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
3. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test
4. Exclusion test Ho: the coefficient of the bolded variable is zero. The associated test statistic is the log likelihood ratio. If the null cannot be rejected, the variable is dropped from the analysis and the model is re-estimated

Table 7A(contd): The Linear Error Correction Model – Specification Testing

Dependent Variable: Δ $lognph_t$	Model					
Independent Variable	(13)		(14)		(15)	
	Coefficient	*Std Error*	*Coefficient*	*Std Error*	*Coefficient*	*Std Error*
Δ $lognph_{t-1}$	-0.425	0.109	-0.419	0.105	-0.419	0.105
Δ $lognph_{t-2}$						
Δ $lognph_{t-3}$						
Δ $lognph_{t-4}$						
Δ $loginc_t$	0.674	0.164	0.671	0.162	0.688	0.158
Δ $loginc_{t-1}$	0.362	0.185	0.359	0.183	0.359	0.182
Δ $loginc_{t-2}$						
Δ $loginc_{t-3}$						
Δ $loginc_{t-4}$						
Δ MR_t	-0.004	0.005	-0.004	0.005	-0.005	0.005
Δ MR_{t-1}	-0.006	0.005	-0.006	0.005	-0.006	0.005
Δ MR_{t-2}	0.014	0.005	0.014	0.005	0.014	0.005
Δ MR_{t-3}	-0.011	0.005	-0.011	0.005	-0.011	0.005
Δ MR_{t-4}	0.006	0.005	0.006	0.005	0.006	0.005
Δ $logphs_t$	3.772	3.512	3.719	3.476	3.521	3.437
Δ $logphs_{t-1}$	-8.645	3.379	-8.630	3.349	-8.644	3.328
Δ $logphs_{t-2}$	-4.525	3.393	-4.512	3.363	-4.778	3.310
Δ $logphs_{t-3}$						
Δ $logphs_{t-4}$						
Δ $loghf_t$						
Δ $loghf_{t-1}$	0.116	0.041	0.115	0.040	0.120	0.039
Δ $loghf_{t-2}$	0.101	0.028	0.100	0.028	0.101	0.028
Δ $loghf_{t-3}$						
Δ $loghf_{t-4}$	**-0.006**	**0.022**				
Δ CGT_t	-0.001	0.001	-0.001	0.001	-0.001	0.001
Δ CGT_{t-1}	-0.002	0.001	-0.002	0.001	-0.002	0.001
Δ CGT_{t-2}						
Δ CGT_{t-3}	-0.001	0.001	-0.001	0.001	**-0.001**	**0.001**
Δ CGT_{t-4}	-0.001	0.001	**-0.001**	**0.001**		
$Dumrpt_t$	-0.026	0.011	-0.026	0.011	-0.026	0.011
$Dumib_t$	0.039	0.015	0.040	0.015	0.039	0.015
$Dums2327_t$	0.013	0.011	0.013	0.011	0.012	0.010
ecm^D_{t-1}	0.057	0.030	0.057	0.029	0.060	0.029
ecm^S_{t-1}	0.021	0.004	0.021	0.004	0.021	0.004
Dq_{1t}						
Dq_{2t}	0.017	0.008	0.017	0.008	0.016	0.008
Dq_{3t}	0.012	0.008	0.012	0.008	0.011	0.008
Constant	1.562	0.297	1.559	0.294	1.592	0.287
Adjusted R squared		0.60		0.61		0.62
SC information criteria		-6.20		-6.25		-6.30
AR(1-4) test p value		0.34		0.31		0.31
Normality test p value		0.53		0.54		0.48
ARCH (1-4) test p value		0.89		0.89		0.89
Exclusion test p value		0.75		0.49		0.42

Notes: 1. Figures in bold refer to the most insignificant regressor in each model
2. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
3. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test
4. Exclusion test Ho: the coefficient of the bolded variable is zero. The associated test statistic is the log likelihood ratio. If the null cannot be rejected, the variable is dropped from the analysis and the model is re-estimated

Table 7A(contd): The Linear Error Correction Model – Specification Testing

Dependent Variable: Δ lognph$_t$	Model					
Independent Variable	(16)		(17)		(18)	
	Coefficient	*Std Error*	*Coefficient*	*Std Error*	*Coefficient*	*Std Error*
Δ lognph$_{t-1}$	-0.416	0.104	-0.410	0.104	-0.421	0.103
Δ lognph$_{t-2}$						
Δ lognph$_{t-3}$						
Δ lognph$_{t-4}$						
Δ loginc$_t$	0.677	0.157	0.686	0.156	0.681	0.156
Δ loginc$_{t-1}$	0.347	0.180	0.347	0.179	0.348	0.179
Δ loginc$_{t-2}$						
Δ loginc$_{t-3}$						
Δ loginc$_{t-4}$						
Δ MR$_t$	**-0.004**	**0.005**				
Δ MR$_{t-1}$	-0.005	0.005	-0.007	0.004	-0.007	0.004
Δ MR$_{t-2}$	0.014	0.005	0.016	0.005	0.015	0.005
Δ MR$_{t-3}$	-0.011	0.005	-0.012	0.005	-0.011	0.004
Δ MR$_{t-4}$	0.007	0.004	0.007	0.004	0.006	0.004
Δ logphs$_t$	3.524	3.420	**3.132**	**3.390**		
Δ logphs$_{t-1}$	-8.970	3.277	-9.596	3.204	-9.218	3.174
Δ logphs$_{t-2}$	-4.920	3.287	-5.315	3.256	-4.872	3.216
Δ logphs$_{t-3}$						
Δ logphs$_{t-4}$						
Δ loghf$_t$						
Δ loghf$_{t-1}$	0.123	0.038	0.124	0.038	0.129	0.038
Δ loghf$_{t-2}$	0.102	0.028	0.104	0.028	0.106	0.027
Δ loghf$_{t-3}$						
Δ loghf$_{t-4}$						
Δ CGT$_t$	-0.001	0.001	-0.001	0.001	**-0.001**	**0.001**
Δ CGT$_{t-1}$	-0.002	0.001	-0.002	0.001	-0.002	0.001
Δ CGT$_{t-2}$						
Δ CGT$_{t-3}$						
Δ CGT$_{t-4}$						
Dumrpt$_t$	-0.024	0.010	-0.024	0.010	-0.023	0.010
Dumib$_t$	0.036	0.014	0.032	0.013	0.031	0.013
Dums2327$_t$	0.012	0.010	0.013	0.010	0.016	0.010
ecm^D_{t-1}	0.065	0.028	0.065	0.027	0.069	0.027
ecm^S_{t-1}	0.022	0.004	0.021	0.004	0.021	0.004
Dq_{1t}						
Dq_{2t}	0.016	0.008	0.017	0.008	0.018	0.008
Dq_{3t}	0.012	0.008	0.013	0.008	0.014	0.008
Constant	1.621	0.282	1.518	0.259	1.577	0.251
Adjusted R squared		0.62		0.62		0.62
SC information criteria		-6.35		-6.39		-6.43
AR(1-4) test p value		0.31		0.31		0.27
Normality test p value		0.49		0.58		0.60
ARCH (1-4) test p value		0.88		0.85		0.78
Exclusion test p value		0.27		0.28		0.12

Notes: 1. Figures in bold refer to the most insignificant regressor in each model
2. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
3. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test
4. Exclusion test Ho: the coefficient of the bolded variable is zero. The associated test statistic is the log likelihood ratio. If the null cannot be rejected, the variable is dropped from the analysis and the model is re-estimated

Table 7A(contd): The Linear Error Correction Model – Specification Testing

Dependent Variable: Δ lognph$_t$	Model					
Independent Variable	(19)		(20)		(21)	
	Coefficient	*Std Error*	*Coefficient*	*Std Error*	*Coefficient*	*Std Error*
Δ lognph$_{t-1}$	-0.435	0.103	-0.471	0.101	-0.480	0.102
Δ lognph$_{t-2}$						
Δ lognph$_{t-3}$						
Δ lognph$_{t-4}$						
Δ loginc$_t$	0.717	0.155	0.778	0.150	0.769	0.151
Δ loginc$_{t-1}$	0.408	0.175	0.477	0.170	0.476	0.171
Δ loginc$_{t-2}$						
Δ loginc$_{t-3}$						
Δ loginc$_{t-4}$						
Δ MR$_t$						
Δ MR$_{t-1}$	-0.007	0.004	**-0.005**	**0.004**		
Δ MR$_{t-2}$	0.015	0.005	0.011	0.004	0.010	0.004
Δ MR$_{t-3}$	-0.012	0.004	-0.010	0.004	-0.008	0.004
Δ MR$_{t-4}$	**0.006**	**0.004**				
Δ logphs$_t$						
Δ logphs$_{t-1}$	-9.516	3.189	-8.627	3.158	-9.535	3.109
Δ logphs$_{t-2}$	-5.782	3.168	-6.530	3.154	-7.335	3.121
Δ logphs$_{t-3}$						
Δ logphs$_{t-4}$						
Δ loghf$_t$						
Δ loghf$_{t-1}$	0.134	0.038	0.134	0.038	0.133	0.039
Δ loghf$_{t-2}$	0.110	0.027	0.110	0.028	0.110	0.028
Δ loghf$_{t-3}$						
Δ loghf$_{t-4}$						
Δ CGT$_t$						
Δ CGT$_{t-1}$	-0.002	0.001	-0.002	0.001	-0.002	0.001
Δ CGT$_{t-2}$						
Δ CGT$_{t-3}$						
Δ CGT$_{t-4}$						
Dumrpt$_t$	-0.024	0.010	-0.027	0.010	-0.027	0.010
Dumib$_t$	0.035	0.013	0.036	0.013	0.035	0.013
Dums2327$_t$	0.014	0.010	0.014	0.010	**0.017**	**0.010**
ecm$^D_{t-1}$	0.079	0.026	0.079	0.027	0.075	0.027
ecm$^S_{t-1}$	0.023	0.004	0.022	0.004	0.022	0.004
Dq$_{1t}$						
Dq$_{2t}$	0.016	0.008	0.015	0.008	0.014	0.008
Dq$_{3t}$	0.014	0.008	0.014	0.008	0.014	0.008
Constant	1.645	0.248	1.547	0.241	1.600	0.239
Adjusted R squared		0.62		0.61		0.60
SC information criteria		-6.45		-6.47		-6.50
AR(1-4) test p value		0.44		0.81		0.61
Normality test p value		0.68		0.63		0.44
ARCH (1-4) test p value		0.78		0.60		0.49
Exclusion test p value		0.10		0.12		0.05

Notes: 1. Figures in bold refer to the most insignificant regressor in each model
2. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
3. Probabilities are distributed around a chi-squared distribution with 2 degrees of freedom for the normality test
4. Exclusion test Ho: the coefficient of the bolded variable is zero. The associated test statistic is the log likelihood ratio. If the null cannot be rejected, the variable is dropped from the analysis and the model is re-estimated

Chapter 8:
Testing for the Presence of a Bubble in the Dublin Housing Market

8.1. Introduction

As the existence of a bubble could potentially account for the part of house price determination that is left unexplained by the smooth transition regression model presented in Chapter 7, this chapter proposes to examine the Dublin marketplace for evidence of speculative behaviour. In doing so, a regime-switching approach, which enables testing of the hypothesis that house prices are being driven solely by fundamental market factors against the alternative of a bubble, is adopted.

The chapter is organised as follows[1]: Section 8.2 briefly describes the general regime-switching framework and outlines the nature of the coefficient restrictions required for acceptance of the bubbles model. Fundamental house prices are modelled in Section 8.3 while Section 8.4 presents a measure of the non-fundamental price or bubble component. Estimation of the switching regression model, along with a discussion of the empirical results, is undertaken in Section 8.5. Section 8.6 concludes.

8.2. The Statistical Model

The empirical analysis begins with the decomposition of the price of the housing asset into two components, namely, a 'fundamental' and a 'non-fundamental' part as in equation (8.1):

[1] Econometric modelling undertaken in the course of this chapter was carried out in GiveWin 2.20 / PcGive 10.0, EViews 3.1 and Gauss 5.0.

$$nph_t = nph^f_t + nph^{nf}_t \tag{8.1}$$

where *nph* refers to real new house prices. The fundamental component of the equation implies that house price determination in the Dublin market relates solely to underlying economic factors, whereas the non-fundamental aspect captures any deviation of house prices away from these conditions. If it behaves in a random manner, it follows that the current non-fundamental price will be unhelpful in forecasting the return from investing in housing in the next period. Thus, house prices will on average reflect fundamental values. Conversely, it may hold that expectations of rising prices and associated short-term capital gains attract speculative investors into the housing market. As price movements increasingly reflect such expectations, they become self-fulfilling and house price determination is said to exhibit evidence of non-fundamental behaviour.

With respect to specifying the precise nature of such behaviour, Blanchard (1979) and Blanchard and Watson (1982) propose a stochastic bubbles model in which two states are considered – state S in which the bubble survives and continues to grow, and state C in which it crashes. The probability of state S occurring is assumed to be constant while the expected value of the bubble following a collapse is set equal to zero. However, as discussed in Chapter 5, van Norden (1996) and Schaller and van Norden (1997) argue that these assumptions are overly restrictive. Instead, they suggest a process that allows for a partial collapse in state C and for the probability of the bubble's continued growth to fall as the bubble expands[2].

From a modelling perspective, this idea of a partially collapsing bubble can be readily encompassed within the following regime-switching framework[3]:

[2] The underlying argument against an instantaneous and complete collapse of the bubble relates to the fact that financial institutions or the government would be expected to intervene to prevent such an occurrence. For a detailed description of the typical life cycle of bubbles, see Kindleberger (1989).
[3] For a more technical treatment, please refer to Section 5.5 or van Norden (1996).

$$(R_{t+1}|C) = \beta_{co} + \beta_{cb}b_t + \varepsilon_{c,t+1} \quad (8.2)$$

$$(R_{t+1}|S) = \beta_{so} + \beta_{sb}b_t + \varepsilon_{s,t+1} \quad (8.3)$$

$$\text{Prob}(\text{State}_{t+1}=S) = q(b_t) = \Phi\,(\beta_{q0} + \beta_{qb}b_t^2) \quad (8.4)$$

where:

R	=	the return from investing in the housing asset (constructed as the change in the log of real new house prices);
b	=	the non-fundamental / bubble component of real new house prices[4];
C	=	the collapsing regime;
S	=	the survival regime;
q	=	the probability of the bubble surviving. This probability is bounded between 0 and 1 using the logit function;
Φ	=	the logistic cumulative distribution function;
β_{cb} and β_{sb}	=	coefficients;
β_{co} and β_{so}	=	intercept terms;
ε_c and ε_s	=	the error terms which are assumed independently and identically distributed with mean zero and standard deviations σ_c and σ_s.

If house prices contain a bubble, equations (8.2) to (8.4) should support coefficient restrictions of the form $\beta_{c0} \neq \beta_{s0}$ and either $\beta_{cb} < 0$, $\beta_{sb} > 0$, $\beta_{qb} > 0$, or $\beta_{cb} > 0$, $\beta_{sb} < 0$, $\beta_{qb} < 0$. In addition, the general switching regression should be able to reject a range of simpler nested models, including those of a mixture of normal distributions and an error correction specification. The former imposes the restrictions $\beta_{cb} = \beta_{sb} = \beta_{qb} = 0$, which if upheld, have the interpretation that fundamentals are driving house prices. On the other hand, the error correction model specifies a fad whereby prices deviate temporarily from underlying fundamentals but will eventually mean revert[5]. Such a situation requires $\beta_{co} = \beta_{so} = \beta_o$,

[4] Section 8.4 provides further details.
[5] This type of non-fundamental price behaviour is discussed in detail in Summers (1986).

$\beta_{cb} = \beta_{sb} = \beta_b$, $\beta_{qb} = 0$ and $\beta_b < 0$. All of these restrictions can be tested using standard likelihood ratio tests.

Equations (8.2) to (8.4) therefore represent the baseline statistical model estimated in Section 8.5. The associated sample period begins in the third quarter of 1980 and ends in the first quarter of 2001.

8.3. Modelling Fundamental House Prices

Prior to estimating the regime-switching model outlined above, certain preparatory steps must be undertaken. The first of these relates to the modelling of fundamental house prices. Unfortunately, a unique model of underlying market fundamentals does not exist in a housing context, and it is therefore necessary to proxy the fundamental component of price. As fundamental values depend on the long run equilibrium situation, one means of doing so is to take the fitted value from an estimated long run price relation. Theoretical and empirical considerations suggest that this relation should be between house prices and demand and supply fundamentals [6]. On the demand side, such fundamentals include income, mortgage interest rates and household formation. From a supply perspective, the relevant variables are mortgage interest rates, building costs and the stock of zoned housing land[7].

The house price relation postulated above is essentially a combination of the long run demand and supply relationships considered in Chapter 6. While these do not define equilibrium for the Dublin housing market, it may hold that a reduced form relation normalised on price does. To test for this possibility and to obtain a measure of the fundamental price, the cointegration analysis undertaken earlier is repeated below.

---

[6] See expression (2.4) presented in Chapter 2 and Roche (2003).

[7] The stock of housing is not included here as the variables which influence its long run evolution are modelled directly.

As discussed previously, the objective of the Johansen cointegration technique is to uncover stationary relationships amongst a set of non-stationary data[8]. The results of Dickey-Fuller and Augmented-Dickey Fuller tests for unit roots in the Dublin housing dataset are reported in Table 6.1. These tests reject the null hypothesis of a unit root for each of the variables once the data is differenced. Consequently, any finding of cointegration implies that these non-stationary variables are linearly related in a stationary manner, and as such, can be interpreted as an equilibrium relationship.

Finally, a priori expectations regarding the coefficient signs of the potentially cointegrating variables suggest a positive relationship between house prices and demand-side push factors such as income and demographics. Similarly, a positive sign is expected on the building costs variable, whereas the stock of zoned housing land should theoretically exhibit a negative relationship with the price of housing. However, recalling Chapter 6's empirical finding of a positively signed land coefficient, this theoretical relation is not expected to hold. Likewise, the probable nature of the relationship between house prices and interest rates is unclear given contradictory responses on the part of the demand (negative) and supply (positive) sides of the market to interest rate changes. In this case, the end result depends on which effect wins out.

8.3.1. *The Fundamentals Model*

The estimated vector autoregressive (VAR) model of Dublin house price fundamentals is defined as:

$$\begin{aligned} X_t = {} & \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \Phi_{tr1} Dum_1 bc_t + \Phi_{p1} Dum_1 ls_t + \Phi_{p2} Dum_2 ls_t + \Phi_{p3} Dum_3 ls_t \\ & + \Phi_{p4} Dum_4 ls_t + \Phi_{tr2} Dum_2 bc_t + \Phi_{tr3} Dummr_t + \Phi_{p5} Dum_3 bc_t \\ & + \Phi'_q Dq_t + \mu + \varepsilon_t \end{aligned} \tag{8.5}$$

[8] Section 5.2.1 and Johansen (1995) discuss this technique in detail.

where[9]:

X = a vector comprising the log of real new house prices, the log of real personal disposable income per capita, nominal mortgage interest rates, the log of real building costs, the log of household formation and the log of the stock of zoned housing land;

Dum_1bc = a transitory shock dummy = 1 for 2000:4 and 0 otherwise;

Dum_1ls = a permanent shock dummy = 1 for 1982:2 and 0 otherwise;

Dum_2ls = a permanent shock dummy = 1 for 1993:2 and 0 otherwise;

Dum_3ls = a permanent shock dummy = 1 for 1998:3 and 0 otherwise;

Dum_4ls = a permanent shock dummy = 1 for 1981:3 and 0 otherwise;

Dum_2bc = a transitory shock dummy = 1 for 1981:2, 0 for 1981:3, -1 for 1981:4 and 0 otherwise;

Dummr = a transitory shock dummy = 1 for 1992:4, -1 for 1993:1 and 0 otherwise;

Dum_3bc = a permanent shock dummy = 1 for 1985:2 and 0 otherwise;

Dq = three centred seasonal dummies Dq_1, Dq_2 and Dq_3, where $Dq_i = 0.75$ in quarter 1 and -0.25 in quarters i+1, i+2 and i+3[10];

Π and Φ = unrestricted parameters;

μ = a vector of unrestricted constants;

ε = a vector of error terms.

The sample period (adjusted for a lag length of two) extends from the third quarter of 1980 to the last quarter of 2000.

---

9 A linear trend was included as part of the initial VAR specification. However, subsequent testing indicated that this could be excluded from the analysis.

10 As noted in Chapter 6, Johansen (1995) suggests using centred rather than standard seasonal dummies. Centred dummy variables shift the mean but do not affect the trend.

Re-writing the above model in vector error correction (VECM) format gives rise to the following expression:

$$\Delta X_t = \Pi_1 X_{t-1} + \Gamma_1 \Delta X_{t-1} + \Phi_{tr1} Dum_1bc_t + \Phi_{p1} Dum_1ls_t + \Phi_{p2} Dum_2ls_t + \Phi_{p3} Dum_3ls_t$$
$$+ \Phi_{p4} Dum_4ls_t + \Phi_{tr2} Dum_2bc_t + \Phi_{tr3} Dummr_t + \Phi_{p5} Dum_3bc_t$$
$$+ \Phi'_q Dq_t + \mu + \varepsilon_t \qquad (8.6)$$

where:

Γ = a matrix of short term dynamics.

8.3.2. *Misspecification Testing*

The validity of the vector autoregressive specification presented in equation (8.5) relies on the satisfaction of a number of underlying assumptions as outlined in Section 6.5. Although this entire set of assumptions is not essential, testing for model misspecification is nevertheless important[11]. Reflecting this assertion, the following tables report the information criteria that motivated the choice of VAR lag length and the outcome of tests for residual normality, autocorrelation, homoscedasticity and parameter constancy.

8.3.2.1. Information Criteria and Diagnostic Checking

The multivariate properties of the VAR specification are illustrated below in Tables 8.1 to 8.5. Initial modelling excludes the dummy variables and begins with the selection of an appropriate lag length. In this respect, overall findings are conclusive with the Schwartz (SC) and Hannan-Quinn (HQ) information criteria, along with the Likelihood Ratio test for model reduction, all choosing a lag of one.

[11] For example, parameter constancy and a lack of autocorrelation are considered crucial whereas less weight is typically attached to distributional form and the absence of ARCH.

Table 8.1: Information Criteria for Choice of VAR Lag Length – No Dummies

Lag Length	*SC*	*HQ*
1	-32.668	-33.739
2	-31.280	-32.992
3	-30.377	-32.732
4	-29.056	-32.052

Notes:
1. To ensure validity, the models are nested

Table 8.2: Likelihood Ratio Test for Model Reduction – No Dummies

Lag Length	*Test Statistic*	*Probability*
Lag 4 to Lag 3	F(36,209) = 0.928	0.59
Lag 4 to Lag 2	F(72,261) = 1.350	0.05
Lag 4 to Lag 1	F(108,276) = 1.266	0.06
Lag 4 to Lag 0	F(144,282) = 25.533	0.00***
Lag 3 to Lag 2	F(36,235) = 1.803	0.01**
Lag 3 to Lag 1	F(72,294) = 1.450	0.02**
Lag 3 to Lag 0	F(108,310) = 35.057	0.00***
Lag 2 to Lag 1	F(36,261) = 1.033	0.42
Lag 2 to Lag 0	F(72,326) = 51.854	0.00***
Lag 1 to Lag 0	F(36,288) = 151.01	0.00***

Notes:
1. To ensure validity, the models are nested
2. Ho: Model reduction is acceptable
3. The F version of this test is used as it corrects for small sample size
4. **(***) denotes rejection of Ho at the 5% (1%) significance level

Referring next to the results of the standard diagnostic checks, the residuals appear homoscedastic and the null hypothesis of non-autocorrelation can be accepted for each lag length with the exception of the second - the LM(1) test indicates that serial correlation is present at this lag. Unsurprisingly, residual non-normality is also a strong feature of the model.

Table 8.3: Testing for Serial Correlation in the VAR Residuals – No Dummies

Lag Length	*Test Statistic*	*Probability*
1	LM (1) = 37.851	0.38
	LM (4) = 37.559	0.40
2	LM (1) = 70.443	0.00***
	LM (4) = 35.107	0.51
3	LM (1) = 40.430	0.28
	LM (4) = 33.352	0.60
4	LM (1) = 47.324	0.10
	LM (4) = 37.335	0.41

Notes:
1. Ho: No serial correlation at lag order h
2. Probabilities are distributed around a chi-squared distribution with 36 degrees of freedom
3. A small sample correction is included
4. *** denotes rejection of Ho at the 1% significance level

Table 8.4: Testing for Normality in the VAR Residuals – No Dummies

Lag Length	*Test Statistic*	*Probability*
1	126.096	0.00***
2	111.260	0.00***
3	80.728	0.00***
4	69.195	0.00***

Notes:
1. Ho: Residuals are multivariate normal
2. Probabilities are distributed around a chi-squared distribution with 12 degrees of freedom
3. A small sample correction is included
4. *** denotes rejection of Ho at the 1% significance level

Table 8.5: Testing for Heteroscedasticity in the VAR Residuals – No Dummies

Lag Length	*Test Statistic*	*Probability*
1	F(252,419) = 1.212	0.04**
2	F(504,382) = 0.799	0.99
3	F(756,131) = 0.327	1.00
4	Chi^2(1008) = 1051	0.17

Notes:
1. Ho: Residuals are homoscedastic
2. ** denotes rejection of Ho at the 5% significance level

While a lack of normality is not necessarily desirable, studies by Cheung and Lai (1993) and Gonzolo (1994) have served to demonstrate the robustness of the Johansen procedure to such a situation. However, as pointed out in Section 6.5, accounting for outliers is not only likely to improve the baseline specification, but in addition, may be justified on the grounds that it is a mistake to treat significant interventions and reforms as mere statistical nuisances (Juselius, forthcoming). Given this, it seems appropriate to augment the above VAR(1) specification by the inclusion of the permanent and transitory shock dummies defined in equation (8.5). Note that these dummies are identical to those modelled as part of the VAR analysis presented in Chapter 6. Moreover, as the location and dummy type were chosen using the approach applied in the earlier analysis, the reader is referred back to Section 6.5 for a detailed discussion.

Tables 8.6 to 8.10 report the information criteria for lag selection and the results of the re-applied misspecification tests for the model inclusive of dummy variables.

Table 8.6: Likelihood Ratio Test for Model Reduction – Dummies

Lag Length	*Test Statistic*	*Probability*
Lag 4 to Lag 3	F(36,174) = 0.914	0.61
Lag 4 to Lag 2	F(72,217) = 1.151	0.22
Lag 4 to Lag 1	F(108,230) = 1.404	0.02**
Lag 3 to Lag 2	F(36,200) = 1.411	0.07
Lag 3 to Lag 1	F(72,250) = 1.673	0.002***
Lag 2 to Lag 1	F(36,226) = 1.872	0.003***

Notes:
1. To ensure validity, the models are nested
2. Ho: Model reduction is acceptable
3. The F version of this test is used as it corrects for small sample size
4. **(***) denotes rejection of Ho at the 5% (1%) significance level

Table 8.7: Information Criteria for Choice of VAR Lag Length – Dummies

Lag Length	*SC*	*HQ*
1	-34.306	-36.232
2	-33.477	-36.045
3	-32.497	-35.708
4	-31.286	-35.139

Notes:
1. To ensure validity, the models are nested

Subsequent findings indicate an element of indecisiveness amongst the different criteria with regard to lag length selection. The SC and HQ methods continue to favour a lag of one whereas the Likelihood Ratio test for model reduction suggests that a minimum of two lags is needed. Irrespective of which of these lag lengths is actually chosen, the residuals of the model are homoscedastic and non-normal as before, though the extent of the latter has now been considerably reduced.

Table 8.8: Testing for Serial Correlation in the VAR Residuals – Dummies

Lag Length	*Test Statistic*	*Probability*
1	LM (1) = 56.593	0.02**
	LM (4) = 55.333	0.02**
2	LM (1) = 36.210	0.46
	LM (4) = 45.127	0.14
3	LM (1) = 35.704	0.48
	LM (4) = 46.356	0.12
4	LM (1) = 34.031	0.56
	LM (4) = 46.827	0.11

Notes:
1. Ho: No serial correlation at lag order h
2. Probabilities are distributed around a chi-squared distribution with 36 degrees of freedom
3. A small sample correction is included
4. ** denotes rejection of Ho at the 5% significance level

Table 8.9: Testing for Normality in the VAR Residuals – Dummies

Lag Length	*Test Statistic*	*Probability*
1	40.076	0.00***
2	40.285	0.00***
3	49.366	0.00***
4	37.521	0.00***

Notes:
1. Ho: Residuals are multivariate normal
2. Probabilities are distributed around a chi-squared distribution with 12 degrees of freedom
3. A small sample correction is included
4. *** denotes rejection of Ho at the 1% significance level

Table 8.10: Testing for Heteroscedasticity in the VAR Residuals – Dummies

Lag Length	*Test Statistic*	*Probability*
1	F(252,335) = 0.817	0.95
2	F(504,255) = 0.409	1.00
3	Chi^2(756) = 734.4	0.71
4	Chi^2(1008) = 984.4	0.70

Notes:
1. Ho: Residuals are homoscedastic

On the whole, a lag length of two is marginally preferred given its ability to render serial correlation insignificant with a much higher probability than that of a lag of one. Moreover, two lags fit better with initial expectations and should prove less restrictive than a model estimated using only one. As such, a lag length of two is assumed from this point forward.

8.3.2.2. Analysis of the Residuals of the Estimated VAR Model

The following tables present the outcome of a range of multivariate and univariate diagnostic tests associated with the estimated VAR(2) model. Combined, these findings imply a statistically well-behaved specification.

Table 8.11: Estimated VAR(2) Model – Multivariate Statistics

Test Type	*Test Statistic*	*Probability*
Information Criteria:		
SC	-33.553	
HQ	-36.083	
Serial Correlation:		
Lagrange Multiplier	LM(1) = 28.779	0.80
	LM(4) = 43.482	0.18
Vector AR(1-4) Test	F(144,177) = 0.877	0.79
Normality	43.480	0.00***
Heteroscedasticity	F(504,287) = 0.415	1.00

Notes:
1. Ho: Residuals are respectively serially uncorrelated, multivariate normal and homoscedastic
2. Probabilities are distributed around a chi-squared distribution with 36 degrees of freedom for the LM test and 12 degrees of freedom when testing for normality
3. Where applicable, F versions of the above tests are used to correct for small sample size. The LM test for serial correlation and the normality test also include small sample corrections
4. *** denotes rejection of Ho at the 1% significance level

With respect to the multivariate framework, Table 8.11 indicates that the residuals of the estimated model are neither serially correlated up to lag order four nor heteroscedastic. In contrast, the null hypothesis of residual normality appears not to hold in this setting. As is evident from Table 8.12, much of this non-normality is attributable to the kurtosis problems exhibited by the univariate household formation and land stock series. However, as highlighted in Section 6.5.1 the presence of non-normal residuals does not render the methodology invalid.

Finally, the correlation matrix outlined in Table 8.13 depicts a reasonably high, though not unexpected, degree of correlation between some of the variables, particularly those of house prices and income.

Table 8.12: Estimated VAR(2) Model – Univariate Statistics

Variable	*Mean*	*Std Dev*	*Skewness*	*Kurtosis*
lognph	0.000	0.034	-0.082	2.917
loginc	-0.000	0.016	-0.094	2.782
MR	0.000	0.552	-0.316	3.580
logbc	-0.000	0.008	-0.177	3.785
dloghf	-0.000	0.101	-0.148	4.426
logls	0.000	0.007	0.334	6.427

Variable	*Test Type*	*Test Statistic*	*Probability*
lognph	AR(1-4)	F(4,54) = 0.737	0.57
loginc		F(4,54) = 1.349	0.26
MR		F(4,54) = 1.178	0.33
logbc		F(4,54) = 2.637	0.04**
loghf		F(4,54) = 0.180	0.95
logls		F(4,54) = 2.885	0.03**
lognph	Normality	0.249	0.88
loginc		0.143	0.93
MR		3.176	0.20
logbc		4.416	0.11
loghf		9.810	0.01**
logls		30.569	0.00***
lognph	ARCH (1-4)	F(4,50) = 1.347	0.27
loginc		F(4,50) = 0.496	0.74
MR		F(4,50) = 2.202	0.08*
logbc		F(4,50) = 0.606	0.66
loghf		F(4,50) = 0.824	0.52
logls		F(4,50) = 1.773	0.15

Notes:
1. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
2. Probabilities for the normality test are distributed around a chi-squared distribution with 2 degrees of freedom
3. *(**) (***) denotes rejection of Ho at the 10% (5%) (1%) significance level

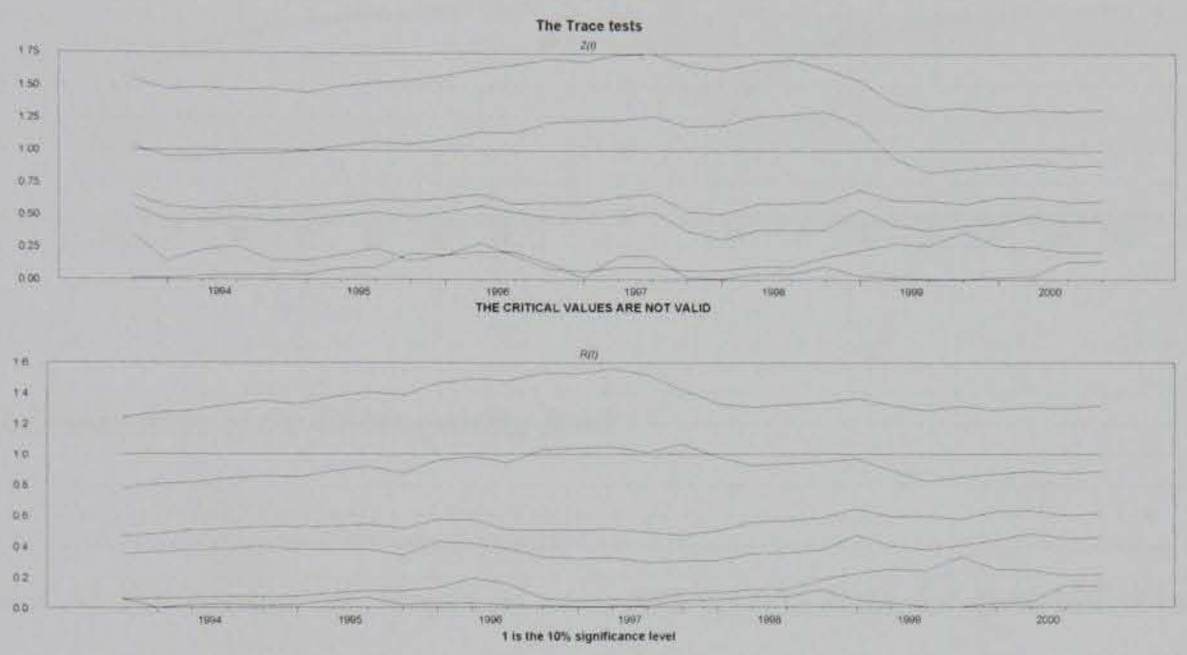
Table 8.13: Estimated VAR(2) Model – Correlation Matrix

Variable	*dlognph*	*dloginc*	*dMR*	*dlogbc*	*dloghf*	*dlogls*
dlognph	1.000					
dloginc	0.442	1.000				
dMR	-0.264	-0.090	1.000			
dlogbc	0.271	0.079	-0.177	1.000		
dloghf	-0.048	-0.024	-0.363	0.272	1.000	
dlogls	0.098	0.222	0.227	0.049	0.120	1.000

8.3.2.3. Recursive Analysis

Recalling that the chosen specification must also satisfy the assumption of parameter constancy in order for the Johansen cointegration technique to be valid, recursive estimation was undertaken on a base sample covering the period 1980:3 to 1993:4[12]. The resultant graphics are presented below in Figures 8.1 and 8.2.

Figure 8.1: Recursive Analysis – Trace Test (Base Sample 1980:3 to 1993:4) [13]

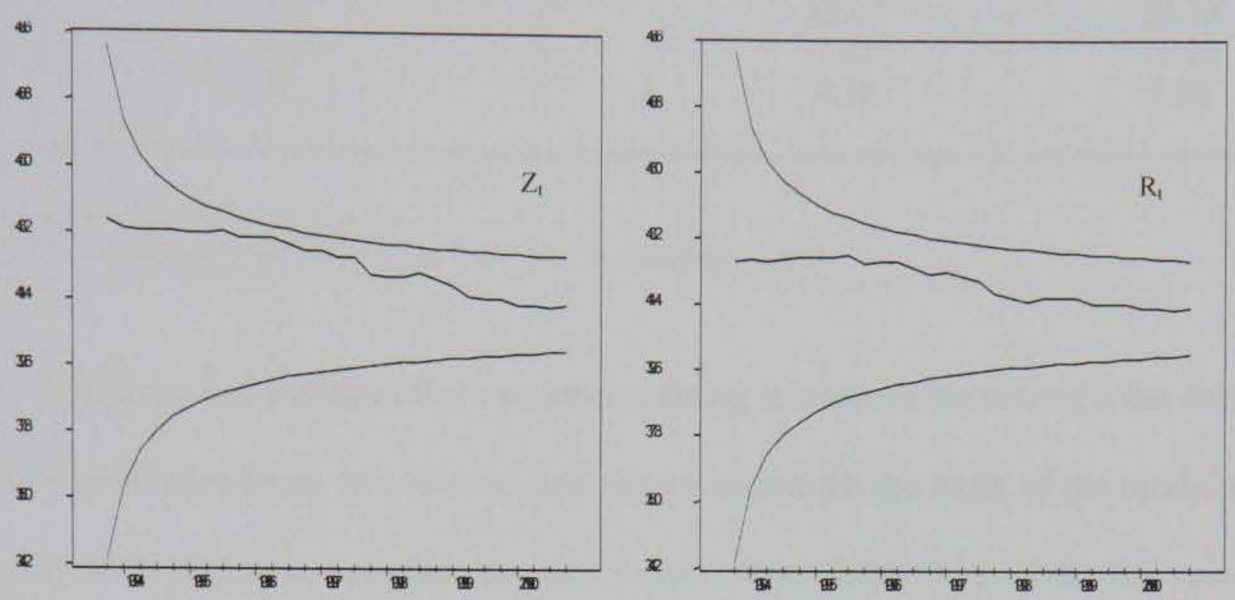


Referring first to the recursively calculated Trace test, Figure 8.1 suggests five unit roots and one stationary relationship amongst the variables. As discussed in Section 6.5.3, the presence of a unit root can be rejected / accepted for components of the Trace statistic which lie above / below the critical value of one; though for models containing dummy variables, a critical value in excess of one is considered to be more appropriate. Secondly, the recursively calculated log likelihood value, graphed in Figure 8.2, lies within the 95% confidence bands

[12] To test the sensitivity of these results, a variety of time periods were modelled as the base sample. As the associated graphs proved similar to those presented in Figures 8.1 and 8.2, they are not reproduced here.
[13] As the estimated model contains a number of dummy variables, the critical values given in this figure are inappropriate. The presence of such dummies implies that the line at one should be higher.

for the full (Z_t) and concentrated versions (R_t) of the model as desired[14]. Taken together, these findings are indicative of a model with relatively constant parameters[15].

Figure 8.2: Recursive Analysis – The Log Likelihood Value (Base Sample 1980:3 to 1993:4)



8.3.3. *Determination of the Cointegrating Rank*

Conditional on the chosen lag length of two, this section tests for possible cointegration amongst the variables in the above vector autoregressive model. Similar to the approach adopted in Chapter 6, the standard test used to identify the cointegrating rank (r) - the Trace test - is applied along with more informal methods such as an examination of the roots of the companion matrix and the t values of the alpha coefficients of the r^{th}+1 cointegrating vector. As discussed earlier, the fact that 'the asymptotic distribution [of the Trace statistic] can be a poor approximation to the true distribution when the sample size is small resulting in substantial size and power distortions' implies that as much additional information as possible be used when determining the cointegration rank (Juselius, forthcoming)[16]. Tables 8.14 to 8.16 detail the findings of the above methodologies.

---

[14] Note that the concentrated version of the model removes all dynamic influences.
[15] It is also unlikely that the tests for autocorrelation would have been passed if the parameters of the model are not constant.
[16] [ ] parenthesis inserted.

Table 8.14: Testing for the Cointegrating Rank – The Trace Test

Eigenvalues	*Ho:r*	*p-r*	*Trace Statistic*	*95% Critical Value*
0.5212	0	6	117.78	93.92
0.3127	1	5	57.39	68.68
0.1627	2	4	26.64	47.21
0.1068	3	3	12.07	29.38
0.0293	4	2	2.81	15.34
0.0046	5	1	0.38	3.84

Notes:
1. Ho: r = q (where q = 0, 1, 2, 3, 4, 5)
2. The 95% critical values are taken from Table 15.3 in Johansen (1995)

Formally, the Trace test indicates that one cointegrating relation exists amongst the data series at the 5% significance level. In contrast, rank determination on the basis of the modulus roots of the companion matrix is less clear cut as is evident from Table 8.15. If the full rank of six is selected, at least two unit roots and therefore four cointegrating vectors appear to be present. However, if the next largest roots of 0.92 and 0.88 are regarded as being sufficiently close in size to one as to constitute unit roots, it follows that five such roots and a single cointegrating relation are actually present.

Table 8.15: Testing for the Cointegrating Rank – Roots of the Companion Matrix

p-r	*r=6*	*r=4*	*r=1*
6	1.00	1.00	1.00
5	1.00	1.00	1.00
4	0.92	0.98	1.00
3	0.88	0.86	1.00
2	0.88	0.86	1.00
1	0.43	0.43	0.82

Notes:
1. p refers to the number of variables in the analysis. p-r = 6 unit roots corresponds to zero cointegration while p-r = 5 unit roots implies the presence of one cointegrating relation amongst the data series

Given the mixed nature of these findings, Table 8.15 also reports the results of a sensitivity analysis that sets the rank equal to four and then one. This analysis reveals that the imposition of a rank of four leaves an unacceptably high root of 0.98 in the model, whereas setting r = 1 gives rise to a much smaller remaining root of 0.82. Thus, a rank choice of one is deemed appropriate and is in line with the implications of the Trace test.

Table 8.16: Testing for the Cointegrating Rank – Unrestricted Estimates of the VAR(2) Model

Beta	*lognph*	*loginc*	*MR*	*logbc*	*loghf*	*logls*
1	-6.798	12.976	0.547	-16.440	9.093	6.408
2	-8.534	15.406	-0.452	-12.821	11.677	-2.945
3	3.591	-16.587	0.197	13.951	3.287	16.178
4	-0.761	3.707	0.524	16.008	4.207	3.787
5	0.833	4.232	-0.120	-28.803	0.917	5.727
6	9.598	-13.691	-0.108	8.940	-1.693	-2.150
Alpha						
dlognph	-0.019	0.005	-0.003	0.003	0.002	-0.002
dloginc	-0.005	-0.002	0.003	0.003	-0.000	-0.000
dMR	-0.170	0.171	0.048	-0.037	-0.070	0.011
dlogbc	0.001	0.001	0.001	-0.001	0.000	-0.000
dloghf	-0.016	-0.056	-0.004	-0.017	0.001	-0.002
dlogls	0.001	-0.000	-0.001	0.000	-0.001	-0.000
Alpha T Values						
dlognph	-5.013	-1.426	-0.731	0.803	0.520	-0.415
dloginc	-2.801	-1.413	1.930	1.987	-0.199	-0.279
dMR	-2.786	2.810	0.785	-0.613	-1.152	0.182
dlogbc	1.657	0.735	1.595	-1.604	0.328	-0.430
dloghf	-1.403	-4.990	-0.326	-1.555	0.094	-0.142
dlogls	1.299	-0.610	-1.302	0.294	-1.189	-0.328

Notes:
1. Beta corresponds to the cointegrating vectors while alpha reflects the adjustment coefficients

Lastly, the alpha t values of the unrestricted estimates are considered in Table 8.16. Interestingly, the first and second vectors are significant, which suggests a rank of two as opposed to the previous findings of a rank of one[17].

[17] With respect to the significance levels of the alpha coefficients, as their distribution is unknown a t value of roughly 3.5 (in absolute terms) is regarded as more appropriate than the standard rule of thumb figure of 2. In addition, note that the significance of a given beta vector requires a high t value for only one of the related alpha coefficients.

Drawing together the information provided by the above formal and informal criteria, the final choice of rank lies between one and two cointegrating vectors. Overall, the results are slightly more skewed towards the former than the latter and consequently, a rank of one is selected. Besides, the existence of a single equilibrium relation amongst the chosen variables fits better with a priori theoretical and empirical expectations than that of two vectors. As such, further modelling on this basis is assumed to be valid.

8.3.4. *Model Specific Data Properties*

This section is concerned with examining the statistical properties of the housing data conditional on the chosen rank of one. In particular, the variables included as part of the estimated VAR(2) model are tested for potential long run exclusion, stationarity and weak exogeneity. The resultant findings are outlined in Table 8.17 overleaf.

The first of these tests considers whether the variable in question can or cannot be omitted from the cointegration space. As is apparent from Table 8.17, the corresponding null hypothesis of long run exclusion is rejected at the 5% significance level for all variables.

Secondly, Table 8.17 indicates that none of the variables are individually stationary, thereby confirming the results of the earlier unit root tests. As such, the uncovered cointegration relation cannot be accounted for by the presence of a stationary variable.

The third test differs from those discussed above in that it imposes restrictions on the alpha rather than on the beta coefficients. In doing so, the possibility that a variable may influence the long run stochastic path of other variables in the system without being influenced by them - a 'no levels feedback' effect - is allowed for. The results presented in Table 8.17 fail to reject the null hypothesis of weak exogeneity for the building costs, household formation and land stock variables; though the reasonably high probabilities pertaining to the last two series are of most concern. These probabilities point towards acceptance of both the household formation and land stock series as weakly exogenous. It then follows that these two variables constitute two of the model's common driving trends[18].

Table 8.17: Estimated VAR(2) Model – Model Specific Data Properties

Exclusion Test (Chi squared (1) = 3.84)					
lognph	*loginc*	*MR*	*logbc*	*loghf*	*logls*
9.97 (0.00)	9.19 (0.00)	15.62 (0.00)	9.56 (0.00)	11.82 (0.00)	6.76 (0.01)
Stationarity Test (Chi squared (5) = 11.07)					
lognph	*loginc*	*MR*	*logbc*	*loghf*	*logls*
49.47 (0.00)	44.28 (0.00)	24.94 (0.00)	35.34 (0.00)	43.08 (0.00)	44.81 (0.00)
Weak Exogeneity Test (Chi squared (1) = 3.84)					
lognph	*loginc*	*MR*	*logbc*	*loghf*	*logls*
18.01 (0.00)	5.70 (0.02)	5.23 (0.02)	2.30 (0.13)	0.93 (0.33)	1.51 (0.22)

Notes:
1. The above tests are conditional on a cointegrating rank of one
2. Respectively, the null hypothesis states that the variable under consideration should be excluded, is individually stationary and is weakly exogeneous
3. p values are given in brackets

8.3.5. *The Partial Model*

Building on the above findings of weak exogeneity, this section of the study adopts a partial systems approach to the modelling of housing market fundamentals. Such an approach implies that valid inference on beta can be obtained from a four dimensional system

[18] This refers to the fact that the unexplained variation of $\varepsilon_{loghf,t}$ and $\varepsilon_{logls,t}$ cumulate in the system so that $\Sigma_{i=1}^{t} \varepsilon_{loghf,t}$ and $\Sigma_{i=1}^{t} \varepsilon_{logls,t}$ are two of the common trends. In this sense, household formation and the stock of zoned housing land 'drive' the system.

comprising house prices, income, interest rates and building costs conditioned on the household formation and land stock variables (Juselius, forthcoming). Moreover, moving to a partial model generally has the added advantage of leading to an improved statistical specification. In addition, it gives rise to the prospect of eliminating insignificant dummies relating to the weakly exogenous variables.

Table 8.18: The Partial Model – Short Run Matrices

Variable	Dum_1bc	Dum_1ls	Dum_2ls	Dum_3ls	Dum_4ls	Dum_2bc	*Dummr*	Dum_3bc
dlognph	0.004	-0.080	-0.009	-0.049	-0.052	0.009	0.028	0.050
dloginc	-0.027	-0.076	-0.031	-0.041	-0.020	0.059	0.015	0.029
dMR	0.369	-1.645	-6.003	-0.899	1.895	-0.453	2.705	0.054
dlogbc	0.100	-0.027	0.029	0.005	-0.029	0.063	0.002	0.041
T Values								
dlognph	0.105	-1.163	-0.139	-0.907	-0.958	0.320	0.973	1.379
dloginc	-1.556	-2.405	-1.031	-1.654	-0.806	4.644	1.128	1.734
dMR	0.665	-1.631	-6.153	-1.124	2.369	-1.108	6.469	0.102
dlogbc	11.908	-1.748	1.984	0.374	-2.445	10.272	0.335	5.072

Table 8.18 presents the short run matrices associated with the partial model. With regard to the results, it appears that three of the four land dummy variables are insignificant following estimation of the conditional VAR(2) model, and can therefore be dropped from the analysis[19]. As such, the empirical model discussed from now on is as defined in equation (8.5), but with the household formation and land stock series assumed to be weakly exogenous and with Dum_1ls, Dum_3ls and Dum_4ls excluded.

Finally, as discussed in Section 6.8, the asymptotic distribution of the rank test statistics for the partial model differs from that of the full system. Given this, rank selection may prove problematic (Johansen, 1995). However, as the rank has already been determined within the

[19] The rule of thumb that dummy variables should have an absolute t value of at least 3.73 in order to be considered significant is applied here.

full framework, and the model subsequently tested for weak exogeneity, Hansen and Juselius (1995) suggest that there is no further need to test for rank in the partial system. Their recommendation, and the method adopted below, is to proceed on the basis of the previously chosen rank.

8.3.5.1. *Analysis of the Residuals of the Partial Model*

Tables 8.19 to 8.21 subject the residuals of the partial system to the various diagnostic checks imposed previously on the full model.

Table 8.19: The Partial Model – Correlation Matrix

Variable	*dlognph*	*dloginc*	*dMR*	*dlogbc*
dlognph	1.000			
dloginc	0.442	1.000		
dMR	-0.308	-0.092	1.000	
dlogbc	0.303	0.106	-0.137	1.000

Table 8.20: The Partial Model – Multivariate Statistics

Test Type	*Test Statistic*	*Probability*
Information Criteria:		
SC	-21.455	
HQ	-23.072	
Serial Correlation:		
Lagrange Multiplier	LM(1) = 6.729	0.98
	LM(4) = 11.633	0.77
Vector AR(1-4) Test	F(64,158) = 0.837	0.79
Normality	4.667	0.79
Heteroscedasticity	F(200,283) = 0.656	1.00

Notes:
1. Ho: Residuals are respectively serially uncorrelated, multivariate normal and homoscedastic
2. Probabilities are distributed around a chi-squared distribution with 16 degrees of freedom for the LM test and 8 degrees of freedom when testing for normality
3. Where applicable, F versions of the above tests are used to correct for small sample size. The LM test for serial correlation and the normality test also include small sample corrections

Table 8.21: The Partial Model – Univariate Statistics

Variable	*Mean*	*Std Dev*	*Skewness*	*Kurtosis*
lognph	0.000	0.034	-0.061	2.943
loginc	-0.000	0.016	-0.049	2.736
MR	0.000	0.543	0.103	3.285
logbc	-0.000	0.008	0.056	3.058

Variable	*Test Type*	*Test Statistic*	*Probability*
lognph	AR(1-4)	F(4,55) = 0.566	0.69
loginc		F(4,55) = 1.161	0.34
MR		F(4,55) = 0.824	0.52
logbc		F(4,55) = 1.327	0.27
lognph	Normality	0.264	0.88
loginc		0.038	0.98
MR		1.465	0.48
logbc		0.548	0.76
lognph	ARCH (1-4)	F(4,51) = 1.380	0.25
loginc		F(4,51) = 0.550	0.70
MR		F(4,51) = 3.062	0.02**
logbc		F(4,51) = 0.656	0.63

Notes:

1. Ho: Residuals are respectively serially uncorrelated, normal and contain no ARCH effects
2. Probabilities for the normality test are distributed around a chi-squared distribution with 2 degrees of freedom
3. ** denotes rejection of Ho at the 5% significance level

With respect to the outcomes, the correlation matrix and homoscedasticity findings are for the most part unchanged, while the multivariate statistics indicate higher values for the probabilities of the serial correlation tests. Furthermore, the residuals now approximate a normal distribution. Given these improvements in the statistical framework, it may be concluded that the re-specification of the VAR(2) model as a partial system is not only appropriate but also, highly desirable.

8.3.6. *Long Run Identification*

By definition, the single cointegrating vector uncovered from the housing dataset is identified and unique. However, the interpretability of this vector as a plausible house price relation may benefit from the imposition of additional restrictions. As such, a series of over-identified models is derived and compared with the exactly identified model using the likelihood ratio test described in Johansen and Juselius (1994). Tables 8.22 and 8.23 present the empirical results[20].

Table 8.22: The Exactly Identified Model[21]

Beta	*lognph*	*loginc*	*MR*	*logbc*	*loghf*	*logls*
	1.000	-1.430	-0.178	3.071	-1.817	-2.969
Standard Errors of Beta						
	0.000	0.425	0.033	1.317	0.429	0.628
Alpha						
	0.071	0.013	0.891	-0.006		
Standard Errors of Alpha						
	0.016	0.008	0.257	0.004		
Alpha T Values						
	4.494	1.594	3.468	-1.578		

Notes:
1. Log likelihood = 601.856106
2. Beta corresponds to the cointegrating vector while alpha reflects the adjustment coefficients

Referring to Table 8.23, it appears that the placing of unit restrictions on the coefficients of the income, interest rate and building cost variables is valid. The resultant long run price relation takes the following form[22]:

Price relation: lognph – loginc – MR – logbc – 4.941 loghf – 12.039 logls

$$\text{lognph} = \text{loginc} + \text{MR} + \text{logbc} + 4.941\ \text{loghf} + 12.039\ \text{logls} \tag{8.7}$$

where ε_t - I(0).

[20] While a wide range of hypotheses were considered and tested, only the chosen over-identified model is discussed in the following.
[21] Notice that as the housing stock is not included in the above analysis, the model presented here is similar to the just identified supply relation set out in Table 6.23.
[22] These restrictions are in line with those imposed on the demand and supply relations in Chapter 6.

Table 8.23: The Over-Identified Model

Beta	*lognph*	*loginc*	*MR*	*logbc*	*loghf*	*logls*
	1.000	-1.000	-1.000	-1.000	-4.941	-12.039
Standard Errors of Beta						
	0.000	0.000	0.000	0.000	2.067	1.638
Alpha						
	0.015	0.002	0.201	-0.001		
Standard Errors of Alpha						
	0.004	0.002	0.056	0.001		
Alpha T Values						
	4.182	0.944	3.565	-0.751		

Notes:
1. Log likelihood = 598.32937
2. Likelihood ratio test: Chi^2 (3) = 7.054 with a probability of 0.07
3. Beta corresponds to the cointegrating vector while alpha reflects the adjustment coefficients

In line with expectations, the above findings indicate a positive relationship between house prices and income and household formation. Similarly, the mortgage interest rate variable is positively signed which suggests that the reduction in supply and the upward price pressure associated with the higher opportunity cost of borrowing, outweighs the downward pressure placed on prices from dampened demand. The nominal nature of the interest rate variable also plays a role in this context. In particular, the presence of high uncertainty and inflation often leads to the use of housing as a hedge against future inflation and thus, as the nominal mortgage interest rate increases, the demand for and the price of housing rise. Therefore, it is possible that some of the observed positive interest rate elasticity can be attributed to hedging.

With respect to the supply-side factors, both building costs and the stock of land enter expression (8.7) bearing positively signed coefficients. As discussed in Section 6.9, the former finding indicates that increased expenditure on the part of suppliers is fully passed onto consumers in the Dublin market, whereas the latter may point to an element of hoarding on the part of landowners, or alternatively, could reflect delays in zoned housing land coming on stream due to a lack of services or planning permission.

Figure 8.3: Long Run House Price Relation

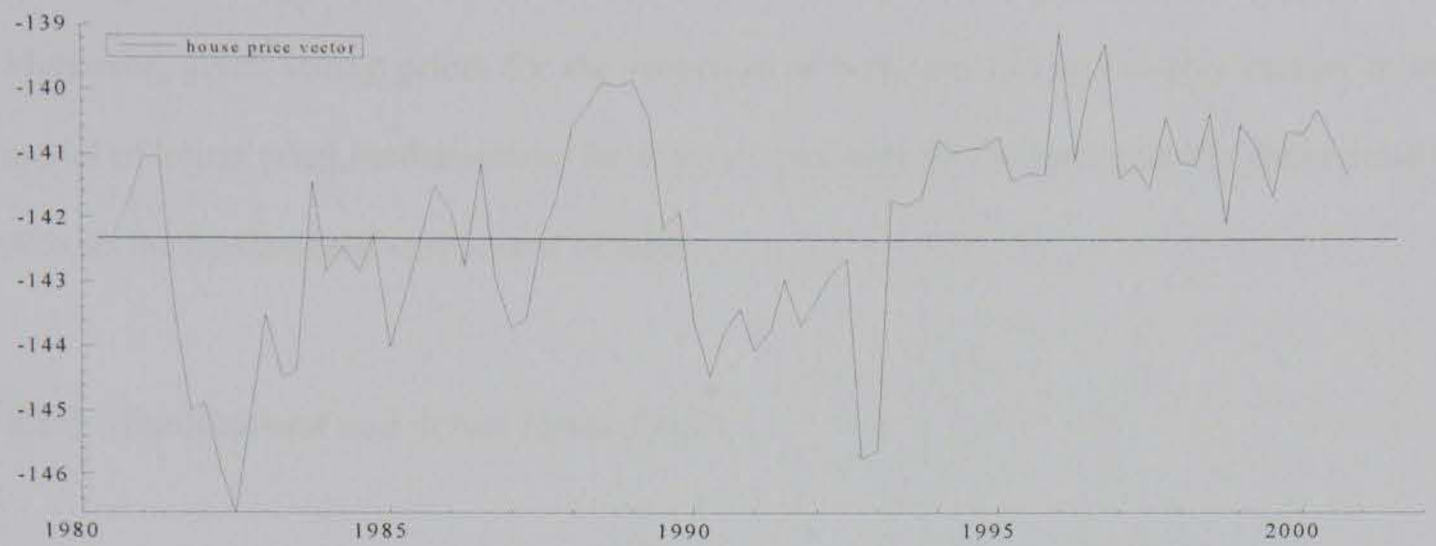


Figure 8.3 illustrates the identified long run house price relation. The graphical depiction shows that from the end of 1987 to 1989, and from approximately 1994 onwards, house prices in Dublin were high relative to the prevailing levels of income, mortgage interest rates, building costs, household formation and land. Accordingly, the elimination of disequilibria in these periods would have required downward price adjustment. Of further interest is the close relationship between this vector and the supply relation presented in Figure 6.5. These similarities are largely explained by the coefficient on the mortgage interest rate which is one in both models.

Before accepting the above relation as an equilibrium relationship, it is essential to consider its adjustment properties, namely the alpha value of the house price variable[23]. This should bear a minus sign reflecting the elimination of steady state deviations and a return to the long

[23] As noted in Chapter 5, the Johansen approach allows for long run effects and short term adjustment patterns to be modelled simultaneously. The sequential manner in which the long run relation and its adjustment properties are discussed here is only for convenience.

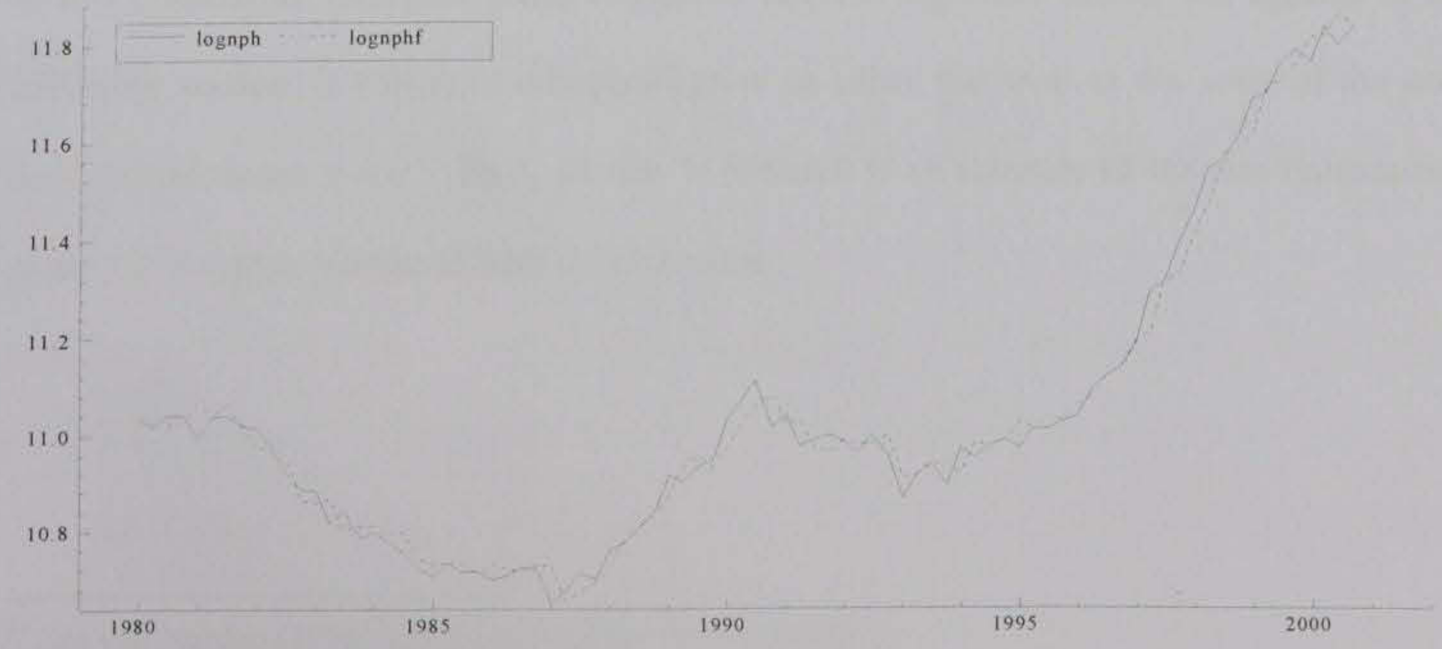
run path. However, the associated coefficient (see Table 8.23) is positively signed, indicative of a move away from equilibrium rather than adjustment towards it.

To the extent that house prices may adjust in a non-linear manner to deviations from the long run path, the hypothesis of no cointegration is not necessarily acceptable at this point. Moreover, given strong priors for the inclusion of both demand and supply factors in any model of house price fundamentals, the analysis proceeds on the basis that the above relation proxies the fundamental component of price.

8.3.7. *Fundamental and Actual House Prices*

Plotting the fitted values of the model presented in Section 8.3 against observed prices enables a comparison of fundamental values and actual house prices. Figure 8.4 presents a graph of the two series. As is evident from the plot, the fundamental component of price and actual house prices were close in terms of fit for much of the sample period, though interestingly; actual prices typically exceeded fundamental values around 1990 and from 1997 up to the beginning of 2000. This finding lends support to the notion that there may have been an element of speculative behaviour at play in the Dublin marketplace during these years.

Figure 8.4: The Fundamentals Model – Actual and Fundamental House Prices



The above measure of the fundamental price is somewhat noisy given that it is estimated from a vector autoregressive model which comprises lagged values of the house price variable. If it holds that house prices contain a bubble or a fad element, then the use of lagged prices in the baseline model implies that such elements will consequently have been incorporated into the fundamental proxy of price. As a result, test findings will tend to be biased against the presence of speculative behaviour (Cutler, Poterba and Summers, 1991). However, Schaller and van Norden (1997), referencing Cutler, Poterba and Summers (1991), state that 'if we use a noisy measure of fundamental prices and still find evidence of either fads or bubbles .... the noisiness of our measure of fundamentals strengthens the case that non-fundamentals matter'. As such, the present analysis proceeds using the fundamental measure of Dublin house prices derived hitherto.

8.4. The Non-Fundamental Component of House Prices

Prior to estimating the regime-switching model outlined in equations (8.2) to (8.4), certain preparatory steps must be undertaken of which modelling fundamental house prices represented the first. The second step is concerned with measuring the non-fundamental component of price. Taking the difference between actual house prices and the fitted value of the model presented in Section 8.3, a proxy for the non-fundamental price is readily obtained.

In this context, an important point to note is that the regime-switching test applied in the following section, is robust to misspecification in either the level or the scale of the non-fundamental house price[24]. Thus, all that is required is an estimate of the non-fundamental price that is highly correlated with the true value.

[24] See van Norden (1996) and Section 5.4 for further details.

8.5. The Regime-Switching Model

The switching regression model outlined earlier in Section 8.2 is estimated below. Overall, the findings are not suggestive of a stochastic bubble in the Dublin housing market. Instead, Tables 8.24 to 8.26 provide evidence of fad type behaviour whereby prices deviate temporarily from underlying fundamentals but will eventually mean revert.

Bearing in mind that the bubbles model requires the satisfaction of various restrictions on the coefficients of equations (8.2) to (8.4), namely those of $\beta_{c0} \neq \beta_{s0}$ and either $\beta_{cb} < 0$, $\beta_{sb} > 0$, $\beta_{qb} > 0$, or $\beta_{cb} > 0$, $\beta_{sb} < 0$, $\beta_{qb} < 0$; consider Table 8.24 and the results of the estimated regime-switching model.

Table 8.24: Estimated Regime Switching Model – Coefficient Restrictions

Parameters	*Coefficient*	*T Value*
β_{co}	-183	0.92
β_{so}	0.017	1.67
β_{cb}	$3.47e^{005}$	12.98
β_{sb}	0.003	1.28
β_{qb}	0.055	0.90
Wald Test	*Test Statistic*	*Probability*
$\beta_{co} = \beta_{so}$	2.006	0.16

Notes:
1. T and Wald tests are based on standard errors taken from the inverse of the Hessian

Looking to begin with at the first set of restrictions, it is evident that the coefficient of β_{cb} is significant, although it is incorrectly signed. In contrast, the estimates of β_{sb} and β_{qb} are signed as expected but prove insignificant at conventional levels. Given equivalence, these findings are reversed when the second set of restrictions are considered; β_{cb} is now correctly signed whereas the estimates of β_{sb} and β_{qb} are not. With respect to the Wald test, the data fails to reject the hypothesis that $\beta_{c0} = \beta_{s0}$ despite the large numerical difference between the two coefficients. This finding similarly conflicts with the predictions of the bubbles model. As such, it may be concluded that the results presented in Table 8.24 provide little evidence in support of bubble type behaviour in the Dublin housing market.

Table 8.25: Estimated Regime Switching Model – Nested Specifications

Model Type	*Test Statistic*	*Probability*
Mixture of Normal Distributions	71.020	0.00***
Error Correction Model	1.306	0.52

Notes:
1. Log likelihood of the regime switching model = 2.35950
2. Ho: Do not reject the model in question in favour of the variable regime switching model
3. An assumption of varying probability underlies both the mixture of normal distributions and the error correction models
4. *** denotes rejection of Ho at the 1% significance level

Furthermore, acceptance of the bubbles model requires the rejection of a range of simpler models nested within the general regime-switching framework. Table 8.25 presents the outcomes of the standard likelihood ratio tests applied for this purpose. The subsequent empirics indicate rejection of the null hypothesis in the case of the mixture of normal distributions model, but fail to do so for the error correction alternative. While the former finding suggests that a bubble component may be favoured over fundamentals, the latter points to the rejection of a stochastic bubble and in place, implies the existence of a fad in the Dublin housing marketplace.

Lastly, the diagnostic tests reported in Table 8.26 reveal a lack of ARCH and Markov switching effects, but find some evidence of serial correlation in the second regime.

Table 8.26: Estimated Regime Switching Model – Misspecification Testing

Test Type	*Test Statistic*	*Probability*
Serial Correlation:		
AR(1) in regime 1	3.674	0.06*
AR(1) in regime 2	15.286	0.00***
Heteroscedasticity:		
ARCH(1) in regime 1	0.385	0.53
ARCH(1) in regime 2	2.444	0.12
Markov Effects	1.002	0.32

Notes:
1. Ho: Residuals are respectively serially uncorrelated and contain no ARCH or Markov switching effects
2. The tests for AR and ARCH are Lagrange Multiplier tests
3. *(***) denotes rejection of Ho at the 10% (1%) significance level

8.6. Conclusion

Building on the analyses presented in Chapter 7, this chapter sought to examine the Dublin new private sector housing market for evidence of speculative behaviour. In doing so, a two-pronged approach was adopted. Firstly, the price of the housing asset was decomposed into two components, namely a 'fundamental' and a 'non-fundamental' part. The former suggests that house price determination relates solely to underlying economic factors, whereas the non-fundamental aspect captures any deviation of house prices away from these conditions. As the latter may prove useful in forecasting the return from investing in housing in the next period, the second prong tested for this possibility using the regime-switching approach proposed by van Norden (1996) and Schaller and van Norden (1997).

The estimated fundamental price includes both demand and supply variables and when graphed against actual house prices, is found to be close in fit for much of the sample period. However, actual prices typically exceeded fundamental values around 1990 and from 1997 up to the beginning of 2000. This suggests that there may have been an element of speculative behaviour at play in the Dublin marketplace during these years.

The outcome of the regime-switching tests confirms that this was the case. Interestingly, the empirical findings indicate the presence of a fad in the marketplace as opposed to a stochastic bubble. Fads give rise to a situation whereby prices deviate temporarily from, but will eventually revert to, underlying fundamentals. Thus, to the extent that a fad existed and house prices exceeded fundamental values towards the end of the sample period, a 'soft landing' may be anticipated for the future path of the Dublin housing market rather than the collapse type scenario associated with a bubble.

Chapter 9:
Conclusion

Applying a range of time series econometric techniques, this thesis has sought to comprehensively model the Dublin market for new, private sector housing over the period 1980-2000. As discussed in Chapter 1, a clear understanding of the empirical conditions that underlie the housing market is a necessary prerequisite for solving disequilibria. However, as borne out by Chapter 3's review of the literature, little research into the factors that lie behind developments in the Dublin marketplace has been carried out to date. By modelling the Dublin housing market within a long run and adjustment type framework, the present study has attempted to overcome this dearth of research. A summary of the approach taken and the resultant findings is provided below.

As any empirical analysis must be conducted within the framework of economic theory, Chapter 2 began by considering the theoretical functioning of the housing market. In doing so, attention was drawn to the special attributes of housing, and to the manner in which these features are typically modified so as to enable time series modelling of the housing market. The key theoretical relationships applicable in a housing context were also set out.

To facilitate the identification of appropriate factors for inclusion in the empirical models of the Dublin marketplace, Chapter 3 then detailed the findings of selected Irish and international studies that have adopted a time series, macroeconomic approach to modelling housing. The limitations of these studies were also considered. Of particular importance in this respect was the incomplete modelling of the supply-side of the market, and in many cases, the failure to examine the time series properties of the data and consequently, the use of inappropriate modelling techniques.

Overall, there appears to be a broad consensus on the part of the literature as to the significance in the housing market of price, income, interest rates, building costs, the existing housing stock, demographics, land and government policies. Moreover, the empirical findings in this respect also fit well with a priori theoretical expectations as to the determinants of housing demand, supply and price inflation.

Building on the analysis in Chapter 3, Chapter 4 sought to provide a description of the data series that underlie the empirical sections of the study and to discuss their time series properties. The findings indicate considerable limitations with respect to the available information. In particular, it proved necessary to refine much of the data and to generate a measure of the stock of zoned housing land. Fortunately, these limitations were for the most part circumvented, leading to the compilation of a comprehensive dataset of Dublin housing market variables for the period 1980-2000.

Next, to aid understanding, Chapter 5 presented a theoretical overview of the modelling techniques adopted in the course of the study, namely those of Johansen cointegration, smooth transition regression modelling and regime-switching tests for stochastic bubbles. These techniques draw on the time series properties of the Dublin housing market data.

With respect to summarising the study's empirical findings, the starting point is the outcome of the long run analysis undertaken in Chapter 6. In line with expectations, the results identify the presence of two cointegrating vectors. The demand relation, which is normalised on the housing stock variable, shows that the stock of housing was below its desired level – that is the level consistent with the prevailing level of house prices, income, mortgage interest rates and demographics – over the time periods 1986-1989, 1993-1997 and 1999. The supply vector is normalised on house prices and suggests that from the end of 1987 to 1989, and from 1994 onwards, the price of housing exceeded its equilibrium relationship with mortgage interest rates, building costs and land.

With the exception of the land stock series, all of the above variables are signed as anticipated a priori. Interestingly, the latter enters the supply-side relation with a positive as opposed to a negatively signed coefficient. As such, the initial expectation that an increase in the stock of zoned housing land would translate into an expansion in supply, and a subsequent fall in the price of housing, does not materialise. A possible explanation, and one that would seem to fit well with on the ground opinion in the Dublin market, is that of hoarding on the part of landowners. However, it should be remembered that while the land concerned may have been zoned for housing purposes, it might not be in receipt of services or planning permission, which would also result in a delay in its coming on stream.

Moreover, an analysis of the adjustment properties of the above relations reveals a failure on the part of both the stock and price of housing to adjust to steady state deviations and restore the market to equilibrium. This lack of error correction is of concern as it implies that the identified demand and supply relationships may not in fact define equilibrium for the Dublin housing market. Alternatively, it could be the case that adjustment behaviour in this marketplace is non-linear in nature, as opposed to linear. Hence, prior to accepting a hypothesis of no cointegration, this latter possibility is tested.

In terms of adding to the literature, Chapter 6 represents the first attempt to include a land availability variable in a long run model of the housing market. The potentially important role of land is often referred to in international studies, but due to a lack of data on costs and availability, is typically not modelled. While recent work on the national market has incorporated somewhat crude measures of land prices, no effort has been made to consider land in empirical studies of the Dublin marketplace. Besides, the shortage of zoned and serviced housing land in the Dublin area implies that availability, as modelled in this study, is a more appropriate land measure than price.

Turning next to the adjustment analysis, the linear error correction model estimated in Chapter 7 pinpoints a range of significant determinants of house price inflation - current and lagged changes in income and household formation are positively related to prices in the Dublin market, capital gains taxation and the housing stock are found to exhibit a negative relationship with the current rate of house price growth while the 'Section 23/27' relief, the tax deductibility of interest on borrowings and the property tax dummy variables all enter the model significantly and bearing the anticipated sign. In contrast, the parameter estimates of lagged mortgage interest rates are significant but unexpectedly signed. The overall positive response of current house price changes to an increase in interest rates is largely explainable by the fact that in times of high uncertainty and inflation, housing acts as a hedge against future inflation and thus, as the nominal mortgage interest rate rises, the demand for housing increases exerting upward pressure on prices.

Furthermore, Chapter 7's explicit modelling of non-linearity provides additional insights into price behaviour in the Dublin housing market. Firstly, a significant negative relationship between current and past house price changes is evident in the non-linear part of the smooth transition regression model. This reflects a tendency for expectations of further price increases to fall after a certain point, namely once the value of the location parameter has been exceeded. Secondly, house prices are found not to correct in either of the extreme regimes of the STR model. As such, it may be concluded that the long run demand and supply relations identified in Chapter 6 do not define equilibrium for the Dublin housing market.

With respect to its contribution to the existing literature, it should be noted that Chapter 7's estimation of error correction models goes considerably beyond the scope of current empirical work on the Dublin housing market, and in the case of the non-linear model, also beyond that at national level. Indeed, the study's use of the smooth transition regression modelling approach is only the second time in which this technique has been applied in a housing context. Furthermore, by including a range of variables designed to capture the effectiveness of government interventionist policies, this chapter allows for a more complete modelling of house price determination than has been undertaken to date.

Building on the preceding analysis, the findings of the regime-switching test employed in Chapter 8 suggest the presence of a fad in the Dublin housing market as opposed to a stochastic bubble. Fads give rise to a situation whereby prices deviate temporarily from, but will eventually revert to, underlying fundamentals. On this basis, a 'soft landing' may be anticipated for the future path of the Dublin housing market rather than the collapse type scenario associated with a bubble.

The contribution of Chapter 8 to the wider literature centres on the attention paid to modelling the fundamental component of house prices. Whereas existing empirical research has modelled underlying fundamentals in the Dublin marketplace on a partial basis, this study includes both demand and supply-side variables. Most notably, the fundamentals model presented in this chapter is the first to incorporate a measure of land availability.

Overall, this thesis adds to the current body of literature by facilitating an improved understanding of the empirical conditions that characterise the Dublin market for new, private sector housing. A particularly striking result in this respect is the failure on the part of both the price and stock of housing to eliminate deviations and restore the market to equilibrium. A possible explanation for this lack of error correction is that adjustment may be taking place in the *de facto* area of Dublin, rather than within Dublin per se. From a policy perspective, such a possibility has important implications and indeed, merits further research.

In addition to the above, there are a number of other ways in which the work presented here could be extended. Foremost amongst these would be a more detailed examination of the land market and the nature of its relationship with the housing market.

References

Anderson, T.W. (1971) *A Statistical Analysis of Time Series*. New York: John Wiley.

Anderson, T.W. (1984) *An Introduction to Multivariate Statistical Analysis.* New York: John Wiley.

Bacon, D.W. and Watts, D.G. (1971) 'Estimating the Transition between Two Intersecting Straight Lines. *Biometrika*, **58**, pp 525-534.

Bacon, P. and MacCabe, F. (2000) *The Housing Market in Ireland: An Economic Evaluation of Trends and Prospects*. Dublin: Stationery Office.

Bacon, P., MacCabe, F. and Murphy, A. (1998) *An Econometric Assessment of Recent House Price Developments*. Dublin: Stationary Office.

Blanchard, O.J. (1979) 'Speculative Bubbles, Crashes and Rational Expectations'. *Economic Letters*, **3**, pp 387-389.

Blanchard, O.J. and Watson, M. (1982) 'Bubbles, Rational Expectations and Financial Markets' in Wachtel, P. (ed.) *Crises in the Economic and Financial Structure*. MA: Lexington Books.

Cheung, Y.W. and Lai, K. S. (1993) 'Finite - Sample Sizes of Johansen 's Likelihood Ratio Tests for Cointegration'. *Oxford Bulletin of Economics and Statistics*, **55**, pp 313-328.

Cutler, D.M., Poterba, J.M. and Summers, L.H. (1991) 'Speculative Dynamics'. *Review of Economic Studies,* **58**, pp 529-546.

Davidson, R. and MacKinnon, J.G. (1993) *Estimation and Inference in Econometrics.* Oxford: Oxford University Press.

Doornik, J.A. and Hansen, H. (1994) 'A Practical Test for Univariate and Multivariate Normality'. Nuffield College, Discussion Paper.

Doornik, J.A. and Hendry, D.F. (2001) *Empirical Econometric Modelling using PcGive*. London: Timberlake Consultants Ltd.

Drake, L. (1993) 'Modelling UK House Prices using Cointegration: An Application of the Johansen Technique'. *Applied Economics*, **25**, pp 1225-1228.

Dublin Corporation, Dun Laoghaire-Rathdown County Council, Fingal County Council and South Dublin County Council (1999) *Housing in Dublin. A Strategic Review by the Dublin Local Authorities*. Dublin: Dublin Corporation.

Economist, The (2003) 'Close to Bursting: A Survey of Property'.

Eitrheim, Ø. and Teräsvirta , T. (1996) 'Testing the Adequacy of Smooth Transition Autoregressive Models'. *Journal of Econometrics,* **74**, pp 59-75.

Engle, R.F. and Granger, C.W.J. (1987) 'Co-integration and Error Correction: Representation, Estimation and Testing'. *Econometrica,* **55**, pp 251-276.

Escribano, A. and Mira, S. (1995) 'Nonlinear Time Series Models: Consistency and Asymptotic Normality of NLS under New Conditions'. Universidad Carlos III de Madrid, Statistics and Econometric Series 14, Working Paper 95-142.

Evans, G. (1991) 'Pitfalls in Testing for Explosive Bubbles in Asset Prices'. *The American Economic Review*, **4**, pp 922-930.

Fitzgerald, J. (2005) 'The Irish Housing Stock: Growth in Number of Vacant Dwellings'. ESRI Quarterly Economic Commentary Spring, pp 42-63.

Flood, R.P. and Hodrick, R.J. (1986) 'Asset Price Volatility, Bubbles and Process Switching'. *Journal of Finance,* **41**, pp 831-842.

Flood, R.P. and Hodrick, R.J. (1990) 'On Testing for Speculative Bubbles'. *Journal of Economic Perspectives,* **4**, pp 85-101.

Friedman, M. (1957) *A Theory of the Consumption Function.* Princeton, N.J.: Princeton University Press.

Giussani, B. and Hadjimatheou, G. (1990) 'House Prices: An Econometric Model for the U.K.'. The Apex Centre, Economics Discussion Papers 90/1.

Godfrey, L.G. (1988) *Misspecification Tests in Econometrics. The Lagrange Multiplier Principle and other Approaches.* Cambridge: Cambridge University Press.

Gonzolo, J. (1994) 'Five Alternative Methods of Estimating Long – Run Equilibrium Relationships'. *Journal of Econometrics,* **60**, pp 203-233.

Granger, C.W.J. (1969) 'Investigating Causal Relations by Econometric Models and Cross-Spectral Methods'. *Econometrica,* **37**, pp 424-438.

Granger, C.W.J. (1981) 'Some Properties of Time Series Data and their Use in Econometric Model Specification'. *Journal of Econometrics,* **16**, pp 121-130.

Granger, C.W.J. and Weiss, A.A. (1983) 'Time Series Analysis of Error Correction Models' in Karlin, S., Amemiya, T. and Goodman, L.A. (eds.) *Studies in Econometrics, Time Series and Multivariate Statistics.* New York: Academic Press. pp 255-278.

Granger, J.W. and Teräsvirta, T. (1993) *Modelling Nonlinear Economic Relationships.* Oxford: Oxford University Press.

Gujarati, D.N. (1995) *Basic Econometrics.* Singapore: McGraw-Hill Inc.

Hadjimatheou, G. (1976) *Housing and Mortgage Markets.* UK: Saxon House.

Hall, S., Psaradakis, Z. and Sola, M. (1997) 'Switching Error Correction Models of House Prices in the United Kingdom'. *Economic Modelling,* **14**, pp 5517-5527.

Hamilton, J.D. (1990) 'Specification Testing in Markov-Switching Time Series Models'. Department of Economics, University of Virginia, Discussion Paper No 209.

Hamilton, J.D. (1994) *Time Series Analysis.* New Jersey: Princeton University Press.

Hamilton, J.D. (1996) 'Specification Testing in Markov-Switching Time Series Models'. *Journal of Econometrics,* **70**, pp 127-157.

Hansen, B.E. (1996) 'Inference When a Nuisance Parameter is Not Identified Under the Null Hypothesis'. *Econometrica*, **64**, pp 413-430.

Hansen, H. and Johansen, S. (1993) 'Recursive Estimation in Cointegrated VAR-Models'. Institute of Mathematical Statistics, University of Copenhagen, No. 1.

Hansen, H. and Juselius, K. (1995) *CATS in Rats: Cointegration Analysis of Time Series.* Evanston: Estima.

Harvey, A.C. (1990) *The Econometric Analysis of Economic Time Series.* Cambridge, MA: MIT Press.

Hendry, D.F. (1984) 'Econometric Modelling of House Prices in the United Kingdom' in Hendry, D.F. and Wallis, K.F. (eds.) *Econometrics and Quantitative Economics.* Oxford: Blackwell, pp 211-252.

Hendry, D.F. (1995) *Dynamic Econometrics.* Oxford: Oxford University Press.

Hendry, D.F. and Juselius, K. (2000) 'Explaining Cointegration Analysis: Part II'. Department of Economics, University of Copenhagen, Discussion Paper 00-02.

International Monetary Fund (2003) *IMF Article IV Staff Report on Ireland.* IMF Country Report No. 03/242.

Johansen, S. (1988) 'Statistical Analysis of Cointegration Vectors'. *Journal of Economic Dynamics and Control*, **12**, pp 231-254.

Johansen, S. (1991) 'Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models'. *Econometrica,* **59**, pp 1551-1580.

Johansen, S. (1994) 'The Role of the Constant and Linear Terms in Cointegration Analysis of Non-Stationary Variables'. *Econometric Reviews,* **13**, pp 205-229.

Johansen, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford: Oxford University Press.

Johansen, S. (2002) 'A Small Sample Correction for the Test of Cointegrating Rank in the Vector Autoregressive Model'. *Econometrica*, **70**, pp 1929-1961.

Johansen, S. and Juselius, K. (1990) 'Maximum Likelihood Estimation and Inference on Cointegration – with Applications to the Demand for Money'. *Oxford Bulletin of Economics and Statistics*. **52**, pp 169-210.

Johansen, S. and Juselius, K. (1994) 'Identification of the Long Run and Short Run Structure: An Application to the ISLM Model'. *Journal of Econometrics*, **63**, pp 7-36.

Johansen, S. and Nielsen, B. (1993) 'Asymptotics for Cointegration Rank Tests in the presence of Intervention Dummies – Manual for Simulation Programme DISCO'. Institute of Mathematical Statistics, University of Copenhagen.

Juselius, K. (forthcoming) *The Cointegrated VAR Model: Econometric Methodology and Macroeconomic Applications.*

Kenny, G. (1998) 'The Housing Market and the Macroeconomy: Evidence from Ireland'. Central Bank of Ireland, Technical Paper 1/RT/98.

Kenny, G. (1999) 'Asymmetric Adjustment Costs and the Dynamics of Housing Supply'. Central Bank of Ireland, Technical Paper 3/RT/99.

Kindleberger, C.P. (1989) *Manics, Panics and Crashes: A History of Financial Crises*. New York: Basic Books.

Lee, L.F. and Porter, R.H. (1984) 'Switching Regression Models with Imperfect Sample Separation Information – With an Application of Cartel Stability'. *Econometrics*, **52**, pp 391-418.

Luukkonen, R., Saikkonen, P. and Teräsvirta , T. (1988) 'Testing Linearity Against Smooth Transition Autoregression'. *Biometrika*, **75**, pp 491-499.

McQuinn, K. (2004) 'A Model of the Irish Housing Sector'. Central Bank of Ireland, Research Technical Paper 1/RT/04.

Meen, G.P. (1993) 'The Treatment of House Prices in Macroeconometric Models: A Comparison Exercise'. Department of the Environment Great Britain, Discussion Paper.

Miles, D. (1994) *Housing, Financial Markets and the Wider Economy*. United States: John Wiley and Sons.

Modigliani, F. and Ando, A. (1963) 'The Life Cycle Hypothesis of Saving: Aggregated Implications and Tests'. *American Economic Review*, **53**, pp 55-84.

Murphy, A. and Brereton, F. (2001) 'Modelling Irish House Prices: A Review'. Paper presented at the Irish Economic Association Annual Conference.

Muth, R.F. (1960) 'The Demand for Non-Farm Housing' in Harberger, A. (ed.) *The Demand for Durable Goods*. Chicago: University of Chicago Press, pp 29-96.

Muth, R.F. and Goodman, A.C. (1989) *The Economics of Housing Markets*. Great Britain: Harwood Academic Publishers.

Olsen, E.O. (1969) 'A Competitive Theory of the Housing Market'. *The American Economic Review*, **59**, pp 144-171.

Poterba, J.M. (1984) 'Tax Subsidies to Owner – Occupied Housing: An Asset Market Approach'. *Quarterly Journal of Econometrics*, **99**, pp 729-752.

Quandt, R.E. (1958) 'The Estimation of Parameters of a Linear Regression System Obeying Two Separate Regimes'. *Journal of the American Statistical Association*, **53**, pp 873-880.

Rao, C.R. (1973) *Linear Statistical Inference and its Applications*. New York: John Wiley and Sons.

Roche, M. (1999) 'Irish House Prices; Will the Roof Fall In?'. *The Economic and Social Review*, **30**, pp 343-362.

Roche, M. (2001) 'The Rise in House Prices in Dublin: Bubble, Fad or just Fundamentals'. *Economic Modelling*, **18**, pp 281-295.

Roche, M. (2003) 'Will there be a Crash in Irish House Prices?'. ESRI Quarterly Economic Commentary Winter, pp 57-72.

Rosen, S. (1974) 'Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition'. *Journal of Political Economy*, **82**, pp 34-55.

Saikkonen, P. and Luukkonen, R. (1988) 'Lagrange Multiplier Tests for Testing Non-Linearities in Time Series Models'. *Scandinavian Journal of Statistics*, **15**, pp 55-68.

Salo, S. (1994) 'Modelling the Finnish Housing Market'. *Economic Modelling*, **11**, pp 250-265.

Schaller, H. and van Norden, S. (1997) 'Fads or Bubbles?'. Bank of Canada, Working Paper No 97-2.

Shenton, L.R. and Bowman, K.O. (1977) 'A Bivariate Model for the Distribution of $\sqrt{b_1}$ and b_2'. *Journal of the American Statistical Association,* **72**, pp 206-211.

Smith, L.B., Rosen, K.T. and Fallis, G. (1988) 'Recent Developments in Economic Models of Housing Markets'. *Journal of Economic Literature*, **XXVI**, pp 29-64.

Stock, J.H. (1987) 'Asymptotic Properties of Least Squares Estimates of Cointegration Vectors'. *Econometrica,* **55**, pp 1035-1056.

Summers, L.H. (1986) 'Does the Stock Market Rationally Reflect Fundamental Values?'. *The Journal of Finance,* **61**, pp 591-602.

Teräsvirta , T. (1994) 'Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models'. *Journal of the American Statistical Association*, **89**, pp 208-218.

Teräsvirta , T. (1998) 'Modeling Economic Relationships with Smooth Transition Regressions' in Ullah, A. and Giles, D.A. (eds.) *Handbook of Applied Economic Statistics.* New York: Marcel Dekker Inc., pp 507-552.

Teräsvirta , T. (2004) 'Smooth Transition Regression Modeling' in Lutkepohl, H. and Kratzig, M. (eds.) *Applied Time Series Econometrics.* Cambridge: Cambridge University Press, pp 222-242.

Thom, R.D. (1983) 'House Prices, Inflation and the Mortgage Market'. *The Economic and Social Review*, **15**, pp 57-68.

Tiebout, C.M. (1956) 'A Pure Theory of Local Expenditures'. *Journal of Political Economy,* **64**, pp 416-424.

van Norden, S. (1993) 'Regime Switching and Exchange Rate Bubbles' in Murray, J. and O'Reilly, B. (eds.) *The Exchange Rate and the Economy*. Ottawa: Bank of Canada.

van Norden, S. (1996) 'Regime Switching as a Test for Exchange Rate Bubbles'. *Journal of Applied Econometrics*, **11**, pp 219-251.

van Norden, S. and Schaller, H. (1993a) 'Speculative Behaviour, Regime-Switching and Stock Market Fundamentals'. Bank of Canada, Working Paper No 93-2.

van Norden, S. and Schaller, H. (1993b) 'The Predictability of Stock Market Regime: Evidence from the Toronto Stock Exchange'. *The Review of Economic and Statistics*, **75**, pp 505-510.

van Norden, S. and Vigfusson, R. (1996a) 'Avoiding the Pitfalls: Can Regime-Switching Tests Detect Bubbles'. Bank of Canada, mimeo.

van Norden, S. and Vigfusson, R. (1996b) 'Regime-Switching Models. A Guide to the Bank of Canada Gauss Procedures'. Bank of Canada, Working Paper No 96-3.

Verbeek, M. (2000) *A Guide to Modern Econometrics*. England: John Wiley.

White, H. (1987) 'Specification Testing in Dynamic Models' in Bewley, T.F. (ed.) *Advances in Econometrics, Fifth World Congress*. Cambridge: Cambridge University Press.

Wooldridge, J.M. (1994) 'Estimation and Inference for Dependent Processes' in Engle, R.F. and McFadden, D.L. (eds.) *Handbook of Econometrics*. Amsterdam: North Holland, pp 2639-2738.