Homelessness Task Force Research Series

Structural Trends and Homelessness: a Quantitative Analysis

SCOTTISH EXECUTIVE

Central Research Unit

STRUCTURAL TRENDS AND HOMELESSNESS: A QUANTITATIVE ANALYSIS

Peter A Kemp, Emily Lynch and Daniel Mackay
University of Glasgow

Scottish Executive Central Research Unit
2001

Further copies of this report are available priced £5.00. Cheques should be made payable to The Stationery Office Ltd and addressed to:

The Stationery Office Bookshop
71 Lothian Road
Edinburgh
EH3 9AZ

Tel: 0131-228-4181
Fax: 0131-622 7017

STRUCTURAL TRENDS AND HOMELESSNESS: A QUANTITATIVE ANALYSIS

Peter A. Kemp, Emily Lynch and Daniel Mackay

University of Glasgow

Scottish Executive Central Research Unit
2001

ACKNOWLEDGEMENTS

The authors would like to express thanks to all of the people who have helped us with this research project. In particular, we would like to thank the members of the project's advisory group, the various people in the Scottish Executive and elsewhere who kindly provided us with data, and Richard Evans, our project manager in the Executive. Thanks also to Julie Morgan, who converted the report into CRU house style.

CONTENTS

SUMMARY OF FINDINGS

INTRODUCTION

It is a widely held view within the research community that the growth of homelessness in Scotland during the 1980s and 1990s was caused by a series of structural trends, such as unemployment, rising rents and the decline in the supply of social rented housing. However, to date the existence of a relationship between homelessness and these structural trends has been frequently asserted but not tested in Scotland.

The aim of the research reported here was to examine whether structural factors can explain statistically the level of homelessness in Scotland.

For the purpose of this research, 'homelessness' was defined as the number of people applying to local authorities for assistance on the grounds of homelessness (homelessness *applications*) or the number who have been assessed as homeless by local authorities (homelessness *acceptances*). The analysis looked at applications and acceptances by all households and at applications by lone parents and by young people under 25.

METHODS

The research team reviewed the literature on the causes of homelessness. Drawing on this literature, an empirically testable model of the structural causes of homelessness was developed. Multivariate statistical techniques were then used to test this model. The statistical analysis had 2 main components. The first comprised an analysis of trends in homelessness in Scotland over a 19-year period, from 1980 to 1998 (the time-series analysis). Second, the analysis examined variations in homelessness between Scottish local authorities at 6 different points in time over the period from 1981 to 1999 (the cross-sectional analysis) and for groups of years taken together (the pooled cross-sectional analysis).

THE CAUSES OF HOMELESSNESS

Explanations of the causes of homelessness tend to fall into 2 main schools of thought. On the one hand, there are those who argue that homelessness is caused by 'structural' factors such as poverty, unemployment, and a lack of affordable housing. On the other hand, some authors argue that homelessness is caused, not by structural factors, but rather by the 'behavioural' characteristics of the individuals concerned; in other words, people become homeless because of their personal failings, bad luck or inability to cope with adverse events.

However, it is important to distinguish between the presenting or *proximate* causes of homelessness and the *underlying* structural causes. The latter provide the context within which particular individuals are unable to cope with an adverse event in their lives. Thus, personal factors and individual behaviour may determine who becomes homeless under unfavourable structural conditions. While behavioural factors may influence why any one individual becomes homeless, structural factors determine the aggregate level of homelessness. In other words, homelessness is likely to result from the complex inter-play between structural and behavioural factors.

The consensus within the literature is that the causes and the growth of homelessness reflect structural factors, yet very little research has been done to support or refute this claim. At present, this consensus amounts to little more than indirect inferences from trends in potential causal factors. Prior to this study, there had been no systematic evaluation in Scotland of these structural explanations. To some extent, this gap reflects the fact that most of the research on homelessness in Scotland has been qualitative rather than quantitative.

A STRUCTURAL MODEL OF HOMELESSNESS

The growth in homelessness over the past 2 decades has been variously ascribed to structural factors such as rising unemployment, social security cutbacks, labour market restructuring, the decline of private renting, the sale of council houses, rising house prices, relationship breakdown. In general, there is little attempt in this literature to move beyond lists and explain how or why particular factors may have affected the overall *level* of homelessness.

A review of the literature suggests that the various structural factors thought to affect the level of homeless can be grouped under 4 main headings:

- Housing demand and supply
- Affordability
- Unemployment
- De-institutionalisation

These 4 groups of factors may cause homelessness by affecting:

- the *availability* of sufficient and suitable accommodation
- the *affordability* of available accommodation, and
- the need for *social support* to enable potentially homeless people to live in the available and affordable accommodation.

HOMELESSNESS OVER TIME

The analysis of trends in homelessness applications over the last 2 decades (1980 to 1998) found support for a structural explanation of the level of homelessness. There was a long-run statistical relationship between homelessness and the housing market (right to buy sales, the number of public sector lettings). There was also a long-run statistical relationship between homelessness and affordability (the number of tenants in rent arrears). Likewise, there was a long-run statistical relationship between homelessness and the unemployment rate and the level of employment in manufacturing. Finally, there was a long-run statistical relationship between homelessness and a proxy measure ‘deinstitutionalisation’ (the number of recorded crimes).

VARIATIONS IN HOMELESSNESS

The analysis of variations in homelessness between Scottish local authorities also provided some support for the view that the level of homelessness is affected by structural factors.

Housing demand and supply

The analysis provided support for the hypothesis that the level of homelessness is affected by the state of the housing market, though the nature of the relationship was not always consistent with that implied by the structural model. In 1981 homelessness was negatively related to the number of vacant properties available for letting. This implies that homelessness was higher where local authority vacancies were lower. In other words, 'tight' housing markets had more homelessness than 'slack' ones.

However, by the 1990s, the relationship between local authority vacancies and homelessness was found to be the reverse of that posited in the model and found for 1981. Homelessness was now positively related to vacancies. In other words (other things being equal) the more empty dwellings, the higher the level of homelessness. It is possible that this change in relationship reflected some underlying change in the nature of social housing that is not captured in the housing demand and supply variables available to the research team. Vacancies may be an 'intervening variable' that is related to a factor not included in the analysis. That is to say, the relationship may reflect a structural trend, the effects of which were concentrated in local authorities with high vacancies.

In 1999 there was a positive relationship between the proportion of the population aged under 25 and the level of homelessness. Thus, local authorities with high proportions of *young people under 25* also had high levels of homeless.

Unemployment

Variations in homelessness between Scottish local authorities were positively related to unemployment, which is in line with the structural model. This relationship applied to various different measures of unemployment, such as the *claimant count*, the *ILO* (International Labour Organisation) definition of unemployment - which is much broader than the claimant count - and *long-term unemployment*.

Local authorities with relatively high levels of unemployment also tended to have high levels of homelessness and vice versa. In other words (other things being equal) as unemployment increases, so too does homelessness.

Affordability

The analysis also found support for the hypothesis that the level of homelessness is affected by the affordability of housing. In general, homelessness applications and acceptances were positively related to the level of *house prices*. In other words (other things being equal) local authorities with high house prices also tended to have high levels of homelessness and vice versa. Homelessness was also positively related to *mortgage arrears*. Areas with high levels of mortgage arrears also tended to have high levels of homelessness. Likewise, there was a positive relationship between *private rents* on unfurnished lettings and homelessness. Other things being equal, the higher the rent, the higher the level of homelessness.

Except for 1981, there was an inverse relationship between *local authority rents* and the level of homelessness. That is to say, other things being equal, the lower the council rent, the higher the level of homelessness. As with local authority vacancies, it is not clear why this inverse relationship between local authority rent levels and homelessness should exist.

De-institutionalisation

The analysis revealed a number of statistically significant relationships between homelessness and various measures of what, for the purpose of this research, was termed 'de-institutionalisation', though the results were not fully consistent.

In 1981 and 1991, there was a positive relationship between homelessness and the number of *psychiatric inpatient discharges* of people with a diagnosis of alcohol misuse. In other words, the greater the number of such discharges, the higher the level of homelessness in the local authority and vice versa. By 1996 and 1997 the relationship had reversed. For these 2 years, the relationship was an inverse one in that homelessness was now lower in areas with high levels of discharge (and vice versa). It is possible that local authorities had by this time adjusted their homelessness practices to provide accommodation for inpatients with a history of alcohol misuse who were being discharged from psychiatric wards and hospitals. However, in 1998 there was a positive relationship between the 2 (except for applications by young people, which continued to have an inverse relationship).

In 1999, there was a positive relationship between the number of *inpatient discharges from long-stay hospitals* and the level of homelessness. There was also a positive statistical relationship between the level of homelessness and the number of *children in care*.

CONCLUSIONS

Taking the results of the various analyses together and bearing in mind the limitations of the data, the results lend support for the hypothesis that there are structural causes of homelessness. At the very least, it can be concluded that the results do not provide sufficient evidence to reject the model. In other words, the level of homelessness over the past 2 decades has to some extent been affected by wider structural trends in society. Unemployment, and to a lesser extent housing affordability and de-institutionalisation, appear to be powerful forces affecting the incidence of homelessness in Scotland.

Thus, while behavioural factors may be important in explaining *individual cases* of homelessness, the analysis indicates that structural trends do affect the *aggregate level* of homelessness. The evidence suggests that this is the case, not only for homelessness acceptances; but also for applications by all households, by young people, and by lone parents.

CHAPTER ONE INTRODUCTION

1.1 The causes of homelessness have been debated for many years but never satisfactorily resolved. Much of the early literature on, and policy responses to, homelessness either implicitly or explicitly assumed that it was caused by the personal failings of the individuals concerned. The appropriate response was therefore what may be described as a social work approach in which the homeless were given temporary accommodation while their personal problems or difficulties, such as relationship breakdown or drug misuse, were addressed. This was largely how the problem was tackled in Scotland and elsewhere in Britain until the mid-1970s (Watchman and Robson, 1983).

1.2 During the mid-1970s, however, it became increasingly accepted that homelessness was essentially a housing problem rather than a social work problem. People were homeless because they lacked accommodation rather than because they were in need of social work assistance. This change of perspective was, of course, reflected in the passage of the Housing (Homeless Persons) Act 1977, which placed a statutory duty on local authorities to provide accommodation for people in 'priority need' who were unintentionally homeless.

1.3 This shift in perspective on the nature of homelessness, from seeing it as a social work concern to one in which it was largely viewed as a housing problem, reflected an important change in the way in which the 'blame' for homelessness was apportioned. If homelessness was a housing problem, rather than a social work problem, this seemed to imply that it was caused, not by the personal failings of the individuals concerned, but rather by the lack of sufficient affordable housing. Expressed in terms of the academic debates on the subject, homelessness was not due to individual *agency* (personal misfortune or failings) but rather to *structural* reasons beyond the individual and over which they had no control (Johnson *et al*, 1991; Neale, 1997).

1.4 Thus, a new consensus emerged among researchers in Britain that homelessness was a housing problem reflecting wider structural factors and not a social work problem reflecting individual inadequacies or personal misfortune (Johnson *et al*, 1991; Third and Yanetta, 2000). The policy implication of this perspective was clear: the way to tackle homelessness was to provide more affordable housing.

1.5 However, subsequent experience in homelessness practice has shown that simply providing housing is not always sufficient to prevent or solve individual cases of homelessness (Dane, 1998; Dant and Deacon, 1989; Vincent, *et al*, 1995). Many homeless people experience an array of social problems - including poor health, alcohol and drug misuse – which make it difficult for them to sustain a tenancy even when it is provided for them with full Housing Benefit. Thus, it has more recently become clear that homelessness is part of a complex of problems which go beyond simply lack of housing (Fitzpatrick, Kemp and Klinker, 2000).

1.6 This new perspective quickly leads to appeals for more 'joined up' policy and practice and to see the problem within a wider social exclusion (Pleace, 1998) or social justice agenda. The problem is not just lack of suitable and affordable housing, but also problems reflecting the wider disadvantages faced by socially excluded groups. Many of these disadvantages are themselves a product of wider structural factors such as low educational attainment, divorce, unemployment, and poverty. The growth in homelessness in the 1980s and 1990s is believed by many people to reflect changes in the incidence of these wider structural forces (e.g., Third and Yanetta, 2000).

1.7 However, there has been little systematic attempt in Britain to examine the validity of these structural explanations. At present they remain largely untested assumptions or assertions.

AIMS OF THE RESEARCH

1.8 The background to the research presented in this report was the appointment of a Homelessness Task Force by the Scottish Executive in 1999. The terms of reference of the Task Force were "to review the causes of homelessness in Scotland; to examine current practice in dealing with cases of homelessness; and to make recommendations on how homelessness in Scotland can best be prevented and, where it does occur, tackled effectively". The research reported here was one of a number of studies commissioned by the Task Force to help inform its deliberations.

1.9 The objective of the research project was to test the hypothesis that the level of homelessness is affected by macro level factors such as demographic, economic and social trends. It aimed to explore whether these structural factors differed for different groups of the population or different types of area in Scotland. Finally, the research aimed to draw on past change to predict the impact of future changes and to consider whether and where policy intervention is possible or appropriate to reduce homelessness.

1.10 For reasons discussed in the report, the analysis of different groups of the population was confined to young people and lone parents, in addition to all homeless people. Limitations of data and of time meant that it was not possible to examine in depth differences between different parts of Scotland.

RESEARCH METHODS

1.11 There were 2 stages to the research project: (1) hypothesis development and (2) statistical modeling.

1.12 The first stage of the work involved a review of the existing research literature on homelessness in order to develop informed hypotheses about the structural factors that are likely to affect the level of homelessness. This aspect of the work drew on the recent literature review on single homelessness conducted by the Department of Urban Studies at the University of Glasgow (Fitzpatrick, Kemp and Klinker, 2000) and on a summary of research evidence on homelessness in Scotland (Third and Yanetta, 2000). It also drew on the quantitative analytical work on the structural causes of homelessness, much of which has been undertaken in the USA. In order to place these hypotheses within an explanatory framework, we followed Burt (1992) in devising a 'structural model' of the causes of homelessness. This model is outlined in Chapter Three.

1.13 The second stage of the project involved statistical modeling of the structural factors that may affect the level of homelessness. In order to identify which, if any, of these factors affect the level of homelessness we used a statistical technique known as multiple regression analysis. This technique makes it possible to measure the influence of a number of explanatory or hypothesised variables on a dependent variable. The data and statistical methods used in the second stage of the research are outlined at length in Chapter Four and in more detail in Appendices 1 and 2.

1.14 For the purpose of this research, 'homelessness' was measured using data from the statistical forms that local authorities return to the Scottish Executive. These returns document the number and types of household applying for assistance under the homelessness provisions of the housing legislation. While these data have a number of well-known limitations, they provide the best available measure of homelessness and, in particular, of *changes* over time in expressed homelessness in Scotland.

STRUCTURE OF THE REPORT

1.15 The structure of the report is as follows:

Chapter Two Trends in Homelessness presents a brief statistical overview of homelessness in Scotland. It focuses on the trends in homelessness over the past 2 decades. The purpose of this overview is to describe the phenomenon, the causes of which this research sought to explain.

Chapter Three A Structural Model of Homelessness examines the factors that are believed to cause homelessness. It reviews previous research on the causes of homelessness and outlines the structural model tested in the research.

Chapter Four Data and Methods describes the data and methods employed in the study.

Chapter Five Testing the Structural Model presents the main results of the extensive statistical modeling that was undertaken to test the structural model of homelessness.

Chapter Six Predicting Homelessness presents some very tentative forward estimates of the future level of homelessness.

Chapter Seven Conclusions sets out the main conclusions of the research and the implications for policy.

Annexes 1 to 4 provide further details of the variables used in the analysis, the statistical methods employed, and the results of the analysis.

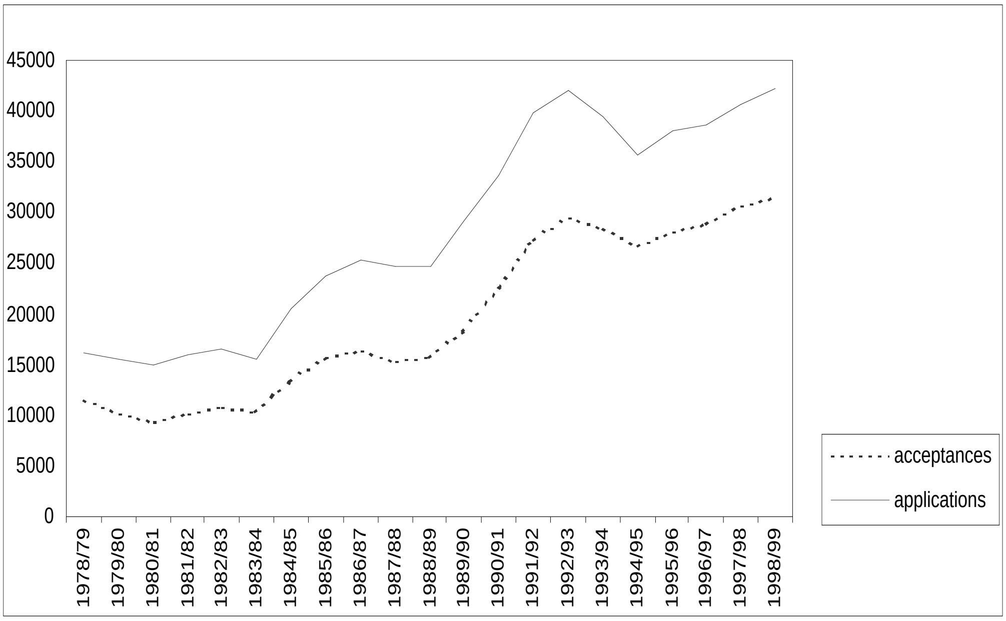
CHAPTER TWO TRENDS IN HOMELESSNESS IN SCOTLAND

2.1 This chapter presents a brief statistical overview of the level of 'statutory homelessness' in Scotland over the past 20 years. By statutory homelessness we mean applications to local authorities for assistance under the homelessness provisions of the Housing Act (see Third and Yanetta, 2000). More detailed analysis of the homelessness statistics for Scotland has been conducted by Pawson (2000) as part of the Homeless Task Force research programme. Meanwhile, Anderson and Tulloch (2000) have reviewed the evidence on people's pathways through homelessness.

HOMELESSNESS IN SCOTLAND

2.2 It is well documented that the pattern of statutory homelessness in Scotland at the national level has been growing over time (see Pawson, 2000). Figure 2.1 illustrates this very well where it can be seen that in terms of HL1 returns from local authorities there has been a marked upward trend in the number of homelessness applications and acceptances in Scotland. However, it is worth pointing out that, while the trend is noticeably upward, the number of applications and acceptances were in fact fairly stable over the late 1970s and early 1980s. Both measures of homelessness actually *fell* between 1992/93 and 1994/95, before rising steadily again up until 1998/99. Nonetheless, from 1978 to 1998 the numbers of homeless applications rose by approximately 163 per cent or 26,110 and the number of acceptances by 175 per cent or 20,060.

Figure 2.1 Homelessness applications and acceptances for Scotland



2.3 Dis-aggregating the HL1 data into rural and urban local authorities (as defined by the Scottish Executive), there are a number of differences between the homeless applications figures. First, the urban series exhibits a number of peaks and troughs or cycles (approximately 5-6 years) throughout its course, in contrast to the series for rural homelessness, which is much smoother. This implies that there are periods of reductions in the number of homeless people for the cities; and, as can be seen from the graph, this is indeed the case, the most marked reduction occurring around 1992.

Figure 2.2 Urban and rural homelessness applications

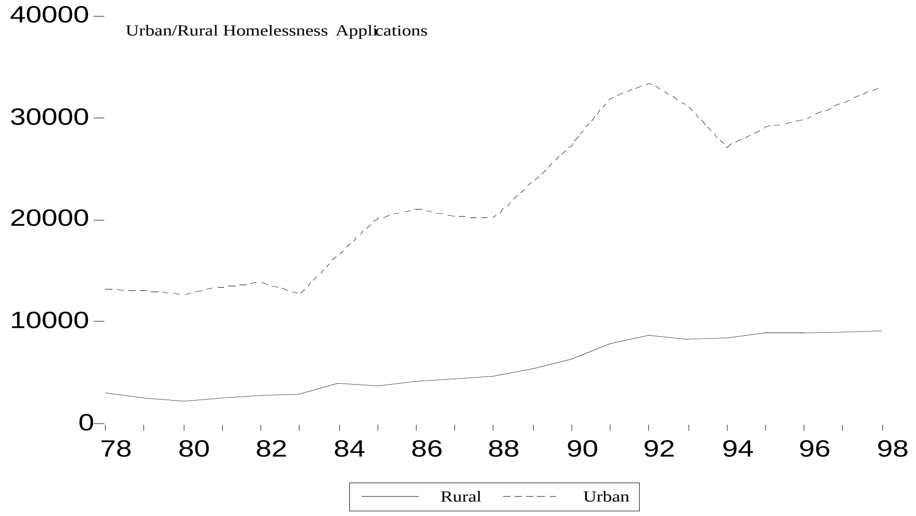
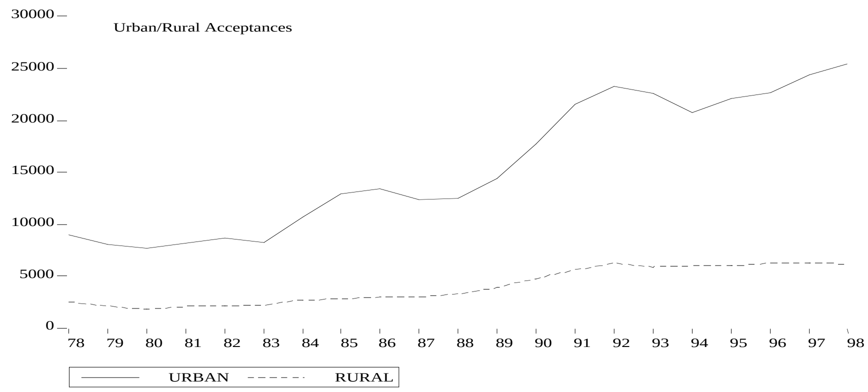


Figure 2.3 Urban and rural homelessness acceptances



2.4 Both series exhibit an upward trend, with the rate of change for the urban series being much greater than that for the rural. However, this hides an important quantitative difference. Over the 20 year period for which we have data, urban homelessness *applications* have increased by 152% (or 19,967) and rural homelessness applications by 208% (or 6,144). Thus, relatively speaking, rural homelessness has increased much more than urban homelessness in the last 20 years but it is far less volatile than urban homelessness applications (Figure 2.2). For homelessness *acceptances*, the increase for rural areas has been 144% compared to 184% for urban areas.

HOMELESSNESS IN THE CITIES

2.5 Figures 2.4 through 2.7 present homelessness applications and acceptances figures for Scotland's 4 main cities. There are some striking differences between these graphs. Looking initially at the City of Aberdeen one can see that there was a marked increase in the number of applications and acceptances from 1984/85 to 1988/89, the numbers trebling in magnitude. For the other cities this effect is much more muted and, indeed, Glasgow's applications/acceptances were more or less constant over this 4 year period, not really accelerating until 1989/90. The figures for Dundee and Edinburgh, however, exhibit a more modest increase. Aberdeen would also appear to be more volatile than the other cities with quite obvious peaks and troughs in the numbers from the early 1980s to the late 1990s. Edinburgh and Glasgow's volatility, although obvious, is much less pronounced than Aberdeen's, with the exception of 1994 and 1995 where the numbers fell quite significantly. Again, applications and acceptances are a mirror image of each other and, with the exception of Aberdeen, the trend is one of increasing homelessness.

Figure 2.4 Aberdeen

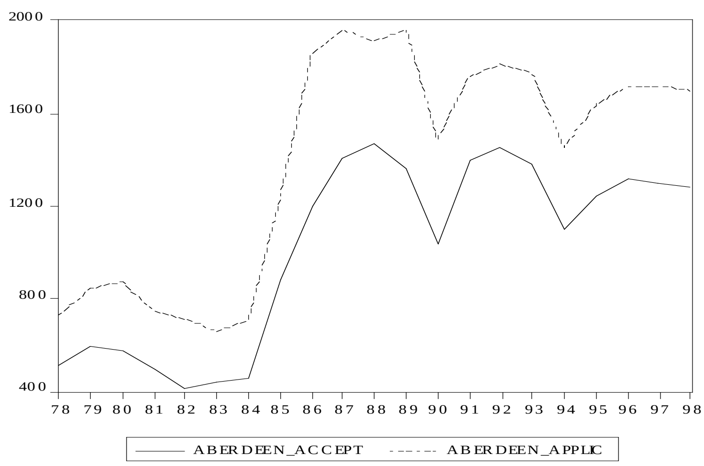


Figure 2.5 Dundee

Figure 2.6 Edinburgh

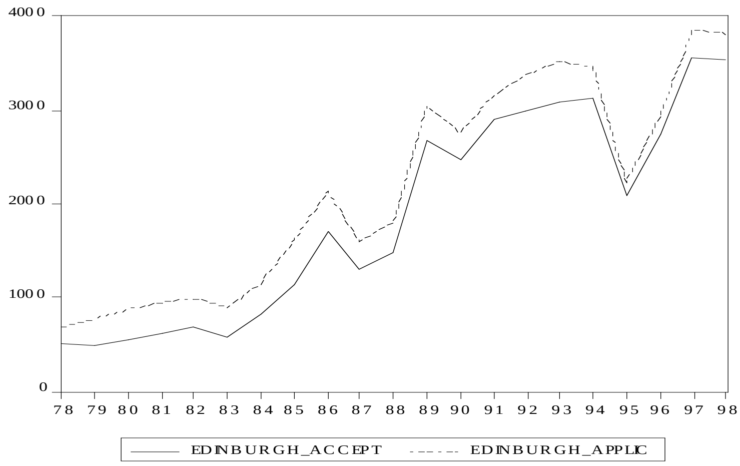
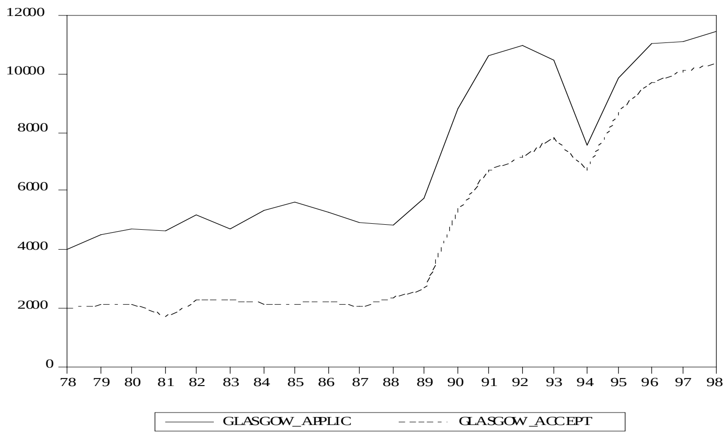


Figure 2.7 Glasgow



SUMMARY

2.6 Thus, while there have been differences between the 4 main cities and between urban and rural local authorities, there has in all areas been a substantial increase over the past 2 decades in the level of homelessness in Scotland. It is not surprising, therefore, that there is considerable interest in what may have caused this increase in homelessness and in why there are variations in homelessness between different local authorities.

CHAPTER THREE A STRUCTURAL MODEL OF HOMELESSNESS

3.1 This chapter outlines a structural model of homelessness. It begins by briefly reviewing the academic debate in Britain over structural and behavioural explanations of homelessness. It then moves on to discuss previous statistical analysis of the structural causes of homelessness, most of which has been conducted in the USA. Finally, drawing on the existing literature about the causes of the growth of homelessness in Scotland and Britain over the past 2 decades, it presents a simplified and testable structural model of homelessness.

STRUCTURAL VERSUS BEHAVIOURAL CAUSES OF HOMELESSNESS

3.2 Explanations of the causes of homelessness tend to fall into 2 main schools of thought (Johnson, *et al*, 1991). On the one hand, there are those who argue that homelessness is caused by 'structural' factors such as poverty, unemployment, and a lack of affordable housing. On the other hand, some authors argue that homelessness is caused, not by structural factors, but rather by the behavioural characteristics of the individuals concerned. They have become homeless because of their personal failings, bad luck or inability to cope with adverse events.

3.3 In the British research community, the weight of opinion has generally been in the structuralist camp (see, for example, Greve 1990). However, while many commentators have argued that homelessness and its growth in recent decades is a function of structural factors, other authors have argued that individual or behavioural factors are more important. For example, in her study of older homeless people, Crane (1999) argued that poverty, unemployment and the availability of affordable housing had little influence on the onset of homelessness in her sample of respondents. She found that personal inadequacies including factors such as coping abilities, mental illness and emotional distress were more important as causes of homelessness.

3.4 However, it is important to distinguish between the presenting or *proximate* uses of homelessness and the *underlying* structural causes. The latter provide the context within which particular individuals are unable to cope with an adverse event in their lives (Wolch, Dear and Akita, 1988). Thus, personal factors and individual behaviour may determine who becomes homeless under unfavourable structural conditions (Elliott and Krivo, 1991). While behavioural factors may influence why any one individual becomes homeless, structural factors determine the aggregate level of homelessness.

3.5 In other words, homelessness is likely to result from the complex inter-play between structural and behavioural factors (Neale, 1997; Fitzpatrick, Kemp and Klinker, 2000). It should be stressed, however, that drawing a line between structural and behavioural factors is not straightforward. For example, it could be argued that family breakdown and drug abuse are themselves related, to some extent, to structural factors such as unemployment and poverty. In practice, therefore, it is an over-simplification to distinguish very sharply between structural and behavioural causes of homeless (Neale, 1997), though analytically the distinction is important.

PREVIOUS RESEARCH ON STRUCTURAL CAUSES

3.6 While it is commonly held in the British literature that the causes and the growth of homelessness reflect structural factors, very little research has been done to support or refute this claim. At present, this consensus amounts to little more than indirect inferences from trends in potential causal factors. With a few exceptions (e.g. Bramley, 1993) there has been no systematic evaluation in Britain of these structural explanations. To some extent this failing reflects the fact that most research in Britain on homelessness has been qualitative rather than quantitative (Fitzpatrick, Kemp and Klinker, 2000).

3.7 In the USA, however, there has been a growing body of quantitative research on the causes of homelessness and, in particular, whether it is a structural problem or not (e.g., Burt, 1991, 1992; Elliott and Krivo, 1991; Quigley *et al*, 2000). Some contributors to this U.S. literature on the 'structural theory of homelessness' (Main, 1996) focus upon whether homelessness is due to broad societal factors or to changes in the housing market and in income distribution. Those who take the latter line argue that the growth in homelessness in the USA reflects a shortage of housing or an inability to pay for it. In contrast, those who argue that it is due to wider societal factors focus on developments such as the de-institutionalisation of the mentally ill and rising drug abuse. However, the 2 are not necessarily incompatible, because the ability of mentally ill people to maintain a tenancy (assuming it is available and affordable) partly turns on whether sufficient social support and care is made available to them. We return to this point in the next section.

3.8 Research in the USA using statistical methods to examine the causes of homelessness has confirmed that housing market factors are related to variations in homeless between different cities or metropolitan areas, but the findings are mixed in relation to de-institutionalisation, poverty and unemployment.

3.9 For example, Elliott and Krivo (1991) examined the effect of 4 commonly posited structural causes of homelessness: lack of low cost housing, high poverty rates, poor economic conditions, and lack of community health care facilities. They also included demographic factors in their statistical model. They found that unemployment and the poverty rate were not significantly associated with the homelessness rate. However, low rent housing, low skill jobs and residential mental health care expenditures were statistically associated with homelessness. These 3 variables accounted for 38 per cent of the variance in homelessness rates across US metropolitan cities.

3.10 Green and Malpezzi (n.d.) analysed homelessness rates across major US metropolitan areas. They found, as expected, that rent levels were positively related to homelessness, that is, that the higher the first quartile rent, the higher the homelessness rate. Although they had expected to find that higher vacancy rates were associated with lower homelessness rates, in fact this variable was insignificant in their model.

3.11 Quigley and Raphael (2000) also analysed variations in homelessness rates across US urban areas. They found that de-institutionalisation could account for at most one-fifth to one-third of the increase in homelessness in the 1980s. Like Elliott and Krivo (1991) they found little evidence that homelessness varied with the local unemployment rate or with the number of recipients of disability pensions. Unlike Green and Malpezzi (n.d.), however, they found that the vacancy rate was negatively related to homelessness. They found that median rents and rent-to-income ratios were positively related to homelessness. Finally, they also found that homelessness was lower in areas with colder weather.

3.12 Martha Burt's (1991, 1992) study is one of the most comprehensive attempts to examine the causes of homelessness. She used data on shelter bed spaces for 147 US cities and 35 suburbs in 1981 and 1989 to examine the causes of the growth on homelessness in the 1980s. The most important variables in her model proved to be: the unemployment rate and the employment structure of cities, population change, the proportion of one-person households, the availability of General Assistance (an income support benefit), the cost of living, and the failure of benefit levels to maintain their purchasing power. Her model explained 50 per cent of the variation in homelessness rates in 1989 and 52 per cent of the change in rates between 1981 and 1989.

3.13 Of course, these results may not carry over to the British and still less the Scottish context. Closer to home, Bramley (1993) analysed statutory homelessness acceptances by local authorities in England. His model explained 60 per cent of the variation in homelessness rates. However, because of problems of 'multi-colinearity' (i.e., statistical correlation between the explanatory variables) some of the results in the model were inconsistent. He found that the most important explanatory factors were socio-demographic variables (such as the number of lone parents and 'New Commonwealth' households), low income, social housing supply (e.g., relets), the size of the private rented sector, rural/urban differences, and the homelessness policy stance of the local authority. Using data for 1981/2, 1986/7 and 1990/91, Bramley found that economic factors became more important over this 10-year period, while social housing supply, private renting and rural/urban differences became less important.

STRUCTURAL CAUSES OF HOMELESSNESS

3.14 In the British literature, the growth in homelessness over the past 2 decades is variously ascribed to structural factors such as rising unemployment, social security benefit cutbacks, labour market restructuring, the decline of private renting, the sale of council houses, rising house prices, increasing inequality, relationship breakdown, and so on (see Fitzpatrick, Kemp and Klinker, 2000). For example, Greve (1990: 16) has argued that 'Structural factors involving the housing market, employment opportunities, wage levels, demographic change, and policies relating to social security and housing benefit…contribute directly and indirectly to the homelessness of single people as well as families".

3.15 In general, there is little attempt in this literature to move beyond lists and explain how or why particular factors may have affected the overall *level* of homelessness.

3.16 A reading of the recent literature reviews (Fitzpatrick, Kemp and Klinker, 2000; Third and Yanetta, 2000) and of the US research, suggests that the various structural factors thought to affect the level of homelessness can be grouped under 4 main headings. These 4 categories are set out below, with some of the causal factors mentioned in the literature listed under each of them. The list is illustrative rather than exhaustive.

1. Housing supply and demand
    - The flow of new lettings
    - The size of the council housing waiting list

- Landlord allocation policies and practices
- New household formation and dissolution
- Rent levels
- House prices

2. Employment
    - Unemployment
    - Availability of low-skill jobs

3. Poverty
    - Earnings
    - Social security benefit levels
    - Housing Benefit (including rent restrictions)

4. De-institutionalisation
    - Psychiatric hospital discharges
    - Provision of mental health care in the community
    - Children leaving care
    - Ex-offenders leaving prison
    - People leaving the armed forces

3.17 These various factors might be expected to have a differential impact on different groups in the population. For example, benefit levels and restrictions in Housing Benefit are likely to have had a greater impact on young people seeking accommodation in the privately rented sector than among other age groups.

3.18 Figure 3.1 draws on Burt's (1991, 1992) model of the causes of homelessness. It seeks to illustrate how these various groups of factors may influence the level of homelessness. It can be suggested that these factors can cause homelessness by affecting:

- the *availability* of sufficient and suitable accommodation
- the *affordability* of available accommodation, and
- the need for some kind of *social support* to enable potentially homeless people to live in the available and affordable accommodation.

Availability of accommodation

3.19 Homelessness can occur because there is insufficient housing of an acceptable quality available to potentially homeless households. For example, there might be an absolute shortage of suitable housing to rent in the locality, perhaps because demand exceeds supply. Or housing might be available, but certain types of household (such as young people) might be a low priority for re-housing by social landlords (Anderson, 1999). Alternatively, private landlords might be reluctant to let their accommodation to certain types of tenant, such as Housing Benefit recipients (Kemp and Rhodes, 1994). Housing might be available, but in locations that homeless people do not want to live or it might be of such poor quality that they do not wish to live in it (Pawson and Bramley, 2000).

3.20 The structural approach to homelessness would imply that, other things being equal, the level of homelessness will be inversely related to the flow of new social housing lettings.

In other words, the smaller is the flow of new lettings, the higher the level of homelessness. Likewise, homelessness is likely to vary inversely with the number of vacant dwellings: other things being equal, the lower the vacancy rate, the higher the level of homelessness.

Affordability

3.21 Homelessness might also occur because, even if there is sufficient accommodation, it is not affordable to potentially homeless people. Such people might not be able to afford to rent the available accommodation because their earnings are too low. People who are not in work may have difficulty paying rent because social security benefits are too low or because their Housing Benefit has been restricted (for example, due to the Local Reference Rent). Or, even if benefits are not too low, people may be reluctant to claim them or face too many obstacles in submitting a claim (Corden, 1995).

3.22 For example, the structural theory of homelessness would lead us to expect that, other things being equal, homelessness will be positively related to the rate of unemployment. In other words, if unemployment rises, so too will homelessness.

Need for social support

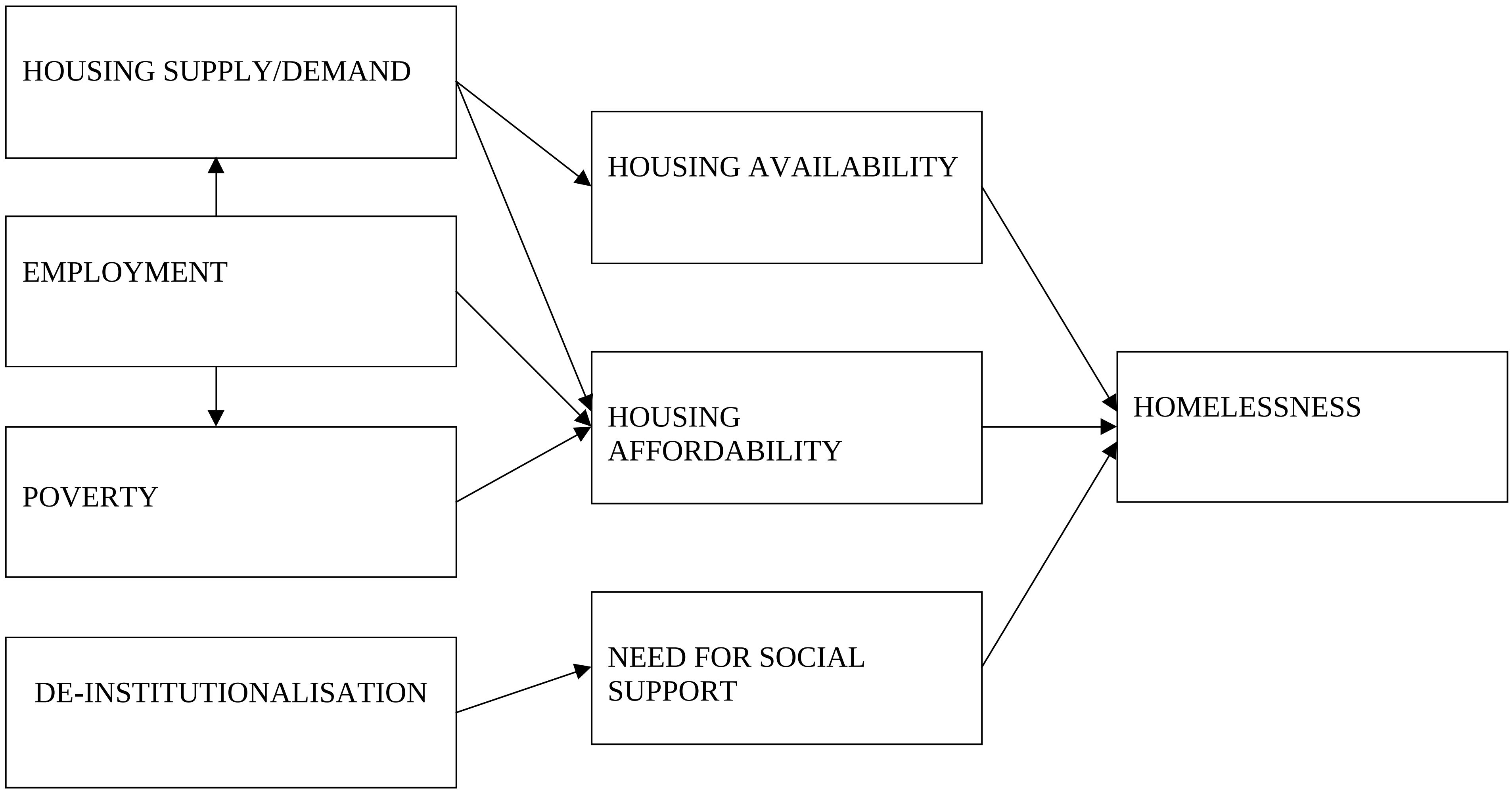
3.23 Finally, people might be homeless, even where accommodation is both available and affordable, because they need some kind of social support in order to maintain the tenancy. This might apply, for example, to young people leaving care, people who formerly were sleeping rough, and mentally ill people including those leaving psychiatric hospitals (Biehal and Wade, 1999; Pleace, 1995; Quilgars and Pleace, 1999).

3.24 For example, a structural understanding of the causes of homelessness would lead us to expect that, other things being equal, there is a positive relationship between the number of children leaving care and homelessness among people aged 16-18. Or alternatively as the number of psychiatric patients leaving hospital goes up, so too will homelessness, unless this is matched by community mental health care provision.

SUMMARY

3.25 This chapter has presented an outline model of structural causes of homelessness. The aim has been to develop a model that can be tested empirically using statistical methods. The next chapter describes the statistical techniques that were employed to test whether structural causes can explain, at least in part, the growth of homelessness over the past 2 decades and variations in levels of homelessness between different local authorities in Scotland.

Figure 3.1: A structural model of homelessness



CHAPTER FOUR DEFINITIONS, DATA AND METHODS

4.1 This chapter presents the data and methods used to test the structural model outlined in Chapter Three. The first section describes the 2 definitions of 'homelessness' that are used in this research. The second section sets out the variables used in the analysis and the sources from which we obtained them. It also introduces some important limitations of the data. Finally, the third section gives a description of each of the 3 statistical methods used: cross section regression analysis, pooled regression analysis and time series analysis.

DEFINING 'HOMELESSNESS'

4.2 Before one can test whether the level of homelessness can, at least in part, be attributed to structural causes, it is first necessary to define that which is to be explained. It is a commonplace in housing research that there is no universally agreed definition of 'homelessness'. Indeed, the concept of homelessness is both politically contested and socially constructed (Clapham, Kemp and Smith, 1990). Moreover, homelessness is also difficult to measure. At best, all that can be produced are estimates of the scale of the problem (Fitzpatrick, Kemp and Klinker, 2000).

4.3 For the purpose of this research project, homelessness has been defined in terms of what is often referred to as 'statutory homelessness' (Third and Yanetta, 2000) arising from the operation of the homelessness provisions of the Housing Act. We have used 2 different but related definitions of the level of homelessness:

- First, we have used as a measure of homelessness the number of people applying to local authorities for assistance under the homelessness provisions of the Housing Act. We have referred to these as 'homelessness applications'.

- Second, we have used the number of applications that have been accepted for re-housing by local authorities under these provisions. We have referred to these as 'homelessness acceptances'.

4.4 In the second category we have included applicants who have been re-housed even though found to be intentionally homeless as well as those found unintentionally homeless. It should be noted that the latter greatly outnumber the former.

4.5 The statutory homelessness statistics are not an ideal measure of the scale of homelessness. As Fitzpatrick, Kemp and Klinker (2000: 13) have pointed out:

> The statutory homeless statistics are the product of a bureaucratic process involving applicants and local authorities. To be recorded in the statutory homelessness statistics, it is necessary for households to express a 'felt need' (Bradshaw, 1972), namely that they are homeless or threatened with homelessness. Some households who feel themselves to be homeless may contact the local authority, some may not know they could do so, while others may not do so because they believe that they have no prospect of receiving assistance. This clearly places an important limitation on the usefulness of the statutory homeless statistics as a means of estimating the number of people who are homeless.

4.6 Using the statutory homeless statistics as the ‘dependent variable’ in our statistical modeling therefore means that we have excluded those homeless people who have not made an application to their local authority for re-housing on the grounds of homelessness. Hence, our measure almost certainly excludes many, but not all, people sleeping rough. Unfortunately, there are at present no reliable data on the number of people sleeping rough in each of the local authorities and over time in Scotland.

4.7 The statutory homelessness statistics also exclude what is sometimes referred to as ‘hidden homeless’ people, namely those who are homeless but either not in contact with statutory and voluntary agencies or not sleeping rough in known sites such as city centres (Webb, 1994). However, the number of ‘hidden homeless’ people is almost by definition unknown and no reliable estimates were available that we could use in the analysis.

DATA SOURCES

4.8 This section identifies the variables that have been used to test the structural model identified in Chapter Three. Data was collected for many more variables than were actually used in the final statistical analysis. We short-listed variables for the analysis based on their accuracy as a measure of the structural model and their availability.

4.9 The main source of the data provided was the Scottish Executive, including the departments of education statistics, local government finance, housing statistics, courts and law statistics, and recorded crime and prisons statistics. Other important sources included Census data (1981 and 1991), the General Register Office for Scotland, the Department of Social Security, the Information and Statistics Division of the NHS Scotland and the Northern Ireland Economic Research Council. We drew upon data collected in journals or publications such as *Regional Trends*, New Earnings Survey, *Scottish Housing Review 1988–1998*, and the University of York Index of Private Rents and Yields. We also gained access to the NOMIS database to retrieve unemployment data from the Labour Force Survey and the Office of National Statistics.

4.10 The variables are arranged here according to the 4 structural categories identified in Chapter Three. These are housing market factors, unemployment, affordability and de-institutionalisation. Table 4.1 shows all independent variables that were used for the statistical analysis, detailing the variable name, the variable description, the source, and the years for which each variable is available. Each variable is available at local authority level, except those marked with an * which are available at Scotland level only. A full description of each variable can be found in Annex 1.

Table 4.1 Data sources for variables used in the statistical analysis

Structural categories	Variable	Definitions	Sources	Years Available
Housing Market	VAC	Number of vacant accommodation, available, ready to let	Census	1981, 1991
	OVERCR	Number of residents in households with more than 1.5 person per room	Census	1981, 1991
	LP	Total households with lone parents	Census	1981
	VACANT	Vacant Local Authority stock	Scottish Executive	1986 – 1998
	NEWBUILD	New build completions by social landlords	Scottish Executive	1980 – 1999
	R2B	Number of right to buy sales	Scottish Executive	1980 – 1999
	<25	Number of under 25's	GRO	1981 – 1999
	LONE-PARENT	Number of unmarried parents (joint registration with different addresses, or sole registration)	GRO	1997 – 1999
	PVTHSN	Private Rented Sector Housing	Scottish Executive	1980 – 1999 *
	SOCHOUS	Social Rented Sector Housing	Scottish Executive	1980 – 1999 *
Unemployment	EA_EMP	Number of economically active people not in employment	Census	1981, 1991
	CLCOUNT	Number of claimant unemployed	Regional Trends	1992 – 1999
	NOMISLT	Number of claimant unemployed for more than 6 months	Nomis	1983 – 2000
	NOMISCC	Number of claimant unemployed	Nomis	1983 – 2000
	ILOUN16	Number of people over 16 who are unemployed	Labour Force Survey	1992 – 1999
	URATE	Unemployment Rate	Regional Trends	1980 – 1999 *
	LMANSC	Manufacturing Employment Index	Northern Ireland ERC	1980 – 1998 *
Affordability	LARENT	LA average weekly rents	Scottish Executive	1988 – 2000
	HPRICES	Average house prices	Scottish Housing Review	1989 – 1997
	HBPRCL	Number of HB private claims	DSS	1994 - 1999
	HBLACL	Number of HB LA claims	DSS	1994 – 1999
	UNF2BED	Average private rent for an unfurnished 2 bedroom flat	University of York Rent Index	1995 – 1999
	MORTARR	Number of actions in court for repossession of property	Scottish Executive	1994 – 1999
	ROREDUC	Rent officer statistics – the difference between mean reference rent and local reference rent	Scottish Executive	1996 – 2000
	RORATIO	Rent officer statistics – mean reference rent as a ratio of local reference rent	Scottish Executive	1996 – 2000
	RENTARRS	Number of decrees granted for evictions (inc. rent arrears)	Scottish Executive	1980 – 1999 *
Deinstitutionalisation	PSYCH	Number of people living in psychiatric homes/hospitals	Census	1981, 1991
	PRISN	Number of people living in prison establishments	Census	1981, 1991
	ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	ISD (NHS)	1980 – 1999
	CRIMES	Number of crimes and offences recorded (top 30 crimes)	Scottish Executive	1980 – 1999
	SCHEXCL	Average number of half days of temporary exclusions per 100 pupils (primary and 2ndary)	Scottish Executive	1995, 1996, 1998, 1999
	INPDSCH	Number of inpatient discharges from long-stay hospitals	ISD (NHS)	1998 – 2000
	DRGCNVC	Number of drugs related offences recorded by Scottish police.	Scottish Executive	1988 – 1998
	CHLDCAR	Number of children in care, or children looked after	Scottish Executive	1997 – 1999
	PSYDSCH	Number of psychiatric discharges	ISD (NHS)	1999

4.11 The variables selected as measures of the ***housing market*** were the number of vacant local authority properties, level of overcrowding, numbers of lone parents, number of under 25-year-olds, numbers of new build completions by social landlords, right to buy sales, the number of private rented dwellings and the number of social rented dwellings. We decided that the best measures of the housing market were vacancies and overcrowding, as they encapsulate both housing demand and supply in one measure. However, as these measures were not available for all years, we also included other housing market variables. The number of right to buy sales was included because it is sometimes claimed that they have been one of the factors behind the increased level of homelessness. The numbers of social and of private rented dwellings were both available only at Scotland level and were used for the time series analysis.

4.12 For ***unemployment***, 7 measures were chosen. These were the number of people economically active but not employed, 2 different counts of the number of claimant unemployed (one from *Regional Trends* and one from the Nomis database), the numbers of long term claimant unemployed (Nomis), the number of unemployed as defined by the International Labour Office (obtained from the Labour Force Survey), the official unemployment rate (Regional Trends) and Manufacturing Employment Index. Different measures were available for different years, but for each year we tested all unemployment measures available to us. The manufacturing employment index and the unemployment rate were available for the years 1980 – 1999, but only at Scotland level, and were used for the time series analysis.

4.13 The measures selected for ***affordability*** were local authority rent, average house prices, private sector Housing Benefit claims (which include housing association as well as private tenants), average private rents for unfurnished flats, the numbers of actions for possession on the grounds of mortgage arrears, rent officer statistics, and the numbers in rent arrears. The number of council tenants in rent arrears was only available at Scotland level, but again for 19 years, and was selected for use in the time series analysis.

4.14 We had identified rent officer data as being a very important measure of affordability. This data reflects the difference between the average private rent referred to the rent officer service and the local rent ceiling for Housing Benefit purposes (called the Local Reference Rent). Our aim was to provide a measure of the shortfall between these 2 figures, i.e. the difference between the mean referred rent and the local reference rent. It has been claimed that such shortfalls could cause homelessness by reducing the ability of Housing Benefit recipients living in privately rented accommodation to afford to pay their rent. However, we encountered a variable construction problem in transforming the variable for analysis, and a significant amount of the information was 'lost'. This suggests that this variable may no longer be an accurate or reliable measure of affordability. However, it was still included in the statistical analysis, along with a number of other measures of affordability.

4.15 Finally, for ***deinstitutionalisation***, the variables selected for use included the number of people in psychiatric institutions, numbers living in prison, number of psychiatric discharges (including those with diagnosis of alcohol misuse), number of crimes committed, number of school exclusions, number of inpatient discharges, number of drug convictions and the number of children in care. We hoped to obtain figures on army discharges and prison releases, as research has found a relationship between them and single homelessness (Anderson, Kemp and Quilgars, 1993). Unfortunately, however, this information was not available to us. Consequently, most of the deinstitutionalisation variables we have are measures of hospital discharges of one kind or another.

4.16 The measures we have selected in our shortlist represent each of the 4 structural categories. For each category, there is a selection of variables available, although the choice varies from year to year. For some structural categories we identified which was the ‘best’ or most important measure, however for other categories, there were no ‘best’ measures and the procedure was simply to test all variables that were available. We were satisfied that the measures available would allow us to robustly test our structural hypotheses using a variety of statistical methods, i.e. cross sectional analysis, pooled analysis and time series analysis.

Homelessness data

4.17 The homelessness data available for use in the statistical analysis comprised HL1 returns was provided to us by the Scottish Executive. We aimed to test the structural hypotheses for different homelessness groups, namely homelessness applications, homelessness acceptances, lone parent applications and young people applications. Table 4.2 shows the availability of such data.

Table 4.2 Homelessness data: HL1 returns

Year	Applications	Acceptances	Lone parents (app)	Young people (app)
1980	✓	✓		
1981	✓	✓		
1982	✓	✓		
1983	✓	✓		
1984	✓	✓		
1985	✓	✓		
1986	✓	✓		
1987	✓	✓		
1988	✓	✓		
1989	✓	✓		
1990	✓	✓		
1991	✓	✓	✓	✓
1992	✓	✓	✓	✓
1993	✓	✓	✓	✓
1994	✓	✓	✓	✓
1995	✓	✓	✓	✓
1996	✓	✓	✓	✓
1997	✓	✓	✓	✓
1998	✓	✓	✓	✓
1999	✓	✓	✓	✓

Limitations of the data

4.18 Our objective had been to obtain 20 years of data for each independent variable. Such data would have allowed us to consider the impact of our hypothesised factors on homelessness over the past 20 years. However, most variables were simply not available for long periods of time, or where they were available, the variable definitions or the methods of recording the information had changed. However, we are satisfied that we have collected data on accurate measures of our structural model and that the data was available for a sufficient number of years to produce reasonably robust results.

4.19 The availability of homelessness data also created problems. It can be seen in Table 4.2 that data on lone parent homelessness applications, and young people homelessness applications was only available for the years 1991 to 1999. This is particularly significant for the time series analysis, as 9 years is not sufficient to perform such an analysis.

4.20 There were also some general 'quality' issues regarding the data. There was missing data for many variables, which reduced the number of data points available for analysis. This is particularly significant for the cross sectional analysis, where we were working with such a small number of observations (32 local authorities). Missing observations in such circumstances cause serious problems for the robustness of the analysis. There are also questions about the accuracy of data collected at local authority level, for example, different local authorities may have used different recording techniques, which may obviously lead to inconsistent data.

4.21 Another potential problem concerned 'repeat applications', that is, applications for re-housing by people who have already been re-housed because of homelessness on a previous occasion. Repeat applications cannot be screened out of HL1 returns and are therefore included in the figures we have analysed. As a result, the same households may show up in the statistics on more than one occasion. However, it is a moot point as to whether the structural causes of homelessness are better tested on the homelessness data gross or net of repeat applications.

4.22 The local government reorganisation in 1996 also posed certain problems. In the first place, reorganisation substantially reduced the number of local authority areas (from 56 to 32). Second, in order to get a long time series, we had to aggregate the data (in respect of the years prior to 1996) for the old local authorities into their new, post-reorganisation counterparts. This aggregation posed the risk of possible loss of information. We collected information about the reorganisation, which allowed us to do the mappings ourselves, and converted all data for 56 old districts into the new 32 unitary local authorities.

Numbers versus rates

4.23 A final methodological issue that had to be considered was whether to use actual numbers for each of our variables or to standardise them by using 'rates' (for example the number of homeless people per 10,000 population). There is much to be said for using numbers when examining the increase in the level of homelessness over time, but the issue is more problematic in respect of comparisons between local authorities.

4.24 Rates or ratios have the advantage that they take into account the size of the population from which the data has been collected. However, there are a number of practical disadvantages to using rates. First, in order to calculate rates we would have to find appropriate denominators for every variable (apart from variables such as rents and house prices are in effect rates rather than volumes). This would have been very time-consuming and difficult to accomplish within the time constraints of the project. Second, for some variables (for example, school exclusions) there is room for disagreement about which denominators should be used, but the choice made can have a significant impact on the rate thereby calculated. Third, transforming data into rates can result in a loss of information and this can be especially problematic where the number of observations (as in this project) is very small. Fourth, debates about levels of homeless, and especially about the trend in homelessness over the past 2 decades, are invariably expressed in terms of the number of applications and acceptances, not in terms of rates.

4.25 An additional potential difficulty with rates is that it can lead to spurious correlation. This could occur when the numerators do not have any relationship to each other, but upon division by some denominators the ratios become correlated. The result may be a significant statistical relation between the variables even though there is no causal relationship (see Kendall and Buckland, 1972; Gujarati, 1995; and Kuh and Meyer, 1955). Using numbers should avoid this problem because the variables are chosen from different structural categories, which in theory should bear no relationship to each other, though some degree of correlation is inevitable.

4.26 Less of an issue, but nevertheless a problem, is the interpretation of regressions when using rates. Changes in rates may result from changes in *either* the numerator *or* the denominator or *both*. For example, suppose the homelessness rate is positively related to the sharing household rate. This means that if the sharing rate goes up then so does the homelessness rate. But the increase in sharing could be due to an increase in the numbers of households sharing or a decline in the total number of households or both. Using numbers avoids this ambiguity, as there are no other variables to blur the implications of the result.

4.27 Thus, there are arguments both for and against using numbers and rates in the analysis. We reviewed other research that has been carried out in this area and found that, while much of it used rates or ratios, some used numbers. After much consideration, we chose to conduct both the time series and cross section analysis (see Statistical Methods below) using numbers, though we did undertake a limited amount of analysis using rates, based on a small subset of variables (see Annex 4).

4.28 Although the analysis was largely based on numbers rather than rates, we did use other methods to address the fact that there are significant variations in the population size of Scottish local authorities. First, we transformed our data into natural logarithms, which has the effect of significantly minimising the dispersion in the data. Second, we used a statistical software package that can correct the results to allow for the fact that the sample contains some large as well as small values of key variables (known as 'heteroskedasticity' – see Annex 2).

STATISTICAL METHODS[1]

4.29 In order to undertake a quantitative analysis of the collected data and test the hypotheses of the structural model outlined in Chapter Three, we applied a statistical technique called 'regression analysis' at a number of different spatial and temporal levels; national, regional and yearly for example.

4.30 In essence, multiple regression analysis is a method for determining the importance of a number of factors (or group of factors) - termed explanatory or independent variables – for the variable whose values we wish to 'explain', called the dependent variable (see Williams and Monge, 2001). Statistical software packages compute the significance or importance of these independent variables and provide various measures of how well the chosen variables as a group explain the dependent variable. The influence of each individual variable can then be assessed from its probability value, as calculated by the software. The lower the probability value, then the more significant or important is that variable. Social researchers tend to assume that a probability of 0.10 or lower[2] is a good indicator of the importance of an independent variable.

4.31 In the context of this project, examples of explanatory variables would be unemployment and local authority dwelling vacancies, while the dependent variable would be the number of homelessness applications. Regression analysis would then ascertain the importance of each (or all) of these variables for the numbers of homelessness applications. The statistical software we used, *LIMDEP* v7.0 for Windows, produces a very comprehensive set of 'diagnostics' that allow the importance of variables in the regression to be assessed as well as providing an overall measure of the 'goodness of fit' of the regression.

4.32 As outlined above, we collected data on numerous potential variables at a number of different spatial scales over a wide range of years though not all data was available for every unitary local authority (ULA) for every year. This inconsistency in scale and time implies that not all the hypotheses that can be deduced from the 'model' may be testable for all years and/or regions. In order to maximise the usefulness of the information collated, we applied 3 variations of regression analysis. Specifically, these were:

- Cross section regression analysis
- Time series regression analysis
- Pooled Cross Section Analysis

Cross section regression analysis

4.33 Our first approach involved conducting regression analysis for a number of years. As there are nearly 20 years of data it became fairly apparent early on in the process that cross section regressions for each year would be both unfeasible and too costly in terms of computation time. Rather, the team decided to focus on the Census years (this provided us with a wide range of very reliable data) and on the years 1996 to 1999 inclusive, again

[1] This sub-section is intended to give a very basic non technical overview of the statistical methods employed by the research team. A fuller description of the technicalities of the methods and related issues can be found in Annex 2.
[2] Typically, only probability values of 0.05 tend to be reported. It was felt that owing to the various data/statistical problems concerned with this research that a value of 0.10 be used instead.

because these years provided us with the largest and most reliable range of potential variables. Thus, in terms of the analysis, regressions 'explaining' homelessness were carried out for the following years: 1981, 1991, 1996, 1997, 1998 and 1999. Our dependent variables were homelessness applications, homelessness acceptances, lone parent applications and young people applications.

4.34 Following the reorganisation of Local Government in 1995 the number of local authorities fell from 56 to 32. Thus, we had only 32 data points, one for each ULA. This is a ***very*** small number for reliable statistical analysis. Consequently, this has ramifications for the number of variables (or 'hypotheses') that can be tested reliably in multiple regression analysis. From a statistical point of view, the researcher cannot include more explanatory variables than there are observations. In other words, one *could* have up to 32 explanatory variables and no more, but this would lead to very unreliable and unstable results. Consequently, the team decided that for the research to stand up to scrutiny and be a *robust* investigation into the structural determinants of homelessness, that a limit be applied to the number of explanatory variables.

4.35 The structural model outlined in Chapter Three has 4 classifications (housing market, unemployment, affordability, de-institutionalisation) and each classification has a number of variables which, it could be argued, are representative of the classification to which it belongs. However, this was less true of de-institutionalisation because of the more limited variables available to us in this category. It was decided that one variable from each classification be included in the regression equation giving us 4 variables at most with which to test the structural model[3]. By adopting this approach we would mitigate the problem of multi-collinearity[4], where it can often be impossible to disentangle the effects of 2 variables which are implicitly measuring the same effect. For example including 'age' and 'work experience' in the same equation to explain, say, 'income', would be 'incorrect' as both variables are very closely related and so interpretation of the effect of 'age' would be complicated by the presence of the 'work experience' variable. We wanted to avoid this situation occurring and, conscious of the small sample size – called 'micronumerosity' by Goldberger (1991) - felt that one variable from each classification would be enough to test the structural model robustly.

4.36 It is important to note that the purpose was *not* to test whether certain *specific* variables had an influence on homelessness but rather to test the over arching framework of the model. Thus, the aim was to test whether 'housing market', 'unemployment', 'affordability', and 'deinstitutionalisation' factors affect homelessness and not whether the 'claimant count' or the 'ILO' measure of unemployment affects homelessness; otherwise there would be no limit to the number of regressions that could be run and not all of them would be statistically meaningful for the reason alluded to above. For this reason, a wide range of variables was selected and tested one at a time within each of the four categories. The results of the cross section analysis are presented in the first section of Chapter Five.

Time series analysis

4.37 Although a series of cross section regressions is very useful in indicating important variables in the determination of homelessness over time, it does not make full use of the

[3] Obviously, if the structural model had five classifications then five variables would have been used.
[4] The concept of multi-collinearity is explained in more detail in Annex 2.

time dimension of the data. This is because cross sections are merely snap shots of what is 'happening' at one period (year) in time. However, time series analysis, applied to a series of observations on the same variable over a specific period of time, will allow the researcher to ascertain whether a *long run* relationship exists between homelessness and its hypothesised factors. We used the latest statistical technique available within the *EViews* v3.0 statistical software to ascertain whether there is a long run relationship between the structural variables and homelessness applications.

4.38 Additionally, time series regression analysis allows us to explore the possibility of forecasting what the number of homeless people will be for the immediate short run, typically 2 - 3 years hence, which was a requirement of the research brief. Techniques do exist which will allow us to forecast into the future and a brief description of these are outlined in Annex 2. Our forecasts are based on HR2 returns.

4.39 For time series analysis to be insightful, however, we require a fairly lengthy series with the relevant variables available for *each year in each local authority*. Our data audit exercise indicated that some form of time series analysis would be possible for those variables recorded since 1980, though again it was not feasible to test every hypothesis outlined in section 3. Indeed, it again became apparent that any time series analysis would have to be conducted at a national level due to the strict data requirements needed for a reliable and robust analysis. The required data was either simply not available at the sub-national level or the series were extremely short so as to render estimation impossible. Thus, the time series analysis in the third section of Chapter Five is confined to the national level only with a very restricted set of structural variables.

Pooled cross section regression analysis

4.40 Lack of suitable, good quality data along with a limited number of data points was an ongoing problem throughout the period of the analysis. The data suitability and data quality problems were beyond our control but the 'limited number of data points' issue could be tackled to a certain extent. It can be seen from the tables in the previous section, that certain key variables are only available for a very short time period. This means that time series regression analysis will *not* be the most appropriate method of analysis owing to the small number of data points.

4.41 However, other statistical techniques exist which allow us to circumvent this problem whilst utilising the 'time' aspect of the data. Specifically, we can 'pool' short time series together to give us a greater number of data points. In other words we stack data for each ULA for those years for which data is available. For example, if data were available for 1991-1993 (3 years) for 32 local authorities, then instead of attempting a time series regression on 3 years worth of data we would 'pool' the data, effectively stacking one year on another, and obtain, 3*32=96 potential data points. These data points would then be used to test the hypotheses of our structural model the results being, in general, more robust the larger the number of data points. However, not all of the data was available for all of those years (or all of the ULAs) so the actual number of data points used, in reality, was considerably less than the maximum outlined above. The results of these regressions are presented in the second section of Chapter Five.

4.42 In fact, missing data proved to be a serious problem throughout the analysis. Unfortunately, nothing can be done about this but it necessarily results in the deletion of that particular data point (ULA) from the analysis. It does not take too many missing values to render analysis unreliable or intractable. Hence, our insistence that analysis should be conducted on *as large a number of data points and as wide a possible range of variables* as feasibly possible. These restrictions however, came at a price; a smaller number of cross section regressions and a more restricted set of variables to test the structural model detailed in Chapter Three. Nevertheless, the aim was to provide a *reliable and robust* analysis of the hypothesised structural variables and we believe that this could only be achieved through our insistence on the requirements outlined above.

CHAPTER FIVE TESTING THE STRUCTURAL MODEL

5.1 This chapter presents the results of the statistical analysis that was carried out in order to test the validity of the structural model. The first section presents the results of the cross-sectional analysis for each of a number of years, namely 1981, 1991, and 1996 to 1999 inclusive. The second section sets out the results from the pooled regression analysis in which the data for a period of years are grouped together for analysis. The third section outlines the main results from the time series analysis.

CROSS-SECTION REGRESSION ANALYSIS

5.2 This approach involved running regression analyses for individual years in order to test the structural model. Multiple regressions were run for 1981, 1991, 1996, 1997, 1998 and 1999 to ascertain the importance of selected variables on the numbers of homelessness applications and acceptances. Table 5.1 shows the variables used in each of the cross sectional analyses. A description of each of these variables is included in Annex 1 of the report.

5.3 The variables selected for our shortlist (Table 5.1) represent each of the four structural categories outlined in Chapter Three (housing market, unemployment, affordability, de-institutionalisation). Each of these categories has a number of variables which, it could be argued, are representative of it in some way. It was decided that one variable from each category should be included in each regression equation, giving us four variables at most with which to test the structural model. By adopting this approach we would mitigate the problem of multi-collinearity, where it can often be impossible to disentangle the effects of two variables which are implicitly measuring the same effect. All variables listed in Table 5.1 were tested. For some structural categories we identified which was the 'best' or most important measure, but for other categories, there were no 'best' measures and the procedure was simply to test all variables that were available. Adhering to the rule that each structural classification must be represented, every possible combination of four variables was calculated and tested.

Results for 1981

5.4 Table 5.1 shows the variables available for the 1981 analysis. For this year only, we had no measure of affordability and therefore had to select variables according to different groups from those suggested by the model. We chose one factor from housing supply and one from housing demand, rather than one overall housing market variable and one from affordability. The measures selected for ***housing supply*** were vacancies and overcrowding. For ***housing demand***, lone parents, right to buy and new build, were used. For ***unemployment,*** the measure used was the number of people economically active but not employed. Finally, the variables for ***deinstitutionalisation*** were the numbers of psychiatric inpatients (including those with diagnosis of alcohol misuse) and numbers in prison establishments. Only homelessness applications and acceptances were available as dependent variables for this analysis.

5.5 The model that best explains variations in homelessness between Scottish local authorities is listed and defined below (see Table 5.2). The table also lists the sign on the variable coefficients (the direction of the relationship) and whether or not they were statistically significant. Note that * denotes that a variable is significant at the 10 per cent level, ** at the 5 per cent level, and *** at the 1 per cent level. This means that, for example, if a variable is significant at the 1 per cent level, then the probability is that 99 times out of 100 the relationship observed in the data reflects the actual importance of the variable in question rather than being due to chance. Full details of the regression results (regression coefficients, adjusted R^2 figures, etc) can be found in Annex 3.

Table 5.1 Variables selected for analysis

Structural category	**Variables**	**1981**	**1991**	**1996**	**1997**	**1998**	**1999**
Housing Market	VAC	✓	✓				
	OVERCR	✓	✓				
	LP	✓					
	NEWBUILD	✓	✓				✓
	R2B	✓	✓				✓
	VACANT			✓	✓	✓	
	<25						✓
	LONEPARENT						✓
Unemployment	EA_EMP	✓	✓				
	CLCOUNT			✓	✓	✓	✓
	NOMISLT			✓	✓	✓	✓
	NOMISCC			✓	✓	✓	✓
	ILOUN16			✓	✓	✓	✓
Affordability	LARENT		✓	✓	✓	✓	✓
	HPRICES		✓	✓	✓		
	HBPRCL			✓	✓	✓	✓
	HBLACL			✓	✓	✓	✓
	UNF2BED			✓	✓	✓	✓
	MORTARRS			✓	✓	✓	
	ROREDUCT			✓	✓	✓	✓
	RORATIO			✓	✓	✓	✓
Deinstitutionalisation	PSYCH	✓	✓				
	PRISN	✓	✓				
	ALCO	✓	✓	✓	✓	✓	✓
	CRIMES	✓					
	SCHEXCL			✓		✓	✓
	INPDSCH					✓	✓
	DRGCNVC			✓	✓	✓	
	CHLDCAR				✓	✓	✓
	PSYDSCH						✓

5.6 In 1981, the factors that best explained variance in homelessness applications and acceptances were the number of vacant dwellings, the number of households with lone parents, the number of people who are economically active but not employed, and finally the number of psychiatric discharges with any diagnosis of alcohol misuse. Only vacancies and psychiatric discharges with a diagnosis of alcohol misuse were significant. However, together, all 4 factors were able to explain 78 per cent of the variation in homelessness applications, and 77 per cent of the variation in acceptances. This can be also be described in terms of an 'adjusted R^2' of 0.78 for applications, and 0.77 for acceptances.

Table 5.2 Best fit models for 1981: homelessness applications and acceptances

Variable	Variable label	Sign	Significance
Applications			
VAC81	Number of vacant accommodation that is available and ready to let	-	*
LP81	Total households with lone parents	+	
EA_EMP	Number of economically active people, minus those in employment (a measure of unemployment)	+	
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	+	***
Acceptances			
VAC81	Number of vacant accommodation that is available and ready to let	-	*
LP81	Total households with lone parents	+	
EA_EMP	Number of economically active people, minus those in employment (a measure of unemployment)	+	
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	+	***

5.7 Table 5.2 also shows the direction of impact each of the variables has on homelessness figures, i.e. the sign on the coefficient. For 1981, the signs are identical for both applications and acceptances. The vacancies variable has a negative ('-') sign. This indicates that local authorities with *high* numbers of vacancies had *low* numbers of homelessness applications and acceptances. This is fully consistent with the structural model and is what one would expect in a market system. The other 3 variables all have a positive relationship with homelessness figures ('+' sign). This implies that local authorities that had *high* numbers of lone parents, of people economically active but not employed, or of psychiatric discharges with diagnosis of alcohol misuse, also tended to have *high* numbers of homelessness applications and acceptances. Again, these results are consistent with the structural model.

Results for 1991

5.8 Table 5.1 also shows the variables selected for the 1991 analysis. The measures chosen to represent the ***housing market*** were vacant dwellings, overcrowding, new build and right to buy sales. For ***unemployment,*** we used the number of people economically active but not employed. For ***affordability***, local authority rents and house prices were selected, and for ***deinstitutionalisation***, the number of psychiatric inpatients (including those with diagnosis of alcohol abuse) and numbers in prison establishments.

5.9 The models that explained the most variance in homelessness applications and acceptances are shown in Table 5.3.

Table 5.3 Best fit models for 1991: homelessness applications and acceptances

Variable	Variable label	Sign	Significance
Applications			
VAC91	Number of vacant accommodation that is available and ready to let	+	
EA_EMP	Number of economically active people, minus those in employment (a measure of unemployment)	+	***
LARENT	Average weekly local authority rents	-	
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	+	**
Acceptances			
VAC91	Number of vacant accommodation that is available and ready to let	+	
EA_EMP	Number of economically active people, minus those in employment (a measure of unemployment)	+	***
HPRICES	Average house prices	+	***
PSYCH91	Numbers living in psychiatric homes and hospitals	-	

5.10 The variables that were best able to explain the variance in homelessness applications were the number of vacant dwellings, the number of people who are economically active but not employed, local authority rent levels, and the number of psychiatric discharges with a diagnosis of alcohol misuse. Only the number economically active but not employed and the number of psychiatric discharges with diagnosis of alcohol misuse were significant.

5.11 As before, the signs on the variables indicate the direction of impact they have on homelessness figures. In line with the structural model, economically active but not employed, and psychiatric discharges with diagnosis of alcohol both possess '+' signs. In other words, local authorities in 1991 that had high numbers of economically active but unemployed people or high levels of psychiatric discharges also tended to have high numbers of homelessness applications.

5.12 However, the '+' sign for vacant dwellings is counter intuitive, as it indicates that authorities with high numbers of vacant dwellings also had high numbers of homelessness applications. The structural model predicts that, if local authority vacancies are high, homelessness will be low. Also counter intuitive is the negative relationship between local authority rents and applications figures, as it suggests that local authorities with high rents in 1991 had low levels of homelessness applications. Again, the structural model would hypothesise that high rents are associated with high levels of homelessness. However, since neither of these 2 relationships is significant, they do not offer evidence against the structural model.

5.13 The factors that best explain the variance in homelessness acceptances in 1991 were the numbers of vacant dwellings, numbers economically active but not employed, average house prices, and numbers living in psychiatric institutions. Only numbers economically active but not employed and average house prices were significant.

5.14 Table 5.3 shows the '+' coefficients for economically active but not employed, and average house prices, which indicate that rises in either of these factors will cause a rise in homelessness acceptances. This is consistent with the structural model. Vacant dwellings again has a '+' coefficient, but this time the variable is not significant either for applications or acceptances. The variable for the numbers living in psychiatric institutions is also not significant. The lack of significance in these counter intuitive findings means they are not able to refute the hypothesised model. Instead, some support for the model can be found in the fact that the significant variables have the direction of impact on homelessness acceptances that the model predicts.

5.15 The models produced for 1981 and 1991 are capable of explaining, in a statistical sense, somewhere between 73 per cent and 78 per cent of the variation in homelessness figures across Scottish unitary local authorities. However, perhaps partly due to the second-best nature of some of the variables, no model contains 4 significant variables from each of the structural categories.

Analysis for 1996 to 1999

5.16 For the years 1996 - 1999 a wider selection of independent variables was available than was the case for 1981 and 1991. The selection of variables for these years can be seen in Table 5.1. The measures selected for the ***housing market*** element of the structural model were the number of local authority vacancies, right to buy sales, new build completions by social landlords, the number of lone parents, and the number of people aged under 25. For ***unemployment***, the measures included International Labour Office (ILO) definition of unemployment, 2 measures of the number of unemployed claimants (one from Nomis and one from Regional Trends), and the ONS measure of the number of long term claimant unemployed. The measures selected to represent ***affordability*** were private sector Housing Benefit claims (which include housing association as well as private tenants), public sector Housing Benefit claims, rent officer statistics, the number of owner-occupiers in mortgage arrears, local authority rent levels, house prices, and private rents. Finally, for ***deinstitutionalisation***, psychiatric discharges (including those with diagnose of alcohol misuse), children in care, school exclusions, drug convictions, and inpatient discharges were selected.

5.17 Not all of these variables were selected for 1996 to 1999 inclusive and indeed not all were available. For 1996, 1997 and 1998, the single variable selected for the ***housing market*** category was the number of local authority vacancies. This variable was identified as the best measure of the housing market as it encapsulates both housing demand and supply. However, since vacancies was not available in 1999, other variables from this category were included (such as right to buy sales and new build by social landlords), but were used in the analysis for this year only.

Results for 1996

5.18 For 1996, the models that were best able to explain the variance in homelessness are listed in Table 5.4.

Table 5.4 Best fit models for 1996: applications, acceptances, lone parents and young people

Variable	Variable label	Sign	Significance
Applications			
VACANT	Number of local authority vacancies	+	
CLCOUNT	Claimant count	+	***
LARENT	Average weekly local authority rent	-	**
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	-	***
Acceptances			
VACANT	Number of local authority vacancies	+	*
NOMISLT	Long term claimant count	+	***
RORATIO	Ratio of mean reference rent to local reference rent	-	*
SCHEXCL	Numbers of half days lost in temporary school exclusions	-	**
Lone Parents			
VACANT	Number of local authority vacancies	+	*
CLCOUNT	Claimant count	+	***
HBENPRC	Housing benefit claims for private rent	+	***
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	-	***
Young People			
VACANT	Number of local authority vacancies	+	*
CLCOUNT	Claimant count	+	***
LARENT	Average weekly local authority rent	-	*
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	-	**

Applications

5.19 The variables that explained the most variance in homelessness applications are vacancies, claimant unemployed (from *Regional Trends*), local authority rents and psychiatric discharges with a diagnosis of alcohol misuse. All these factors are significant except vacancies. Overall, the model explains 84 per cent of the variance in applications figures (i.e. has an 'adjusted R^2' of 0.84).

5.20 Table 5.4 also shows the direction of impact the factors have on application figures. The '+' coefficients for vacancies and claimant unemployed indicate both of these factors have a positive relationship with homelessness applications. The positive relationship between claimant unemployed and homelessness applications is consistent with the structural model. It indicates that, in 1996, local authorities that had high numbers of the claimant unemployed also had high numbers of homelessness applications, holding other things constant.

5.21 However, once again the '+' coefficient for vacancies is counter intuitive and not consistent with the structural model. While the coefficient is not significant for applications, it is for the other 3 dependent variables in 1996 (acceptances, applications by lone parents, and applications by young people). This raises the question of why local authorities with high numbers of vacant dwellings should also experience high levels of homeless. One possible explanation could be that local authorities with high levels of vacancies were experiencing falling demand for local authority housing, perhaps due to the poor quality or undesirability of the housing. In this situation, the effective level of vacancies might be much higher than the nominal level because much of the stock is, in practice, unlettable.

Alternatively, the social malaise associated with problematic council estates might be responsible for generating both difficult to let housing (and hence high levels of vacancies) *and* high levels of homelessness. The connection between high levels of vacancies and homelessness may also reflect the phenomenon of repeat homelessness applications discussed in the previous chapter.

5.22 The '-' coefficients for local authority rent and psychiatric discharges with alcohol misuse indicate that both factors have a negative relationship with application figures. The negative relationship between local authority rent and applications figures is at first sight counter intuitive, as it implies that local authorities which had relatively high rents in 1996 also had relatively low levels of homelessness – and vice versa. However, it may be that areas with high vacancies also have low rent levels. This would again relate to the low demand for local authority housing.

5.23 Also counter intuitive is the negative relationship between psychiatric discharges with diagnosis of alcohol misuse and homelessness applications. The structural model suggests that the higher the number of such discharges, the higher the number of homelessness applications. However, the reverse is true in this regression, and this may possibly be explained by local authority homelessness practices. It is possible local authorities were now taking action to ensure people being discharged from psychiatric institutions did not become homeless.

Acceptances

5.24 The model that explains most variance in homelessness acceptances includes vacancies, long term claimant unemployed, the ratio of the mean referred rent (MRR) to the local reference rent (LRR) on private rents referred to the Rent Officer Service by local authority Housing Benefit officials, and the number of school exclusions (see Table 5.4). All 4 variables are significant and the model is able to explain 80 per cent of the variance in acceptance figures (i.e. it has an adjusted R^2 of 0.8).

5.25 The vacancies and long-term claimant unemployed variables both have '+' coefficients, indicating a positive relationship with acceptances. The positive relationship between long term claimant unemployed and homelessness acceptances is in line with the structural model. It indicates that local authorities with higher unemployment will have higher homelessness acceptance figures.

5.26 Once again, there is a '+' coefficient for vacancies. The '-' coefficients for the ratio of MRR to LRR and the number of school exclusions imply a negative relationship with acceptance figures. The negative relationship between the ratio of MRR to LRR and acceptances is counter intuitive. This measure was constructed so that a higher ratio (or difference) reflected a larger shortfall between the referred rent and the rent accepted for Housing Benefit purposes. According to the structural model, the hypothesis would be that the higher the ratio, the higher the homelessness figures. However, this is not the case here, or indeed in any model where rent officer statistics appear as a significant variable in explaining homelessness figures. The relationship is always a negative one. This could be because in transforming the rent officer data into a suitable variable for analysis, a significant loss of information may have occurred. This implies that, in practice, both the ratio of MRR to LRR, and the difference between MRR and LRR, may be poor measures of affordability for this analysis.

5.27 The '-' coefficient for school exclusions indicates that local authorities with high levels of school exclusions had low levels of homelessness acceptances. This is not consistent with the structural model. Since school exclusions are related to deprivation, unemployment, poor housing, and poor educational attainment, one would expect higher school exclusions to be accompanied by higher, not lower, levels of homelessness. However, for all years except 1999 (see Table 5.7) higher numbers of school exclusions predict lower numbers of homelessness acceptances. There is no immediately obvious explanation for this finding.

Lone parents

5.28 The factors best able to explain the variance in lone parent homelessness applications were vacancies, claimant unemployed (obtained from *Regional Trends*), private sector Housing Benefit claims, and psychiatric discharges with diagnosis of alcohol misuse. Again, all variables are significant, and the model produces an adjusted R^2 of 0.88.

5.29 In terms of the direction of impact each of these factors has on lone parent applications, vacancies, unemployed and private sector Housing Benefit claims all have positive relationships (+ sign). These results for the unemployed and Housing Benefit variables are in line with the structural model, but the result for vacancies is not.

5.30 Only psychiatric discharges with diagnosis of alcohol misuse has a negative relationship (- sign) with lone parent applications. This is contrary to the structural model as it suggests that rises in the numbers of such discharges will cause a fall in the number of lone parent homelessness applications. The structural model predicts that increases in such discharges are likely to cause increased homelessness. Again, it is possible that this negative relationship can be explained by local authority practice in relation to the re-housing of people being discharged from institutions.

Young people

5.31 Finally, the model that is most able to explain the variance in young peoples' homelessness applications includes vacancies, claimant unemployed, local authority rents, and psychiatric discharges with diagnosis of alcohol misuse. All variables are significant and the model has an adjusted R^2 of 0.71.

5.32 The signs on the variable coefficients show that vacancies and claimant unemployed once again have a positive relationship with young peoples' homelessness applications. As with applications, acceptances and lone parent applications, the number of vacancies is positively related to the number of young peoples' applications. Meanwhile, local authority rents and psychiatric discharges with diagnosis of alcohol misuse once again have '-' coefficients, which implies an inverse relationship with young peoples' applications.

Results for 1997

5.33 Table 5.5 shows the models that are best able to explain variance in homelessness figures across local authorities in 1997.

Table 5.5 Best fit models for 1997: applications, acceptances, lone parents and young people

Variable	Variable label	Sign	Significance
Applications			
VACANT	Number of local authority vacancies	+	***
ILOUN16	Number of people over 16 who are unemployed	+	***
HPRICES	Average house prices	+	***
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	-	***
Acceptances			
VACANT	Number of local authority vacancies	+	***
ILOUN16	Number of people over 16 who are unemployed	+	***
HPRICES	Average house prices	+	***
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	-	***
Lone Parents			
VACANT	Number of local authority vacancies	+	**
ILOUN16	Number of people over 16 who are unemployed	+	***
HPRICES	Average house prices	+	***
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	-	**
Young People			
VACANT	Number of local authority vacancies	+	***
ILOUN16	Number of people over 16 who are unemployed	+	***
HPRICES	Average house prices	-	***
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	-	***

5.34 For all homeless groups, the factors that explain most variance are vacancies, the ILO measure of unemployment, average house prices and psychiatric discharges with a diagnosis of alcohol misuse. All 4 variables are significant in each model. The adjusted R^2 applications is 0.86, for acceptances, 0.82, for lone parents, 0.89 and for young people it is 0.81.

Applications, acceptances and lone parents

5.35 As before, the signs on the coefficient in Table 5.5 describe the direction of the impact each factor has on homelessness figures. The relationships are the same for applications, acceptances and lone parent applications.

5.36 Vacancies, ILO unemployment and house prices all have positive relationships with homelessness (+ sign). This implies that rises in the number of vacancies, ILO unemployment, and house prices will be associated with a rise in homelessness applications, acceptances and lone parent applications. For unemployment and house prices, this is exactly as predicted by the structural model. However, the positive relationship for vacancies and the negative relationship for psychiatric discharges with a diagnosis of alcohol misuse, are not consistent with the structural model. Both results are, however, consistent with the findings for 1996 presented above, where possible explanations for this were discussed.

Young people

5.37 The one difference for this group is the negative relationship between house prices and application figures. This indicates that a rise in house prices will lead to a fall in homelessness applications by young people, and is contrary to what is hypothesised by the structural model.

Results for 1998

5.38 The models that explain most variance in homelessness figures for 1998 are outlined in Table 5.6.

Table 5.6 Best fit models for 1998: applications, acceptances, lone parents and young people

Variable	Variable label	Sign	Significance
Applications			
VACANT	Number of local authority vacancies	+	**
NOMISLT	Long term claimant count	+	***
LARENT	Average weekly local authority rent	-	*
SCHEXCL	Number of half days missed through temporary exclusions	-	*
Acceptances			
VACANT	Number of local authority vacancies	+	*
NOMISLT	Long term claimant count	+	*
RORATIO	Mean reference rent as a ratio of local reference rent	-	**
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	+	*
Lone Parents			
VACANT	Number of local authority vacancies	+	*
NOMISLT	Long term claimant count	+	***
RORATIO	Mean reference rent as a ratio of local reference rent	-	*
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	+	*
Young People			
VACANT	Number of local authority vacancies	+	***
NOMISLT	Long term claimant count	+	***
LARENT	Average weekly local authority rent	-	*
SCHEXCL	Number of half days missed through temporary exclusions	-	**

Applications and young peoples' applications

5.39 The factors that best explain variance in all applications and young peoples' applications are vacancies, long-term claimant unemployed, local authority rents and school exclusions. All variables are significant. For applications figures, the model has an adjusted R^2 of 0.71, and for young peoples' applications, there is an adjusted R^2 of 0.68.

5.40 For both of these homelessness groups, the direction of impact of each of the variables is the same. Rises in vacancies and long-term unemployed will be accompanied by rises in homelessness figures (+ sign), while rises in local authority rents, and school exclusions will result in falls in homelessness figures (- sign).

5.41 The positive relationship between long term unemployed and homelessness figures is intuitive, and what we would expect according to the structural model. This was also the finding in 1996, and 1997, and indicates that rises in long term unemployment will cause a rise in the numbers of total applications and young people applications.

5.42 However, as before, the '+' coefficient on vacancies is contrary to the structural hypotheses, as it suggests that as the number of local authority vacancies rise, there will be a rise in applications figures. This is the same result as that found using the 1996 and 1997 data. Likewise, the negative relationship between local authority rents and applications is counter intuitive, as it suggests that as rents rise, homelessness applications will fall. Again, this was also the finding with the 1996 and 1997 data.

5.43 The negative relationship between school exclusions and acceptances figures is also contrary to the structural model (and replicates that found with the 1996 data). This finding implies that as the number of school exclusions rises, homelessness applications falls. There is no obvious explanation for this finding.

Acceptances and lone parent applications

5.44 The models which best explained the variance in homelessness acceptances and lone parent applications included vacancies, long-term claimant unemployed, the ratio between the referred rent and the local reference rent, and psychiatric discharges with diagnosis of alcohol misuse. Again, all variables are significant. The model for acceptances produced an adjusted R^2 of 0.66, and for lone parents it was 0.74.

5.45 The signs denoting direction of impact on homelessness were the same for acceptances and lone parent applications. For both these groups, rises in vacancies, long-term unemployed and psychiatric discharges with diagnosis of alcohol misuse will be accompanied by rises in homelessness figures. Whereas, an increase in the ratio between the referred rent and the local reference rent will lead to a decrease in homelessness figures.

5.46 The finding that rises in long term unemployment will cause rises in homelessness acceptances and lone parent applications is predicted by the structural model. This finding has been consistent throughout the cross sectional analyses. Likewise, the positive relationship between vacancies and homelessness figures, while counter intuitive, has also been consistent (except for 1981 when it was a negative relationship).

5.47 The finding that a rise in psychiatric discharges with diagnosis of alcohol misuse will be accompanied by a rise in homelessness figures, although intuitive, is a change to previous years findings. Previous analyses have found that rises in such discharges have been found to result in lower levels of homelessness. This has been a consistent, although counter intuitive finding. However, for homelessness acceptances and lone parent applications, the findings indicate that increases in psychiatric discharges with diagnosis of alcohol misuse will lead to increases in the numbers of acceptances and lone parent applications. This is what the structural model predicts.

Results for 1999

5.48 In 1999, different combinations were selected for the regressions because some of the variables used for 1996 to 1998 were not available for analysis in this later year. This means that the models that best explain the variance in homelessness figures for 1999 (see Table 5.7) will be markedly different from those for previous years.

Table 5.7 Best fit models for 1999: applications, acceptances, lone parents and young people

Variable	Variable label	Sign	Significance
Applications			
<25	Numbers under 25 years old	+	*
NOMISCC	Claimant count	+	**
UNF2BED	Average private rent for an unfurnished 2 bedroom flat	+	***
INPDSCH	Number of inpatient discharges from long-stay hospitals	+	**
Acceptances			
LONEPT	Number of lone parents	+	**
NOMISCC	Claimant count	+	*
UNF2BED	Average private rent for an unfurnished 2 bedroom flat	+	**
SCHEXCL	Number of half days missed through temporary exclusions	-	*
Lone Parents			
NEWB	New build completions by social landlords	+	***
CLCOUNT	Claimant count	+	***
UNF2BED	Average private rent for an unfurnished 2 bedroom flat	+	***
INPDSCH	Number of inpatient discharges from long-stay hospitals	+	***
Young People			
R2B	Number of right to buy sales	+	**
ILOUN16	Number of people over 16 who are unemployed	+	***
ROREDUCT	Difference between mean reference rent and local reference rent	-	**
SCHEXCL	Number of half days missed through temporary exclusions	+	**

Applications

5.49 The factors that explained most variance in homelessness applications were the number of under 25 year olds, claimant unemployed, private rents and inpatient discharges. All variables were significant and the model produced an adjusted R^2 of 0.85.

5.50 All variables show a positive relationship with applications (+ sign), implying that rises in any of the 4 factors will result in a rise in applications figures. This is what would be expected according to the structural model. It indicates that those local authorities with the highest number of under 25-year-olds, the highest level of claimant unemployed, the highest private rent levels and highest number of inpatient discharges are likely to have the highest number of homelessness applications.

Acceptances

5.51 The model that best explained the variance in acceptance figures included numbers of lone parents, claimant unemployed, private rents and numbers of school exclusions. Again all factors were significant in this model, producing an adjusted R^2 of 0.82.

5.52 The '+' coefficients for lone parents, claimant unemployed and private rents indicate that rises in any of these factors will be accompanied by rises in acceptances figures. This is consistent with the structural model, implying that those local authorities with high numbers of lone parents, high claimant unemployment and high private rents are likely to have higher numbers of homelessness acceptances.

5.53 The '-' coefficient for school exclusions implies that a rise in the number of school exclusions will result in a fall in homelessness acceptances. This is counter intuitive, but is consistent with findings from previous years.

Lone parents

5.54 The factors most significant in explaining the variance in lone parent homelessness applications included new build completions by social landlords, claimant unemployed, private rents and inpatient discharges. All variables were significant and the adjusted R^2 for this model was 0.88.

5.55 Each of these 4 factors has a positive relationship with lone parent applications, implying a rise in any of these factors will result in a rise in lone parent homelessness applications. These results are consistent with the structural model for claimant unemployed, rent and inpatient discharges, but not so for new build. The model predicts that local authorities with high unemployment, high private rent levels and high numbers of inpatient discharges will have higher levels of lone parent homelessness applications, but that those with high levels of new build will have low levels of homelessness applications. This apparently perverse result may relate to the finding for previous years that local authority vacancy rates were positively related to homelessness. Authorities with high vacancy rates are perhaps less likely to be building new stock. Alternatively, it may reflect a policy *response* to high levels of homelessness in the local authority area.

Young people

5.56 Finally, the model most able to explain the variance in young peoples' homelessness applications included number of right to buy sales, the ILO measure of unemployment, the difference between the mean referred rent (MRR) and the mean local reference rent (LRR), and the number of school exclusions. All variables were significant, and the model produced an adjusted R^2 of 0.86.

5.57 The numbers of right to buy sales, ILO unemployment and school exclusions have a positive relationship with homelessness applications by young people. This implies that rises in these factors will be accompanied by a rise in homelessness applications by young people. This finding is consistent with the structural model, and suggests that local authorities with high numbers of right to buy sales, unemployment and school exclusions are likely to have high levels of homelessness applications by young people. Once again, however, the negative ('-') coefficient for the difference between the referred rent and the local reference rent is not consistent with the structural model.

5.58 Most interesting here is that for the first time, the school exclusion variable has a positive relationship with homelessness figures. It is unclear why this relationship has changed compared with previous years.

Cross section analysis 1996 – 1999

5.59 So far, we have discussed only the models that are ***best*** able to explain the variance in homelessness figures across local authorities, i.e. have the highest adjusted R^2. However, for each year there were additional models, which also had high adjusted R^2 and contained 4 significant variables. The variables that were significant in these additional models are also important, as they provide further support for the structural model of homelessness.

5.60 Tables 5.8 and 5.9 show all such variables, by homeless group and by year. These tables allow us to see quickly what variables were significant each year and whether they had a positive or negative relationship with the homelessness figures. It also allows us to see the consistency of the results between different years and different homeless groups. (These results, together with the best fit results presented earlier in this chapter, are summarised in the final table in Annex 3. Although there are some inconsistencies, in most cases the direction of the relationships- whether positive or negative - is generally constant for most years for which the variables are significant in the regression models.)

5.61 Table 5.8 shows the variables from 1996, 1997 and 1998. Since similar variables were entered in the regressions for these years, comparisons between the years can be made.

5.62 Table 5.8 shows that the measure used for the ***housing market*** category is number of vacancies. This variable is significant for all 3 years and consistently has a positive (though counter intuitive) relationship with homelessness.

5.63 At least one of the different measures of ***unemployment*** is also significant each year and consistently has a '+' coefficient, indicating that a rise in unemployment will lead to a rise in homelessness figures. This is in line with the structural model.

Table 5.8 Variables from all significant models in 1996, 1997 and 1998

Structural categories	Variables	1996				1997				1998			
		App	Acc	LP	YP	App	Acc	LP	YP	App	Acc	LP	YP
Housing market	Vacant	+	+	+	+	+	+	+/-	+	+	+	+	+
Unemployment	Claimant count	+	+	+	+			+					
	Nomis LT	+	+							+	+	+	+
	Nomis CC								+				
	ILO U/E					+	+	+	+	+		+	+
Affordability	HB Pvte Claims	+	+	+									
	Private Rent					+		+	+			+	
	LA rent	-			-					-			-
	House Prices					+	+	+	+	n/a	n/a	n/a	n/a
	Mortgage Arrears					+	+	+	+	+		+	+
	Rent officer (ratio)		-							-	-	-	-
	Rent officer (reduct)		-										-
De-institution -alisation	Alcohol	-	-	-	-	-	-	-	-	+/-	+	+	-
	School Exclusions		-			n/a	n/a	n/a	n/a	-			-
	Inpatient Discharges	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a		+		+
	Drug Convictions							-				-	
	Childcare	n/a	n/a	n/a	n/a			-					

Key:
n/a = data not available for this year
App = Applications, Acc = Acceptances, LP = Lone parents, YP = Young people

5.64 Measures of ***affordability*** were also significant for 1996, 1997 and 1998, although they were less consistent. Contrary to the structural model, local authority rents and rent officer statistics both consistently showed a negative relationship with homelessness. However, other indicators of affordability - such as the number of Housing Benefit claims, private rent levels, house prices, and the numbers of people in mortgage arrears - consistently showed positive relationships, implying that rises in any of these would lead to an increase in homelessness. This supports the structural model.

5.65 There was a lack of suitable data available for ***deinstitutionalisation***, which contributed to less consistent findings than for some of the other structural categories. Although every year there was a measure of deinstitutionalisation which was significant, the signs on the coefficients were often counter intuitive or were changeable between years (i.e. school exclusions and psychiatric discharges with diagnosis of alcohol misuse).

Table 5.9 Variables from all significant models in 1999

Structural categories	Variables	Applications	Acceptances	Lone parents	Young people
Housing	Vacant	n/a	n/a	n/a	n/a
	Under 25s	+		+	+
	R2B	+	+		+
	Newbuild	+		+	
	Loneparent	+	+		+
Unemployment	Claimant count	+	+	+	+
	Nomis LT	+		+	
	Nomis CC	+	+	+	
	ILO U/E	+	+	+	+
Affordability	HB Pvte Claims	+	+	+	+
	HB caseload				+
	Private Rent	+	+	+	
	LA rent			-	
	House Prices	n/a	n/a	n/a	n/a
	Mortgage Arrears	n/a	n/a	n/a	n/a
	Rent officer (ratio)	-	-		-
	Rent officer (reduct)	-	-	-	-
De-institution-alisation	Alcohol	n/a	n/a	n/a	n/a
	School Exclusions	+	+/-	+	+
	Inpatient Discharges	+	+	+	+
	Drug Convictions	n/a	n/a	n/a	n/a
	Childcare	+	+	+	
	Psychiatric Discharges	-		+/-	-

5.66 The list of variables used in the analysis for 1999 is shown in Table 5.9. In this year, due to data restrictions, several different variables were used compared with those used for the years from 1996 to 1999. Vacancies, psychiatric discharges with diagnosis of alcohol misuse, house prices, numbers in mortgage arrears and drug convictions were not available. Consequently, the results for 1999 are not directly comparable to those from 1996, 1997 and 1998.

5.67 The number of vacant local authority dwellings was not available in 1999 as a measure of the state of the ***housing market***. This variable had been identified as our best measure of the state of the housing market because it encapsulated both housing demand and supply. However, it was unavailable in 1999 and consequently other measures were used instead. These included (on the demand side) the numbers of under 25-year-olds and the number of lone parents and (on the supply side) right to buy sales and new building. Each of these was significant in explaining homelessness in 1999 and each had a positive relationship with homelessness figures. The finding that increases in the number of people aged under 25, in the number of lone parents and in right to buy sales, would cause higher levels of homelessness was as predicted by the structural model. However, the positive relationship between new build and homelessness was not consistent with the structural model. Possible reasons for this relationship were discussed above.

5.68 The results for ***unemployment*** in 1999 are consistently in line with the structural model. Other things being equal, rises in unemployment will lead to rises in homelessness for each of the homeless groups.

5.69 The findings for ***affordability*** also offer some support for the structural model. There is a statistically significant measure of affordability in each model explaining homelessness in 1999. The positive coefficients for Housing Benefit claims by private and housing association tenants and for private rent levels are findings that support the structural model. The results indicate, therefore, that local authorities that have large numbers of Housing Benefit claimants or high private rent levels tend to have high levels of homelessness. However, the Rent Officer statistics (the ratio of the mean referred to rent to the mean Local Reference Rent) are negatively related to homelessness. This counter-intuitive result is consistent with previous years' findings.

5.70 Finally, for ***deinstitutionalisation*** in 1999, Table 5.9 shows that school exclusions, inpatient discharges and numbers of children in care are all positively related to homelessness. These are consistent relationships and offer further support for the structural model. There is some inconsistency, however, in the relationship between homelessness and psychiatric discharges.

Summary

5.71 The cross sectional analyses from 1981, 1991, 1996, 1997, 1998 and 1999 have produced models which are capable of statistically 'explaining' between 68 per cent and 89 per cent of the variance in homelessness figures (i.e., in technical terms the adjusted R-squares varied from 0.68 to 0.89). The analysis also identified the individual factors within each of these models. For recent years, the analysis has produced models that not only explain a high percentage of the variance in homelessness figures, but where all 4 variables included in the model, are statistically significant.

5.72 Since the variables are direct measures of the 4 categories within the structural model, they can be interpreted as providing evidence to support this model. The structural model predicts not just whether a variable will be significant, but also what direction of impact it will have on homelessness. The coefficient signs are broadly consistent with the hypothesised model, although some relationships are counter intuitive, and others change from one year to the next. However, there is evidence that a structural model such as the one outlined previously is capable of explaining variation in homelessness between local authorities in Scotland.

POOLED REGRESSION ANALYSIS

5.73 As discussed in Chapter Four, the limited number of observations (i.e., 32 local authorities) available on which to conduct the cross sectional analysis was a serious constraint on the regression modeling. In order to overcome this problem, short time series of data were 'pooled' together to provide a greater number of data points. In practical terms, this meant that each local authority counted as a separate observation or data point for each year for which the data were pooled.

5.74 Three pooled cross sectional analyses were conducted: 3 years (1996-1998), 5 years (1994-1998) and 10 years (1989-1999). Due to time and practicality restrictions, only those variables that appeared in significant models in the cross sectional analysis were selected (see Tables 5.8 – 5.9). Table 5.10 shows the variables that were used in each pooled analysis.

Table 5.10 Variables available to be tested in pooled regression

Structural category	**Variables**	**3 years (1996–1998)**	**5 years (1994– 1998)**	**10 years (1989–1998)**
Housing Market	Vacant	✓	✓	✓
Unemployment	Nomis LT	✓	✓	✓
	Nomis CC	✓	✓	✓
	Claimant Count	✓	✓	
	ILO Unemployed	✓	✓	
Affordability	LA Rent	✓	✓	✓
	Private Rent	✓		
	Rent Officer data	✓		
	House Prices	✓	✓	✓
	HB Private Claims	✓	✓	
	Mortgage Arrears	✓	✓	
Deinstitutionalisation	Alcohol	✓	✓	✓
	Drug Convictions	✓	✓	✓
	School Exclusions	✓		
	Children in care	✓		

5.75 The adjusted R^2 figures for the pooled regressions were extremely high and should be regarded with considerable caution. For this reason, they have not been included in the main body of the report, but can be found in Annex 3 along with the other detailed results of the regressions.

Three years' pooled analysis

5.76 Data from 1996, 1997, and 1998 were 'pooled' together, giving 96 observations. This is still fewer observations than one would ideally hope to run in a regression analysis. However, since a wide choice of variables was available for only these 3 years (see Table 5.10) it was decided to go ahead and test the 3-year data. The large number of missing data in the variables caused serious computational problems. This, and the small number of observations, implied that results from this analysis were not robust. Therefore, these results are not included in this discussion.

Five years' pooled analysis

5.77 Data for the 5 years from 1994 to 1998 were also 'pooled' together, providing 160 observations (i.e. 32 local authorities * 5 years). Table 5.10 shows that there were fewer variables available for this analysis than for 3 years, but more data observations were available. Combinations of 4 variables - one variable from each of the 4 structural categories

- were entered into the regression analysis. To find the best model, all possible combinations of 4 variables were run for all 4 dependent variables. Table 5.11 shows the models that were best able to explain the variance in homelessness figures for applications, acceptances, lone parents and young people.

Table 5.11 Best fit models using pooled data from 1994 – 1998

Variable	Variable label	Sign	Significance
Applications			
VACANT	Number of local authority vacancies	+	*
ILOUN16	Number of people over 16 who are unemployed	-	*
HPRICES	Average house prices	+	**
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	+	
Acceptances			
VACANT	Number of local authority vacancies	+	**
ILOUN16	Number of people over 16 who are unemployed	+	
LARENT	Average weekly local authority rents	+	**
DRGCNVC	Number of drugs related offences recorded by Scottish police.	+	
Lone Parents			
VACANT	Number of local authority vacancies	-	
NOMISLT	Long term claimant count	+	
MORTARRS	Number of actions raised in court for repossession of property	-	**
DRGCNVC	Number of drugs related offences recorded by Scottish police.	-	*
Young People			
VACANT	Number of local authority vacancies	+	
NOMISCC	Claimant count	-	***
LARENT	Average weekly local authority rents	+	
ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse	-	

Technical note: Fixed effects for local authority are included in each regression

Applications

5.78 The factors that are most able to explain the variance in homelessness applications using 5 years' of pooled data include the number of vacancies, ILO unemployment, house prices, and psychiatric discharges with diagnosis of alcohol misuse. All factors are significant except psychiatric discharges with diagnosis of alcohol misuse.

5.79 All variables except ILO unemployed have a positive relationship with homelessness applications. According to the structural model, high levels of house prices and psychiatric discharges with diagnosis of alcohol misuse should have a positive relationship with the numbers of homelessness applications. However, the relationship between vacancies and homelessness is contrary to the structural model, but consistent with the findings of the cross section regressions. The '-' coefficient for ILO unemployed is also contrary to the findings from the cross sectional analysis presented above.

Acceptances

5.80 For homelessness acceptances, the variables that best explain the variance are vacancies, ILO unemployment, local authority rent, and numbers of drug convictions. Only the number of vacancies and local authority rent are significant.

5.81 All variables show a positive relationship with homelessness acceptance figures. For vacancies, while counter intuitive, this is consistent with the findings in previous analyses. For unemployment, local authority rents and drug convictions the findings that rises in any of these factors will lead to a rise in the number of homelessness acceptances supports the structural model.

Lone parent applications

5.82 The model that best explains the variance in lone parent homelessness applications includes number of vacancies, long term claimant unemployed, numbers in mortgage arrears and number of drug convictions. Only the numbers of owner-occupiers in mortgage arrears and the numbers of drug convictions are significant.

5.83 There are some differences in the direction of impact for lone parent applications, compared with those for all homelessness applications and acceptances. Unemployment continues to show a positive relationship with homelessness numbers. So too do vacancies, but the relationship is not statistically significant. The inverse relationship between homelessness applications by lone parents and both mortgage arrears and drug is significant but counter intuitive.

Young people applications

5.84 Finally, the model which best explains the variance in homelessness applications by young people includes number of vacancies, claimant unemployed (as recorded in the Nomis database), local authority rents and psychiatric discharges with a diagnosis of alcohol misuse. However, only claimant unemployed is statistically significant.

Ten years' pooled analysis

5.85 Data from 1989 to 1998 was 'pooled', providing 320 observations. Table 5.10 shows there are very few variables available for this analysis, but again all combinations were tested to find the model that could explain the most variance. Again only combinations of 4 variables were entered in our regressions. Only applications and acceptances were available for this analysis. Table 5.12 shows the best models using 10 years of pooled data.

Table 5.12 Best fit models using data pooled from 1989 - 1998

Variable	Variable label	Sign	Significance
Applications			
VACANT	Number of local authority vacancies	+	
NOMISCC	Claimant count	+	**
HPRICES	Average house prices	+	***
DRGCNCV	Number of drugs related offences recorded by Scottish police.	+	***
Acceptances			
VACANT	Number of local authority vacancies	+	
NOMISCC	Claimant count	+	*
HPRICES	Average house prices	+	***
DRGCNVC	Number of drugs related offences recorded by Scottish police.	+	***

Technical note: Fixed effects for local authority are included in each regression

5.86 The factors that were able to explain the most variance in application and acceptance figures were vacancies, claimant unemployed, house prices and drug convictions. The number of vacancies is not significant for either homelessness groups.

Applications and acceptances

5.87 All variables have a positive relationship with homelessness applications and acceptances. For vacancies, this relationship is counter intuitive, but is consistent with almost all findings from previous analysis. Numbers of unemployed, house prices and drug convictions also have positive relationships with homelessness figures, which is consistent with the structural model. These results indicate that rises in the number of claimant unemployed, house prices and drug convictions will be associated with rises in homelessness applications and acceptances.

5.88 Since there were a far greater number of data observations included in the 10-year pooled analysis, greater weight can be placed on these results than on the 5-year pooled analysis results. The 10-year pooled analysis provides the most robust analysis so far of the relationship between structural factors and homelessness, and has been able to offer strong support for the structural model.

Summary

5.89 The pooled regression analysis was performed with the aim of providing a more robust analysis than that produced by the cross section analysis. In theory at least, pooled analysis is more robust as it combines several years of data, thus providing more observations for analysis. The pooled analyses for 3 years only had 96 observations, but a wide choice of independent variables to test. However, missing data was a serious problem for several of these variables and as a result this analysis was less robust than the cross sectional analysis. Therefore, the results for 3-year pooled analyses were not presented.

5.90 Using 5 years of pooled data, there was a degree of inconsistency in the direction of impact of the variables. Overall, the findings from the 5-year analysis provided support for the structural model, but also some inconsistent results.

5.91 The regression analyses with 10 years of pooled data involved 320 observations and should provide more robust results. The findings were consistent with single-year cross section analyses and provided strong support for the structural model.

TIME SERIES ANALYSIS

Testing for a long run relationship

5.92 The purpose of conducting time series analysis on the homelessness data was to investigate whether there was a long run relationship between homelessness applications and structural variables. If such a relationship exists, that would imply that the structural variables have influenced the 'path' of homelessness applications over the period from1980 to 1998, the years for which the data are available.

5.93 This section presents the results of this analysis using the latest time series methods, which attributable to both Johanson (1988, 1991) and Stock and Watson (1988). This technique differs from traditional regression analysis in that the objective is not to obtain a set of coefficients *per se*. Rather, the method exploits the mathematical properties of matrices to determine whether two or more variables are related to each other over a period of time, termed 'co-integration'. Co-integration can be said to be present when, over a period of time, movements in one data series (such as homelessness applications) reflect movements in another data series (such as unemployment). As one would expect, there are advantages and disadvantages to this approach (see Annex 2).

5.94 Before one can conduct a time series analysis, the properties of the series under consideration should be investigated. This is to ensure that only those series that have the same basic 'structure' are included, so as to avoid spurious results between the dependent variable and the independent variables or, in this case, structural variables. Again, Annex 2 elaborates on the technicalities of the investigation procedure. Depending on the outcome of the investigations, the researcher may well be faced with the problem of not being able to use those variables that at first sight would probably have been included in the co-integration testing procedure. This was the case here. Not only did the different series have very different properties, but they also covered a fairly short time period; both problems necessitating their removal from the analysis. Once the inappropriate series were eliminated, there were 6 potential variables left with which to test the structural model (see Table 5.13).

5.95 The number of usable observations (i.e., years) is 19. This covers the period from 1980 to 1998. As it happened this gave us just enough observations to allow the co-integration test to proceed. Attempting analysis with series of smaller length resulted in failure of the testing procedure.

Table 5.13: Time series structural variables

Series	Length of series in years
Unemployment Rate	19
Manufacturing employment index	19
Right to Buy Sales	19
Public lettings	19
Persons in rent arrears	19
Crimes	19

5.96 The available series allows us to test all 4 categories of the structural model. For the ***unemployment*** classification we have the choice of 2 variables: the unemployment rate and manufacturing employment and for the ***housing market*** we have 2 variables: right to buy sales, and public lettings. For ***affordability*** we have one variable: persons with rent arrears. Under ***de-institutionalisation*** we have just one variable: recorded crimes. This is intended as a crude proxy for the numbers of people being released from prison (data for which were not available).

Results

5.97 For all 6 tests, the results show that there is a long run relationship between homelessness applications and the structural variables (see Table 5.14). In other words, the variables included in the analysis have, from a statistical point of view, influenced the number (and path) of homelessness applications. This provides further support for the structural model outlined in Chapter Three and helps reinforce the conclusions of the cross-sectional analysis.

Table 5.14: Co-integration tests

Variables	Co-integrated?
Homelessness applications, crimes, unemployment rate, right to buy sales	Yes
Homelessness applications, crimes, unemployment rate, rent arrears	Yes
Homelessness applications, crimes, unemployment rate, public lettings	Yes
Homelessness applications, crimes, manufacturing employment, right to buy sales	Yes
Homelessness applications, crimes, manufacturing employment, rent arrears	Yes
Homelessness applications, crimes, manufacturing employment, public lettings	Yes

Causality

5.98 The above analysis, however, cannot say what causes homelessness - it only says that all 4 variables are related over the long run. To address the issue of causality a different kind of analysis is needed - called Granger Causality - which is only applicable to time series analysis. At this stage, a word of caution is required: when statisticians talk about causality they are *not* using it in the normal sense of the word. They are in fact talking about *statistical* causality, termed Granger Causality, which is a different concept altogether. Koop (2000) gives a clear definition of the concept:

> '...a variable X Granger causes Y if past values of X can help explain Y'

5.99 Granger causality tests were conducted on the variables included in the co-integrating relationships described above. The tests take the form of a series of hypotheses on the direction of the causality for each pair of variables. The F statistic is computed along with its probability level, which tells us whether a relationship is a significant one or not. The results of the tests are presented in Table 5.15.

Table 5.15 Granger causality tests

Null hypothesis	Level of statistical significance
Crimes does not Granger cause homelessness applications	0.53
Homelessness applications does not Granger cause crimes	0.57
Unemployment rate does not Granger cause homelessness applications	0.03
Homelessness applications does not Granger cause unemployment rate	0.46
Right to buy does not Granger cause homelessness applications	0.12
Homelessness applications does not Granger cause right to buy	0.67
Rent arrears does not Granger cause homelessness applications	0.12
Homelessness applications does not Granger cause rent arrears	0.39
Public lettings does not Granger cause homelessness applications	0.31
Homelessness applications does not Granger cause public lettings	0.80
Manufacturing employment does not Granger cause homelessness applications	0.88
Homelessness applications does not Granger cause manufacturing employment	0.08

5.100 The results of the tests are mixed. It appears that the direction of statistical causality for homelessness applications and crimes cannot be determined. A similar result holds for homelessness applications and public lettings. Testing the unemployment rate and homelessness applications does lead to a significant result at the 3% level. Unemployment appears to Granger cause homelessness applications. A test of the reverse causality reveals that we cannot reject the hypothesis that unemployment rate does not Granger cause homelessness applications. Looking at right to buy sales and homelessness reveals a not quite significant causality result (significant at the 12% level): right to buy 'causes' homelessness applications; the reverse hypothesis cannot be rejected. The result for rent arrears and homelessness is also slightly above the 10% level (significant at the 12% level). The reverse hypothesis cannot be rejected. Finally, there would appear to be no causality between homelessness applications and the level of manufacturing employment.

CONCLUSION

5.101 The results of this time series analysis lend support to the structural model outlined in the previous chapters. Long run relationships are revealed between homelessness applications and a number of variables in the structural model. In addition, there is weak evidence to suggest that these structural variables in some sense Granger cause homelessness, though the definition of cause in this context should be *strongly borne in mind.*

CHAPTER SIX PREDICTING HOMELESSNESS

6.1 This chapter of the report addresses the final requirement of the brief: to make some prediction or forecast of the level of homelessness in future years. The literature on forecasting is vast and many techniques are available to the researcher. Below we present 3 sets of forecasts derived from time series analysis, each of which is different from a methodological point of view, to generate predictions of the numbers of homelessness to the first quarter of 2002. We also present an altogether different method of 'prediction' or forecasting by looking at the coefficients on the cross sectional and pooled regression results. These are termed 'elasticities' and can be used to make judgements as to what will happen to the homelessness level if there is an increase in say, unemployment.

METHOD 1 TIME SERIES APPROACH

6.2 One of the strengths of time series analysis is its ability to predict or forecast future values of a series of data. However, as with any prediction, there is much scope for inaccuracy. The data we used for this purpose was HL2 quarterly returns supplied by the Scottish executive. The HL2 returns are the homelessness figures placed on the Scottish Executive website. They are the most frequently used by researchers looking at the homelessness figures over time. They incorporate local authority revisions to the figures and consequently are not always exactly the same as the HL1 returns.

6.3 We obtained quarterly HL2 data going back as far as the second quarter of 1990 and ending in the first quarter of 2000, giving us approximately 40 observation points on which to conduct our analysis. Our method of analysis (forecasting/prediction) comprised 2 approaches. The **first time series approach** involved using previous values of the HL2 series to *explain the series' behaviour* and make predictions as to the numbers of homelessness in the future beyond the first quarter of 2000, a statistical methodology called 'univariate time series' or Box-Jenkins modeling[5] (Box and Jenkins, 1976). The main output from such an approach is a regression model explaining current values of the series in terms of past values. These coefficients can then be used to forecast the series into the future. We conducted 2 variations of the Box Jenkins approach. The first variation involved *removing* the strong seasonal element present in the data (see below) whilst the second variation explicitly took this seasonality into account.

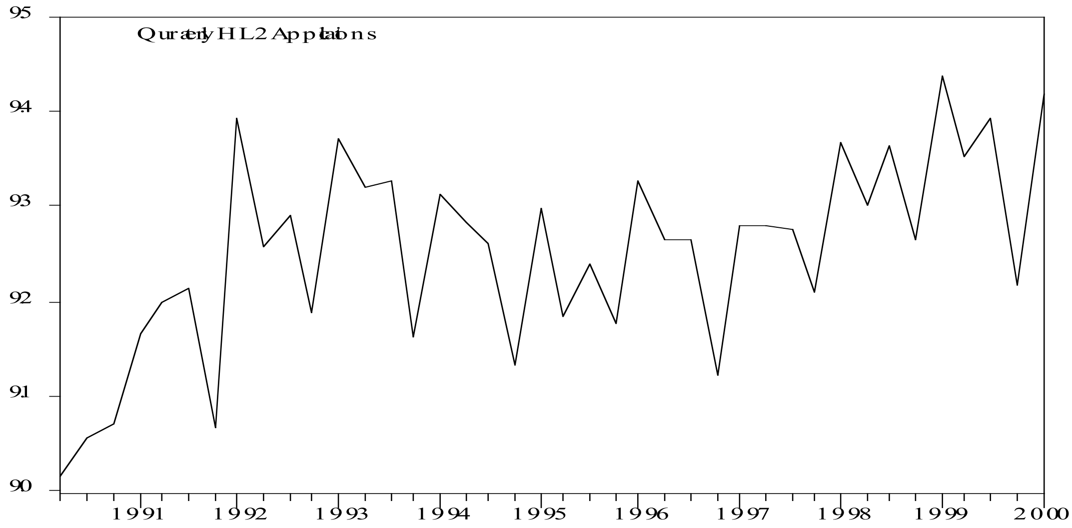
6.4 The **second time series approach** was to use a technique called Holt-Winters Exponential Smoothing. The purpose of this technique is *not* to explain the behaviour of the series *per se*, but to obtain estimates for (i) long run average, (ii) trend and (iii) seasonal properties (which are continuously updated) of the data. These estimates are then used in conjunction with a number of mathematical equations for the purposes of forecasting. Thus, there is no single regression model as is the case for the Box-Jenkins approach. Annex 1 explains the difference between the 2 techniques in slightly more detail.

6.5 The first stage of analysis for both Box Jenkins and Holt-Winters involves plotting the series in a graph as shown in Figure 6.1. What is immediately apparent is the strong

[5] Ideally, we would have preferred to use other additional series to build a forecasting model. However, there are also problems associated with that other approach that are not easily overcome. See Annex 1.

seasonality of the data. Every fourth quarter (October, November, December) the recorded number of homelessness applications took a large dip before increasing in the first quarter of the next year, presumably due to a reluctance to evict people around the festive Christmas and New Year period[6]. As was the case with the HL1 returns, there would appear to be an upward trend in the data over time.

Figure 6.1 Quarterly HL2 returns (logged)



6.6 Annex 1 outlines the exact modeling procedure we used and the significance levels of the variables contained within the Box-Jenkins regression analysis. As should be the case for Box-Jenkins modeling all variables were significant and the amount of variation explained by the regression was 62 per cent for the model with seasonality removed and 79 per cent with the seasonality modelled. These figures suggest a respectable fit for the models. Other statistics were also produced to assess the 'fit' of the regression and all were found to be very favourable. Again, these are provided in Annex 1. Using these acceptable 'models' of the series the following forecasts in Table 6.1 of HL2 returns were made for every quarter up to the first quarter of 2002. Accompanying these forecasts are forecasts obtained from using the Holt-Winters Exponential Smoothing method.

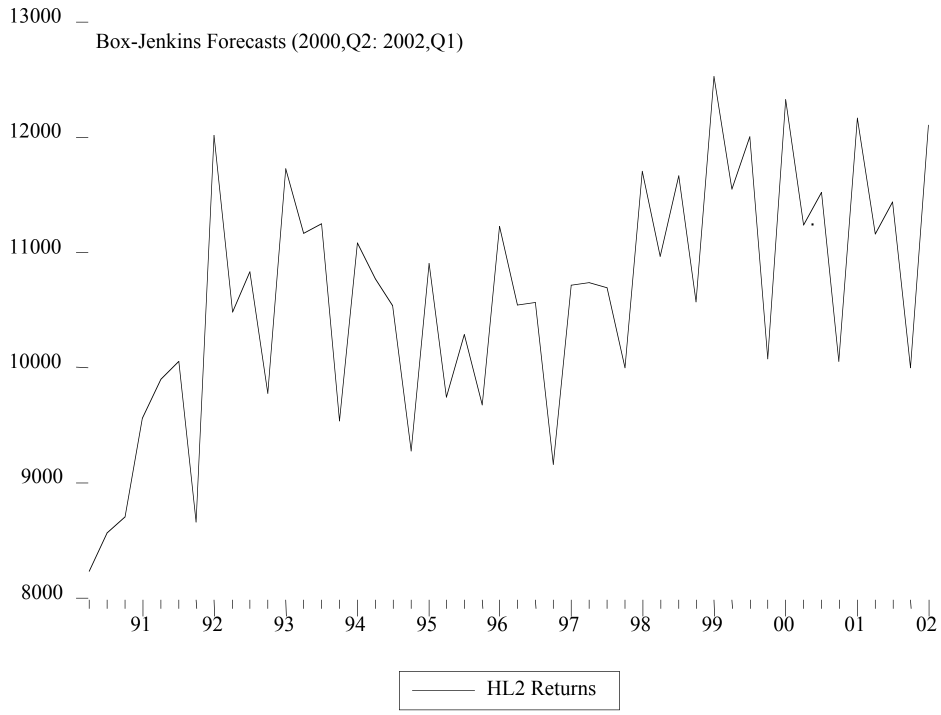
6.7 We have provided 3 sets of forecasts because each method of forecasting has its own merits. As can be seen, there is a broad correspondence between the figures especially between the 2 Box Jenkins variations in quarters 2 and 3 of the year 2000. After that, the estimates begin to diverge more. The Holt Winters method tends to produce higher forecasts than the Box Jenkins models. Plotting the original data along with the forecasts reveals some useful information.

[6] In fact, seasonality is also a characteristic of the HL1 data. When this data was also broken down by month (and quarter) it could be seen that the value in December for every year was much lower than that for all other months of the year.

Table 6.1 Forecasts of HL2 returns[7]

Period	Forecast (Box-Jenkins Approach Seasonality Removed)	Forecast (Box-Jenkins Approach Seasonality Modelled)	Forecast (Holt-Winters Approach)
2000, Q2	11,240	11,129	11,602
2000, Q3	11,523	11,310	11,868
2000, Q4	10,051	9,971	10,659
2001, Q1	12,164	11,784	12,698
2001, Q2	11,160	11,000	11,955
2001, Q3	11,440	11,164	12,229
2001, Q4	9,995	9,890	10,983
2002, Q1	12,106	11,698	13,084

Figure 6.2 HL2 quarterly applications data with forecasts using Box-Jenkins Model (seasonality removed)



[7] As mentioned earlier these forecasts are subject to error, the magnitude of which simply cannot be determined. They should be read as being a "ball park" figure and nothing more.

Figure 6.3 HL2 applications data with forecasts using Box-Jenkins Model (Seasonality Modelled)

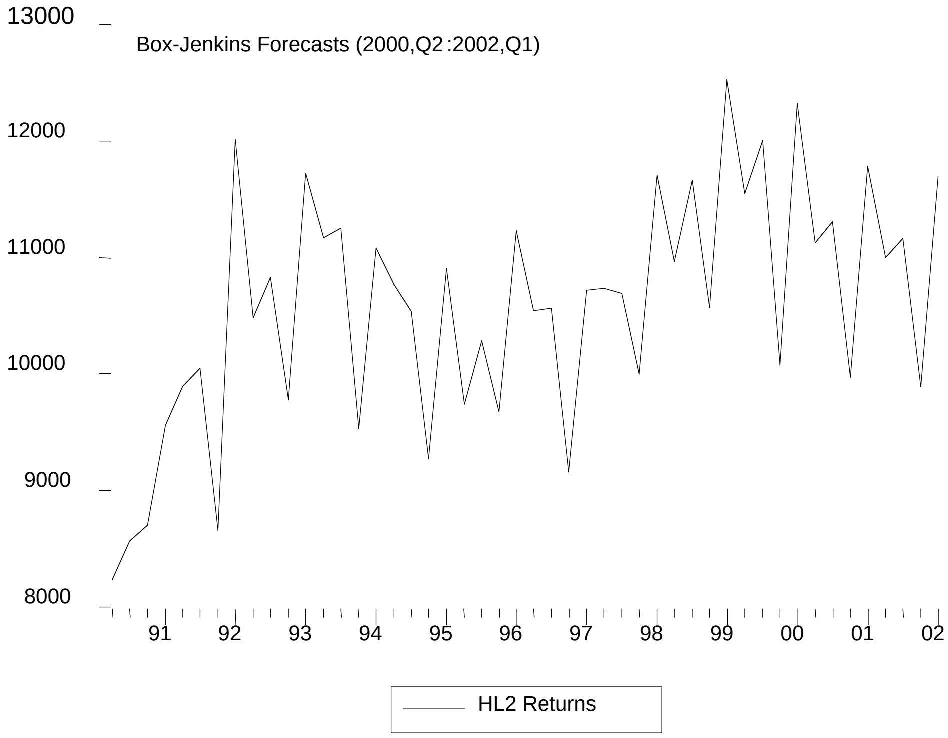
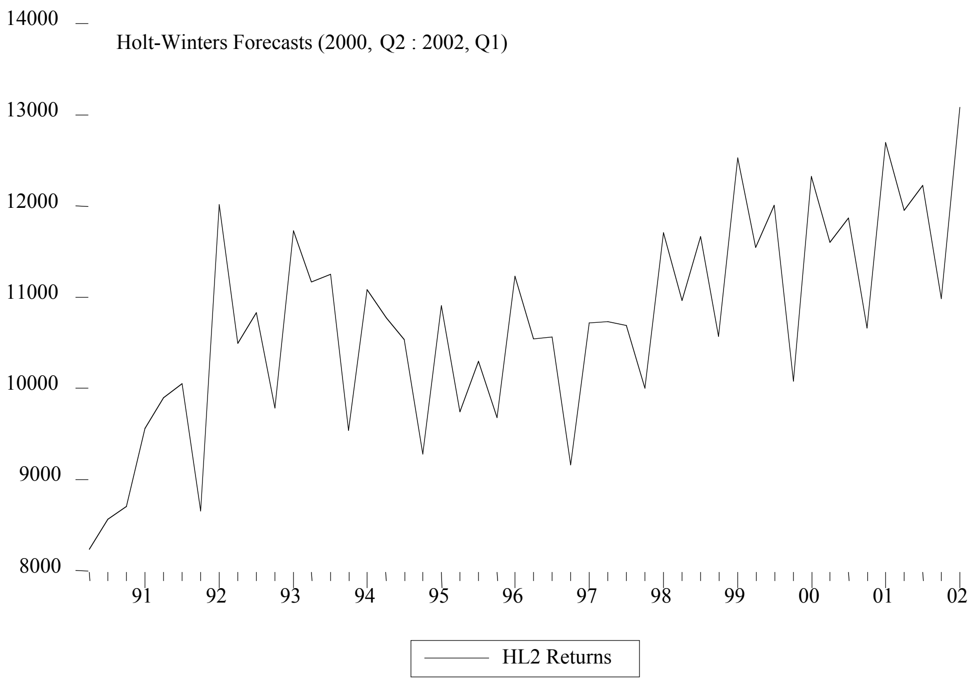


Figure 6.4 HL2 applications data with forecasts using Holt-Winters Exponential Smoothing



6.8 As can be seen in the graphs the Box-Jenkins forecasts imply that for the next 7 to 8 quarters there will be a slight downward trend in the number of HL2 applications. On the other, hand the Holt-Winters method suggests that HL2 returns will continue their upward trend. Unfortunately, it is impossible to say which forecast is most likely to be the most accurate! Rather, as emphasised above, they are presented to allow comparison between a model-based method and a non-model based method. It may well come to pass that neither of the forecasts presented above is 'accurate'. In addition it should be noted that forecasting beyond 4 points in time does increase the unreliability of the forecasts (see Enders, 1995). This is unfortunately an inherent weakness of any forecasting of social or economic phenomena, though the margin of error may vary from one data series to another.

6.9 One method that is commonly used amongst researchers/forecasters is to take a simple unweighted average of all the forecasts for each point in time (Hall, 1994). There is evidence to suggest that forecasts combined in this way are in fact more accurate than the individual forecasts (Chatfield, 2000). Using this method results in the following estimates[8]

Table 6.2 Unweighted average of homelessness forecasts

Period	Combined forecast
2000, Q2	11,324
2000, Q3	11,567
2000, Q4	10,257
2001, Q1	12,215
2001, Q2	11,372
2001, Q3	11,611
2001, Q4	10,289
2002, Q1	12,296

6.10 Figure 6.5 shows that, after combining the forecasts, homelessness levels will remain largely constant or may in fact decline slightly until the first quarter of 2002 which is the end of the forecasting period.

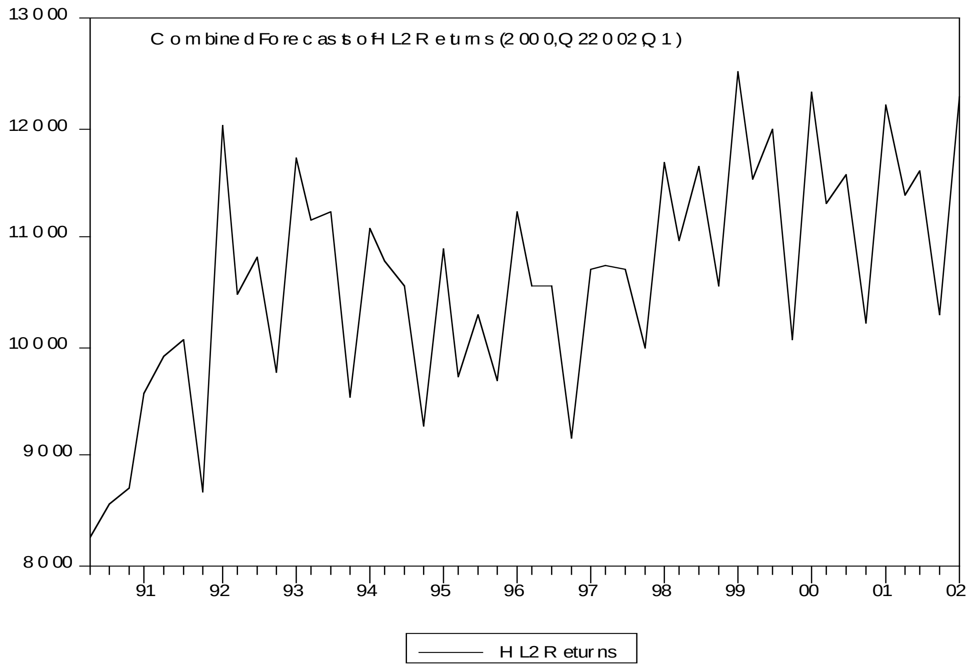
METHOD 2 ELASTICITIES APPROACH

6.11 Predicting what will happen to homelessness applications using the results from the cross section and pooled analysis is more problematic. This is because the data are not time series; they are typically snapshots, or a relatively short series of snapshots, of homelessness applications, acceptances and the structural variables observed at various points in time.

6.12 However, because the data was logged before running these regressions we are able to quantify to some degree what the percentage increase or decrease in homelessness applications/acceptances will be for a 1 per cent increase or decrease in one of the structural variables. To do this we need to use the *calculated coefficient* corresponding to the structural variable in which we are interested. Basically, the calculated figure is what economists call an 'elasticity', which is the degree of 'responsiveness' of one variable to a change in another. Typically, economists assume that if the elasticity exceeds the value one then this is termed as being 'elastic'. Values less than one are termed 'inelastic'.

[8] A simple unweighted average is only one method for combining forecasts. Diebold (2001) and Holden *et al* (1990) review other methods.

Figure 6.5 HL2 data with combined forecasts



6.13 Table 6.3 presents the results of what will happen to homelessness applications/acceptances for a 1 per cent increase in the hypothesised structural variable. It is simply Table A1 from Annex 1. The results should be interpreted as follows:

Homelessness applications (acceptances) will increase (decrease) by X per cent for each 1 per cent increase in the particular structural variable concerned.

6.14 So, within the table we can see that the coefficient on psychiatric inpatient discharges with any diagnosis of alcohol misuse alcohol discharges is 0.6969 or 0.7 after rounding. Thus, a 1 per cent increase in alcohol discharges will result in a 0.7 per cent *increase* in the homelessness level. Similarly, a 1 per cent increase in local authority vacancies will result in a *fall* in homelessness levels of 0.46 per cent. And a 1 per cent increase in unemployment among economically active people will result, other things being equal, in a 0.43 per cent increase in homelessness applications.

6.15 The coefficients from the other tables in Annex 2 can be read in a similar fashion. It is important to note that the precise coefficients for particular variables (unemployment, local authority vacancies, etc) vary somewhat from one year to the next. In drawing policy conclusions from these findings, therefore, it would be a mistake to assume that the precise elasticity for, say, unemployment in 1999, will apply with equal precision to 2001 or any other year. It can, however, be concluded that the unemployment and housing structural variables are the most elastic. In many of the regressions from 1991 onwards, a one per cent increase in these variables commonly shows an increase in homelessness approaching or even exceeding one per cent. Affordability measures also show some evidence of elasticity.

Table 6.3 Regression coefficients for homeless applications in 1981

Variable	Coefficient	Standard error	t-ratio	Probability
Constant	-.2950	1.431	-.206	.8382
VAC81	-.4626	.2485	-1.861	.0736
EA_EMP	.4399	.3229	1.363	.1843
LP81	.3182	.3166	1.005	.3239
ALCO	.6969	.1310	5.321	.0000

CONCLUSION

6.16 This final empirical section of the report has addressed the issue of forecasting homelessness into the future. The approach adopted reflected a number of methods available in practice to researchers. Initially, we presented a series of forecasts using the time series element of the data and concluded that homelessness figures will stabilise over the next 2 years and may even show a small decline. The second approach we adopted was to use the elasticities calculated from the regression equations. We found that, across all the cross section and pooled regressions, housing and unemployment would appear to impact most strongly on homelessness numbers; in effect, they were the most elastic, a result that appears to be quite robust. In practical terms, this means that policies aimed at tackling labour market and housing structural causes will have *proportionately* more effect on homelessness numbers than tackling other structural causes. This does not mean that the other causes are not important; our work has shown that indeed they are. Rather, if an objective is to reduce homelessness efficiently then the elasticities evidence suggests that resources should be directed particularly towards housing and the labour market.

CHAPTER SEVEN CONCLUSIONS

7.1 It is a widely held view within the academic research community in Britain that the growth of homelessness in the 1980s and 1990s was caused by a series of structural forces. However, with the important exception of Bramley (1993), to date the existence of a relationship between homelessness and these structural trends has been frequently asserted but not tested in Britain. In contrast, a number of researchers in the USA have attempted to examine statistically the structural 'causes' of homelessness.

7.2 In this report we have developed a model of the causes of homelessness in order to test the hypothesis that structural factors can, at least in part, 'explain' the level of homelessness in Scotland. We then used multivariate statistical techniques to test this model. We examined variations in homelessness between the different Scottish local authorities for 6 different points in time over the period from 1981 to 1999 (the cross-sectional analysis) and for groups of years taken together (the pooled cross-sectional analysis). In addition, we examined trends in homelessness in Scotland as a whole over the 19-year period from 1980 to 1998 (the time-series analysis).

7.3 The cross-sectional analysis provided some support for the view that the aggregate level of homelessness is affected by structural factors. In particular, we found that homelessness was positively related to ***unemployment***, which is in line with the structural model. Local authorities with relatively high levels of unemployment also tended to have high levels of homelessness and vice versa. In other words, (other things being equal) as unemployment increases, so too does homelessness. (In so far as unemployment contributes to family breakdown, it may also impact upon housing demand and some measures of 'deinstitutionalisation' such as the number of young people in local authority care.)

7.4 The cross-sectional analysis also provided support for the hypothesis that the level of homelessness is affected by the state of the ***housing market***, though the nature of the relationship was often not consistent with that implied by the structural model. For 1981, we found that homelessness was negatively related to the number of vacant properties available for letting. This finding is consistent with the structural model (and the US research findings). It implies that homelessness is higher where local authority vacancies are lower. Expressed differently, 'tight' housing markets have more homelessness than 'slack' ones.

7.5 However, by the 1990s, the relationship between local authority vacancies and homelessness was found to be the reverse of that posited in the model and found for 1981. Homelessness was now positively related to vacancies. In other words (other things being equal) the more empty dwellings, the higher the level of homelessness. This relationship was found for all applications and all acceptances in 1996, 1997 and 1998. It was also found separately for applications by young people and by lone parents for each of these 3 years. For applications and acceptances in 1991 and applications in 1996, this positive relationship was not significant and may therefore be discounted. However, for the other categories and years, it was significant.

7.6 It is not immediately apparent why there should be a positive relationship between local authority vacancies and the level of homelessness. It may be that vacancies are in fact an 'intervening variable' that is related to a factor not included in our data. That is to say, the relationship may reflect a structural trend, the effects of which were concentrated in local authorities with high vacancies. For example, it is possible to speculate that a trend towards less stable households – due to increased relationship breakdown, lone parenthood and step parenting – resulted in higher levels of homelessness and that this was concentrated in local authorities that had a high level of vacancies. This trend in household instability may have coincided with the growth of low demand housing and neighbourhoods during the 1980s and 1990s. However, this is speculation and the relationship between vacancies, family breakdown and homelessness is clearly something that deserves further exploration.

7.7 The statistical analysis also found support for the hypothesis that the level of homelessness is affected by the ***affordability of housing***. In general, we found that homelessness applications and acceptances were both positively related to the level of house prices. In other words (other things being equal) local authorities with high house prices also tended to have high levels of homelessness and vice versa. Homelessness applications were also positively related to court actions for possession by mortgage lenders. Areas with high levels of these court actions also tended to have high levels of homelessness (and vice versa). Likewise, there was a positive relationship between private rents on unfurnished lettings and homelessness. Other things being equal, the higher the rent, the higher the level of homelessness. This is consistent with the US research on the causes of homelessness, which has found a similar relationship between private sector rent levels and homeless rates.

7.8 However, we found an inverse relationship between local authority rents and the level of homelessness. That is to say, other things being equal, the lower the council rent, the higher the level of homelessness. As with local authority vacancies, it is not clear why this inverse relationship between local authority rent levels and homelessness should exist, but it deserves further exploration.

7.9 The analysis revealed a number of statistically significant relationships between homelessness and various measures of '***de-institutionalisation***', though the results were not fully consistent. The lack of data for prison discharges and for people leaving the armed forces meant that the analysis was largely focused on hospital discharges. For 1981 and 1991, we found a positive relationship between homelessness and the number of psychiatric inpatient discharges of people with a diagnosis of alcohol misuse. In other words, the greater the number of such discharges, the higher the level of homelessness in the local authority (and vice versa).

7.10 By 1996 and 1997 the relationship had reversed. For these 2 years, the relationship was an inverse one in that homelessness was now lower in areas with high levels of discharge (and vice versa). It might be tempting to conclude from this that local authorities had by this time adjusted their homelessness practices to provide accommodation for inpatients with a history of alcohol misuse who were being discharged from psychiatric wards and hospitals. However, in 1998 - the last year for which the data was available - there was a positive relationship between the 2 (except for applications by young people, which continued to have a negative sign). In 1999, there was a positive relationship between the number of inpatient discharges from long-stay hospitals and the level of homelessness (all applications and acceptances as well as applications by lone parents and young people).

7.11 The results from the pooled, cross-section statistical analysis were to some extent inconsistent and must therefore be treated with caution. More reliance can be placed on the pooled data for the 10-year period from 1989 to 1998 than on the 5-year data because the former is based on a much larger number of data observations. Analysis of the 10-year pooled data indicated that the level of homelessness acceptances was positively related to unemployment and house prices. It also found that homelessness applications and acceptances were both strongly related to the number of drug offences recorded by the Scottish police service.

7.12 The time-series analysis of trends in homelessness applications over the last 2 decades also found support for a structural explanation of the level of homelessness. We found a long-run relationship between homelessness and both the unemployment rate and the level of employment in manufacturing. We also found a long-run statistical relationship between homelessness and 2 selected housing market variables: right to buy sales and the number of public sector lettings. There was also a similar relationship between homelessness and the number of tenants in rent arrears. Finally, there was a long-run relationship between homelessness and the number of recorded crimes.

7.13 Taking the results of the various analyses together and bearing in mind the limitations of the data, the results lend support for the hypothesis that there are structural causes of homelessness. At the very least, it can be concluded that the results do not provide sufficient evidence to reject the model. In other words, the level of homelessness over the past 2 decades probably has been affected by wider structural trends in society. Unemployment, and to a lesser extent housing affordability and de-institutionalisation, appear to be powerful forces affecting the incidence of homelessness in Scotland.

7.14 Thus, while behavioural factors may be important in explaining *individual cases* of homelessness, the results of our analysis indicate that structural trends do affect the *aggregate level* of homelessness. The evidence suggests that this is the case, not only for acceptances (the number of people assessed as homeless), but also for applications (the number of people submitting applications for re-housing on the grounds of homelessness). It also appears to be the case for applications by young people and lone parents.

POLICY IMPLICATIONS

7.15 This conclusion has important implications for policy and practice. It implies that policy responses that focus on the problems and needs of individual homeless people need to be complemented by action that addresses the structural causes of homelessness if the size of the problem is to be substantially reduced.

7.16 Of the 4 structural categories, it was unemployment that most consistently emerged as having the clearest structural relationship with homelessness. This implies that success in tackling unemployment will have a significant impact on the level of homelessness in Scotland. In this respect the need may be not just for action to tackle unemployment in general, but also for measures that are targeted on localities with high levels of unemployment where homelessness is also high, such as Glasgow. But since there are inevitably limits to the extent to which the level of unemployment can be reduced by government action, it also implies that the degree to which homelessness can be reduced by such action is also highly constrained. It is necessary, therefore, to consider additional avenues to a reduction in the aggregate level of homelessness.

7.17 The positive (rather than the expected, inverse) relationship between homelessness and local authority vacancies (1991, 1996-98) and new build completions by social housing landlords (1999) suggests that raising the level of social housing construction is an ambiguous solution, not least because of the problem of low demand. The complex problem of low demand for social housing is well beyond the scope of this research and its interconnection with homeless is something that merits further investigation.

7.18 However, the cross section regression results for 1999 and the time series analysis indicate that there is a significant relationship between homelessness and some measures of housing demand (the number of lone parent households and young people under 25 years) and supply (right to buy sales). Policy makers have little influence over the number of young adults or lone parents, but they can influence the level of right to buy sales. Given the finding on vacancies, this relationship may be due less to the fact that the total supply of low cost rental dwellings is reduced as a result of the right to buy, but rather to the fact that the supply of the more popular social housing is reduced. However, this is speculation and takes the discussion to the limits of what can be said on the basis of this research project.

7.19 So far as affordability is concerned, the unexpected inverse relationship between local authority rents and homelessness may possibly reflect the same factors that produced the positive relationship between homelessness and vacancies. However, the positive relationships between homelessness and mortgage arrears, house prices and private rents are in the direction that would be expected from the structural model. This suggests that action to improve the affordability of private owned and rental accommodation could help to reduce the level of homelessness. Such action would almost certainly have public expenditure implications and may touch upon policy matters that are reserved for the UK Parliament rather than within the power of the Scottish Executive.

7.20 Finally, the regression analysis showed that there exists a statistical relationship between homelessness and various measures of de-institutionalisation, but that the nature of the relationship was not always consistent. Many of the variables in this category are ones that are susceptible to influence by policy and practice, for example, school exclusions, the number of children in care, and the number of psychiatric patients being discharged from institutions. Moreover, the latter is something to which local authority allocation and homelessness policies and practices can respond. For example, if social landlords work closely with health trusts to ensure that patients being discharged from long-stay institutions are re-housed *and* provided with the support they need to remain in the accommodation, homelessness applications from ex-patients will be lower than if they fail to do so. Thus, there is a feedback loop between the dependent and independent variables in this instance. This appears to be an area where better co-ordination of policy and practice could help to reduce the aggregate level of homelessness.

CONCLUSION

7.21 In conclusion, the analysis reported here suggests that the trend in homelessness over the last two decades, and variation in homelessness between local authorities, in Scotland can to a significant extent be explained by structural factors. It follows that, while action to tackle the difficulties or failings of homeless people themselves may help to prevent or alleviate homeless in individual cases, it is also important address the wider structural factors. The analysis also suggests that these structural factors are not confined to issues of housing supply and demand or of affordability, but also relate to unemployment and what we have termed de-institutionalisation. It follows that policy responses should include action to tackle unemployment and the housing and support needs of people discharged from institutions, and not just the supply of adequate and affordable housing.

REFERENCES

Anderson, I. (1999) 'Young People and Access to Social Housing', in J. Rugg (ed.) *Young People, Housing and Social Policy*, London: Routledge.

Anderson, I., Kemp, P. A. and Quilgars, D. (1993) *Single Homeless People*, London: HMSO.

Anderson, I. and Tulloch, D. (2000) *Pathways Through Homelessness: A Review of the Research Evidence*, Edinburgh: Scottish Homes.

Biehal, N. and Wade, J. (1999) '"I Thought It Would be Easier": The Early Housing Careers of People Leaving Care', in J. Rugg (ed.) *Young People, Housing and Social Policy*, London: Routledge.

Box, G. E. P and Jenkins G. M (1976) *Time Series Analysis, Forecasting and Control*, revised ed., San Francisco: Holden Day.

Bradshaw, J. (1972) 'The Concept of Social Need', *New Society*, 30 March.

Bramley, G. (1993) 'Explaining the Incidence of Statutory Homelessness in England', *Housing Studies*, Vol. 8, pp 128-147.

Burt, M. R. (1991) 'Causes of the Growth of Homelessness During the 1980s', *Housing Policy Debates*, Vol. 2, pp 903-936.

Burt, M. R. (1992) *Over the Edge: The Growth of Homelessness in the 1980s*, Washington DC: The Urban Institute Press.

Chatfield, C. (2001) *Time Series Forecasting*, Boca Raton: Chapman and Hall/CRC.

Chatfield, C. (1996) *The Analysis of Time Series*, London: Chapman and Hall.

Clapham, D., Kemp, P. and Smith, S. J. (1990) *Housing and Social Policy*, Basingstoke: Macmillan.

Corden, A. (1995) *Changing Perspectives on Benefit Take-up*, London: HMSO.

Craig, P. (1991) 'Costs and Benefits: A Review of Research on Take-up of Income-Related Benefits', *Journal of Social Policy*, Vol. 20, pp 537-566.

Crane, M. (1999) *Understanding Older Homeless People*, Buckingham: Open University Press.

Dant, T. and Deacon, A. (1989) *Hostels to Homes*, Aldershot: Avebury.

Dickey, D. and Fuller, W. A. (1979) 'Distribution of the Estimates for Autoregressive Time Series with a Unit Root', *Journal of the American Statistical Association*, Vol. 74, pp 427-431.

Dickey, D. and Fuller, W. A. (1979) 'Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root', *Econometrica,* Vol. 49, pp 1057-1072.

Diebold, F. (2001) *Elements of Forecasting,* 2nd Ed, Cincinnati: South Western.

Elliott, M. and Krivo, L. .J. (1991) 'Structural Determinants of Homelessness in the United States', *Social Problems*, Vol. 38, pp 113-131.

Enders, W. (1995) *Applied Econometric Time Series*, Wiley.

Engle, R. F. and Granger, C. W. J. (1987) 'Cointegration, Error Correction: Representation, Estimation, and Testing,' *Econometrica*, Vol. 55, pp 251-276.

Fitzpatrick, S., Kemp, P. A. and Klinker, S. (2000) *Single Homelessness: An Overview of Research in Britain*, Bristol: The Policy Press.

Franses, P. H. (1998) *Time Series Models for Business and Economic Forecasting*, Cambridge: University Press.

Goldberger, A. S. (1991) *A Course in Econometrics*, Cambridge, Massachusetts: Harvard University Press.

Granger, C. W. J. and Newbold, P. (1986) *Forecasting Economic Time Series*, San Diego: Academic Press.

Granger, C. W. J. and Newbold, P. (1986) 'Spurious Regressions in Econometrics', *Journal of Econometrics*, Vol. 2, pp 111-120.

Green, R. K. and Malpezzi, S. (n.d.) *A Primer on US Housing Markets and Housing Policies*, City?: American Real Estate and Urban Economics Association.

Greene, W. (2000) *Econometric Analysis*, New Jersey: Prentice Hall.

Greve, J. (1990) *Homelessness in Britain*, York: Joseph Rowntree Foundation.

Gujarati, D. N. (1995) *Basic Econometrics*, Singapore: McGraw Hill.

Hall, S. G. (1994) *Applied Economic Forecasting Techniques,* Harvester Wheatsheaf.

Harvey, A. C. (1990) *The Econometric Analysis of Time Series*, Cambridge, Massachusetts: MICT Press.

Hausman, J. (1978) 'Specification Tests in Econometrics', *Econometrica*, Vol. 46, pp 1251-1271.

Holden, K., Peel, D. A. and Thompson, J. L. (1990) *Economic Forecasting: An Introduction*, CUP.

Honig, M. and Filer, R. K. (1993) 'Causes of Intercity Variation in Homelessness', *The American Economic Review*, Vol. 83, pp 248-255.

Hsiao, C. (1986) *Analysis of Panel Data*, New York: Cambridge University Press.

Johanson, S. (1988) 'Statistical Analysis of Cointegration Vectors', *Journal of Economic Dynamics and Control*, Vol. 12, pp 231-54.

Johnson, J. and Di Nardo, J. (1997) *Econometric Methods*, Singapore: McGraw-Hill.

Kemp, P. A. and Rhodes, D. (1994) *Private Landlords in Scotland*, Edinburgh: Scottish Homes.

Koop, G. (2000) *Analysis of Economic Data*, Chichester: Wiley.

Kuh, E. and Meyer, J. R. (1955) 'Correlation and Regression Estimates When the Data Are Ratios', *Econometrica*, Vol. 23, pp 400-416.

Maddala, G. S. (1989) *Introduction to Econometrics*, New York: Maxwell Macmillan International Editions.

Main, T. (1996) 'Analyzing Evidence for the Structural Theory of Homelessness', *Journal of Urban Affairs*, Vol. 18, pp 449-457.

Main, T. (1998) 'How to Think About Homelessness: Balancing Structural and Individual Causes', *Journal of Social Distress and the Homeless*, Vol. 7, pp 41-54.

Makridakis, S Wheelwright, S.C and Hyndman, R.J (1998) *Forecasting: Methods and Applications,* New York: Wiley

McCleary, R. and Hay, R. A. (1980) *Applied Time Series for the Social Sciences*, SAGE.

McKinnon, J. G. (1991) 'Critical Values for Cointegration Tests', in R. Engle and C. W. J. Granger (eds.) *Long Run Economic Relationships*, OUP.

Neale, J. (1997) 'Homelessness and Theory Reconsidered', *Housing Studies*, Vol. 12, pp 47-61.

Niemira, M. P. and Klein, P. A. (1994) *Forecasting Financial and Economic Cycles*, New York: Wiley.

Pawson, H. (2000) *A Profile of Homelessness in Scotland*, Edinburgh: Scottish Homes.

Pawson, H. and Bramley, G. (2000) 'Would You Live Here?', *Roof*, July/August, pp 26-28.

Pleace, N. (1998) 'Single Homelessness as Social Exclusion', *Social Policy and Administration*, Vol. 32, pp 46-59.

Pindyck, R. S. and Rubinfeld, D. L. (1998) *Econometric Models and Economic Forecasts*, Singapore: McGraw-Hill.

Quigley, J. M. and Raphael, S. (2000) 'The Economics of Homelessness: The Evidence from North America', paper presented at the AREUEA/EHNR Conference, Gavle, June.

Quigley, J. M., Raphael, S. and Smolensky, E. (2000) *Homeless in America, Homeless in California*, Program on Housing and Urban Policy Working Paper No. W99-001, Berkeley: University of California.

Quilgars, D. and Pleace, N. (1999) 'Housing and Support Services for Young People', in J. Rugg (ed.) *Young People, Housing and Social Policy*, London: Routledge.

Rugg, J. (1999) 'The Use and "Abuse" of Private Renting and Help with Rental Costs', in J. Rugg (ed.) *Young People, Housing and Social Policy*, London: Routledge.

Stock, J. and Watson, M. (1988) 'Testing for Common Trends', *Journal of the American Statistical Association,* Vol. 83, pp 1097-1107

Third, H. and Yaneta, A. (2000) *Homelessness in Scotland: A Summary of Research Evidence*, Edinburgh: Scottish Homes.

Thomas, R.. L. (1997) *Modern Econometrics: An Introduction*, Harlow: Addison-Wesley.

Webb, S. (1994) *My Address is Not My Home: Hidden Homelessness and Single Women in Scotland*, Edinburgh: Scottish Council for Single Homeless and Glasgow Council for Single Homeless.

Williams, F. and Monge, P. (2001) *Reasoning with Statistics How to Read Quantitative Research,* Orlando, Florida: Harcourt.

Wolch, J. R., Dear, M. and Akita, A. (1988) 'Explaining Homelessness', *Journal of the American Planning Association*, Vol.54, pp 443-453.

Vincent, J., Deacon, A. and Walker, R. (1995) *Homeless Single Men: Roads to Resettlement?*, Aldershot: Avebury.

Yule, G.U. (1926) 'Why Do We Sometimes Get Nonsense Correlations Between Time-Series?', *Journal of the Royal Statistical Society*, Series A, Vol. 89, pp 1-69.

ANNEX 1: DESCRIPTION OF THE VARIABLES USED IN THE STATISTICAL ANALYSES

Structural categories	Variable name	Definitions	Sources
Housing market	VAC	Total vacant household spaces (all types of occupancy). This includes property that is new and ready for occupation but not yet occupied, and also property that is clearly without furniture, or is not occupied, for example, because new tenants were awaited, or that the occupier was deceased. It excludes those properties which are in the course of conversion or improvement (renovation or decoration) and are not occupied.	Census
	OVERCR	Total households with more than 1.5 person per room. The number of persons per room is the ratio of the number of persons in a private household with usual residents to the number of rooms in the accommodation of that household	Census
	LP	Total households with one adult and one or more dependent children (a father or mother together with his or her dependent child(ren)). Dependent children are defined as persons aged 0-15 ; or persons aged 16-18, never married, in full-time education and economically inactive.	Census
	VACANT	Vacant Local Authority stock that is available to let. Excludes stock that is awaiting or undergoing improvement or repair, held to decant households during repair, acquired for and awaiting demolition, dwellings used as temporary accommodation for homeless, dwellings to be sold or demolished within 2 years, dwellings under modernisation or major repairs and dwellings in low demand areas	Scottish Executive
	NEWBUILD	New build completions by social landlords. A dwelling is completed when it is ready for occupation, whether in fact occupied or not. If a dwelling is transferred to another agency after completion, it is considered to have been completed by the first agency.	Scottish Executive
	R2B	The number of sales to sitting tenants in each local authority. Includes right-to-buy, rent-to-mortgage and voluntary sales.	Scottish Executive
	<25	Number of under 25's derived from mid-year population estimates	General Registrars Office
	LONE-PARENT	Number of unmarried parents (joint registration with different addresses, or sole registration)	General Registrars Office
	PVTHSN	Private rented sector housing stock (from estimated stock of dwellings by tenure). Includes stock privately rented, and rented with a job or business. Estimates of the total dwelling stock are based on the 1981 and 1991 Census. It includes a count of the number of dwellings. The tenure was derived using a method which combined information from the Census, the Post Census Survey of Vacant Dwellings, and public authorities' counts of their own stock. This baseline figure is updated using information on new housebuilding, conversion of property to housing use and demolitions, collected on returns submitted to the Scottish Office as well as public authorities' counts of their own stock and Scottish Homes count of housing association stock.	Scottish Executive
	SOCHOUS	Social rented sector housing stock (from estimated stock of dwellings by tenure). Includes stock rented from housing associations, local authorities, new towns and Scottish Homes.	Scottish Executive

Unemployment	EA_EMP	Number of economically active people not in employment. Economically active includes persons in employment (full and part-time employees, self employed with and without employees, and persons on a government scheme), unemployed (those waiting to start a job, or those seeking work) and students (those included above).	Census
	CLCOUNT	Number of claimant unemployed. Claimant count measures those claiming unemployment-related benefits (unemployment benefit, income support and national insurance credits) at Employment Service local area offices. This data is provided by the Employment Department.	Regional Trends
	NOMISLT	Number of claimant unemployed who have been unemployed for more than 6 months. This is unemployment data derived from the Office for National Statistics' (ONS) administrative counts of unemployment. Claimant count measures those claiming unemployment-related benefits (unemployment benefit, income support and national insurance credits) at Employment Service local area offices.	Nomis
	NOMISCC	Number of claimant unemployed. This is unemployment data derived from the Office for National Statistics' (ONS) administrative counts of unemployment. Claimant count measures those claiming unemployment-related benefits (unemployment benefit, income support and national insurance credits) at Employment Service local area offices.	Nomis
	ILOUN16	Number of people over 16 who are unemployed. The ILO unemployment measure includes people who at the time of interview: were without work; were available to start work within the next 2 weeks; had either looked for work in the 4 weeks before interview, or were waiting to start a job.	Labour Force Survey
	URATE	Unemployment Rate. Uses claimant count as the measure of unemployment. The denominator used in calculating the unemployment rate is the appropriate mid-year estimates of total employees in employment, unemployed, self employed and HM Forces and participants in work-related Government training schemes.	Regional Trends
	LMANSC	Index of Manufacturing Employment in Scotland.	Northern Ireland Economic Research Council
Affordability	LARENT	Local Authority average weekly standard rent. These figures are based on HRA returns from local authorities.	Scottish Executive
	HPRICES	Average house prices of unitary authority areas in Scotland	Scottish Housing Review
	HBPRCL	Number of Housing Benefit claims made by private tenants. Includes new claims and renewal claims.	Department of Social Security
	HBLACL	Number of Housing Benefit claims made by Local Authority tenants. Includes new claims and renewal claims.	Department of Social Security
	UNF2BED	Average private rent for an unfurnished 2 bedroom flat at unitary authority level. The value is based on an unfurnished 2 bedroom self-contained flat in the most representative private rented sector locality in the local authority area.	University of York Rent Index
	MORTARR	Number of actions in court for repossession of property (number of actions before the courts at the instance of mortgage lenders seeking recovery of property).	Scottish Executive
	ROREDUC	Rent officer statistics. Calculation used is the difference between mean reference rent and local reference rent. Local reference rent is calculated using the midpoint of the range of rents (excluding exceptionally high and exceptionally low rents) for properties of the same size in the same locality. The objective is to limit housing benefit to the general level of rents for accommodation of the same size (number of rooms) in the locality.	Scottish Executive

	RORATIO	Rent officer statistics. Calculation used is the mean reference rent as a ratio of local reference rent. See above for details on how local reference rent is calculated.	Scottish Executive
	RENTARRS	Total number of decrees granted for evictions. This includes decrees granted due to rent arrears.	Scottish Executive
Deinstitutionalisation	PSYCH	Number of persons present in psychiatric establishments. Includes psychiatric NHS and non-NHS hospitals and homes.	Census
	PRISN	Number of persons present in prison service establishments. Includes prisons, detention centres and young offender institutions. Excludes approved schools, ex-offenders hostels, probation and remand homes, police stations with a lock up, and special hospitals.	Census
	ALCO	Number of psychiatric inpatient discharges with any diagnosis of alcohol misuse (Discharge episodes include multiple discharges for individual patients)	Information and statistics division (NHS)
	CRIMES	Number of crimes and offences recorded by each unitary local authority. (Crimes include non-sexual crimes of violence, crimes of indecency, crimes of dishonesty, fire-raising, vandalism etc, other crimes (crimes against public justice, drugs, other). Offences include miscellaneous offences (petty assault, breach of the peace, drunkenness) and motor vehicle offences)	Scottish Executive
	SCHEXCL	Average number of half days of temporary exclusions per 100 school pupils. Includes both primary and secondary schools.	Scottish Executive
	INPDSCH	Number of inpatient discharges from geriatric long-stay hospitals by local council area	Information and statistics division (NHS)
	DRGCNVC	Number of drugs related offences recorded by sheriff courts in Scotland	Scottish Executive
	CHLDCAR	Number of children in care, under supervision or children looked after by each local authority	Scottish Executive
	PSYDSCH	Number of mental illness hospitals & psychiatric units inpatient discharges by local council area	Information and statistics division (NHS)

ANNEX 2: STATISTICAL METHODS

This annex will outline in more detail than Chapter Four the various methods and issues involved in carrying out any quantitative or statistical analysis such as that conducted in this report. Its purpose is *not* to replicate material that can be found in any good statistical or econometric text- rather it is to provide more detail as to how the research team conducted the analysis and dealt with the various intricacies that an analysis of this kind inevitably throws up. This means that a certain amount of statistical knowledge is assumed on the part of the reader.

1. CROSS SECTION ANALYSIS

As stated in Chapter Four the main statistical tool employed for much of the analysis was regression analysis. This is a method for determining the importance of a number (or group) of variables, called explanatory variables, on the variable whose values we wish to 'explain', called the dependent variable. It is commonly summarised in an equation form as:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \qquad (1)$$

Where is a constant, Y is the dependent variable, X_1 to X_4 are the independent variables and ε is an error term. The regression is estimated using a method known as Ordinary Least Squares, which minimises the sum of squared residuals between the actual observation and the predicted observation (see Williams and Monge, 2001, for details).

Our sample consisted of 32 Unitary Local Authorities with various structural and other data collected in each. Most researchers would agree that a sample of 32 is quite ambitious for the analysis we proposed to conduct. This was further hampered by the quality of the data; in particular, the fact that the dependent variable, homelessness numbers, was not a pure measure of the homelessness problem in that repeat presentations could not be screened out. Thus, our dependent variable was upwardly biased which will inevitably have consequences for the reliability of the analysis. Moreover, conversations with key players in the homelessness field revealed that homelessness data may not have been recorded in a consistent manner across all ULAs over the last 20 years. It was felt that 1990s data was probably the most reliable.

These problems could not be tackled adequately from a statistical point of view and these issues need to be borne in mind when interpreting the results of Chapter Five. In addition to the practical difficulties outlined in the previous paragraphs some statistical problems were also encountered during the analysis, though they were not unanticipated.

Two common statistical problems

Any regression analysis is prone to the twin problems of multicollinearity and heteroskedasticity.

Multicollinearity

From the objectives of this project the most serious of these is 'multi-collinearity'. This particular problem arises when the variables included in regression are 'related' to each other or highly collinear. Basically, one variable is a substitute for another; for example including the variable 'age' and say, a person's work experience in a regression to explain an individual's income would result in a possible multi-collinearity problem. This is because age is closely 'related' to, or correlated with, a person's work experience. i.e age *and* work experience are fulfilling the role that age *or* work experience should be doing. The consequences for the regression analysis is that the researcher *may* be unable to determine with any degree of certainty which of the variables is influencing income or more relevantly in our case, the number of homelessness applications. Unfortunately, nothing can be done to rectify or 'adjust' for multicollinearity, although it is possible to discern from regression analysis output whether multi-collinearity is presenting a problem or not.

Multicollinearity manifests itself through large standard errors of the coefficients in a regression (Thomas, 1998). This then implies that we will have low confidence in the values of the coefficients computed by the software since a large standard error implies considerable variation in the coefficient's value.

Large standard errors also imply a low value of the t statistic (the main determinant of whether a variable is important or not) since a t statistic is merely the estimated coefficient divided by its own standard error. If the standard error is large relative to the coefficient then small values of "t" will be the outcome resulting in the researcher not being able to adequately assess whether a variable is indeed significantly different from 0 or not. Thomas (1998) gives the following advice

> *A high degree of multicollinearity, then, may have an adverse effect on regression results, but this is by no means inevitable. The implication is that, if regression equations have low estimated standard errors and high t ratios, we should not spend too much time worrying about any multicollinearity that might be present.*

Fortunately, the results presented in the main body of the text did have low standard errors and high t ratios so we can be fairly confident that the problems associated with multicollinearity did not arise. Where it did occur during the modeling process, a number of variable(s) were substituted for the original variables such that any multicollinearity previously present was mitigated. Fortunately, the results we are able to present are free of multicollinearity.

Heteroskedasticity

This particular problem is less intuitive to understand than that of multicollinearity. In its simplest sense it occurs when there is a large variation in the values of the explanatory variables in a regression analysis (see Thomas, 1998, for a more detailed discussion). Given that our units of analysis are Scottish ULAs ranging in size from the City of Glasgow to Clackmannanshire, we were not surprised to encounter this particular statistical problem. Again, heteroskedasticity manifests itself through biased standard errors which may or may not be large and thus the calculated t statistics may not be accurate; i.e they too will be biased. This could mean that we may falsely conclude that a variable is insignificant when (in the absence of heteroskedasticity) it may actually be significant.

However, it is possible to request a statistical correction or adjustment to the regression output to allow for the fact that our sample contains some large as well as small values of key variables. This correction is handled automatically by the software and precise statistical details as to the method of correction can be found in Thomas (1998), Greene (2000), or Johnston and Di Nardo (1997) amongst others. Having applied the correction the researcher can then make a better and more informed decision as to whether to accept or reject the hypothesis of a certain variable having a significant role in the determination of the number of homeless individuals. Given that our data has considerable variation in its values, all our regression models were corrected for the influence of heteroskedasticity; thus we can be confident that the t-statistics quoted are an accurate assessment of the significance of the independent variables in the regression equation (1) above. It is worth stressing that the actual coefficients themselves do not change. They are unbiased estimate of the true population coefficients.

Assessing the goodness of fit of the regression model

As part of the regression analysis all statistical software reports a measure of the overall 'goodness of fit' of the regression 'model'. Depending on the package used this can range from the simple coefficient of multiple determination, R^2, to log likelihood values and Chi^2 statistics. The package we used, *LIMDEP* v7.0 for Windows, outputs a fairly comprehensive list of regression diagnostics. The statistics we opted to use for assessing the regression models were the Adjusted R^2, (which measures how much variation is accounted for by the 'model') and the F statistic (which assesses the overall significance of the regression). The adjusted R^2 is very similar to the simple R^2 - the difference being that the adjusted R^2 takes into account the number of coefficients in the regression. The simple R^2 does not and so it is possible to conclude that the model fits the data better than it actually does. This is because adding variables to the right hand side of a regression equation will eventually result in the R^2 value tending to its limiting value, 1, indicating a perfect fit. This will not happen with the adjusted R^2 as the number of explanatory variables is explicitly taken into account when it is calculated (see Thomas, 1998, and Greene, 2000 for details). It should be borne in mind that the adjusted R^2 is only one aspect of assessing a 'model'. While it is desirable that this measure is high it is perfectly feasible that the coefficients may have the incorrect sign or even be insignificant. This would suggest revisiting the theory and if necessary revising the hypotheses on which the model was estimated.

The F statistic is a test of the significance of the regression line based on an analysis of variance. Again, calculated along with the value of the F statistic is a probability value indicating how significant the computed F value is. The lower the probability value then the more significant the regression. Thus, in assessing our regression results we looked for a high adjusted R^2 and a significant F statistic.

These criteria also held for the pooled analysis as this is essentially regression analysis applied to a series of repeated observations for each unit of interest (ULAs in this case) over a number of years.

2. POOLED CROSS SECTION ANALYSIS

Pooled cross section regression analysis (sometimes referred to as 'panel analysis') is an area of statistical/econometric analysis that has expanded rapidly in recent years leading to major advances in theory and estimation. This is because, for most applied areas of research, data is sometimes hard to come by and very often we only have 4-5 years worth of data which is inadequate for a time series analysis. Additionally, simple cross section regression cannot incorporate the effects of time into the estimated coefficients. Thus, pooling enables those 4 or 5 years of data to be stacked into one data set and then allows (depending on the research question in mind) the application of some quite sophisticated regression techniques to exploit its information content whilst incorporating the effects of time and other statistical 'quirks'.

Pindyck and Rubinfeld (1998) suggest 3 basic ways in which one can analyse pooled data.

1. Simple regression analysis on the pooled data set
2. Fixed-Effects Model
3. Random Effects Model

The first method of analysis is a simple regression analysis on the pooled data using a dependent variable and various independent or explanatory variables, as was the case with the simple cross section analysis. The consequences of this are that we are not allowing for the effect of *time* on homelessness nor for any *variation in homelessness across ULAs*[9]. Essentially, we would be ignoring those factors that panel data analysis was expressly designed to investigate.

Consequently, a more efficient method of analysis, which explicitly takes into account the possibility that homelessness will vary across time and ULAs, is a *fixed effects model*. This sounds more complicated than it actually is. Essentially, the same regression is estimated as in the case of a simple pooled analysis with the exception that we include additional terms, called dummy variables, to take into account time and in this case ULA wide variations in homelessness numbers. This technique, although more efficient than the simple method outlined earlier, has a number of drawbacks. The inclusion of a large number of time dummy variables and ULA dummy variables will reduce the number of degrees of freedom by a substantial amount. Additionally, this technique does not *identify* what causes homelessness to vary across time and ULAs; it just allows for the fact that there probably will be variation. Pyndick and Rubinfeld (1998) note that to overcome potential degrees of freedom problems most researchers omit dummy variables for time but do include them for the cross section units, i.e ULAs in this case.

The fixed effects model 'assumes' that we are able to disentangle the effects of time and units (i.e., ULA); in other words the error term in the regression specification is un-correlated across time and ULAs (Johnston and Di Nardo, 1998).

This may or may not be appropriate. In the case where it is not appropriate i.e where we suspect that there is some correlation over time *and* ULAs (called cross section units) then we have to specify the *random effects* model. Again, a regression is estimated but because we are allowing for correlation over time and units one of the assumptions of classical regression

[9] In effect this is tantamount to assuming that the intercept and slope terms in the regression equation (1) do not vary across time or ULAs.

analysis (see Thomas, 1998, for a comprehensive discussion) has been violated. Thus, the method of computation is different from that previously employed. OLS will no longer provide correct estimates of the standard errors and therefore t statistics. Instead a technique known as Generalised Least Squares is used and again most econometric/statistical software packages incorporate this estimator within their specification.

Having outlined the 3 approaches to analysis panel data an obvious question to ask is which method do we use? Are there advantages/disadvantages to choosing one method over the other? Fixed effects has the disadvantage of using up a large number of degrees of freedom but it does allow the researcher to investigate whether there are significant differences between units. The random effects model cannot do this but it has the advantage of using up fewer degrees of freedom as dummy variables for each unit and year are not included in the regression analysis. In addition, intuitively speaking, we would expect some correlation over time and units. Pyndick and Rubinfeld (1998) and Greene (2000) list other pros and cons of the various methods.

How then does one choose between fixed or random effects? Fortunately, it is possible to test whether fixed or random effects is the appropriate method for the data (and problem) in hand using what is called the Hausman Specification test (Hausman, 1978).

For the analysis conducted within this report we stacked the data and opted for a fixed effects model approach to control for variations in homelessness figures across ULAs. We did not control for time as this would have used up too many degrees of freedom. Our approach, then, was what is called a ***cross section fixed effects*** approach, the most popular in econometric literature, as we suspected that homelessness would vary quite systematically (not randomly) across ULA's.

Fortunately, the LIMDEP for Windows software has some very advanced panel data analysis features and estimates, by default, the 3 approaches outlined above as well as conducting a Hausman Specification (HS) test. In all cases the (HS) test result suggested that the cross section fixed effects model was the most appropriate one that described the relationship between the level of homelessness and the hypothesized structural variables[10].

3. TIME SERIES ANALYSIS

In one sense time series analysis is less complicated than that of panel data analysis but vast strides have been made in time series techniques in the last decade especially with regard to multivariate time series analysis. We employed these latest techniques in our analysis of homelessness in this report[11].

Prior to the more complicated modeling process our initial line of inquiry was to investigate the 'stationarity' properties of the series in order to determine their order of integration. This is investigated through unit root tests. To neglect to do so would have ramifications for the more involved analysis later on in the modeling procedure.

[10] The statistical/econometric theory relating to the field of panel data analysis is complicated to say the least. Hsiao (1986) is the authoritative text in this field.
[11] For good introductory texts dealing in more depth with the material outlined in this subsection see Pyndick and Rubinfeld (1998), Koop (2000) and Johnston and DiNardo (1998)

Unit root tests

Unit root tests are a formal test of the behaviour of a series over time, in particular of the behaviour of the mean of the series and its variance. If either or both is increasing, to a certain extent over time, then this will influence the next stage in the analysis of the series[12].

The most common form of unit root test is the augmented Dickey-Fuller Test (ADF) which is conducted by undertaking the following regression.

$$\Delta Y_t = a_0 + \theta t + \gamma Y_{t-1} + \beta_1 \Delta Yt_{-1} + \beta_2 \Delta Y_{t-2} + \beta_3 \Delta Y_{t-3} + \varepsilon_t$$

where Y is the series, _ Y $_t$ is the first difference of the series (i.e $\Delta Y_t = Y_t - Y_{t-1}$) and _ $_t$ is an error term assumed to be white noise (i.e., a random variable having a zero mean and constant variance). Depending on the value of the 't statistic' obtained on Y_{t-1} we will be able to ascertain whether or not the series has a unit root by comparing the t value with the tables constructed by Dickey and Fuller (1979, 1981) or better still McKinnon (1993). If we conclude on the basis of the unit root tests that the series has one root then we say that it is integrated of order 1. If we find 2 unit roots then the series is integrated of order 2. Unit root tests were conducted on all the series we were considering as potential variables for the modeling process.

Testing for the presence of unit root tests can become complicated if the series exhibits what is called seasonality; i.e where the series deviates from a (relatively) smooth pattern on a periodic basis (see Franses, 1998). For example, a series of monthly toy sales may be relatively smooth throughout the year, every year, until December when the series suddenly takes a jump owing to increased sales over Christmas. Obviously, if the toy sales data were yearly then seasonality would not be an issue. Fortunately, the series we used for our analysis were annual

For the multivariate analysis for this part of the study we used annual HL1 returns as our dependent variable. The series was found to be non-stationary and integrated of order 1. For the prediction of homelessness into the future we used quarterly HL2 returns. This series was found to exhibit seasonality with a sizeable drop in applications in every 4th quarter (October-December). We ran 2 univariate (or ARMA) models: one with the seasonality removed by taking the fourth difference[13], see Enders (1995) and Franses (1998) and one explicitly modeling the seasonality. The seasonally differenced series was found to be stationary; in other words it did not have a unit root.

The modeling approach

There are essentially 2 ways to model time series; univariate analysis where only one series is modeled (homelessness applications) on its own past values and multivariate analysis where homelessness applications are analysed along side a number of the posited structural causes. The reason for this twofold approach is related to the research objectives of the

[12] This is very much an over simplification of the concept of unit roots. The definition is, in fact, more refined than this, see Franses (1998) and Diebold (2001)

[13] Other methods also exist to take account of seasonality which do not involve differencing. Typically, they involve seasonal adjustment applied using the Census X- 11 or Census X- 12 ARIMA methods. See Enders (1996) for a brief description of these methods.

study. The univariate approach does not tell us what factors are related to homelessness but it can provide a succinct way of providing a parsimonious 'model' of homelessness that can be used to *predict* what the number of homelessness applications will be in the future. The multivariate analysis on the other hand tells us what *structural factors* are associated with homelessness but cannot be used to predict the number of homelessness owing to the limited information set regarding the future values of these structural causes.

Univariate time series analysis

The most common method of univariate time series analysis is the Box-Jenkins methodology. The aim of this method is to provide an accurate, parsimonious representation of the series under study using only the past values of the series to model its behaviour. The method, which is widely used and well known, consists of a number of steps that have to be applied before it can be claimed that an adequate model for the series has been found. Enders (1995) provides an excellent overview of this methodology.

We applied the Box Jenkins methodology in 2 ways to the logged quarterly HL2 returns:

- Method one involved using the Box Jenkins methodology to model the differenced series (i.e., with the seasonality removed. This helps identify the non-seasonal pattern within the data which is normally of most interest)

- Method 2 involved using the Box Jenkins methodology to model the non-differenced series explicitly taking into account the seasonality of the data.

Method one gave us a parsimonious ARMA representation of the data with the following algebraic form:

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \beta_4 \varepsilon_{t-4} + \varepsilon_t \qquad (2)$$

Where Y_t is quarterly HL2 Returns at time t

i.e an ARMA [2, (4)] model that contains 2 autoregressive components (lagged values of the series) and one moving average component at lag 4. The residuals were found to approximate a white noise process; in other words no useable information about the series' behaviour was left over after applying the above model.

The summary statistics are produced in Table A1 below. All are satisfactory.

Table A1 Univariate analysis with seasonal effect removed

Variable	Coefficient	T-Statistic
AR(1)	0.50	3.01
AR(2)	0.28	3.24
MA(4)	-0.94	-23.64

Adj R^2	0.62
AIC	-3.52
Schwartz Criterion	-3.39

Method 2 consisted of incorporating dummy variables for the different seasons of the year in addition to lags (i.e autoregressive parameters) of the series itself. The most parsimonious model can be represented as:

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \Sigma \delta_i D_{it} + \varepsilon_t$$

Where the variables are as before and D represents a dummy variable for any particular quarter of the data.

Table A2 Univariate analysis incorporating seasonal effects

Variable	Coefficient	T-Statistic
AR(1)	0.30	2.01
AR(2)	0.37	2.78
D1	3.12	2.83
D2	3.04	2.76
D3	3.02	2.73
D4	2.92	2.64

Adj R^2	0.79
AIC	-3.37
Schwartz Criterion	-3.11

Again, the model is quite respectable and is in fact a bit better than the model derived from method 1. Additional tests, using recursive least squares, showed that the seasonal dummy variables are very stable over time as are the other coefficients and that the residuals are white noise.

The purpose of most univariate time series models is for forecasting beyond the available sample and this is easily accomplished using the EViews V 3.0 software we used for this aspect of the analysis. Details of the out of sample forecasts were presented in the main body of the report.

An additional method of forecasting out of sample that does not rely on specifically modeling the series is the Holt -Winters Exponential Smoothing Method. This method is at least as good as the Box Jenkins method when the sample size is small (Chatfield, 2001), as is the case here. The research team felt that to present just one set of forecasts would be potentially misleading. Hence, we opted to also calculate forecasts using a non model based methodology so that the 2 sets of forecasts could be compared. If one set was markedly different to the other then that may have suggested we should rethink our modeling strategy.

This particular method uses as its theoretical base the method of *exponential smoothing* which relies on a set of geometric weights applied to the observations X_n to X_1 in a series (assumed to be non seasonal with no trend) to predict out of sample:

i.e $\hat{X}_{N+1} = C_0X_N + C_1X_{N-1} + \ldots\ldots\ldots\ldots C_1X_1$ (3)

where the C_i are the weights. Geometric weights are chosen as they give more importance to more recent observations:

Thus: $C_i = \alpha\,(1 - \alpha)^i$ (4)

And so:

$$\hat{X}_{N+1} = \alpha X_N + \alpha(1-\alpha)X_{N-1} + \alpha\,(1 - \alpha)^2\, X_{N-2}\ldots\ldots\ldots \text{ etc} \qquad (5)$$

This equation can then be rewritten in recursive form, which is then used for forecasting purposes (see Chatfield, 1996, for further details). Essentially, these forecasts are computed using the latest observation and the previous forecast.

The Holt -Winters method is more refined than that described above in that it can also deal with seasonality and trend (in addition to the level of the variable) which are also subject to exponential smoothing. Thus, the Holt-Winters method uses 3 equations as a basis to forecast out of sample: an equation for the local level of the variable, L, the trend of the variable, T, and the seasonality of the variable, I[14]. These 3 equations are combined into one equation for forecasting purposes, again see Chatfield (1996) for details.

We used the Holt-Winters Exponential Smoothing algorithm with additive seasonality (because data was logged) available with the EViews 3.0 econometric package. As stated earlier there are no diagnostic results to display with this method nor are there any parameter values *per se*. Certainly, none that would add anything to the analysis.

Multivariate time series analysis

Multivariate analysis of time series is very similar to that of basic regression analysis except that now we are dealing with a series of observations on variables and not just a one year snap shot. The dynamic nature of the data means that caution must be used when specifying a multivariate time series model. The reasons for this have been touched on earlier and relate to the behaviour of the series with regard to their mean and variance. If the mean and variance are changing over time then this means that the series are non-stationary or in other words we would expect to find a unit root within the series. If that is the case then they are integrated of order 1.

[14] Seasonality may be multiplicative or additive, see Chatfield (2001) and Makridakis et al (1998) for example.

Non-stationarity causes problems for statistical inference in that the usual t statistics used to assess the significance of coefficients are no longer valid. In addition there is also the problem of 'spurious regression' first discovered by Yule (1926) and developed by Granger and Newbold (1974). This occurs when 2 variables of different orders of integration are regressed on one another and the t statistics reveal that there is a relationship between them when in fact there is not. The same problem can also arise when non-stationary series of the same order are regressed on each other. Johnston and Di Nardo (1998) provide a good resume of the problem of spurious regressions.

The issues discussed above are all related to the topic of "co-integration". A set of non-stationary variables are said to be co-integrated, or in *long run equilibrium*, if a linear combination of the variables is stationary, i.e integrated of order 0. Thus , X_t and Y_t may both be I(1) but $X_t - \lambda Y_t$ *may* be I(0). The parameter λ is called the cointegrating relation (or vector) and is estimated by running OLS on X_t and Y_t and examining the residuals. These, should, by implication be stationary. If they are, then we can say that X_t and Y_t are indeed co-integrated; in other words the series X_t and Y_t have a common trend (Stock and Watson, 1988). Testing for cointegration in this way is often called the Engle-Granger method (Engle and Granger, 1987), but as Enders (1995) highlights it does suffer from a number of inadequacies.

The first problem relates to which variables should be placed on what side of the equation. For example, suppose there are 2 variables (series) X and Y and we want to test whether they are co-integrated. Do we run:

$$X_t = \beta_{10} + \beta_{11} Y_t + e_{1t} \qquad (6)$$

Or do we run:

$$Y_t = \beta_{20} + \beta_{21} X_t + e_{2t} \qquad (7)$$

In theory, the residuals from either can be used to test for co-integration. However, it is possible that using one equation's residuals will indicate co-integration but the other will not. This is a serious problem, for if the 2 variables really are co-integrated, then it should not matter which residuals are used. We should find co-integration.

There is also the problem that the Engle-Granger method is a 2 step procedure. First of all, a co-integrating (or equilibrium) equation like (6) or (7) is estimated then the residuals are used from this equation to test for co-integration via unit root tests. Hence, errors made in the first step are carried over into the second step (Enders, 1995). It should be noted that the appropriate critical values of the t statistics are *not* the same as those used to test for the presence of a unit root in a series. This is because the residuals have themselves been estimated from an 'estimated' equation. Consequently, one must use the critical values provided by Engle and Granger (1987).

Fortunately, recent advances in multivariate time series made by Johanson (1988) and by Stock and Watson (1988) avoid these problems by testing for multiple co-integrating vectors, whilst avoiding a 2 step estimation procedure. Estimation is via maximum likelihood and relies on the mathematical properties of matrices called the rank and characteristic roots.

Enders (1995) gives an accessible account of this procedure, which is known as the Johanson methodology (see also Thomas, 1998).

Note, though, that the presence of more than one co-integrating vector is still problematic because the researcher does not know which is the "correct" co-integrating vector. All that can be said (with some certainty) is that the variables are co-integrated and therefore have a long run relationship with each other.

Granger causality

This sub section will expand more on the concept of causality as applied within statistics. In the main body of the report Granger Causality was defined in terms of past values of a variable X helping to explain a variable Y and the procedure used to test for Granger Causality reflects this. Suppose we have 2 series Y_t and X_t. Then, to test the hypothesis that X does *not* Granger cause Y then the following regression would be estimated:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \varepsilon_t \qquad (8)$$

And then the hypothesis $\beta_1 = \beta_2 = \beta_3 = 0$ would be tested by running:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \varepsilon_t \qquad (9)$$

Then an F statistic would be computed using the residual sum of squares from both equations. If the statistic was significant, we would conclude that X does Granger cause Y.

SUMMARY

This annex has attempted to outline the main issues involved in conducting a rigorous quantitative study of the kind required in this project. Its purpose was not to reproduce material that can be found in many textbooks. For this reason, algebra and equations has been kept to a minimum. Inevitably, it has been impossible to cover all the intricacies involved, but it is hoped that this section will at least give some insight into the complexities and logic involved in conducting the empirical work for this research.

ANNEX 3: REGRESSION RESULTS USING NUMBERS

1981 – Least squares regression – best fit model for homeless APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-.2950	1.431	-.206	.8382	
VAC81	-.4626	.2485	-1.861	.0736	7.6072
EA_EMP	.4399	.3229	1.363	.1843	8.6584
LP81	.3182	.3166	1.005	.3239	6.7491
ALCO	.6969	.1310	5.321	.0000	4.6242
Degrees of Freedom – 27 Adjusted R^2 - .78016					

1981 – Least squares regression – best fit model for homeless ACCEPTANCES

Variable	Coefficient	Standard Error	t-ratio	Probability	Mean of x
Constant	-.1906	1.331	-.143	.8872	
VAC81	-.4148	.2262	-1.833	.0778	7.6072
EA_EMP	.3665	.2797	1.311	.2010	8.6584
LP81	.4018	.2681	1.498	.1456	6.7491
ALCO	.5549	.1348	4.116	.0003	4.6242
Degrees of Freedom – 27 Adjusted R^2 - .77351					

1991 – Least squares regression – best fit model for homeless APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-.6859	1.1386	- .602	.5519	
VAC91	.2002	.2565	.780	.4419	7.4858
EA_EMP	.7997	.3127	2.557	.0165	6.2365
LARENT	-.2422	.1592	- .152	.8802	3.3492
ALCO	.1323	.2763	.479	.6360	4.6481
Degrees of Freedom – 27 Adjusted R^2 - .69370					

1991 – Least squares regression – best fit model for homeless ACCEPTANCES

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-17.5383	5.4518	-3.217	.0035	
VAC91	.3130	.2028	1.543	.1348	7.5225
EA_EMP	1.0635	.1839	5.782	.0000	6.3079
HPRICES	1.3975	.5033	2.777	.0100	10.6518
NEW91PSY	-.8014	.5354	-1.497	.1465	4.7370
Degrees of Freedom – 26 Adjusted R^2 - .75838					

1996 – Least squares regression – best fit model for homeless APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-3.1444	.9001	-3.494	.0018	
VACANT	.1070	.0694	1.541	.1359	4.7391
CLCOUNT	1.4164	.1845	7.675	.0000	8.4166
LARENT	-.2316	.0954	-2.425	.0229	3.6923
ALCO	-.4177	.1375	-3.037	.0055	4.4572
Degrees of Freedom – 25 Adjusted R^2 - .83750					

1996 – Least squares regression – best fit model for homeless ACCEPTANCES

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	1.4458	.9773	1.479	.1539	
VACANT	.1850	.0959	1.929	.0674	4.7716
NOMISLT	.9431	.1605	5.875	.0000	7.0845
RORATIO	-1.7293	.9051	-1.911	.0698	1.1137
SCHEXCL	-.2393	.1086	-2.203	.0389	4.0568
Degrees of Freedom – 21 Adjusted R^2 - .79648					

1996 – Least squares regression – best fit model for homeless LONE-PARENT APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-5.4290	.9401	-5.775	.0001	
VACANT	.1414	.07353	1.923	.0767	4.3451
CLCOUNT	1.0804	.2318	4.660	.0004	8.0997
HBENPRC	.3807	.1264	3.012	.0100	8.2476
ALCO	-.4654	.1209	-3.851	.0020	4.1531
Degrees of Freedom – 13 Adjusted R^2 - .88328					

1996 – Least squares regression – best fit model for homeless YOUNG-PEOPLE APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-4.1343	1.1385	-3.631	.0013	
VACANT	.1911	.1085	1.762	.0903	4.7391
CLCOUNT	1.3730	.2521	5.445	.0000	8.4166
LARENT	-.3159	.1575	-2.006	.0557	3.6923
ALCO	-.4509	.1872	-2.409	.0237	4.4572
Degrees of Freedom – 25 Adjusted R^2 - .71119					

1997 – Least squares regression – best fit model for homeless APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-18.7780	3.1868	-5.892	.0000	
VACANT	.2384	.0608	3.918	.0007	4.7790
ILOUN16	1.1397	.1649	6.912	.0000	8.4664
HPRICES	1.5377	.2860	5.378	.0000	10.8270
ALCO	-.4566	.1162	-3.928	.0007	4.5469
Degrees of Freedom – 23 Adjusted R^2 - .86308					

1997 – Least squares regression – best fit model for homeless ACCEPTANCES

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-18.1211	4.4834	-4.042	.0005	
VACANT	.2411	.0713	3.378	.0026	4.7790
ILOUN16	1.1207	.1837	6.100	.0000	8.4664
HPRICES	1.439	.4081	3.524	.0018	10.8270
ALCO	-.4199	.1285	-3.267	.0034	4.5469
Degrees of Freedom – 23 Adjusted R^2 - .81775					

1997 – Least squares regression – best fit model for homeless LONE-PARENT APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-19.5304	2.8435	-6.868	.0000	
VACANT	.1782	.0652	2.730	.0119	4.7790
ILOUN16	.9743	.1513	6.438	.0000	8.4664
HPRICES	1.5810	.2512	6.293	.0000	10.8270
ALCO	-.2638	.1026	-2.572	.0170	4.5469
Degrees of Freedom – 23 Adjusted R^2 - .88959					

1997 – Least squares regression – Best Fit Model for homeless YOUNG-PEOPLE APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-16.4288	3.1214	-5.263	.0000	
VACANT	.3360	.0839	4.002	.0006	4.7790
ILOUN16	1.1088	.1697	6.532	.0000	8.4664
HPRICES	1.1940	.3006	3.972	.0006	10.8270
ALCO	-.4987	.1351	-3.690	.0012	4.5469
Degrees of Freedom – 23 Adjusted R^2 - .81439					

1998 – Least squares regression – best fit model for homeless APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	6.2128	2.7358	2.271	.0350	
VACANT	.2764	.0977	2.829	.0107	4.8469
NOMISLT	.8289	.1522	5.447	.0000	6.4643
LARENT	-1.4024	.7707	-1.820	.0846	3.4571
SCHEXCL	-.3450	.1980	-1.742	.0977	4.0897
Degrees of Freedom – 19 Adjusted R^2 - .70852					

1998 – Least squares regression – best fit model for homeless ACCEPTANCES

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	5.7561	1.4512	3.966	.0007	
VACANT	.24501	.1283	1.909	.0700	4.8522
NOMISLT	.3343	.1902	1.757	.0934	6.4830
RORATIO	-3.7544	1.3428	-2.796	.0108	1.0698
ALCO	.2717	.1418	1.917	.0690	4.6556
Degrees of Freedom – 21 Adjusted R^2 - .66528					

1998 – Least squares regression – Best Fit Model for homeless LONE-PARENT APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	5.1985	2.8216	1.842	.0811	
VACANT	.3712	.1117	3.322	.0036	4.8468
NOMISLT	.7849	.1769	4.435	.0003	6.4643
RORATIO	-1.4554	.7684	-1.894	.0736	3.4571
ALCO	-.4265	.1969	-2.166	.0433	4.0897
Degrees of Freedom – 19 Adjusted R^2 - .66703					

1998 – Least squares regression – best fit model for homeless YOUNG-PEOPLE APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	5.1984	2.8215	1.842	.0811	
VACANT	.3712	.1117	3.322	.0036	4.8468
NOMISLT	.7848	.1769	4.435	.0003	6.4643
LARENT	-1.4554	.7684	-1.894	.0736	3.4571
SCHEXCL	-.4265	.1969	-2.166	.0433	4.0897
Degrees of Freedom – 19 Adjusted R^2 - .66703					

1999 – Least squares regression – best fit model for homeless APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-5.3739	1.2542	-4.285	.0002	
<25	.4019	.2252	1.785	.0855	10.5312
NOMISCC	.4654	.1829	2.544	.0170	7.9452
UNF2BED	.8525	.3006	2.836	.0086	4.1908
INPDSCH	.1450	.5311	2.731	.0110	3.6654
Degrees of Freedom – 27 Adjusted R^2 - .85108					

1999 – Least squares regression – best fit model for homeless ACCEPTANCES

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-3.5752	2.0215	-1.769	.0887	
LONEPT	.6249	.2499	2.500	.0191	5.2116
NOMISCC	.4285	.2172	1.973	.0592	7.9417
UNF2BED	.9468	.3813	2.483	.0198	4.1827
SCHEXCL	-.1980	.1083	-1.828	.0790	4.0199
Degrees of Freedom – 26 Adjusted R^2 - .81925					

1999 – Least squares regression – best fit model for homeless LONE-PARENT APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-4.7034	.9503	-4.949	.0000	
NEWBUILD	.0954	.0237	4.016	.0004	3.9811
CLCOUNT	.5883	.0613	9.592	.0000	7.9951
UNF2BED	1.0707	.2378	4.503	.0001	4.1908
INPDSCH	.1739	.0394	4.416	.0001	3.665
Degrees of Freedom – 27 Adjusted R^2 - .87874					

1999 – Least squares regression – best fit model for homeless YOUNG-PEOPLE APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
Constant	-4.4091	.5718	-7.711	.0000	
R2B	.2990	.1184	2.525	.0189	5.4985
ILOUN16	.8631	.0989	8.720	.0000	8.3286
ROREDUCT	-.0306	.0121	-2.526	.0189	5.8794
SCHEXCL	.2670	.0966	2.763	.0111	4.1253
Degrees of Freedom – 23 Adjusted R^2 - .86538					

POOLED LEAST SQUARES WITH GROUP DUMMY VARIABLES – 5 YEARS

Best fit model for homeless APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
VACANT	.0358	.0193	1.852	.0668	4.0422
ILOUN16	-.1494	.0865	-1.727	.0871	8.5138
HPRICES	.4847	.2326	2.084	.0396	10.8061
ALCO	.0654	.1203	.544	.5876	4.5708
Degrees of Freedom – 73 Adjusted R^2 - .95031					

POOLED LEAST SQUARES WITH GROUP DUMMY VARIABLES – 5 YEARS

Best fit model for homeless ACCEPTANCES

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
VACANT	.0797	.0381	2.092	.0388	4.2751
ILOUN16	.0705	.0770	.916	.3616	8.4961
LARENT	.0736	.0329	2.232	.0277	4.0872
DRGCNVC	.0662	.0577	1.147	.2539	4.5130
Degrees of Freedom – 80 Adjusted R^2 - .95143					

POOLED LEAST SQUARES WITH GROUP DUMMY VARIABLES – 5 YEARS
Best fit model for homeless LONE-PARENT APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
VACANT	-.0106	.0139	-.767	.4445	4.2239
NOMISLT	.0716	.0459	1.558	.1219	6.9786
MORTARR	-.0519	.0247	-2.096	.0381	4.2404
DRGCNVC	-.0929	.0488	-1.902	.0595	4.4729
Degrees of Freedom – 96 Adjusted R^2 - .95880					

POOLED LEAST SQUARES WITH GROUP DUMMY VARIABLES – 5 YEARS
Best fit model for homeless YOUNG-PEOPLE APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
VACANT	.0128	.0441	.291	.7716	4.1949
NOMISCC	-.4397	.1283	-3.426	.0008	8.2296
LARENT	.0266	.0429	.621	.5357	4.0465
ALCO	-.0457	.1295	-.353	.7243	4.5235
Degrees of Freedom – 106 Adjusted R^2 - .92056					

POOLED LEAST SQUARES WITH GROUP DUMMY VARIABLES–10 YEARS

Best fit model for homeless APPLICATIONS

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
VACANT	.0139	.0201	.692	.4894	4.4621
NOMISCC	.3115	.1253	2.487	.0136	8.3921
HPRICES	.4627	.1560	2.965	.0033	10.6649
DRGCNVC	.2387	.0511	4.668	.0000	4.2533
Degrees of Freedom – 209 Adjusted R^2 - .92895					

POOLED LEAST SQUARES WITH GROUP DUMMY VARIABLES–10 YEARS
Best fit model for homeless ACCEPTANCES

Variable	Coefficient	Standard error	t-ratio	Probability	Mean of x
VACANT	.0196	.0222	.884	.3775	4.4621
NOMISCC	.2089	.1246	1.676	.0950	8.3921
HPRICES	.5483	.1739	3.152	.0018	10.6649
DRGCNVC	.2458	.0480	5.115	.0000	4.2533
Degrees of Freedom – 209 Adjusted R^2 - .92110					

VARIABLES FROM ALL SIGNIFICANT MODELS IN 1996 TO 1999

Structural Categories	**Variables**	**1996**				**1997**				**1998**				**1999**			
		App	Acc	LP	YP	App	Acc	LP	YP	App	Acc	LP	YP	App	Acc	LP	YP
Housing Market	Vacant	+	+	+	+	+	+	+/-	+	+	+	+	+	*n/a*	*n/a*	*n/a*	*n/a*
Unemployment	Under 25s	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	+		+	+
	R2B	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	+	+		+
	Newbuild	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	+		+	
	Loneparent	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	+	+		+
Affordability	Claimant count	+	+	+	+			+						+	+	+	+
	Nomis LT	+	+							+	+	+	+	+		+	
	Nomis CC								+					+	+	+	
	ILO U/E					+	+	+	+	+		+	+	+	+	+	+
	HB Pvte Claims	+	+	+										+	+	+	+
	HB caseload	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*				+
	Private Rent					+		+	+			+		+	+	+	
Deinstitutional	LA rent	-			-					-			-			-	
	House Prices					+	+	+	+	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*
	Possession Actions					+	+	+	+	+		+	+	*n/a*	*n/a*	*n/a*	*n/a*
	Rent officer (ratio)		-							-	-	-	-	-	-		-
	Rent officer (reduct)		-										-	-	-	-	-
	Alcohol	-	-	-	-	-	-	-	-	+/-	+	+	-	*n/a*	*n/a*	*n/a*	*n/a*
	School Exclusions		-			*n/a*	*n/a*	*n/a*	*n/a*	-			-	+	+/-	+	+
	Inpatient Discharges	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*		+		+	+	+	+	+
	Drug Convictions							-				-		*n/a*	*n/a*	*n/a*	*n/a*
	Childcare	*n/a*	*n/a*	*n/a*	*n/a*			-						+	+	+	
	Psychiatric Discharges	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	-		+/-	-

n/a – not applicable as this variable was not tested for this year

ANNEX 4: REGRESSION RESULTS USING RATES

As discussed in Chapter Four, although most of the analysis was conducted with numbers (transformed into the natural logs), some additional analysis was carried out using rates. As will be seen, the results of this analysis are somewhat mixed. However, in comparing the results it should be borne in mind that, strictly speaking, like is not being compared with like (different dependent and independent variables); and hence some differences are inevitable.

In line with Bramley (1993) and Elliot and Krivo (1991) these cross section regressions using rates were weighted (in this case by population) to adjust for the considerable variation in local authority size.

At a late stage in the research, the research team was provided with a number of variables at the local authority level from the Scottish Household Survey, which we incorporated into the regressions undertaken with rates. These additional variables comprised the proportion of households in each local authority that:

- were headed by a lone parent
- contained a step child or step parent
- contained a sharing household
- were renting from a private landlord

In general, the weighted, cross section regressions were able to explain between 40 per cent and 60 per cent of the variation in homelessness rates between Scottish local authorities. As one would expect, this is a lower level of 'explanation' than for the cross section regressions performed with numbers (which were in the region of 60 per cent to 80 per cent), but is respectable for analysis of this kind and probably reflects, in part, the small sample size. By comparison, Elliot and Krivo (1991) were able to explain 38 per cent of the variation in homelessness between US cities, while Bramley (1993) was able to explain 60 per cent of the variation in homelessness between English local authorities. The tables at the end of this Annex show the results for the 'best fit' models.

HOUSING MARKET

Local authority *vacancy rates* were not significant in any of the limited number of regressions performed with rates. However, *young people under 25* (a housing demand variable) proved to be significant when included in the regressions for 1999 (the vacancy rate was not available for this year). Local authorities with a high proportion of people under 25 tended to have a high rate of homelessness applications.

Neither the proportion of households *renting from a private landlord* nor the proportion of households headed by a *lone parent* proved to be significant in the regressions for 1999. However, the proportion of households that were *sharing with another household* - a housing demand variable - was significant (at the 5 per cent confidence level). This indicates that, as the proportion of sharing households increases - through either a fall in the denominator ***or*** an increase in the numerator - the rate of homelessness applications also increases.

UNEMPLOYMENT

Unemployment variables – both the *claimant count* and the *long-term unemployment* rate - were significant in the regressions for 1996, but not for subsequent years.

AFFORDABILITY

The variable used to measure affordability was the average rent for a 2 bedroom private sector flat, taken from the University of York Index of Private Rents and Yields. This variable was generally significant and positively related to homelessness in the regressions performed using numbers (transformed into natural logs). However, in the regressions with rates it proved not to be significant.

DE-INSTITUTIONALISATION

The proportion of *children in care* consistently proved to be significant in the rate regressions. This variable was transformed into a ratio by dividing the number of children in care by the number of children under 15 years in each local authority. Local authorities with high rates of children in care also had high rates of homelessness applications.

The proportion of households containing either a *stepchild or a stepparent* was significant, though the relationship was negative: local authorities with high stepchild/ parent rates had low homelessness rates. Again, it is possible only to speculate as to the nature of the causality of this result.

CONCLUSION

The analysis with rates was not entirely consistent with that using logged numbers. Some variables that were significant in the logged numbers, such as local authority vacancies and private rents, were not when the analysis was conducted with rates. Other variables, such as unemployment, which were consistently significant in the logged number analysis, were significant only for one year and not for the others.

What matters so far as the structural model is concerned are not particular variables so much as the structural category that they represent. The analysis with rates confirmed that homelessness rates are significantly affected by housing market conditions and by what we have inelegantly termed ‘de-institutionalisation’. The results for unemployment only partially confirmed the results with logged numbers. Meanwhile, the results for affordability were inconsistent with those for logged numbers.

However, for the reasons discussed in Chapter Four, for the purposes of this research project, the regression analysis using logged numbers is more robust than that using rates. The results of that analysis were presented in Chapter Five (with statistical details in Annex 3). The results of the analysis using rates have been presented in this Annex for completeness.

Best Fit Models for Regressions using Rates

Variable	Variable label	Sign	Significance
UNF2BED	Private rent for 2 bed unfurnished flat	+	
SINGLEPA	Percentage of households headed by a lone parent	+	
NLTRATE	Long-term claimant unemployed rate	+	
INCARRAT	Percentage of children in care	+	**
	Adjusted R-squared: 0.40		

Note: * denotes variable significant at 10% level, ** at 5% level, and *** at 1% level.

Variable	Variable label	Sign	Significance
UNF2BED	Private rent for 2 bed unfurnished flat	-	
SINGLEPA	Percentage of households headed by a lone parent	+	
NCCRATE	Claimant count rate	-	
INCARRAT	Percentage of children in care	+	*
	Adjusted R-squared: 0.42		

Variable	Variable label	Sign	Significance
UNF2BED	Private rent for 2 bed unfurnished flat	+	
STEPCHIL	Percentage of households containing a stepchild or stepparent	-	*
NLTRATE	Long-term claimant unemployed rate	+	
INCARRAT	Percentage of children in care	+	
	Adjusted R-squared: 0.45		

Variable	Variable label	Sign	Significance
UNF2BED	Private rent for 2 bed unfurnished flat	+	
STEPCHIL	Percentage of households containing a stepchild or stepparent	-	
NCCRATE	Claimant count rate	+	
INCARRAT	Percentage of children in care	+	***
	Adjusted R-squared: 0.42		

Variable	Variable label	Sign	Significance
UNF2BED	Private rent for 2 bed unfurnished flat	+	
SHAREDRO	Percentage of households sharing their accommodation	+	**
NLTRATE	Long-term claimant unemployed rate	+	**
INCARRAT	Percentage of children in care	+	*
	Adjusted R-squared: 0.51		

Variable	Variable label	Sign	Significance
UNF2BED	Private rent for 2 bed unfurnished flat	-	
SHAREDRO	Percentage of households sharing their accommodation	+	**
NCCRATE	Claimant count rate	+	*
INCARRAT	Percentage of children in care	+	
	Adjusted R-squared: 0.50		