palgrave
macmillan

Construction and Real Estate Dynamics

Edited by

Philippe Thalmann and Milad Zarin-Nejadan

Applied Econometrics Association Series

General Editors: **Jean H. P. Paelinck**, Emeritus Professor, Erasmus University, Rotterdam, and **Henri Serbat**, Chamber of Commerce, Paris

The vital importance of econometrics for understanding economic phenomena is increasingly recognised in every field of economics.

The discipline is based on 'scientific processes which aim to identify, explain and forecast economic phenomena using formalised tools to produce, measure, structure and model the information' (Gerard Duru and Henri Serbat, 1991).

The Applied Econometrics Association, established in 1974, seeks to develop the use of econometric tools by producing regular updates on the state of the art and the progress made in each specific field, and so to further the transformation of unconnected facts into pertinent information for the use in analysis and decision-making.

The series was conceived to form a set of working references for advanced students and researchers in each specific field, as well as a guide to development across the discipline more broadly.

This exchange of knowledge will be continued permanently by the opening of a debate-site on the Internet [http://www.lyonl.fr/aea].

Titles include:

Françoise Bourdon and Jean Bourdon (*editors*)
WAGE ECONOMETRICS AND MODELLING

Fabrizio Carlevaro and Jean-Baptiste Lesourt (*editors*)
MEASURING AND ACCOUNTING ENVIRONMENTAL NUISANCES AND SERVICES

Arthur Getis, Jesús Mur and Henry G. Zoller (*editors*)
SPATIAL ECONOMETRICS AND SPATIAL STATISTICS

Siv S. Gustafsson and Danièle E. Meulders (*editors*)
GENDER AND THE LABOUR MARKET
Econometric Evidence of Obstacles to Achieving Gender Equality

Hans Heijke and Joan Muysken (*editors*)
EDUCATION AND TRAINING IN A KNOWLEDGE-BASED ECONOMY

Pavlos Karadeloglou (*editor*)
EXCHANGE-RATE POLICY IN EUROPE

Emile Quinet and Roger Vickerman (*editors*)
THE ECONOMETRICS OF MAJOR TRANSPORT INFRASTRUCTURES

Philippe Thalmann and Milad Zarin-Nejadan (*editors*)
CONSTRUCTION AND REAL ESTATE DYNAMICS

Applied Econometrics Association
Series Standing Order ISBN 0–333–71460–1
(*outside North America only*)

You can receive future titles in this series as they are published by placing a standing order. Please contact your bookseller or, in case of difficulty, write to us at the address below with your name and address, the title of the series and the ISBN quoted above.

Customer Services Department, Macmillan Distribution Ltd, Houndmills, Basingstoke, Hampshire RG21 6XS, England

Construction and Real Estate Dynamics

Edited by

Philippe Thalmann
Federal Institute of Technology
Lausanne
Switzerland

and

Milad Zarin-Nejadan
University of Neuchâtel
Switzerland

palgrave
macmillan

First published 2003 by
PALGRAVE MACMILLAN
Houndmills, Basingstoke, Hampshire RG21 6XS and
175 Fifth Avenue, New York, N.Y. 10010
Companies and representatives throughout the world

PALGRAVE MACMILLAN is the global academic imprint of the Palgrave Macmillan division of St. Martin's Press, LLC and of Palgrave Macmillan Ltd. Macmillan® is a registered trademark in the United States, United Kingdom and other countries. Palgrave is a registered trademark in the European Union and other countries.

ISBN 1–4039–1545–8

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources.

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data.
Construction and real estate dynamics /
edited by Philippe Thalmann and Milad Zarin-Nejadan.
p. cm. – (Applied Econometrics Association series)
Includes bibliographical references and indexes.
ISBN 1-4039-1545-8
1. Construction industry. 2. Real estate business. I. Thalmann, Philippe. II. Zarin-Nejadan, Milad, 1957- III. Series.

HD9715.A2C6475 2003
333.33–dc21 2003049802

10 9 8 7 6 5 4 3 2 1
12 11 10 09 08 07 06 05 04 03

Printed and bound in Great Britain by
Antony Rowe Ltd, Chippenham and Eastbourne

Contents

Notes on the Contributors

Peter ANTONIONI
Lecturer
Business School, University College, London, U.K.
Email: ucznpja@ucl.ac.uk

Michael BALL
Director
Urban and Property Economics Consultancy, London, U.K.
Visiting Professor
Department of Real Estate and Planning, University of Reading, Reading, U.K.
Email: ballm@sbu.ac.uk

Francis CALCOEN
CNRS Research Director
Cresge/Labores, Catholic University of Lille, France
Email: fcalcoen@cresge.fr

Didier CORNUEL
Professor of Economics
MEDEE-IFRESI, University of Lille, France
Email: didier.cornuel@ifresi.univ-lille1.fr

Alastair McFARLANE
Senior Economist
U.S. Department of Housing and Urban Development
Office of Policy Development and Research, Washington, USA
Email: alastair_mcfarlane@hud.gov

George OFORI
Professor and Head
Department of Building, School of Design and Environment
National University of Singapore, Singapore
Email: bdgofori@nus.edu.sg

Gerbert ROMIJN
Researcher Public Finance
CPB Netherlands Bureau for Economic Policy Analysis, The Hague, Netherlands
Email: g.romijn@cpb.nl

Freddie TAN
Adjunct Senior Researcher
Department of Building, School of Design and Environment
National University of Singapore, Singapore
Email: bdgtanf@nus.edu.sg

Philippe THALMANN
Professor of Economics of the natural and built environment
Federal Institute of Technology, Lausanne, Switzerland
Email: philippe.thalmann@epfl.ch

Milad ZARIN-NEJADAN
Professor of Economics
Department of Economics and IRER, University of Neuchâtel, Neuchâtel, Switzerland
Email: milad.zarin@unine.ch

This page intentionally left blank

1
Introduction

Philippe Thalmann and Milad Zarin-Nejadan

The construction industry is a major constituent of European economies. The gross output of the sector in the European Union (EU) represented 10.4 per cent of GDP and 5.4 per cent of total value added (2000 figures). The sector also accounts for one half of gross capital formation. However, the share of construction expenditure in GDP has been declining steadily since 1973, with the exception of a small boom at the end of the 1980s. The evolution is basically the same for residential and non-residential components, paralleling other industrialised countries.

The EU is the principal world exporter of construction services. European companies, among the largest in the world, hold about one half of international markets. The value of cross-border trade is however still relatively low compared to the total value of the construction market. Taking into account indirect imports (that is purchases of

Table 1.1 Construction expenditure as a percentage of GDP

	1970–73	**1974–79**	**1980–89**	**1990–95**	**1996–00**
European Union	14.5	13.3	11.7	11.3	10.4
– residential	6.9	6.4	5.6	5.3	5.2
– non-residential	7.6	6.9	6.1	6.0	5.2
OECD	13.8	13.2	12.3	11.3	11.0
– residential	6.2	6.0	5.3	4.9	4.9
– non-residential	7.6	7.2	7.0	6.4	6.1

Source: OECD, Historical Statistics, Paris, different years.

non-domestic origin through a local subsidiary), overall EU public sector import penetration is between 7 and 10 per cent.

Construction has always been a key sector in terms of employment. In the EU, it provides jobs for 11.9 million workers (2000 figures) representing 7.2 per cent of the total workforce. It is also a powerful engine of job creation. Every job created in the construction sector generates two further jobs in related sectors. Thus, more than 30 million workers in the EU depend, directly or indirectly, on the construction sector. Construction employment is, however, largely hard work, dangerous and volatile. It attracts poorly qualified labour with low educational achievement along with unskilled immigrant workers willing to work below normal wage levels. Over 90 per cent of the workforce is male. Construction workers are self-employed in the proportion of 22 per cent, more than in manufacturing (7.3 per cent) and even in services (17.8 per cent).

Under those conditions, it may not be surprising that construction contributes a disproportionate share of the shadow economy. Further reasons are high non-wage labour costs, which can amount to more than 30 per cent of the wage cost, the high mobility of the workforce, the short duration of contracts, the strong cyclical and seasonal variation in the industry and illegal immigration which can easily find work in the sector requiring little or no qualification or experience.

The construction sector is extremely heterogeneous. Although it can generally be considered as quite labour-intensive, particularly in its final assembly stage on site, it also includes activities with high human capital intensity (for example design and management) as well as high machinery intensity (for example excavation). The construction industry is strongly fragmented. Even if a few global players command high turnovers, construction is, in comparison with other industrial sectors, far from being prone to oligopolistic or dominant tendencies, except in selected sectors such as tunnelling. In 1990, the turnover of the ten largest European construction companies only accounted for less than 6 per cent of the total European market. Craftsmen and small and medium-sized enterprises (SMEs) play a major role in the sector. Indeed, 97 per cent of some 2 million EU companies have less than 20 employees, and 93 per cent have less than 10. As a result, 67 per cent of construction value added is contributed by companies with less than 50 employees, two and a half times the corresponding share in manufacturing. The construction market tends to be segmented both geographically and by field of specialisation.

Research and development expenditure in the European construction sector does not reflect its economic importance. Investment in R&D is limited to 0.3 per cent of the sector's turnover which is quite low in comparison, for example, to Japan (2–3 per cent). This might account for the relatively low productivity gains registered in this sector: during the period 1970–85, productivity in the construction sector increased at an average annual rate of 0.9 per cent, well below the rate of 2.3 per cent for all other industries.

The construction industry plays a major environmental role which goes beyond the mere transformation of landscape and natural habitat. The sector generates an enormous quantity of construction waste and demolition material (more than 270 million tonnes per year). Buildings are further responsible for 42 per cent of energy consumption within the EU, with an expected annual growth rate of 1.5 per cent during the next decade. The sector is the second largest contributor to CO_2 emissions.

Construction activity is highly cyclical, rising in proportion to GDP during expansions and falling during economic downturns (see Figure 1.1). This close link to the economic cycle combined with the considerable weight of the sector in the economy gives a particular poignancy to business fluctuations in construction activity. On the demand side, the main determinants are short-term conditions on financial markets and medium-term business perspectives in other markets as well as available income. On the supply side, the short-run evolution of building activity is often explained in a residual manner (that is by the situation on other markets), but in the long run, demand-related factors such as population, taste and public construction policies tend to be predominant.

After having recovered slowly from the boom and bust around 1990, the European construction sector may now be heading for yet another trough of its secular roller coaster. In the early 1990s, it was the bursting real-estate bubble of the late 1980s that pulled down construction activity. Today it is the general slowdown in the economy and fiscal contraction. The difficulties of construction and real estate in the early 1990s were largely responsible for a decade of low growth (and low inflation). And yet, few in academia seem to care much about the construction sector and real estate. It seems as though

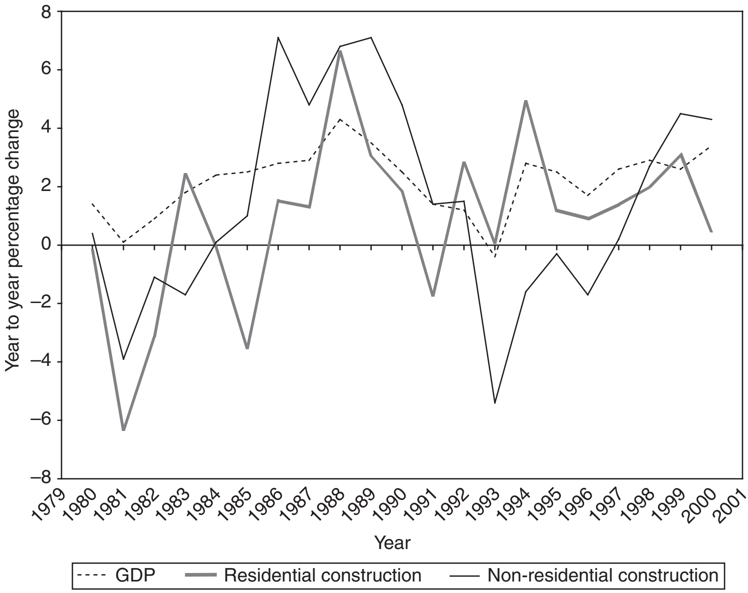

Figure 1.1 Cyclical fluctuations of construction activity in the EU 1980–2000
Source: OECD, *Historical Statistics*, Paris, different years.

university economists had grown tired of attempting to model the volatility of those sectors, and preferred to turn to more modern topics such as financial derivatives, tradable pollution permits and monetary unions. The National Bureau of Economic Research was founded some eighty years ago in the United States to examine the cycles of construction, housing and real estate. A search in the NBER's working paper database yields only four papers dealing with the construction sector over the last fifteen years. No working paper analysing construction, real estate or housing was edited in the NBER program 'Economic Fluctuations and Growth' since 1986. The ISI database contains 1700 social sciences journals, of which 114 deal with the environment, 21 with finance and 7 with building and construction, and the latter are in fact engineering journals.

The difficulty of obtaining good data for such a heterogeneous sector may explain the reluctance of economists to analyse construction activity. The construction sector is often missing from international statistics. It is left out of data sets and analyses both of manufacturing and services, being a bit of both. Productivity, R&D efforts, international trade or simply prices are difficult to define and measure in a sector so fragmented, which produces nearly only one-of's upon demand.

Clearly, not all researchers have given up on the issues of construction and real estate. There are very active associations and journals, particularly in the latter field, but their members and authors seldom find a way into mainstream congresses and journals. This book provides evidence that there are interesting economic questions in construction and real estate, to be handled with the recent economic research technology. Needless to say that these are also very serious questions, as documented by the above-mentioned facts and figures.

How do construction firms cope with the volatility of construction? Do speculative bubbles or market fundamentals drive those fluctuations? Are there better ways to predict construction demand? Why did the office market lead the real-estate cycle? Is regulation responsible for speculative behaviour? Those are the questions addressed in turn by the following contributions.

Michael Ball and Peter Antonioni start by challenging the view that the construction market – construction demand and prices – is particularly volatile. They distinguish and estimate very carefully short-term and medium-term volatility in the UK construction market. They also explain why construction firms care about fluctuations in market activity, in spite of their supposed flexibility. In fact, those firms face considerable sunk costs, so they would rather raid another market and push down tender prices than downsize. It remains to be explained why firms struggle to stabilise their returns instead of letting their shareholders minimise variance over their portfolios. Presumably firms care for other stakeholders – management, workers – who cannot diversify their income.

Ball and Antonioni next show how construction firms can minimise order variance by adopting optimal proportions of work across the three main construction sectors – housebuilding, civil works and other construction. It might be interesting to examine also geographic diversification or diversification of services offered. The authors note that their model does not apply equally to all type of firms. Smaller firms gain more from sectoral specialisation, which saves on overhead costs, particularly in the relatively stable housebuilding sector.

Larger firms gain more from diversification through economies of scale and scope. With such observations, Ball and Antonioni rejoin the description Campinos-Dubernet (1988) has made of the two types of strategies followed by construction firms in Europe. 'Primary firms' are internationally diversified to stabilise their orders; they invest in their reputation, their know-how and their work force. 'Secondary firms' bet on flexibility; they assemble resources from subcontractors on demand and go for the rapid gain in compensation for high risk.

Didier Cornuel and Francis Calcoen thoroughly examine the last cycle of the market for second-hand flats in Paris, in order to decide whether sales prices and transaction volumes were subject to a bubble or to real shocks. Obviously this is only one real-estate market among many, but hopefully analysing it might help in explaining the cycles observed over the last ten years in other markets and regions. What happened in Paris has generally been interpreted as a speculative bubble. As Cornuel and Calcoen remind us, a bubble occurs when expectations of capital gains lead to price increases until doubts emerge in the market. Throughout the bubble, markets are in equilibrium, that is supply and demand move together, with the same actors on both sides, but the prices diverge from the fundamental values, first above then below. This model does not help to explain the observed cycle in trading volumes. The authors could have extended the basic bubble model, to have for instance investors realise rapidly their capital gains and sit on capital losses.

How might one test for a bubble? The authors compute the fundamental value by discounting rents. They divide the current rent by the current yield of public sector bonds minus the average growth rate of rents observed in the following years.[1] The discount factor is surprisingly constant, so that the fundamental value closely follows rents. Rents grew steadily and rapidly since 1979. Both results – the closeness of fundamental value and observed price and the growth of rents – lead Cornuel and Calcoen to conclude that there was no bubble but growth of prices driven by the growth of rents. The growth of rents has not yet stopped, however, so that the decline in the prices of flats after 1991 is not explained (their Figure 3.2).

Why did rents grow so rapidly and steadily after 1985? Cornuel and Calcoen show that population and incomes did not grow in Paris, so they concentrate on decreases in supply. Things become more complicated – and thereby more interesting – when the inter-relations in the housing market are examined. Enter tenure choices, conversion, wholesale and retail trade, transaction chains and household mobility. The danger is that the door is now opened to *ad hoc* explanations: anything can be explained with so many channels of transmission. A formal model is needed, that integrates the real estate and the services sectors. Cornuel and Calcoen's contribution is a nice analysis challenging too rapid reliance on bubble stories, but generalisation of their results to other real estate markets may not be warranted. Extensions of their paper will pay attention to changes on credit markets or to portfolio considerations, which might have played a role on other markets.

Freddie Tan and George Ofori construct a forecasting model for construction demand in Singapore utilising an approach which is still rather novel in our field: neural networks.[2] Simulating neuronal systems of the brain has proven a powerful method for detecting signal–effect relationships when little is known about the system under scrutiny. That approach is non-parametric, but it is also a bit of a black box as regards the nature of the relationships detected. The authors select potential determinants of construction demand from other models and the specificity of Singapore's economy. They show that construction demand (measured by the value of contracts awarded) is predominantly determined by three variables: building costs, manufacturing output and gross fixed capital formation (GFCF). Together those factors account for more than 56 per cent of demand fluctuation. The influence of GFCF suggests that public investment can help stabilise construction demand.

The paper was written in early 1997 with data going to the second quarter of 1996. Thus the forecasts made for 1997 to 2000 are not very interesting, except to show once more the limitation of any such exercise when major turmoil occurs in financial markets. Should neural networks be used more frequently, then a thorough comparison of such models with more standard econometric models is called for. Goh (1996) performs such a comparison with Singapore data for residential construction demand and concludes that the mean absolute percentage error of a neural network model is one-fifth of that of a standard multiple regression.

Gerbert Romijn examines the Dutch office market, one of the most interesting in Europe: not only did it expand dramatically over the last twenty years, but what was built was frequently quite innovative.

Or should we say that the innovation has concentrated on architectural superficialities, patios, facades and fitting into the urban landscape? It seems that too little attention was paid to how the use of office space is changing with the rapid development of communication technology and organisation of work (Veldhoen and Piepers, 1995). As functional requirements change, the market for office buildings is heading for major mutation.

Romijn develops a model of office space demand with lump-sum adjustment costs, where firms move when the difference between actual and desired office space use exceeds a certain threshold. He solves explicitly the intertemporal maximisation programme of the individual firm. Its rigidity has macroeconomic consequences: aggregate office space demand will depend on market conditions through two channels, desired office space and relocation speed.

Romijn tests the predictions of his model concerning aggregate office space use with Dutch data for 1974–95. He finds that actual office space use is only about half as volatile as desired use inferred statistically from the relative cost of rent to labour. He also confirms that his model with adjustment costs much better tracks actual office use. Change in office demand is a weighted average of current and lagged changes in desired space. The weights could be fixed or, better, depend on current and lagged desired office space. His relatively smooth and low frequency data does not give a clear advantage to the more complex model with variable weights. Romijn promises to explore that issue with more precise data for the limited market of Amsterdam.

Alastair McFarlane works in the fields of real estate and urban economics. Urban economists work with stylised models of the development process with a view to obtain conclusions on the timing of land development and the expansion of cities. An approach that is becoming standard is the monocentric urban growth model with durable housing. Assuming that all lots of land are identical except for distance to the centre of the city, their intent is more macro than micro. Such a model will predict the spatial structure of a city over time but is restricted to one type of land use. It will not predict when a particular lot will be developed or for what activity. General results can be obtained on how changes in various parameters such as interest rates and income growth affect equilibrium city size. That determines directly construction activity, when it only takes place at the urban fringe as in McFarlane's model.

McFarlane's model is such a partial equilibrium model (the source of rents growth is exogenous) designed to compare how the stock of housing in a city expands under two sets of conditions: (a) the unregulated housing market, on which rents of occupied housing may follow the rents of new housing; and (b) the regulated market with rent surveillance, on which rents are fixed at their initial (free) level. McFarlane shows that such rent regulation causes delays in development because promoters gain from waiting for higher rents. Rent regulation introduces an irreversibility effect even when there is only one possible type of development and the future is perfectly certain. The delay is reduced when interest rates are higher. We may even see construction increase when the interest rate rises, because of a smaller irreversibility effect! McFarlane has further results on city size dynamics as well as rents and land prices. He also examines what tax on vacant land and what subsidy to construction would hasten development and thus offset the effects of rent regulation.

His model deserves to be extended by introducing risk, by endogenising the rents of new units (it is hard to believe that they would be equal without and with rent regulation), and by allowing for cost pass-through in rent regulation.

All of these papers and many more were first presented at the 55th International Conference of the Applied Econometrics Association (AEA), 'Construction Economics and Econometrics'. The conference was held at the University of Neuchâtel, Switzerland, on February 20 and 21, 1997. Philippe Thalmann chaired the international scientific committee and Milad Zarin-Nejadan led the local organising committee. We wish to thank Henri Serbat of the AEA and the members of the local organising team for organising the congress. In selecting the contributions to the congress, Philippe Thalmann was assisted by a scientific committee comprising R. Albriktsen (Veidekke ASA, Norway), R. Bon (University of Reading, UK), B. De Borger (University of Antwerp, Belguim), E. Deutsch (Tech. University of Vienna, Austria), D. Emmanuel (DEPOS, Greece), P. Englund (Uppsala, Sweden), L. Hjalmarsson (Goteborg University, Sweden), S. Hylleberg (University of Aarhus, Denmark), R. Maquaire (Pont à Mousson, France), S. Pulakka (Tech. Research Center Helsinki, Finland), G. Pult (University of Neuchâtel, Switzerland), E. Quinet (ENPC, France), Z. Raisse (Féd. Par. Bâtiment, France), A. Sadler (Arbed SA, Luxembourg), H. Serbat (AEA), T. Siebe (RWI Essen, Germany),

C. Zimmermann (Université de Québec) and K. Yamada (Kyoto, Japan). Finally, last but not least, we would like to thank Mrs. Kira Facchinetti for preparing the manuscript of this book.

Notes

1. Such an approach, which dispenses with predictions about expected rents beyond the date of analysis, becomes increasingly problematic for fundamental values at the end of the data sample.
2. For other applications in the construction sector, see Goh (1996) and Boussabaine and Kaka (1998). In fact, the business forecasting package NeuroForecaster includes construction demand forecasting as an application.

References

Boussabaine, A.H. and A.P. Kaka (1998), 'A Neural Networks Approach for Cost Flow Forecasting', *Construction Management & Economics*, 16(4), 471–9.

Campinos-Dubernet, M. (1988), 'Diversité des formes de gestion de la variabilité des processus du bâtiment: effet sectoriel ou effet national?', *Actes du Colloque 'Europe et Chantiers'*, Plan Construction et Architecture, Paris.

Commission of the European Communities (1997), *The Competitiveness of the Construction Industry*, Communication from the Commission, November.

Eurostat, *European Facts and Figures, Data 1990–2000*, Office for Official Publications of the European Communities, Luxembourg, 2002.

Goh, B.H. (1996), 'Residential Construction Demand Forecasting Using Economic Indicators: A Comparative Study of Artificial Neural Networks and Multiple Regression', *Construction Management & Economics*, 14(1), 25–34.

Veldhoen, E. and B. Piepers (1995), *The Demise of the Office – The Digital Workplace in a Thriving Organisation*, Uitgeverij 010 Publishers, Rotterdam.

2

Diversification as a Strategy for Minimising Fluctuations in Construction Firm Turnovers

Michael Ball and Peter Antonioni

1 Introduction

Construction firms frequently complain of the volatility of demand in their markets with periods of boom followed by slumps. In a number of countries, the stabilisation of construction demand through more sensitive macroeconomic policies is often on the policy agenda of construction industry lobbies. In this chapter, we wish to explore the empirical validity of the claim that construction markets are exceptionally unstable by examining the evidence from one of the most volatile major construction markets in Europe, that of Great Britain, and, in addition, to explore some potential firm strategies towards market volatility.

When specific markets have cycles that do not perfectly coincide, an obvious strategy to stabilise overall output at the level of the firm is for it to operate across several markets. This strategy can be seen in practice, with many large construction firms stabilising their turnovers by operating across several types of construction market – civils, private housing, office building and so on. Yet, the issue of market spread is a particularly topical one because several large firms have in recent years chosen to narrow sharply the range of markets in which they have a presence, after their experiences in the early 1990s recession. What consequence, if any, is there for turnover volatility when firms reduce their market spread? This issue is examined by developing a simple turnover variance optimisation model. This model is then applied to recent British experience, testing whether diversified market presences would have limited turnover variations for British firms. The results presented here suggest that a diversified market presence does limit turnover fluctuations. The simulations of the model, however, highlight that a targeted strategy, rather than one based on presences in all markets, is the best one for limiting turnover fluctuations. Turnover stability, of course, is not necessarily the same as profit-maximising, so we initially offer some comments on the link between the two for the modern construction industry.

2 Why should construction firms bother about fluctuating turnover?

A common, if somewhat misplaced, view of construction is that it is a labour intensive industry in which knowledge, skills and other inputs can easily be transferred from one sector to another. Construction from this perspective, in other words, is a good example of a highly flexible industry able to cope with considerable variations in total demand and its composition. Firms should, from this perspective, be able to switch to other areas of work when one type declines or to vary their total work levels with ease. As a result, they need no strategy towards market fluctuations.

A more realistic assessment, however, suggests that the actual degree of resource flexibility in the construction industry depends on the type of work being considered and the aspect of the value-added chain under consideration. Some types of work require standardised, routine tasks, whereas others require dedicated pieces of equipment and specialist labour – contrast basic housing repair and maintenance, for example, with civil engineering. Particular types of construction employment, moreover, involve skills that are dedicated to a narrow range of sector specific activities – and the same could be said of plant and machinery. Inflexibilities in resource availability, in addition, may exist at the level of the construction industry as a whole rather than for individual firms or projects. An example would be trained construction labour. Individual firms or project leaders may be able to bid successfully at prevailing prices for the available pool of skilled labour but the size of that pool may be relatively fixed in the short-run. This is particularly so for increases in skilled labour availability as it takes time to hire and train. Increases in construction demand may consequently lead to overall shortages of skilled labour, higher earnings and bottlenecks in a significant proportion of projects.

There is further indirect evidence that construction is not as flexible as it is sometimes portrayed. At times of severe demand recessions, tender prices often tend to fall faster than outputs. This occurs, first, because some firms raid markets in which they are not normally active to limit losses elsewhere and incumbent firms respond to protect market share. Second, incumbent firms may lower bid prices in order to try to win sufficient contracts to cover the overheads necessary to maintain a presence in those markets. These forms of behaviour are unlikely to occur in the purely flexible model, because under its assumptions firms do not have many overheads to worry about and so should be less compelled to cut their rates.

The reason why falling prices occur in general, we would like to suggest, is that most construction firms, particularly the larger ones, face sunk costs when involved in particular areas of business. Sunk costs by definition cannot be recouped if a firm quits a business but rather have to be written off in balance sheets. To avoid losing the benefits of sunk costs, construction firms have a propensity to tender during downturns in work at low prices in order to remain in a sector. The sunk costs faced by construction firms exist for several reasons, and it is worth elaborating on them in order to identify the connections between profitability and turnover volatility.

The most obvious sunk costs for a construction firm are those associated with completing its existing portfolio contracts. If input prices or the complexity of the necessary work are underestimated, a firm is locked into an unprofitable contract. Such sunk costs, however, are temporary. Furthermore, in the absence of systematic forecasting errors, they should not be high for large firms because they should have a sufficient number of projects on hand so that any project forecasting errors should approximately be normally distributed. Unexpected losses on some contracts, therefore, should be compensated for by unexpected gains elsewhere, and the expected value of such sunk costs should approximate to zero.

There are other sunk costs, nonetheless, that may be significant and permanent arising from operating in specific markets. They take several forms. For some markets, firms need to have a team of staff with knowledge of specialist tasks or to own dedicated equipment. Assembling these is not costless nor is laying them off, so their existence contains an element of sunk cost. Other sunk costs are associated with building up market information. For specific types of work or geographic location, firms active in a market can acquire privileged insider information on client behaviour and input suppliers, with the latter especially important when extensive sub-contracting exits. This information quickly dates when a firm quits a sector and it is unlikely to be able to be sold on to a competitor. Yet, it may be costly to acquire in terms of management time and learning by experience and these costs are unrecoverable sunk ones. The final types of sunk cost are associated with the importance to construction firms in an industry renowned for its cowboys of signalling probity to potential clients. This is done by establishing reputations for competence in particular markets. Reputation is an important signal because construction project outcomes are uncertain prospects and it is generally difficult for clients to monitor the detailed abilities of firms. Yet, a firm's reputation in a particular market sector is lost when it leaves it and may be costly to reacquire.

The existence of sunk costs helps to explain why construction firms do adopt long-term strategies with respect to the markets in which they are active. Such behaviour can be seen widely in the literature on firm strategies (for example Constable, 1986; Lansley, 1987; Hillebrandt and Cannon, 1990; Hillebrandt et al., 1995; Betts and Ofori, 1992) and in mergers and acquisitions (Ball, 1988). Sunk costs and other overheads, however, are only part of the picture. Although such costs exist in construction, they are far lower than in some other industries where they play an important part in explaining market structures (Sutton, 1991). In construction, it is still common for firms to make short-term moves into specific sectors, especially when they offer substantial price discounts against incumbents' tender bids. The close correlation of movements over time in tender prices across construction sectors implies that there is movement by firms between them. Nonetheless, construction firm managers when interviewed, according to the literature, in general argue that for medium-term profitability they need to develop medium-term presences in markets.

If firms could costlessly and rapidly move between different types of work and geographic location, the best management strategy would be to bid for all types of work that offer the best current rate of return. The apparent need for a medium-term presence corresponds to the arguments above that sunk costs are of importance in construction.

This suggests that the volatility of construction demand overall is of relevance to profitability. The greater the volatility of demand, the more likely that intense price competition will break out. Firms with sunk costs, moreover, should avoid sectors were turnover is hard to predict because of orders variability. Senior management, thus, has to adopt strategies towards their markets. Specifically, they have to decide which businesses to enter and which to leave and, when incumbent, the optimal scale of activity. There is consequently a real link between profitability and turnover volatility.

3 Evidence on the volatility of construction markets

Britain tends to have a construction market that is more volatile than in many other countries (Ball et al., 1996), so it is a useful guide to the relative volatility of construction compared with other economic sectors. This section will examine the pattern of fluctuations in construction output, orders and tender prices. Volatility is generally measured as fluctuations around a trend. Using fluctuations of de-trended data as the indicator of volatility is both intuitively plausible and has the statistical advantage of rendering the data stationary to avoid serial correlation biases. First differences are a standard method of doing this, so that percentage fluctuations in growth rates become the objects of study. Standard deviations for these modified data are calculated and compared between time periods and types of activity. F-tests enable the statistical significance of these comparisons to be determined. This procedure is adopted here.

First differences, however, may not always provide the ideal indicator of volatility, because they only examine immediate changes in growth rates between time periods, in this case years or quarters. They consequently give no indication of volatility over any longer term period. Yet it may well be the degree of uncertainty in the medium-term that influences construction firm behaviour, and such information will be only poorly picked up by first differences. As a result, another de-trending method is used as well which tends to smooth the data over longer time periods. The detrending method chosen was structural time series models (STSM) utilising the Kalman filter (Harvey, 1989; Ball and Wood, 1996) (see Appendix for details). In the results that follow, therefore, two sets of standard deviations are presented. One is based on first differences, with the aim of identifying 'short-term' volatility. The other, derived from the STSM approach, is the variance of the growth rate of the trend – aimed at illustrating 'medium-term' volatility. The results are presented as tables of standard deviations.

The data used are the 1995 revised construction quarterly time series produced in Construction Statistics. The data series have been recalibrated since 1980 to include an infrastructure category, which consists of both private and public projects, as well as the traditional divisions into public and private housing, other public works, plus private industrial and commercial work. The result is that meaningful volatility comparisons between the years before and after 1980 are only possible for the series not affected by this reclassification.

The time periods chosen reflect the available span of the data and also what seemed logical break points within it. For example, the first quarter of 1981 (1981Q1) was the trough of the early 1980s slump, the lowest point that total construction output reached between 1962 and 1994.

It is difficult to measure the volatility of demand directly, because data on both construction orders and output are the products of the interaction of demand and supply. Only when supply functions are perfectly elastic will changes in output purely reflect changes in demand. Orders data may correspond more closely to demand but, even here, the fit is imperfect as firms will respond to supply capacity constraints by raising prices or by not bidding for work, so that orders are affected by the state of supply as well as demand. It is impossible to avoid this problem, but nonetheless demand fluctuations are likely to have the major influence on both output and orders variations. Supply constraints, for one thing, are only binding when work loads are at or near full capacity. Measuring fluctuations in output is consequently likely to be a reasonable means of testing the hypothesis of excess volatility.

4 Volatility results

4.1 Output

Table 2.1 compares the volatility of construction output with that of two other areas of productive activity. They are the investment and intermediate goods industries, which constitute 70 per cent of

Table 2.1 Construction industry volatility compared with investment and intermediate goods industries

	Standard deviations					
	Medium-term			Short-term		
	Investment goods	Intermediate goods	New construction	Investment goods	Intermediate goods	New construction
1970Q1–1981Q1	3.8	5.9	2.3	5.1	7.1	5.1
1981Q1–1994Q4	3.7	4.2	1.9	5.7	4.8	7.7
F-tests:						
time periods	No	No	No	No	No	No
industries	No	No	No	No	No	No

Note: 'No' indicates no significant statistical difference. Industries are compared with construction.

the value added of Britain's production industries. Construction is an investment goods industry, so unfortunately there is an element of double counting, but as the results show no statistical difference between the series, this is not of major concern. Interestingly, there has been a slight, though statistically insignificant, fall in volatility for most industries shown between the period 1970–1981Q1 and 1981Q1–1994Q4 but a small rise for construction, though again statistically insignificant. Construction, conversely, saw a slight fall in its medium-term output volatility over these years to very low levels.

The data in Table 2.1 are huge aggregates, so it might well be objected that the volatility of individual construction markets is smoothed out in the aggregation. This is true to an extent, as Tables 2.2 and 2.3 show. Housing work, for example, for which long-term comparisons from 1955 to the present day can be made, is six to seven times more volatile in the medium-term than all new construction, and two to three times more volatile in the short-run (Table 2.2). There has also been a significant increase in the medium-term volatility of housebuilding in both the public and private sectors between the 'long boom' years of 1955–69 and the later period, 1970–94. The volatility of public housing output, moreover, deteriorated from the 1970s through to the 1980s/90s, though not the private sector. Results reported elsewhere also show that the UK has had exceptionally

Table 2.2 New construction output volatility, 1955–94

	Standard deviations							
	Medium-term				Short-term			
	Public housing	**Private housing**	**Total housing**	**All other work**	**Public housing**	**Private housing**	**Total housing**	**All other work**
1955–69	5.2	7.8	6.7	3.5	14.1	13.1	11.5	8.1
1970–94	9.3	13.7	12.9	2.1	19.9	15.9	14.0	7.5
1970–1981Q1	4.5	12.9	9.7	2.3	11.0	14.9	11.5	5.1
1981Q2–1994	11.9	14.5	15.1	1.9	24.9	16.4	15.4	7.7
F-tests:								
1955–69 vs 1970–94	**Yes**	**Yes**	**Yes**	No	Yes	No	No	No
1970–81 vs 1981–94	**Yes**	No	Yes	No	**Yes**	No	No	Yes

Note: 'No' indicates no significant statistical difference. 'Yes' indicates a significant difference: in bold at the 5% level, normal typeface at the 10% level. The industries are compared with construction.

Table 2.3 Output for non-housing new work – volatility of sub-categories

	Standard deviations						
	Old definitions			New definitions			
	Public other	Private industrial	Private commercial	Infrastructure	Public other	Private industrial	Private commercial
Medium-term							
1968–1979Q4	3.5	4.8	3.6				
1980Q1–1994				4.6	4.7	7.6	1.9
Short-term							
1968–1979Q4	7.0	9.7	8.2				
1980Q1–1994				13.0	6.3	15.7	14.6

Note: The categorisation of construction work has changed. Up to 1980 there is no separate category for infrastructure; hence the 'old definitions' and 'new definitions' distinction.

high housing investment volatility compared with other advanced economies. All other work, however, as a whole is relatively stable. There is even a statistically significant (at the 10 per cent level) fall in short-term volatility of non-housing work between 1970–1981Q1 and 1981Q2–1994Q4.

A breakdown of this non-housing construction work category into its three (old definition) or four (new definition) sub-components produces conflicting results. As Table 2.3 shows, the medium-term volatility of these sub-categories is still low – private commercial spectacularly so. Short-term volatility conversely is much higher for infrastructure, industrial and commercial work, though the increase hardly makes construction sub-sectors subject to uniquely volatile output fluctuations. One implication of the higher output fluctuations in individual construction sectors in contrast to the industry as a whole is that firms can reduce fluctuations in their own turnovers by diversifying across different sectors.

4.2 Orders

Not surprisingly orders are more volatile than output – compare Table 2.4 with Table 2.2 and Table 2.5 with Table 2.3. Firms in the construction industry temporally smooth the effect of the greater fluctuations in orders by varying the time between signing a contract and starting the work, and also by varying the construction time. Typically construction times as a result increase during booms, other things being equal. Infrastructure looks a particularly volatile sector (Table 2.5). To an extent this is because of the effect of the ordering of the Channel Tunnel in the late 1980s. When the quarter affected by it is removed, infrastructure's standard deviation falls to 21 per cent in the medium-term, from 34 per cent, and the short-term to 29, from 42 per cent.

The data series for orders only goes back to 1968, so it is impossible to see whether the volatility of orders has changed from what it was in the 1950s and 1960s. Comparing the 1970s and post-1981, it can again be seen that there has been no appreciable increase in volatility, apart from public housing.

Once again, however, the degree of volatility depends on the extent of aggregation. This is brought out by using an even greater degree of disaggregation of construction orders, as shown in Table 2.6. The first two columns rank each sector by the degree of the volatility of orders

Table 2.4 Volatility of new construction orders, 1970–94

	Standard deviations							
	Medium-term				Short-term			
	Public housing	Private housing	Total housing	All other	Public housing	Private housing	Total housing	All other
1970–1981Q1	24.8	26.2	20.4	7.2	23.9	25.4	21.6	12.3
1981Q2–1994	25.0	27.4	21.6	9.1	35.3	21.0	19.0	15.9
F-tests:	No	No	No	No	No	No	No	No

Note: 'No' indicates no significant statistical difference. Industries are compared with construction.

Table 2.5 Volatility of new construction orders, 1968–94

	Standard deviations						
	Old definitions			New definitions			
	Public other	Private industrial	Private commercial	Infrastructure	Public other	Private industrial	Private commercial
Medium-term							
1968–1979Q4	15.1	18.1	22.0				
1980Q1–1994				33.1	15.3	17.8	12.0
Short-term							
1968–1979Q4	20.4	23.9	22.5				
1980Q1–1994				41.7	19.0	23.3	19.4

Note: The categorisation of construction work has changed. Up to 1980 there is no separate category for infrastructure; hence the 'old definitions' and 'new definitions' distinction.

Table 2.6 Construction sub-sector orders volatilities, 1985–95 (standard deviations)

Rank	**Quarterly (%)**	**Annual (%)**	**Rank**
Railways	1028	468	Railways
Harbours	130	98	Harbours
Electricity	112	94	Electricity
Gas, comms, air	68	69	Oil, steel, coal
Oil, steel, coal	68	63	Gas, comms, air
Water	64	55	Water
Roads	51	47	Miscellaneous
Agriculture	51	47	Agriculture
Miscellaneous	42	41	Shops
Factories	36	37	Offices
Health	32	34	Roads
Garages	29	34	Factories
Entertainment	28	34	Garages
Sewers	27	31	Health
Schools & universities	25	29	Warehouses
Shops	25	28	Entertainment
Warehouses	23	27	Schools & universities
Offices	21	27	Sewers
New housing	15	19	Total
Total	13	19	New housing

Note: All data are quarterly. 'Annual' refers to the difference between the current quarter and its equivalent quarter a year previously; whereas the quarterly rank is based on quarter by quarter first differences.

between quarters, starting with the most volatile and ending with the least. The two columns on the left hand side repeat the same exercise using annual data. It can be seen that there is considerable variation in the volatility of each of these sub-sectors and that total orders are a smooth, far less volatile, aggregate consequence of much greater sectoral variation. Lower aggregate volatility is a consequence of the fact that fluctuations in sub-sectors are not coincident, highlighting the importance of covariances between sub-sectors when undertaking an exercise in minimising orders volatility.

The time period chosen also has considerable effects on the rank order of volatility, because it can be seen in Table 2.6 that it varies depending on whether quarterly or annual data are used. When ranked by the quarterly information, the quarterly data are more volatile than the annual series for eight of the nine most volatile sub-sectors and are smaller for those of lower volatility. The rank order changes noticeably between the quarterly and the annual data but most of the change arises from within narrow ranges of volatility differences and can be attributed to white noise within the data as much as to any fundamental change in the behaviour of the two series. The two major exceptions to this generalisation are offices and shops whose annual ranking indicates much greater relative volatility than the quarterly data. This is probably accounted for by the commercial building boom of the late 1980s and the ensuing slump which stands out more starkly in the annual series than the quarterly ones.

4.3 Prices

So far only fluctuations in the volume of construction activity has been considered. Price fluctuations give some additional information. The volume data considered above are derived by deflating current price data by what the government construction statisticians call 'output price indices' (Construction Statistics, National Statistics, London). They are meant to approximate the price charged by construction firms for the work they do, and so include wages, materials, sub-contract payments, financing costs and profit-margins. Private housing is the odd one out as it is based on an amalgam of house prices, materials and wage costs. Since the early 1980s, the housing output price index has behaved distinctly because of this. If these output price series are deflated (by the GDP deflator here), it gives some indication of real returns to the industry and the costs to its consumers over time. Price volatility measures the temporal sensitivity of those prices.

Table 2.7 summarises the evidence. Again it is mixed. Medium-term price volatility did increase for all non-housing work from the 1970s onwards compared to the earlier period. But when the 1970–81 period is compared with the time after that, volatility actually 'fell' significantly. There seems to have been a sea change in the volatility of output and prices after 1981. Before then price volatility was greater than output volatility, whereas after the early 1980s output became the most volatile.

What could have caused the change? There are two plausible explanations. First, construction may be becoming more like other modern oligopolistic industries whereby fluctuations in demand are borne

Table 2.7 Changes in real construction costs: new output price volatility

	Standard deviations					
	Medium-term			Short-term		
	Public housing	**Private housing**	**All other**	**Public housing**	**Private housing**	**All other**
1955–69	7.3	11.3	7.8	7.3	10.7	7.7
1970–94	10.1	10.4	11.8	7.2	7.4	8.1
1970–1981Q1	13.3	12.9	14.2	8.7	9.0	9.1
1981Q2–1994	6.5	7.6	9.4	4.5	5.3	5.9
F-tests:						
1955–69 vs 1970–94	No	No	**Yes**	No	No	No
1970–81 vs 1981–94	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

Note: 'No' indicates no significant statistical difference. 'Yes' indicates a significant difference: in bold at the 5% level, normal typeface at the 10% level. The industries are compared with construction.

more in quantities than in prices. However, the third or more average real fall in the real output price indices since the 1989 peak of the last boom does not suggest that this explanation is likely. The reduction in price volatility may instead result from special characteristics of the 1970s, with two short-lived booms during the decade, and through the changing institutional structure of the industry since the early 1980s. The far greater use of sub-contractors since the late 1970s may enable main contractors, from whom the output price indices are principally derived, to impose the effect of short-term price variations onto sub-contractors through variations in payments and their timing. In the medium-term, conversely, they have to bear the impact of recessions as well in their tender bids.

4.4 Summary of volatility results

Taking stock of the conclusions so far:

- Construction does not have exceptionally volatile output fluctuations compared with other UK industries;
- the volatility of individual sectors is greater than for the industry as a whole;
- orders are more volatile than output;
- there has not been an increase in output volatility over the post-war period, with the exception of the now small public housebuilding sector. Construction output volatility is what it is, and does not seem to be greatly affected by different government regimes;
- the volatility of construction output prices has decreased significantly since the early 1980s.

Our results suggest that short-term fluctuations in aggregate construction demand are not exceptionally high. Similar, if less detailed, conclusions were reached in the mid-1970s and 1980s (see the survey in Ball, 1988). There, however, is much greater volatility at the level of individual sectors. The case for public policy to smooth construction demand is weak on this evidence, at least for Britain though similar results probably hold elsewhere; yet, the importance of large construction firms having a strategy towards sectoral fluctuations looks strong. The next section considers some statistical evidence on the form that those strategies should take.

5 A model of orders variance minimisation

For simplicity, it is assumed that construction firms are interested in the volatility of orders alone. The objective of a firm in this model is to minimise fluctuations in its total orders by adopting optimal proportions of work across construction sectors. This is called here an orders variance minimisation (OVM) strategy. Again, for simplicity, only workloads in the UK are considered. 'Sector' here is defined as the categories of disaggregated work shown in the orders data that were used in the earlier analysis, listed in Table 2.6.

The optimisation technique used for minimising the variance of a portfolio of orders is a Markovitz model of variance and covariance terms for each sector. The expression for a portfolio p consisting of two sectors i and j is:

$$\sigma_p^2 = X_i^2\sigma_i^2 + X_j^2\sigma_j^2 + X_iX_j\sigma_{ij} \quad (1)$$

where X_i and X_j are the orders won in sectors i and j respectively, as a proportion of total orders; and the σ items are the respective variances and covariances. It can be shown from this formula that as the number of terms in X increases towards infinity, so the variance of the portfolio converges towards the average covariance of the elements in the portfolio. It is this property that leads to a diversified portfolio in general having a lower level of risk than a less diversified portfolio.

Equation (1) is initially minimised with respect to a constraint:

(i) $\sum_{i=1}^{n} X_i = 1$

where n is the total number of sectors, so that the sum of sector orders has to be equal to total orders – an obvious, but necessary, constraint.

After an initial run of the model, further restrictions are then imposed to test how resources would be allocated when firms are required (1) to stay in a sector or (2) to maintain a presence within a fixed range of the average amount of total construction orders derived from that sector. To these ends, restrictions were imposed upon the feasible values that an optimal solution may take. To capture the need to stay in any given sector it was assumed that constraint (ii) holds:

(ii) $X_i \geq 0$ for all i in the portfolio p

For the second scenario, the restriction was set that any X_i must fall within one standard deviation of its mean value. The reason for this is to constrain the solution to something approximating to its real world average value. Unconstrained solutions may result in all firms doing no work in the most volatile sectors, which is obviously not a feasible outcome. Instead, restricting the solution to within one standard deviation of the mean prevents obviously unrealistic solutions from appearing. As a basis for comparison, the results for the various models are compared with a 'mean' strategy based on achieving the sectoral average share of total orders.

6 Simulation results

6.1 Simulations across the three main sub-sectors of new work

Initially, the OVM procedure was tested using data for the three largest subdivisions of construction: housebuilding, civil engineering and other building.

Part of the volatility of the civils series used in these simulations arose from one major order for the Channel Tunnel in 1987; a similar effect arises for the civils sub-sector of railways considered later. The data were not adjusted for these events because it is to be expected that periodic shocks of large orders are a fact of life in these construction sectors, with the Channel Tunnel being one of an exceptionally high order of magnitude. It was consequently felt that to remove it from the data would unduly smooth out these series given the objective of the exercise.

The results of the unconstrained simulation are summarised in Figure 2.1. As can be seen, the unconstrained variance minimising result differs considerably from the average proportions of work in each sector during the sample period. In particular, the volatility of the 'other building' sector is much higher than for the others with the result that a firm wishing to minimise its exposure to order variability risk, according to the portfolio method, ought to have a lower than average share of its activities there. In fact, an unconstrained OVM approach recommends a negative share for 'other building' of −5 per cent – an impossible result.

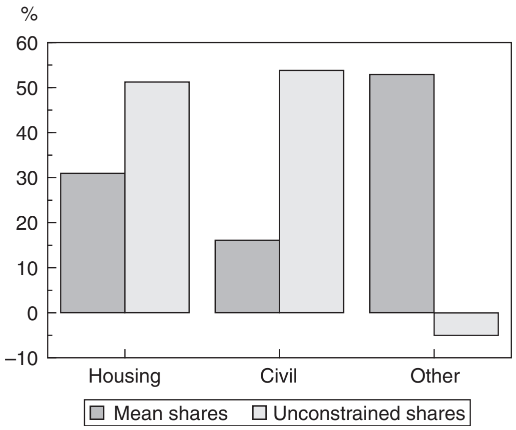

Figure 2.1 Actual versus unconstrained optimising output shares: the average share of total new work compared with the turnover variance minimising shares

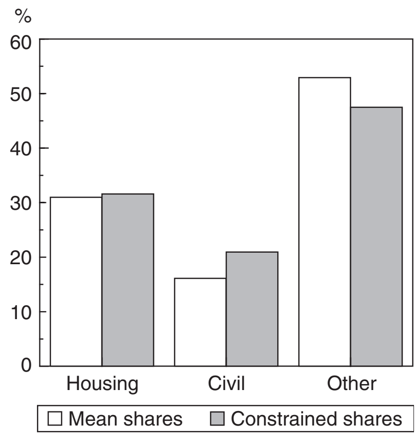

Figure 2.2 Actual versus constrained optimising output shares: 1 sd band constraints

Figure 2.2 shows the results for a more restrictive scenario that constrains a firm to having a presence in all sectors and to vary its sectoral output by one standard deviation either side of that sector's mean share in total orders. In this case, the shift away from 'other building'

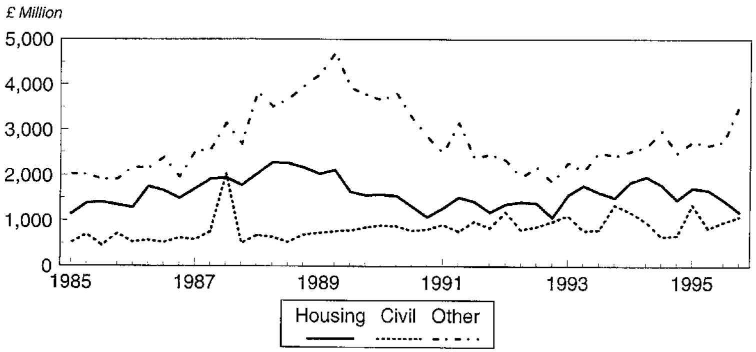

Figure 2.3 Orders: civil, housing, other 1985 to 1995

is far less than the unconstrained case. The high variance of 'other building' is clear when time series for orders are examined for the sample period 1985–95 (Figure 2.3). 'Other building' shows the greatest variance especially because of the boom of the late 1980s and the subsequent downturn. Once again, it would seem unwise to exclude this boom period on the grounds that such infrequent booms and slumps are a fact of life in construction. Although Table 2.1 shows that, for the quarterly data used in the simulation, offices and shops ranked second and fourth least in terms of sub-sector orders volatility, their correlation is highly positive (0.61) intensifying rather than offsetting their joint fluctuations.

The two simulations reported so far show how sensitive the results are to the assumptions used about relevant market shares in sectors. Nevertheless, some important general conclusions can be made about general firm behaviour and the recent strategies of large UK construction firms. First, the OVM optimising approach does suggest that the expected volatility of orders can be minimised by diversifying. It is unreasonable, therefore, to identify construction as an industry prone to excessive orders volatility simply by reference to the variance of orders in one sector alone. Second, firms that specialise in housing or civil engineering alone will have far less volatility in their orders than those concentrating on 'other building'. Although focusing on housing does not represent the optimal way of minimising orders fluctuations, it is still a reasonable smoothing strategy, especially when combined with the need for specialist land and housing market expertise. Housing is the least volatile of the 20 sub-sectors shown in Table 2.1. Smaller housebuilding firms, in particular, are likely to benefit from sectoral specialisation because of the reduced overhead costs implied by such a strategy; whereas larger firms should be more able to reap economies of scale and scope and hence gain greater benefits in turnover stability from diversification. The result conforms to general observation of the UK construction industry as a whole, where there are more housing and civils specialists than commercial building ones. Bovis stands out as the exception with its high concentration in commercial building. However, Bovis probably has uniquely sunk cost advantages in this sector, with its strong brand name and track record, which may outweigh the potential volatility disadvantages of its specialisation.

Third, the results bring into question some of the recent strategies of several major firms in completely withdrawing from the UK housing market. A characteristic observed from these simulations is that the lowest variations in turnover occur for firms that concentrate on housing and civils, and this result corresponds to the traditional strategies and behaviours of the UK construction majors, such as Wimpey, Laing, Costain and Tarmac. Some of them, however, have in recent years withdrawn from the housing market after severe losses in private housebuilding – while Wimpey divested itself of all but its housebuilding operations. Those losses were historic ones, associated with management strategy mistakes in the land market and the scale of expansion plans during the boom years of the late 1980s rather than from a presence in market itself. Such past mistakes give no indication of the future performance of the housing market when strategies are better managed. Some of these firms may consequently in the future come to regret the greater turnover volatility they have imposed on themselves by withdrawing from housing.

6.2 Simulations across civil engineering sub-sectors

A similar exercise was taken by focusing on civil engineering sectors alone, and similar general results were derived. Several variations of the simulations are shown in Figures 2.4, 2.5 and 2.6. Figure 2.4 compares an OVM simulation with the restriction that firms can only vary their activity by one standard deviation from the mean sector shares

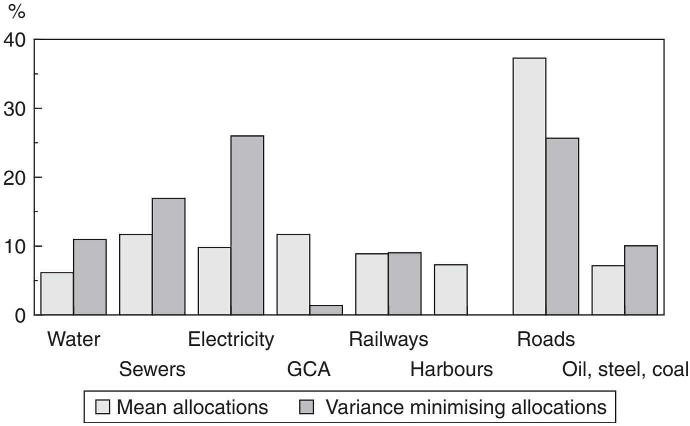

Figure 2.4 Civil engineering: variance minimising versus mean allocations within 1 sd constraint bands, 1985–89

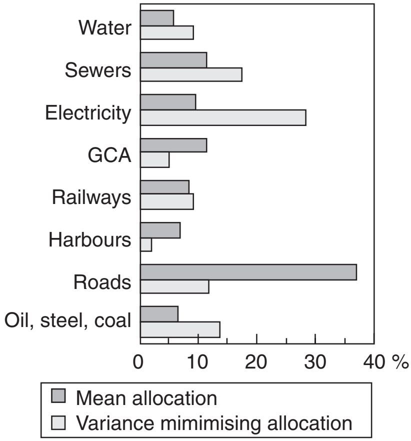

Figure 2.5 Variance minimising versus mean shares with maximum constraints

of total civil engineering orders. Figure 2.5 imposes a somewhat different constraint: firms can have a no greater share of an activity in their portfolio of work than the highest actual quarterly share of total civils work recorded for that type of work. Figure 2.6 finally compares

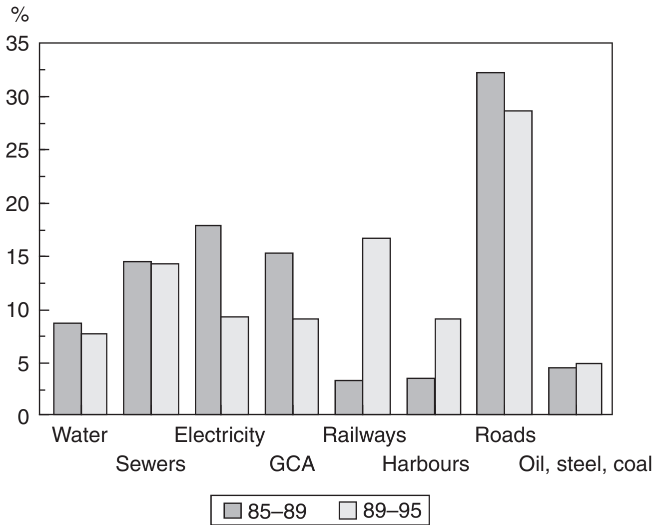

Figure 2.6 Comparison of optimal allocations: 1985–89 vs 1989–95: constrained within 1 sd bands

the stability of the simulations in Figure 2.5 by breaking the time period in two, 1985–89 and 1989–95, roughly the upswing and downswing phases in the overall construction cycle. The sensitivity to the two alternative time periods however should not be exaggerated as the prime result of the difference in the sectoral distributions is the railways sub-sector and the post-1989 impact of the Channel Tunnel on it.

Once again, a strategy based on tracking the mean shares of work is consistently sub-optimal in minimising total orders volatility. In these more highly disaggregated models, the effect of the Channel Tunnel now matters in determining the optimal distribution of work. For instance, the variance of the railways sub-sector is now much more crucially affected by this single large order. Exception peaks and troughs can also be observed in other sectors. Figure 2.7 highlights the one standard deviation bands for road building for the quarters of 1989 to 1995, and it can be seen that the years 1994–95 were exceptional. Isolating the exceptional 'shocks' from the more routine cyclical variations is one of the benefits of the approach suggested here.

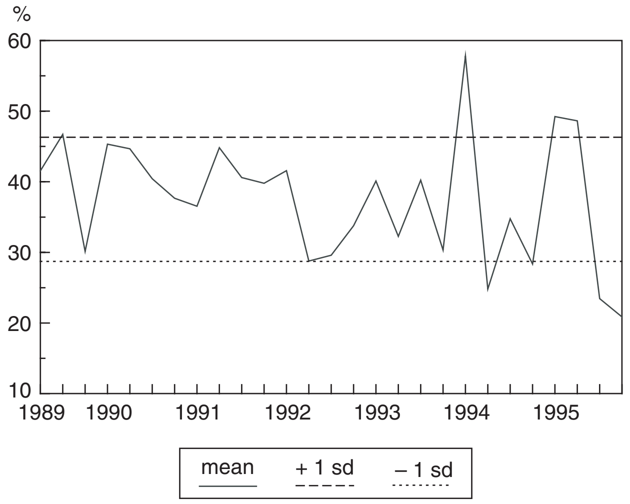

Figure 2.7 Example: actual road building 1989–95 and 1 sd bands

7 Conclusion

This chapter has argued and presented evidence that construction as a whole is not an industry suffering from particularly volatile demand. Nonetheless, construction firms can benefit from strategies of having diverse medium-term presences in construction markets because they minimise fluctuations in the total orders they win. The need for firms to adopt strategies towards orders volatility arises from the existence of sunk costs in the industry. These sunk costs require firms to develop medium-term presences in particular markets, but are not so great that they inhibit cross-market diversification.

Simulations reported above for the three main divisions of new construction work and for civil engineering, using data for British construction markets, indicate the importance of the initial assumptions on which the simulations are based. They suggest that the switch out of housebuilding by several major UK contractors will probably lead to greater turnover volatility for them. They also show that following the herd by adopting market presences that mirror the average distribution of work to a sector is a poor strategy when its comes to turnover smoothing.

This exercise has been a relatively simple one of generating model simulations of orders smoothing through sectoral diversification. Several limitations to its usefulness as a strategic tool should be noted in conclusion. The levels of disaggregation used have been limited by the available published data; better disaggregated information would make for more effective strategic management analysis. The study took no account of either the levels or the variances of the relative profitabilities of individual sectors. An argument was made about why turnover volatility was important for profitability but that does not necessarily translate into a claim that construction as a whole is afflicted by highly unstable demand. The characteristics of construction indicate the importance of construction firms adopting relatively sophisticated techniques when formulating their business strategies. Special pleading to government as an alternative strategy, in contrast, seems doomed by the data as much as by any perceived public expenditure constraints.

8 Appendix: trends, cycles and data

Investigating cycles and trends

The objective is to take long time series data and try to decompose it into seasonal, cyclical, trend and irregular components. A problem is that most attempts to separate out these components from data lead to biases and the spurious creation of cycles. In recent years some techniques have emerged which go a long way to eliminating, or at least minimising, statistical bias. Another problem is that all de-trending methods require judgement by the researchers on what is an acceptable decomposition.

The technique used here has been developed in the last few years. It is based on the Kalman filter and is called Structural Time Series Analysis. It is available as a PC package called STAMP-structural time series analyser, modeller and predictor (Harvey, 1989).

Starting from the premise that an annual time series can be decomposed into two unobserved components, the trend and cycle, the basic statistical model takes the form:

$$y_t = \mu_t + \psi_t + e_t \tag{A-1}$$

where y_t is the observed variable, μ_t its trend, ψ_t its cycle and e_t a white noise error term. The trend and cyclical components are state variables which are allowed to vary stochastically with time, unlike the deterministic models of standard OLS regression methods.

The Kalman filter is a recursive technique which estimates the optimal state of a system (state space models). Starting with initial estimates of the parameters and variances of the system summarised in equation (A-1) the estimates are revised recursively, moving through time, adding additional observations and estimating the optimal position of the trend and cyclical components at time t given the information available at time $t - 1$. Thus at time T the position of the state variables are estimated based on all observations up to that point in time, and then estimates are found for $T - 1$, $T - 2$ and so on. It is these smoothed estimates which we have presented in the form of graphs.

The variances of the level, trend and cycle are subjected to significance tests. If the variance of the cycle is found to be insignificantly different from zero, then it is removed and the variance of y_t is captured entirely by the variance of the trend and the error term. If the variance of the trend is found to be insignificantly different from zero, then the approach resorts to a deterministic trend model, which has been estimated using conventional Ordinary Least Squares techniques.

References

Ball, M. (1988), *Rebuilding Construction*, London, Routledge.

Ball, M. and P. Antonioni (1996), 'Is the UK Construction Industry Afflicted by Exceptionally Volatile Demand?', in S. Gruneberg (ed.), *Responding to Latham: The Views of the Construction Team*, Englemere, Chartered Institute of Building.

Ball, M., T. Morrison and A. Wood (1996), 'Structures Investment and Economic Growth: A Long-run International Comparison', *Urban Studies*, 33, 1687–1706.

Ball, M. and A. Wood (1996), 'Trend Growth in Post-1850 British Economic History: The Kalman Filter and Historical Judgement', *The Statistician*, 45, 142–52.

Betts, M. and G. Ofori (1992), 'Strategic Planning for Competitive Advantage in Construction', *Construction Management & Economics*, 10, 511–18.

Constable, J. (1986), 'Diversification as a Factor in UK Industrial Strategy', *Long Range Planning*, 19(1), 52–8.

Copeland, T. and J. Weston (1988), *Financial Theory and Corporate Policy*, Reading, MA, Addison Wesley Publishing.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, Cambridge University Press.

Hillebrandt, P.M. and J. Cannon (1990), *The Modern Construction Firm*, London, Macmillan.

Hillebrandt, P.M., J. Cannon and P. Lansley (1995), *The Construction Company In and Out of Recession*, London, Macmillan.

Lansley, P.R. (1987), 'Corporate Strategy and Survival in the UK Construction Industry', *Construction Management & Economics*, 5, 141–52.

Levy, H. and M. Sarnat (1984), *Portfolio and Investment Selection: Theory and Practice*, Englewood, NJ, Prentice Hall.
Sutton, J. (1991), *Sunk Costs and Market Structure*, Cambridge, MA, MIT Press.
Williamson, O. (1985), *The Economic Institutions of Capitalism*, New York, Free Press.

3
The 1985–95 Cycle in Real Estate Markets: Bubble or Shock?

Didier Cornuel and Francis Calcoen

1 Introduction

Over the 1985–95 period, the French real estate sector has probably undergone its most outstanding cycle. A similar phenomenon has been experienced by most industrialised countries. This cycle came forward as an explosion in prices and volumes, together with a considerably increased debt for the agents involved in the property market: investors, property developers, the so-called 'marchands de biens'[1] and households. This first stage was followed by a stage of decreasing prices and volumes, leaving the operators with stocks which they could no longer work down, and which fell in value as prices decreased. As these operators were unable to discharge their debts, the lending institutions have been obliged to set up reserves for doubtful or irrecoverable debts, leading some of them to go through a bad patch or to enter into liquidation.

This cyclical phenomenon can be illustrated by the joint movement of prices and volumes[2] in the older housing market in Paris, the only area for which such data are available (Figure 3.1). The last cycle can be considered as beginning in 1985. We shall not decide whether this cycle is actually to be overcome; however this point will be considered below.

A major question with which the operators, and consequently the lending institutions, have to cope is the estimate of the assets at the basis of loans. If this estimation aims at indicating market value, the actual market price cannot be used alone. Indeed, as selling requires time, sometimes up to several years, it is necessary to

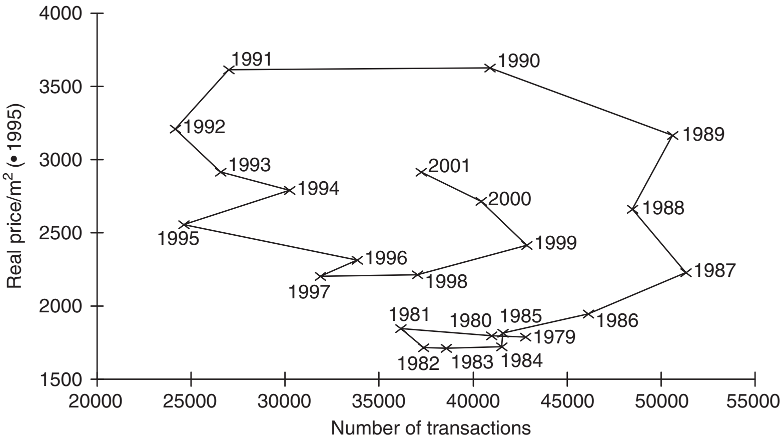

Figure 3.1 Second-hand flats in Paris

know whether the actual price is long-lasting. In other words, is it an equilibrium market price, is it to go on decreasing or is it to recover?

Answering this question requires to identify the economic nature of this phenomenon. Indeed, according to the nature of the cycle, the price at the end of the cycle, that is the equilibrium market price, will differ. It may be higher than the price at the beginning of the cycle in 1985. It also may be equal, or even lower. For example, if the price at the end of the cycle appears to be lower than at the beginning, this would mean for the operators that the worst is still to come. A contrario, it will be seen that the price at the end of the cycle is an indication of the nature of the cycle.

Different hypotheses can be put forward concerning the nature of the cycle. They are characterised on the one hand by the mechanism at the origin the cycle and, on the other hand, by the type of expectations. The mechanism may be a bubble, that is a purely speculative phenomenon, or a cycle resulting from a shock. One or several patterns of expectations can be associated to these mechanisms. Before dealing successively with both hypotheses about the mechanism, we shall present the links between the different dimensions of property business.

2 Property market, service market and building market

Figure 3.1 deals with transactions on the existing housing stock. These are transactions on property, that is on housing asset, either owner-occupied or rented, but not blocks of flats. Housing asset usually provides incomes, rents, effective or imputed. The financial theory argues that when there are competitive markets, the value of an asset is equal to the discounted value of the expected incomes:

$$p_t = \sum_{n=0}^{N} \frac{x_{t+n}}{(1+r)^n} + \frac{p_{t+N}}{(1+r)^N}$$

where p_t denotes the asset price at time t, x_{t+n} the expected rent for the period $t+n$, r the discount rate, often considered as the rate of return of a riskless asset, N the term, p_{t+N} the expected resale price at the date $t+N$. If N tends to infinity, this gives:

$$p_t = \sum_{n=0}^{\infty} \frac{x_{t+n}}{(1+r)^n} \qquad (1)$$

This price is the so-called 'fundamental value'. If an annual variation of rents is expected at a constant rate $x_{t+n} = x_{t+n-1}(1+\alpha)$ and if N tends to infinity, the price will grow at the same rate as rents:

$$p_{t+n} = p_{t+n-1}(1+\alpha)$$

The global return of the asset at the period t, r_t, can be broken up into two terms, the current return and the added-value rate:

$$r_t = \frac{x_t}{p_t} + \frac{(p_{t+1} - p_t)}{p_t}$$

When N tends to infinity, while r_t remains constant, the price tends to:

$$p_t \to \frac{x_t}{(r-\alpha)}$$

In other words, when there is equilibrium in the financial asset and property markets, the returns of financial and real estate investments are equal if there is no risk:

$$r = \frac{x_t}{p_t} + \alpha \qquad (2)$$

Then, the return of real estate investment is equal to the current return, constant over time, plus the rent and price growth.

These equations relating the property market to the service market assume equilibrium markets and imply to have information on the development of markets and consequently of prices in the future. Then, knowing the expectation nature is a crucial issue.

3 Rational bubble hypothesis

We deal first with the mechanism of rational bubble, which implies a hypothesis about the nature of expectations. Next, we examine the quantity variations within a bubble and finally we compare the theoretical predictions with the phenomena observed.

3.1 The rational bubble mechanism

There is a bubble when an asset price moves away from its fundamental value. This value is estimated by the future receipts rationally expected.

In rational expectations the expected value is the mathematical expectation of the variable, given the information available at the date when the anticipation is expressed. This is written: $P_t^e = E(P_t/I_{t-1})$, where P_t^e represents the expected price for period t at period $t-1$, P_t the price at period t, I_{t-1} the information available at date $t-1$ and E the mathematical expectancy. Available information includes all the previous values of variables likely to condition the formation of the expected variable, as well as the knowledge of the price formation mechanism, that is the knowledge of the supply and demand curves and of their determinants. This means that expectations are perfect on the average, in the sense that there cannot be any systematic error. Indeed, if the agents who formulate these expectations notice a systematic error, they include it in their information and correct their expectations. Then, error of expectancy can only be at random, with a zero mathematical expectation, without auto-correlation.

However, there can be deviation between the price and the fundamental value on the basis of rational expectations. This is the phenomenon of rational bubble. The following account comes from Blanchard and Watson (1984).

The asset return is equal to the sum of income plus the expected capital gain:

$$r_t = \frac{(p^e_{t+1} - p_t)}{p_t} + \frac{x_t}{p_t}$$

where r_t is the total return at period t, p^e_{t+1} the expected price for period $t + 1$ and x_t the current rent of period t. The equilibrium price can be expressed as a function of the rent and expectations by considering r_t as given and equal to the return of riskless financial investments, noted r and assumed to be constant. This gives:

$$p_t = \left(\frac{1}{1+r}\right) p^e_{t+1} + \left(\frac{1}{1+r}\right) x_t \tag{3}$$

Assuming expectations to be rational, which means that $p^e_{t+1} = E(p_{t+1}/I_t)$, then the price is:

$$p_t = \left(\frac{1}{1+r}\right) E(p_{t+1}/I_t) + \left(\frac{1}{1+r}\right) x_t \tag{4}$$

This recurrence equation can only be solved by a recursive process up to infinity (see Appendix). One particular solution is:

$$p^*_t = \sum_{n=0}^{\infty} \left(\frac{1}{1+r}\right)^{n+1} E(x_{t+n}/I_t)$$

This solution expresses the fact that p^*_t is the discounted sum of the expected rents in the future. It corresponds to the fundamental value expressed in equation (1) with rational expectation of rents.

Another solution is $p_t = p^*_t + z_t$, with:

$$z_t = \left(\frac{1}{1+r}\right) E(z_{t+1}/I_t) \tag{5}$$

The value of the variable z_t only depends on its value expectations in the future. Then this variable is characterised as a bubble, since it leads to a deviation of the asset value from its fundamental value.

In such a conceptualisation, the existence of a bubble is not related to any exogenous factor. The bubble is a purely endogenous price mechanism. Effective market trade represents a situation of equilibrium between supplied and demanded quantities. The supply and demand variations are only determined by the price. As a bubble is not induced by any exogenous factor, its occurrence is unforeseeable. If a bubble is unforeseeable, it is also insuperable. Indeed, it results neither from an information gap nor from any real phenomenon, for example a shock. Consequently, it was possible neither for public authorities nor for any other agent to intervene to prevent the bubble.

It has been considered that there is a bubble in case of a gap between the actual price and the fundamental value. However, if the fundamental value expresses rents and if only expected rents allow to determine the fundamental value, it is a priori difficult to identify a bubble. It is necessary to observe actual rents or to know the investors' expectations – supposed to be rational – to estimate the fundamental value and consequently measure any possible price deviation from it.

Unforeseeable, insuperable and probably impossible to be identified while it is developing, the bubble mechanism may awkwardly occur in the asset markets. Moreover, even if a bubble is a purely endogenous price mechanism, in reality such a mechanism is often accompanied by high variations in trading, as shown by Figure 3.1. Such variations need now to be explained.

3.2 Volume variation within a bubble

With regard to reproducible assets, as real estate, Blanchard and Watson (1984) refer to an explanation proposed by Poterba (1980). In the case of real estate, the volumes are endogenous and then modify the fundamental value. Here is how the mechanism of real estate pricing is stated:

'Real estate goods consist in two parts: the plot of land and the structures. The supply of land for building is a growing function of land price. The structure supply is inelastic in the short run, elastic in the long run. In a situation of long-term equilibrium and in the absence of bubble, home price equals the present value of the expected services, of rents in the future. At the same time, it must be such a price that the property stock remains constant; more precisely, the difference between property price and land price, that is the structure price, must be such that real estate investment equals the depreciation of the existing stock.'

'Let us assume now a bubble arising in this market, and buyers ready to pay a price higher than the fundamental value. This higher price implies a higher price of structures, then a higher rate of production and finally a larger property stock. This larger expected stock, combined with an unchanged demand for property services, implies the expectations of lower prices for these services and lower prices of rents. Such decreased expected rents reduce the present value of rents. Then the bubble results in an increased price and a decreased fundamental value.'

'What occurs over time? When it is not a stochastic bubble, property price goes on increasing and the fundamental value goes on decreasing. If land supply becomes increasingly inelastic, the property stock growth slows down. The scenario is similar when the bubble is stochastic. When the bubble bursts, the price falls down far below its initial level because of the excessive size of the property stock.'

Then, the volume movement consists of a phenomenon of production aiming at fitting the speculative demand resulting from the price dynamics. This explanation only holds for reproducible assets. It does not account for the growth of transactions in a market exclusively made of stock as the older property market or the market for some works of art. In other words, it does not give any explanation for what the price and quantity path represented in Figure 3.1 refer to the following analysis is suggested: as the reasoning is based on an equilibrium market, supply and demand are equal. Moreover, in a bubble mechanism, quantities only change under the mere influence of price. In other words, the supply and demand curves do not shift; however, there are shifts along these curves. Consequently, supply and demand curves are the same and the price-quantity path represents both demand and supply curves. This is not surprising insofar as the same agents most often intervene on the supply and demand sides.

Thus, under the assumption of a bubble, the price-quantity path on the market of older apartments in Paris is the demand curve and the supply curve at the same time. And as both curves are the same, this means that the run of one of them is different from what is usually expected. When the slope is positive, the demand run is unusual with a positive price-elasticity. The sign of this elasticity expresses the fact that households buy more because they expect increased prices; and this increase does actually occur because their expectations are rational. As it is a cyclical phenomenon, this is a short-term characteristic. In the long run, it may be the usual behaviour of negative price-elasticity. When the slope is negative, the price-elasticity of supply is unusual. When prices are increasing, housing holders restrict their supply.

However, a crucial issue remains to be coped with: to find out whether the phenomenon of upsurge in real estate values followed by their fall, in France and abroad over the decade 1985–95, results from a bubble phenomenon.

3.3 Was there a bubble in the property markets?

This is a frequent assertion (see Renard, 1993; Revue d'Economie Financière, 1993; Granelle, 1996). For France, only Nappi (1994) provides us with a test the relevance of which will be discussed further on. In other countries the method to determine the fundamental value starts from the hypothesis of rational or perfect expectations. The estimation of the fundamental value is based on macro-economic variables, and not on rents. It is assumed that the actual prices before the explosion in prices represent the fundamental value, and they are adjusted to the exogenous variables to determine the parameters used to estimate the fundamental value when prices shoot up. It can be noticed that prices move away from the fundamental value. Deductively, one says that there is a bubble. Such is, for example, the approach adopted by Case (1986) who uses the growth of employment, the growth of population, the interest rates, the income, the construction costs and a certain number of other variables as fundamental determinants. He states that housing price in the region of Boston, Massachusetts, should have increased by 15 per cent between 1983 and 1986 according to these variables, while in fact it practically doubled. This is also the method used by Noguchi (1991, quoted by Aveline, 1995) for Japan.

Let us notice first that this approach is correct only if the market price previous to the price explosion is an equilibrium price. In matters of housing, there is no reason why it could not be the case in most of the developed countries. On the other hand, the situation may be different in the case of offices. In that respect, Renaud (1995) noticed that in the various countries hit by the explosion in home prices, this explosion was all the more strong since it occurred simultaneously to a deregulation of the real estate markets. Such is especially the case of the Nordic countries and, in France, of the Paris region where the agreement by public authorities required to build offices was abolished in 1985. This leads us to suppose that the real estate demand was rationed before the explosion in prices. In that case, the rationing price was superior to the competitive equilibrium price. Deregulation would have allowed supply to meet this rationed demand. In such circumstances, the price should have fallen down; the fact that it increased can give substance to the bubble hypothesis.

On the other hand, this method to determine the fundamental value does not take the external demand into consideration. However, according to Renaud, the demand shock would have come from Japan. Thus, it is not sure that the fundamental value determination mode is accurate.

Finally, this method is relevant only if the model is relevant. The poor adequacy of the model to the phenomena observed can be interpreted as due to a bubble or to the modelling inadequacy. This model would not represent the short-run dynamics, but the long-run dynamics instead. The difficulties of macro-economics to account for both long-run and cyclical dynamics within a single model can explain why this has not yet been reached in real estate models.

A simple method allows to point out the existence of a bubble, by looking at the price at the end of the cycle. Indeed, Poterba's analysis shows that in the case of a bubble, the price at the end of the cycle is not equal but inferior to what it was at the beginning, since the upsurge in prices led to the production of new housing without any relation to demand, finally resulting in lower prices. In other words, a bubble could be identified when the price at the end of the cycle can be observed. For some authors, the cycle is actually ending. However, the actual real estate prices are slightly higher than their 1985 levels. This goes against the existence of a bubble. This price at the end of the cycle does not imply that there was no bubble. It rather means that if there was a bubble, another phenomenon has contributed to increase prices. And this phenomenon must have been powerful enough to more than compensate the fall in prices which should have led the bubble to burst. And vice versa, it means that if the real estate values in Paris are undergoing a bubble, they still have to decrease.

The most satisfying method to test the bubble hypothesis consists in estimating the fundamental value directly from rents.[3] Relation (2) can be written as $p_t = x_t/(r - \alpha)$, giving the fundamental value at a date t as a function of the current rent and of the discount and growth rates of the expected rents. The calculation is based on the

Table 3.1 Values of housing assets under rational expectations

	1979	1984	1988	1992	1996
Average rent €/m²/year	34.17*	73.07	102.75	131.11	148.91
$\bar{r}/t - \bar{\alpha}/t$ (rational expectations)	0.0406	0.0438	0.0418	0.0413	0.061***
calculated fundamental value	842	1668	2456	3172	
observed price €/m² (by notaries)	806	1245	2259	3036	2348
calculated fundamental value/observed price	1.045	1.340	1.087	1.045	

* 1978 rent.
** Source OLAP, observatory of the rents in the Parisian agglomeration.
*** calculated value, see the text for explanations.

data available for Paris from the National Housing Surveys (Table 3.1). The income x_t is the rent in Paris of privately rented housing, r is the yield of public sector bonds, and α the growth rate of rents. As we are on the assumption of rational expectations, the values of r and α chosen for each date are the averages of the actual values from the date considered. The following table gives the values of these variables at the dates of the Housing Surveys.

The results show that, as a whole, rents correctly account for prices, and generally speaking, that prices are slightly below the fundamental value. In particular, the level of rents does account for the 1992 peak prices. The underestimation is more important for 1984; this tends to show that there was a price-lag to fulfil at the beginning of the cycle. A slight part of the price progress is due the decreased discount rate that mechanically increases the values of assets. Such a decrease results itself from the diminished interest rates almost fully compensated by the lower growth rate of rents.

The investors' expectations for 1996 are still unknown, because the data of the 2000 Housing Survey are not yet available; so the fundamental value cannot be estimated. However, the reverse calculation can be done: it is possible to determine the expectations of the rent trend by assuming that the present price does correctly reflect them. Thus, the discount rate is equal to 0.061. If the rate of return on alternative investment remained at its 7.5 per cent level, the expected growth rate of rents would be 1.4 per cent. Actually, the average rate from 1996 up to 2001 has been a bit higher, 1.6 per cent. So the 1996 price underestimated the rise of rents and it had to grow at an accelerated pace to offset this underestimation. And so it did.

However, it cannot be expected that rents and prices vary at the same rate. Indeed, there are market defects. First, rents are contractually determined by rental leases for three, six or nine years during which rents are most often indexed to the building price index (ICC, Cost of Construction Index). Then, rents can increase by more than the ICC only when new renters move in, and even then, under certain conditions. Therefore, leases act as a brake upon rent variations, upward as well as downward variations. By contrast, housing prices can increase as soon as an owner puts real estate on sale. Thus, real estate prices are logically more volatile than rents.

Second, the period 1981–86 was marked by a strong public intervention on rents through the implementation of the 'Quilliot Act' limiting their increase. The question is then to know whether the fundamental value calculation is based on the regulated rents or whether the landlords also anticipate the changes in law.

Finally, the bubble hypothesis must not hide the fact that there really was an explosion in rents. For the whole of France, rent increase was 20 points higher than inflation over the last decade. As a consequence, even if there was a bubble, it would have occurred at the same time as a rent cycle. Such a cycle cannot be due to a bubble, it results from a real phenomenon. And as bubbles are random phenomena, it would be surprising to have had both together: a bubble and a real phenomenon, unless it can be shown that the real phenomenon was generated by the bubble. Therefore, the fluctuation was caused by a real event, that is a shock. But what is the nature of this shock? This problem is relevant whatever the expectation hypothesis. The nature of the shock must be identified. We will deal with this issue in the next part.

On the other hand, in the bubble model as in any dynamic model, the price formation model – in this case, the bubble phenomenon – and the (rational) expectation scheme are both simultaneously tested. As the bubble hypothesis is null, one does not know whether to get rid of the expectation scheme, the model, or both. Then it is possible to consider the bubble hypothesis together with another expectation scheme.

Another expectation hypothesis is the hypothesis of naive expectations. On this basis, Nappi (1994) calculated the fundamental value to point out the existence of a bubble. However, naive expectations are by definition partly erroneous expectations. In other words, the fundamental value anticipated from naive expectations is not the true fundamental value. By definition, any scheme of naive expectations leads to a deviation of the actual price from the fundamental value based on these expectations. Then, the gap between the price and this fundamental value is due to expectation errors and also possibly to a bubble. In such a case, a bubble means that the gap between the actual price and the fundamental value has two sources: first, the gap between the true fundamental value and the expected fundamental value based on naive expectations; second, the gap between the actual price and the expected fundamental value. The issue can be written as an equation in two unknowns, giving the price (known) as the sum of the fundamental value and one gap. That shows how it is impossible to decide on the existence of a bubble in a scheme of naive expectations when it is not an endogenous scheme. Indeed, the anticipation scheme must be given a priori to estimate the fundamental value – as Nappi does by choosing Koyck-Nerlove's – while it should be estimated by the model.[4] It is always possible to choose an expectation scheme showing up a bubble. Seeking to make the expectation scheme endogenous leads to an indetermination between misexpectations and the bubble.

Finally, the bubble hypothesis is invalidated by the facts, especially by the explosion in rent prices over the last decade. Thus, let us put it aside and consider the hypothesis of a cycle due to a shock.

4 Fluctuations due to a shock

As, from an economic point of view, the real-estate business is a service, an asset and a product at the same time, the three corresponding markets must be taken into account to identify the shock(s) undergone.

4.1 The cycle of real estate service

The service market has undergone a shock because the growth of rents has been about 20 points higher than the growth of retail prices

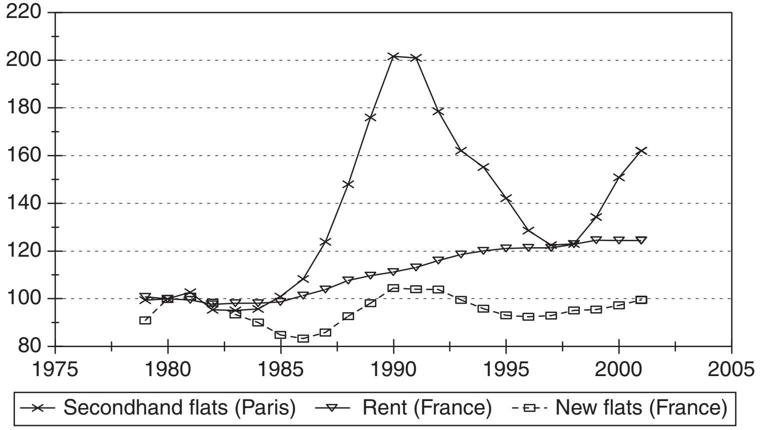

Figure 3.2 Real price indexes

between 1985 and 1995.[5] They increased rapidly up to 1993, and more slowly since then (see Figure 3.2).

The growth of rents can result from two – not necessarily incompatible – explanations. It can be due to a growth in rental demand. It can also be caused by shrinking rental supply.

On the rental demand side, a variation in quantities may be attributable to income growth, to public interventions or to demographic phenomena. One major reason incites us to put aside the demand growth hypothesis: it has been noticed almost all over the world. However, the political, demographic phenomena, and to a lesser extent the income variations, are local factors. It may be a significant increase of migration in some countries or agglomerations, as in Montreal. It may be a growth in the number of students in France and the settlement of rent allowances. Such purely local factors cannot account for an almost world-wide phenomenon.

In the case of Paris the data give no evidence for a rise in demand. The population of the city of Paris has decreased from 2.31 million to 2.2 between 1988 and 1992 and so did the number of households, from 1.155 million to 1.1. The number of all renters (including people in the social sector and accommodated gratis) has decreased from 844 000 to 798 000. The evolution of the net disposable income, though positive, is far from being important: in France as a whole, the average rate of increase between 1988 and 1992 is 2.5 per cent a year, very close to the values of the preceding years. This evolution should not have involved a rise of rental demand.

Hence, the explanation for the phenomenon must be sought on the rental supply side. Indeed, if investors are rational, their determinants are the same whatever the location of their assets, and at a given expected demand. If prices increased, it means that rental supply, that is the rental stock, has shrunk.

To measure the rental stock movements, we have data about the number of rented housing provided by the National Housing Surveys, ENL.[6] The observation of the rented stock shrinking was based on this information. The shrinking of the private rented stock was the subject of numerous and more or less conflicting estimates, leading first to its overvaluation[7] before its estimation at a lower level (see Taffin, 1992 and Lacroix, 1994).

For France as a whole, the decrease in the number of rented housings was by more than 6 per cent between 1984 and 1988; it was compensated by about an equivalent increase between 1988 and 1993 (see Table 3.2). However, this measure is only a partial indicator of the volume movement. It must also take into account other characteristics of the housing stock, that is especially the location, size and quality. The general movement of these characteristics, including the number, finds its expression in the movement of real rents. Then, a more satisfying method to assess the volume movement consists in deflating nominal rents by the corresponding rent index. Table 3.2 shows the results of this calculation based on the ENL data. By taking all the stock volume dimensions into account, a stabilisation can be noticed, with a slight increase by 1.3 per cent over four years. This figure represents the total variation of the stock. To appraise the cyclical dimension, it would be necessary to compare this figure to the trend of the rental stock. From this point of view, it probably represents a decrease against a slightly increasing trend.

Another way of assessing volume can be based on the data provided by the Housing Accounts (Ministry of Equipment). It provides the amount of rents per year according to the nature of the owner (natural person, public sector, other legal persons). These data are obtained by propping up the series on the data from housing inquiries, and by adjusting the inter-survey trend to data from employment surveys.

Table 3.2 Estimation of the non-regulated rented real estate in volume (whole of France)

	1984			**1988**			**1992**		
	Nbr	**Rent/ unit**	**Rent *x* Nbr***	**Nbr**	**Rent/ unit**	**Rent *x* Nbr***	**Nbr**	**Rent/ unit**	**Rent *x* Nbr***
1948 law	708	9 419	6 669	522	14 336	7 483	442	15 452	6 830
Non-regulated sector	3 862	14 348	55 411	3 769	20 263	76 371	4 118	26 942	110 947
Total	4 570	23 767	62 080	4 291	34 599	83 854	4 560	42 394	117 777
Amount of deflated rents			42 375			42 936			52 206
Variation between 2 surveys				–0.061		0.013	0.063		0.216

* in thousands of francs.

Source: 1992–93 ENL.

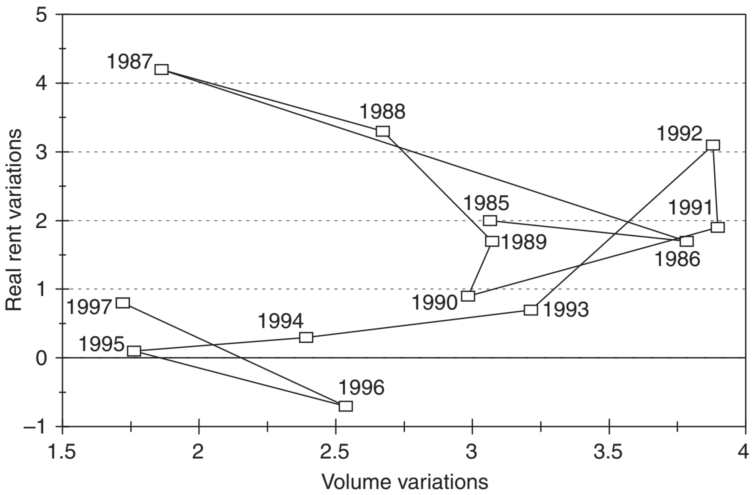

Figure 3.3 Private rental stock (France)

By deflating these rent data by the rent index, we obtain a series of the private rented stock volume (see Figure 3.3).

Unlike the ENL data, the variation of quantities over the period 1985–95 does not indicate any decrease in quantities; however, as our argument deals with the cyclical evolution, it is necessary to reason in terms of deviation from the trend, which is growing. Then, it can be considered that a lower growth accompanied by higher prices consists in a negative supply shock. This phenomenon seems to be more general than the mere French case. Montreal also experienced a phenomenon of that kind with the 'conversions' of rented units into owner-occupied housing units.

Is this relative supply shock sufficient to account for the price variation? As it is a supply shock, the impact of volume variation on price depends on the price elasticity of demand. Only estimates of price elasticity calculated from cross-section data are available (Cornuel, 1985 and Nichèle, 1989); they are roughly −0.5. With a price elasticity of demand at −0.5, the impact of volume variation equals $-1/0.5 = -2$. A 10 per cent increase in rents requires the supply shock to amount to 5 per cent. The value of the price elasticity rather corresponds to a long-run elasticity. In the short run, demand is obviously even less elastic. Then, the volume variation required to explain price variation is lower if the short-term price elasticity is considered inferior to 0.5 in absolute value. The figures for these variations in quantities (−6 per cent in number and +1.3 per cent in volume) do not seem to correspond to the required value. However, these variations in quantities are for France as a whole. The rented stock shrinking was more important in Paris. The extent of the shock appears to be compatible with the data of the rented stock.

From 1987 in France, the loop has first shown a slowing down of rents followed by an increase in volumes. This trend indicates the existence of a (positive) counter-shock – at least temporary – which might be ascribed to public incentives in favour of rental investment into new housing (Méhaignerie-Quilès provisions). From 1992, the curve shows a more erratic trend.

We still have to account for the rented stock contraction in the first part of the cycle in France. Considering the investors' portfolios as a whole, that is including real estate assets, the contraction of the rented stock is due to a shift of the investments towards financial assets. Real estate assets are considered as good investments when inflation is high, because their value vary according to inflation. From the mid-1980s, the inflation fall reduced the interest in real estate investment (see Figure 3.4). At the same time, nominal interest rates decreased, but at a slower pace than inflation, so that real rates remained high, or even higher than in the period of high inflation. Thus, financial investment became more profitable. And as these phenomena were generalised all around the world, their effects were global.

Concerning the other countries, Case (1992) for example indicates that in Massachusetts, rents doubled as did prices. As far as offices are concerned, the available data show that the real estate services market has also undergone a considerable price growth. This is the case in Europe as shown by the data of the John Lang Wootton's European Office Index. Nominal rents were multiplied by three and sale prices by 3.5 in European agglomerations from the beginning to the end of the 1980s. The multiplier is higher in Paris. This is also the case in Japan as it is shown in the Mitsui Fudôsan directory mentioned by Aveline (1995).[8] Office rents in Tokyo were multiplied by 2.5 from 1983 to 1992. Therefore, a real shock did occur in the market of office real estate services. As we do not have data about the movement of the

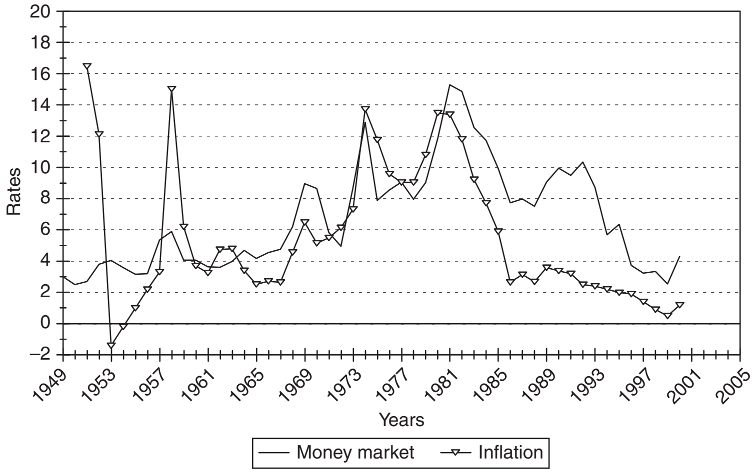

Figure 3.4 Interest rate and CPI

rented stock, it is impossible to estimate the relevance of the hypothesis of a price explosion caused by a supply squeeze. The hypothesis of a demand shock is more plausible in the case of offices than in the case of housing real estate, in so far as the needs of firms and their economic determinants are more homogeneous at the international level than they are for the households.

These phenomena which affected the rented market also affected the market of real estate assets.

4.2 Real estate asset cycle

If the asset price trend follows the fundamental value, the growth of rents tends to increase prices. Rents alone can account for prices, which makes a bubble impossible. On the other hand, the squeeze of the rented stock at the origin of the rent increase can only occur where there is a private rental market subject to the squeeze. In the housing sector, this is the case of Paris and Montreal and not the case in Great Britain. As far as these countries and regions are concerned, other factors are required. Moreover, in every case, the variations in the number of transactions, and especially the precedence of volume variation over price variation observed on Figure 3.1, have still to be explained. Then, the volume variations on the asset market must be explained.

Three phenomena may have had an influence. First, if there was a rented stock shrinking in absolute values, supply in the asset market must have increased in a similar intensity, unless the landlords keep their premises vacant when they cease to supply them on the rental market. However, this phenomenon that would lead to a positive supply shock on the asset market, should have meant lower prices in the asset market. And this could be the case at the beginning of the cycle since the price is below the fundamental value in 1984 revealing a positive supply shock (see Table 3.1).

Second, as the accommodation is a necessity, the (possibly relative) contraction of the rental supply has obliged the upward trend of rental demand to move either towards property demand, or towards the outskirts in the case of Paris, after a redistribution of the households in the rented stock. If households would stay in Paris and cannot rent, they buy. In such conditions, it can be considered that there was a move by part of the rented stock towards the property stock, followed by a transfer of rental demand towards property demand. Hence, the cycle would have meant a move of the line between rented housing and owned housing. It is possible to speak about a ‘cycle of the tenure’. The cycle would have been the process of demand adaptation to this change in the type of holders, given the necessary time of acquisition.

Once this process was over, one would come back to the initial situation in terms of rent and price since in perfect markets, both tenures are equivalent for the occupier. This would make the phenomenon appear as a bubble. The price at the end of the process might even be lower than at the beginning, as in the case of a bubble. Indeed, the fact that producers interpret the transitory price increase as an insufficient supply is enough to lead the production growth to reduce prices. However, the phenomenon might not come back to its initial situation. Indeed, the fluidity is not perfect between the rental and ownership sectors. Some households do not want or are not able to afford to purchase. A lower growth or a rented stock shrinking which would not be compensated afterwards may lead rents to stay durably at higher levels than what they were initially.

Finally, this second phenomenon would consist in a temporary demand shock, since the decrease – or more precisely, the slowing down – in the rental stock growth was over in 1990. The shift in part of the rental demand in excess has contributed to increase the quantities demanded in the asset market, and then to increase prices.

Third, a more durable shock in demand, due to another factor, may have occurred, if we consider that at the end of the cycle, the price is slightly higher than at the beginning. Of course this increased demand of real-estate assets is a demand for owner-occupation and not a demand from landlords since the latter prefer to sell. On the other hand, this additional demand is different from the rental demand which has partly turned to a demand of assets. Then, the question is to know why what has become a relatively bad investment for a landlord is now a good investment for an occupier.

One explanation might be that each type of investor makes different expectations. This may be possible since the landlords are professionals in investment, informed about the real estate and financial markets, while households only punctually intervene in the real estate market.

A more relevant explanation is based on the interest rate. The drop in inflation and nominal rates have made financial investment more advantageous than real estate investment. The same drop of nominal rates may have increased the indebtedness capacity of households subject to liquidity constraint to purchase their housing unit. In such a case, they are more sensitive to the nominal amount of annuity than to the real amount. The rate path may have made more attractive real investment for personal use and increased the demand in that type of investment.

Thus, the price-quantity dynamics pointed out in Figure 3.1 corresponds to a cobweb mechanism with purchaser's naive expectations. The anticlockwise rotation is due to the fact that expectations are made by demanders who need time to find the funds required to purchase, through resaling a house for example. This also explains why quantities increase before prices.

4.3 Cycle of the real estate product

With improvements, real estate production constitutes a way of adjusting the real estate stock after an increased demand, whatever the tenure of the occupiers. However, on the assumption of a bubble, new production can result from speculative demand.

We will first describe the evolution of prices and volumes before trying to provide an explanation.

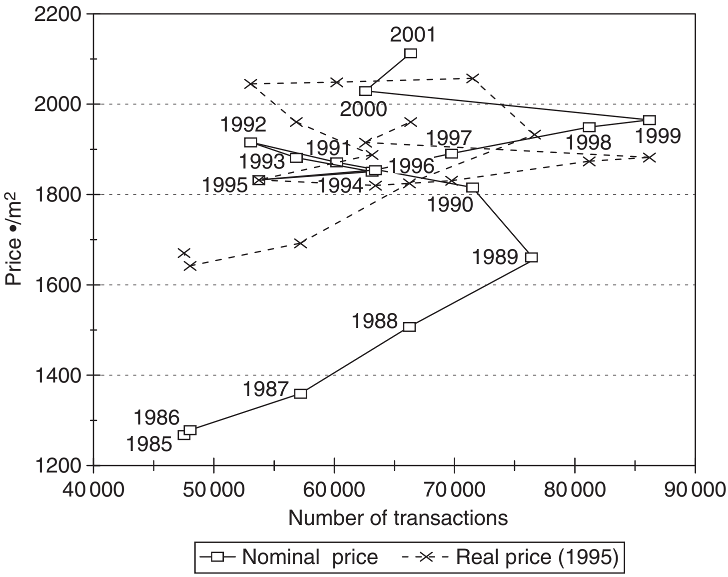

Figure 3.5 New flats (France)

The price-quantity loop is based on data coming from the ECLN (survey on the marketing of new housing units) (Figure 3.5). These data give the number and the price per square metre of retail sales of non-subsidised flats, for owner-occupation or to let.

The general run of the price-quantity loop is characterised by a positive move of quantities up to 1989 and prices up to 1990. It shows an increasing demand. Globally speaking, between 1986 and 1990, the growth of real prices was by 18.4 per cent, that is comparable to the rent growth. This is an additional argument to reject the bubble hypothesis.

From 1992, the growth in quantities can be ascribable to the subsidies offered to favour investment in new rental housing, the number of which cannot be identified in the statistical data.

Over the whole cycle, and in contrast to what happens in the older housing market, the loop is much tighter. It shows a simultaneous growth of volumes and prices in the phase of rise. This corresponds to a rapid adjustment and expresses the higher reactivity of demanders and suppliers to the market conditions. On the supply side, this can be understood insofar as they are professionals. Their correct expectations of market equilibrium price must find its expression in an adequate correlation between the putting on sale and the selling of the unit. Indeed, if producers overestimate prices, they initiate more housing starts than what can be absorbed by the market; if they underestimate prices, they do not produce enough and must withdraw from the stocks. The correlation between units sold and units put on sale amounts to 0.860 in annual data.[9] The suppliers' expectations can be supposed to be quite correct.

On the demand side, and contrary to the second-hand market, they are most often first time purchasers who mostly borrow the capital. This is the reason why they react more rapidly to market changes.

As for the origin of demand, the 'cycle of the tenure' finds its expression in a move of rental demand towards property demand. From a quantitative point of view, there is no reason why there should be an impact on new construction since the additional demand of stock means a demand of services in excess. However, a qualitative inadequacy is possible since property demand does not deal with the same products as rental demand. Very generally speaking and except in the specific case of Paris, apartments are rented and houses are bought. Moreover, more housing volume is purchased than rented. This may result from transaction costs higher in the case of acquisition than in the case of renting. Thus, redistribution in the stock is less frequent when you are an owner than when you are a tenant. The life cycle dynamics and the fact that purchases are made during phase of income increase lead the income increase expectation to be more relevant for purchasing than for renting. In addition, some property ownership subsidies are more purchasing-incentive than renting-incentive.

5 Conclusion

It is now possible to sum up the mechanisms at work, in the French housing markets and probably in other countries.

We started from the hypothesis of a bubble which consists of the disconnection between prices and the fundamental value, the discounted amount of the future rents rationally expected. The fundamental value calculated from the actual rents in the housing market in Paris shows that there wasn't any bubble. The high growth of the housing asset price is explained by the high growth of rents between 1985 and 1990.

Then, the high increase of rents, 20 points higher than the inflation rate for France as a whole, and even more in Paris, has to be explained. It is due to the relative shrinking of the rental stock. As for this lower increase, which may even be a decrease in absolute value, it is due to portfolio arbitrages between real estate and financial investment.

The rental demand in surplus was partly transferred to the asset markets, generating a 'cycle of the tenure', that is the shift of the line separating tenants and owners. This phenomenon does not fully account for the increase in volume on the asset market.

Then the hypothesis of a demand shock in the asset market must be accepted. Such a shock can be ascribed to the drop in interest rates in the 1980s. It would mean that new owners are more sensitive to nominal rates than to real rates, or that they undergo a liquidity constraint smoothed by the rate decreases. In such a case, this shock of property demand can only be met by increasing construction. This mechanism may also have been affected by expectation errors made by households, explaining the upward and downward phases of the cycle.

The end of the cycle that real estate professionals seem to perceive today might occur at rent and price levels lower than the previously reached maxima, but higher than at the beginning of the cycle, that is in 1985–86 in France.

Prices at the end of the cycle will first depend on the trend of rental supply, which is now stimulated by public aid (Méhaignerie-Quilès's and Périssol's provisions). In a second place, they will depend on the more or less durable effect of lower nominal rates on the indebtedness capacity of households.

6 Appendix

Analytical formulation of the rational bubble model (Blanchard and Watson, 1984)

The equation giving the effective actual asset price as a function of its income x_t and of the expected price p^e_{t+1} is:

$$p_t = \left(\frac{1}{1+r}\right) p^e_{t+1} + \left(\frac{1}{1+r}\right) x_t \tag{A-1}$$

where $p^e_{t+1} = E(p_{t+1}/I_t)$. At the order $t+1$ equation (A-1) is:

$$p_{t+1} = \left(\frac{1}{1+r}\right) E(p_{t+2}/I_{t+1}) + \left(\frac{1}{1+r}\right) x_{t+1}$$

By taking the conditional mathematical expectation of $p_{t+1}, E(p_{t+1}/I_t)$, we get:

$$E(p_{t+1}/I_t) = \left(\frac{1}{1+r}\right) E(E(p_{t+2}/I_{t+1})/I_t) + \left(\frac{1}{1+r}\right) E(x_{t+1}/I_t) \qquad \text{(A-2)}$$

One uses the property that $E(E(p_{t+2}/I_{t+1})/I_t) = E(p_{t+2}/I_t)$, which results from the fact that $I_t \subseteq I_{t+1}$ and that agents are supposed not to forget any part of the information previously available. Under these conditions, (A-2) gives:

$$E(p_{t+1}/I_t) = \left(\frac{1}{1+r}\right) E(p_{t+2}/I_t) + \left(\frac{1}{1+r}\right) E(x_{t+1}/I_t)$$

Similarly we have:

$$E(p_{t+2}/I_t) = \left(\frac{1}{1+r}\right) E(p_{t+3}/I_t) + \left(\frac{1}{1+r}\right) E(x_{t+2}/I_t)$$

And so on for increasing values of t.

Substituting successively the value of $E(./I_t)$ into the previous equation, one obtains a solution of the recurrence equation:

$$\begin{aligned} p^*_t &= \left(\frac{1}{1+r}\right) \left[E(x_t/I_t) + \left(\frac{1}{1+r}\right) E(x_{t+1}/I_t) \right. \\ &\qquad \left. + \left(\frac{1}{1+r}\right)^2 E(x_{t+2}/I_t) + \cdots \right] \\ &= \sum_{n=0}^{\infty} \left(\frac{1}{1+r}\right)^{n+1} E(x_{t+n}/I_t) \end{aligned}$$

This solution expresses that p^*_t is the discounted sum of the expected future rents. It constitutes what is called in finance the fundamental value.

This is not the only possible solution. The general solution is of the kind $p_t = p^*_t + z_t$, as soon as:

$$z_t = \left(\frac{1}{1+r}\right) E(z_{t+1}/I_t) \qquad \text{(A-3)}$$

Indeed, if p_t^* is a solution of (A-1), adding equation (A-1) to equation (A-3) gives:

$$p_t^* + z_t = \left(\frac{1}{1+r}\right) p_{t+1}^e + \left(\frac{1}{1+r}\right) E(z_{t+1}/I_t) + \left(\frac{1}{1+r}\right) x_t$$

$$p_t^* + z_t = \left(\frac{1}{1+r}\right) [p_{t+1}^e + E(z_{t+1}/I_t)] + \left(\frac{1}{1+r}\right) x_t$$

$$= \left(\frac{1}{1+r}\right) (p_{t+1}^e + z_{t+1}^e) + \left(\frac{1}{1+r}\right) x_t$$

As a consequence, $p_t^* + z_t$ is also a solution of equation (A-1) where z_t is a bubble.

The bubble hypothesis can be checked by comparing rent and price variances. If there isn't any bubble, the asset price must be correlated to the rents which determine the fundamental value. The fundamental value as a solution of the recursive equation can be written:

$$p_t = \sum_{n=0}^{\infty} \left(\frac{1}{1+r}\right)^{n+1} x_{t+n} + u_t \quad \text{with} \quad E(u_t/I_t) = 0$$

Taking the non-conditional variance of p_t leads to:

$$V(u_t) = E\left\{\left[\sum_{n=0}^{\infty} \left(\frac{1}{1+r}\right)^{n+1} x_{t+n} + u_t\right](p_t)\right\}$$

$$= \sum_{n=0}^{\infty} \left(\frac{1}{1+r}\right)^{n+1} Cov(p_t, x_{t+n}) + E(u_t \cdot p_t)$$

However, the construction of u_t gives $E(u_t \cdot p_t) = 0$. Hence:

$$\frac{\sum_{n=0}^{\infty}(1/1+r)^{n+1}\ Cov(p_t, x_{t+n})}{V(u_t)} = 1$$

that can be written:

$$\left(\frac{\sigma_x}{\sigma_p}\right) \sum_{n=0}^{\infty} \left(\frac{1}{1+r}\right)^{n+1} \rho(p_t, x_{t+n}) = 1$$

Such a result is obtained when there isn't any bubble. In case of bubble, the correlation between rents and prices are reduced and the variance of prices increases, leading to a lower left-hand member in the equation.

7 Acknowledgement

The chapter was translated by Cécile Soudan (IFRESI-CNRS FU3). The authors wish to thank Philippe Thalmann for helpful comments on a previous version of this text.

Notes

1. These are operators who buy second-hand rented buildings, to resell them as condominiums, eventually after some improvements.
2. These data do not match the economic concepts of price and volume. The distribution between price and volume is not satisfying. Prices by square metre keep a volume dimension corresponding to the housing quality. However, as the numbers of transactions increased with the price by square metre, the same can be thought about quality. That shows how the movement of prices by square metre may overvalue the real movement of prices.
3. The relation between price and fundamental value can be tested differently, by using the variances of the rent and price series (see Appendix). This method cannot be used here because rent and price series are not long enough.
4. The existence of bubbles and the expectation scheme are tested simultaneously (Blanchard and Watson, 1984).
5. This trend concerns the general rent index, including rents in the social sector. There is also a private sector index. It is not systematically publicised. Elements about it can be found in Cases (1995). It deals with the whole of France. For each category of district, and especially the agglomeration of Paris, we only know a general rent index.
6. These surveys were carried out in 1992, 1988, 1984, 1978 and 1970. The 1988-ENL data were restated when the 92-ENL was published. The 1984 data were restated only for France as a whole; then the trend in Paris is impossible to estimate.
7. This overestimation of the rented stock shrinking has possibly urged the landlords to increase their rents higher than what would have done the effective shrinking.
8. It is significant to notice that the authors did not use these data on rents for the bubble test, except Nappi in the case of France.
9. The correlation is not improved by the introduction of time-lags between two series.

References

Aveline, N. (1995), *La bulle foncière au Japon*, ADEF.
Blanchard, O.J. and M.W. Watson (1984), 'Bulles, anticipations rationnelles et marchés financiers', *Annales de l'INSEE*, 54.
Blanchard, O.J. and S. Fischer (1989), *Lectures on Macroeconomics*, Cambridge, Massachusetts, MIT Press.
Case, K. (1992), 'The Real Estate Cycle and the Economy: Consequences of the Massachusetts Boom of 1984–1987', *Urban Studies*, 29(2).
Case, K. (1986), 'The Market for Single Family Homes in the Boston Area', *New England Economic Review*, May/June.
Cases, L. (1995), 'De moindres hausses de loyer en 1994', *INSEE Résultats*, No. 375, May.
Comby, J. (1996a), 'Evaluations de convenance', *Etudes Foncières*, No. 71, June.
Comby, J. (1996b), 'Nouveau décrochage à Paris', *Etudes Foncières*, No. 71, June.
Comby, J. (1994), 'En attendant la reprise', *Etudes Foncières*, No. 62, March.
Comby, J. (1992), 'Le cycle de l'immobilier', *Etudes Foncières*, No. 56, September.
Cornuel, D. (1985), *L'évaluation des politiques du logement locatif*, PhD dissertation in Economics, University of Lille 1, June.
Granelle, J.J. (1996), 'Le marché des bureaux: une revue des modèles théoriques', *Revue de l'OFCE*, No. 59, October.
Hort, K. (2000), 'Prices and Turnover in the Market for Owner-occupied Homes', *Regional Science and Urban Economics*, 30, 99–129.
Lacroix, T. (1994), 'Tassement de la propriété et redressement du locatif privé', INSEE Première, No. 313, May.
Ministère du Logement (Ministry of Housing), *Le compte du logement, Rapport à la Commission des Comptes du Logement*, Paris, Economica.
Muet, P.A. (1993), *Croissance et cycles, théories contemporaines*, Paris, Economica, Economie Poche.
Nappi, I. (1994), *Le marché des bureaux: analyse des acteurs de l'offre*, Doctoral dissertation, University of Paris XII.
Nerlove, M. (1958), 'Adaptative Expectations and Cobweb Phenomena', *Quarterly Journal of Economics*, 73.
Nichèle, V. (1989), 'Effets redistributifs et gains d'efficacité de l'allocation de logement en France: une analyse micro-économique', *Economie et Prévision*, No. 91.
Poterba, J.M. (1980), *Inflation, Income Taxes and Owner Occupied Housing*, PhD thesis, Harvard University, March.
Renard, V. (1993), 'Bulles spéculatives, prix immobiliers, prix fonciers', in Granelle, J.-J. and Th. Vilmin (eds.), *L'articulation du foncier et de l'immobilier*, ADEF.
Renaud, B. (1995), 'Le cycle global de l'immobilier', *L'Observateur de l'Immobilier*, No. 31, May.
Revue d'économie financière (1993), *La crise financière de l'immobilier*, special edition.

Roehmer, B.M. (1999), 'Spatial analysis of real estate price bubbles: Paris 1984–1993', *Regional Science and Urban Economics*, 29, 73–88.

Taffin, C. (1992), 'Le parc locatif privé: une espèce en voie de disparition?', *L'Observateur de l'Immobilier*, Revue du Crédit Foncier de France, No. 22.

Tutin, C. (1990), 'Pourquoi la hausse? Quatre hypothèses sur une déconnexion', *Etudes foncières*, No. 47.

4
Estimating Construction Demand in Singapore: Potential of Neural Networks

Freddie Tan and George Ofori

1 Introduction

In Singapore, the construction sector's share in GDP has steadily climbed from 5.4 per cent in 1989 to 7.1 per cent in 1995. Its output has been growing at 11.9 per cent per annum since 1989. A growing outward trend has been the 'regionalisation' of the local construction industry with the result that there has been an increase in construction firms competing for projects abroad. Contracts won abroad will bring in export earnings to the economy that can offset, in part, the leakage due to imports of foreign services and building materials. With the maturing of Singapore's economy, we shall see increasing refurbishment and restoration work in the future. As construction output is a derived demand, it also reflects the importance of inter-sectoral linkages and associative growth. This notion is supported by the high percentage which capital formation in construction contributes to the Gross Fixed Capital Formation (GFCF) in Singapore.

The ability to anticipate construction demand enables growth opportunities in terms of emerging markets to be identified. It also has the potential to facilitate the upgrading and strategic re-structuring of the industry. It should enhance Singapore's market economy via aiding the fine-tuning of the industry. Finally, the industry's performance over time can be ascertained by establishing the level of production to facilitate comparison with other economic sectors and to enable suitable measures to be adopted to prepare the industry to meet any future changes in the size or nature of demand. The close relationship between Singapore's economy and its construction industry has been established chronologically by Ofori (1988). Therefore, it is advantageous to be able to estimate demand and nurture construction activity to enable it to play its expected role in economic development. Turin (1973), Hillebrandt (1984) and many others have shown the importance of the construction industry to the national economy.

2 Existing models

Existing econometric models used to predict construction demand in Singapore rely on statistical approaches to understand the relationships among the influencing factors and for forecasting. Most of them have used ordinary and simultaneous least squares methods of estimation for model building purposes, with particular contributions from Koh (1987) and work jointly carried out by the Construction Industry Development Board (CIDB) of Singapore and Toh to develop a construction demand model comprising 20 structural equations and 20 identities. In the latter, the equations were estimated individually via ordinary least squares. However, the model is of a proprietary nature and has not been updated over the last three years owing to its complexity. The existing econometric models have been constrained by real-world problems which make it difficult to develop an algorithm to forecast construction demand. Besides, there is the need to re-validate the variables in the model in response to sharp corrections in the time series data.

3 History of neural networks

The concept of neural networks first received attention in 1943, when the binary McCulloch-Pitts Model showed that even simple types of neural networks comprising two-state threshold elements could, in principle, perform any imaginable computation. This led to research on learning laws by Hebb (1949). Subsequently, the original experimental and modelling work of Hodgkin and Huxley (1952) on the giant squid axon provided the foundation for a series of new models. Rosenblatt (1958) invented the Perceptron, and showed that 'given linearly separable classes, a Perceptron will, in a finite number of training trials, develop a weight vector that will separate the classes [...] independent of the starting value of the weights'. In 1960, a device called the ADALINE was constructed by Widrow and Hoff and it was equipped with a new powerful learning law known as the Least Mean Squares or 'delta' rule. However, in 1969, Minsky and Papert proved that the single-layer Perceptron Architecture could not solve Boolean exclusive or (XOR) problems. The direction shifted to Expert Systems. Hopfield (1982) rekindled interest in neural networks and multilayer perceptrons were developed to overcome the XOR constraint. Since then, numerous artificial neural network architectures have been developed. The Back-propagation model is among the most popular models in use for two reasons:

1. the learning strategy incorporates minimisation of least mean squares (LMS) error across all training patterns whereby this LMS error technique is traditionally accepted; and
2. it is a supervised learning and the network's performance accuracy can be compared with the target training set.

In the field of construction and real estate in Singapore, neural networks have been used in the mass appraisal of private housing using the Back-propagation Network (BPN) approach; modelling of the overall private housing price index using the General Regression Neural Network (GRNN); and empirical modelling of buildings' indoor air quality using Neuro-Fuzzy Network (NFN).

4 Objectives of this study

This chapter intends to demonstrate the capability of the state-of-the-art technologies, neural network solutions, to explain the variables influencing construction demand. It seeks to develop a demand model using neural network technologies which is explainable. The model should also be robust, and adjustable to changes in government regulations, land constraints and others.

To date, structural or explanation-based models in econometrics have been the norm in various studies to establish the relationships among economic variables (Bergstrom, 1967). These statistical techniques belong to a group of traditional programming techniques. The difficulty here is to develop an algorithm that can simulate real-world complexities. Furthermore, there are inherent problems with time series data, namely, sparse data, non-stationary data, serial correlation, and multicollinearity. Corresponding corrective techniques have been developed to resolve most of these problems to ensure the robustness of these structural econometric models. However, such econometric approaches are constrained by the need to make assumptions about the time series data and sometimes may limit the parametric analysis to a certain number of possible interactions.

5 Research methodology

The first stage of the study involved a comprehensive literature review of construction demand models and identification of the indicators that are often used in the prediction of construction demand levels and patterns. Past quarterly data were drawn from national statistics published by a number of public authorities. The data between 1981Q1 and 1996Q2 were used for our analysis.

The second stage comprised the training and testing of the neural network model with the selected indicators. To demonstrate the capability of neural networks, the quarterly data between 1981 and 1996 were used for training and developing the model, while 'ex post' forecasts were being made over a historical period between 1994Q4 and 1995Q3. Prediction for the period 1997 to 2000 was also made with assumptions on the future economic conditions of Singapore. The last stage involved the ranking of the various selected indicators in order of their magnitude of influence on the demand factor.

6 Conceptual framework of research

The neural network model seeks to examine the effects of the economy on the construction sector which in turn affects the growth of the economy. Construction demand is deemed to be affected by demand factors, supply factors and cyclical factors. The conceptual framing of the research problem is shown in Figure 4.1.

A study of the conditions of demand and supply factors relating to the construction industry reveals that its organisation is largely the response to economic factors. This is because an increase in construction efficiency accelerates real estate development as it will be relatively cheaper to provide a replacement building, and facilitates more intensive development as more cheaper capital can be substituted for land (DiPasquale and Wheaton, 1995). This leads to the belief that focus should be placed on the demand aspects of

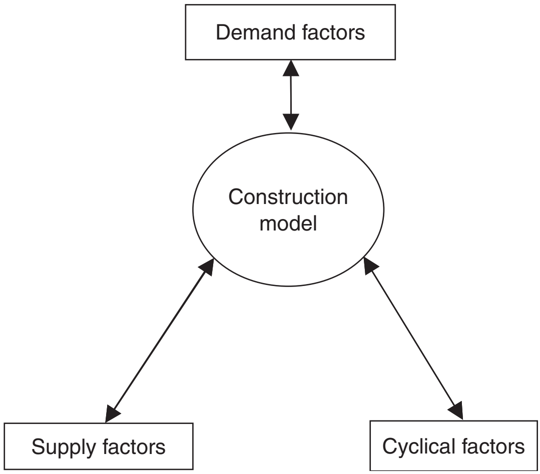

Figure 4.1 Conceptual framework of research

the construction industry to enhance economic efficiency. However, the considerations outlined below should be reviewed carefully in the development of the construction demand model.

(i) The method of pricing depends on the specifications of the construction project and its components. This reduces opportunities for standardisation and mass production, which in turn also depends on the availability of materials, labour and plant.
(ii) The supply of new real estate assets by the construction sector depends on the prices of those assets relative to the cost of replacing or constructing them. In the long run, the asset market should equate market prices with replacement costs that include the cost of land. In the short run, however, the two may diverge significantly because of the lags and delays that are inherent in the construction process. Rent is a key decision factor and demand for space depends on rent and factors such as income levels, firm's production levels and number of households.
(iii) The available data show that at least one-third of the value of the output of the construction industry in Singapore is on repairs and maintenance of the existing stock because buildings are durable.

(iv) Construction demand is a derived demand and is subjected to changes in business expectations and fluctuations in the economy.

(v) Gross Fixed Capital Formation (GFCF) can be largely accounted for by output from the construction industry during recession period such as in 1985 (Ofori, 1993). The theory of national income approach is meaningful to demonstrate that the changing pattern of private spending (I) or/and government expenditure (G) under different economic conditions can influence construction output and therefore, its demand level and pattern.

(vi) As an investment goods industry, construction is prone to fluctuations in demand resulting from changes in expectations, a rise in the cost of borrowing and induced changes related to the national income.

(vii) In a resource-scarce country such as Singapore, leakage via importation of foreign labour, materials and plant significantly affects the economy, and hence, the demand for domestic construction services.

(viii) Government's budgetary policy plays an important role in construction demand. In Singapore government spending normally offsets the decrease of private investment during an economic down-turn such as in 1985, in order to stabilise economic activity. On the other hand, a reduction in public spending on capital projects to lessen inflationary overheating reduces overall construction demand.

(ix) Construction demand is influenced by the cost of credit and the availability of money which are administered by the Monetary Authority of Singapore (MAS). Increasing money supply and lowering of credit cost stimulates more investment, and thereby increases construction demand.

(x) Fiscal policy involving changes in tax and subsidies can also affect the rate of real estate development and thereby the derived demand for construction.

(xi) Land use policies and control mechanisms by the government, for example, the Land Release Programme of the Urban Redevelopment Authority (URA), also play a role in moderating the construction demand.

Using the above conceptual framework, the behaviour and response of each variable (factor) admitted into the model must give an in-depth understanding in order to measure their real impact on the construction demand.

7 Significant indicators for predicting construction demand

From the above considerations, the list of significant indicators that influence the level of demand for construction in Singapore and may be used to build the model are:

1. Prime Lending Rate
2. Money Supply (M2)
3. Gross Domestic Product (GDP)
4. Gross Fixed Capital Formation (GFCF)
5. Consumption Expenditure
6. Increase in Stock
7. Manufacturing Output
8. Building Cost
9. Value of Contracts Awarded

Using these variables, modelling was based on the relationship between the indicators and a suitable demand proxy (that is the dependent variable). The Value of Contracts Awarded was chosen as the demand proxy.

Dependent variable

One possible measure of construction demand is the number of development or planning permits issued by the public authorities. However, permits may not translate into actual construction due to changes in demand conditions or escalation of costs beyond thresholds that allow for profitable returns from investment. For similar reasons, space or occupancy may not either serve as reliable indicators of construction demand unless the market is sufficiently perfect in information dissemination for developers and planners to perceive market performance, expectation and demand for space (Koh, 1987). Another pertinent consideration is that space commenced is concomitantly an indicator of supply although not necessarily specific as space completed. Both Goh (1996) and Tang et al. (1990) chose the Value of Contracts Awarded as the demand proxy. In our model, value of contracts awarded is also admitted as the demand proxy for the simple reason that a dollar change in the value of contracts awarded is a close reflection of the change in the level of construction demand in Singapore.

Independent variables

The remainder of this section is devoted to a discussion of the significant indicators that would be admitted into suitable neural networks as independent variables, in order to examine their influence on construction demand.

Interest Rate Firms invest in plant and equipment in pursuance of the goal to maximise the present value of expectations of future income, and being subjected to the costs of obtaining information, production function constraints, factor supply and product demand functions. Assuming a perfect capital market, a firm will invest in all projects with an internal rate of return exceeding the market rate of interest (Hirshleifer, 1958).

Money Supply, M2 The model considers M2 as a variable because it is the aggregate amount of money made available to meet societal need. The velocity of money flow can moderate swings in capital investment and spending patterns. The money supply has increased during the last ten years in Singapore and this strong demand for money can be logically explained by the rapid pace of economic growth. The purpose of incorporating this indicator is to examine the effect of monetary measures adopted by the MAS on the level of construction demand.

Gross Domestic Product (GDP) Throughout the period under review – 1981 to 1996 – the national economy experienced growth except in 1985. This economic growth has resulted in a higher level of affluence among the people, and has been accompanied by more real property developments.

Gross Fixed Capital Formation (GFCF) Capital investment by the government in the form of infrastructure and buildings has a direct impact on the level of construction demand in Singapore. This has provided many large-scale projects such as the Mass Rapid Transit (MRT) System, the public housing programme, airport and seaport facilities, and expressways/roads.

Consumption Expenditure As more private and public spending is injected into the economy, economic activities become more vibrant. This in turn stimulates an increase in the output of the construction industry through new developments such as retail and office space.

Increase in Stock The measurement of the percentage change in stock over previous years is important in two ways. Firstly, the repairs and maintenance account for at least one-third of the value of the output of the construction industry in Singapore and the other is the generation of new capital assets to accelerate economic growth. Secondly, during the period under review, the increase in stock has also been accompanied by asset appreciation of buildings in Singapore.

Manufacturing Output The performance of the manufacturing industry, especially the electronic and electrical sector, has a strong bearing on the output of the Singapore economy, which in turn affects the performance of the construction industry. The perceived boom stage of manufacturing generally fuels asset investment to produce more output while a slow-down in the performance of this industry may result in shrinkage of asset investment. In essence, construction output is affected by the lagged effect of such a situation.

Building Cost In the context of the study, Building Cost is a measure of the percentage change in the cost of construction over previous years. The close relationship between Building Cost and Value of Contracts Awarded can be easily understood by studying the conditions of demand and supply of various factors of production. In Singapore, materials account for approximately two-thirds of the building cost. Most building materials are imported as Singapore has no natural resources. Due to a severe shortage of labour faced by the industry, there is also a heavy reliance on relatively inexpensive foreign workers. Plant and equipment are generally imported too.

8 Theory of neural networks

A generalised neural network is now presented to explain the theory. Neural networks is a computational technology from the artificial intelligence discipline whose architecture emulates the network of nerve cells in the human brain. A neural network is a parallel distributed information-processing structure consisting of processing elements (PEs) which contains local memory. The PEs can also carry out localised information processing operations interconnected via unidirectional signal channels called connections (Hecht-Nielsen, 1989).

Figure 4.2 shows how a neural network architecture such as a standard Back-propagation Neural Network can be developed by using the

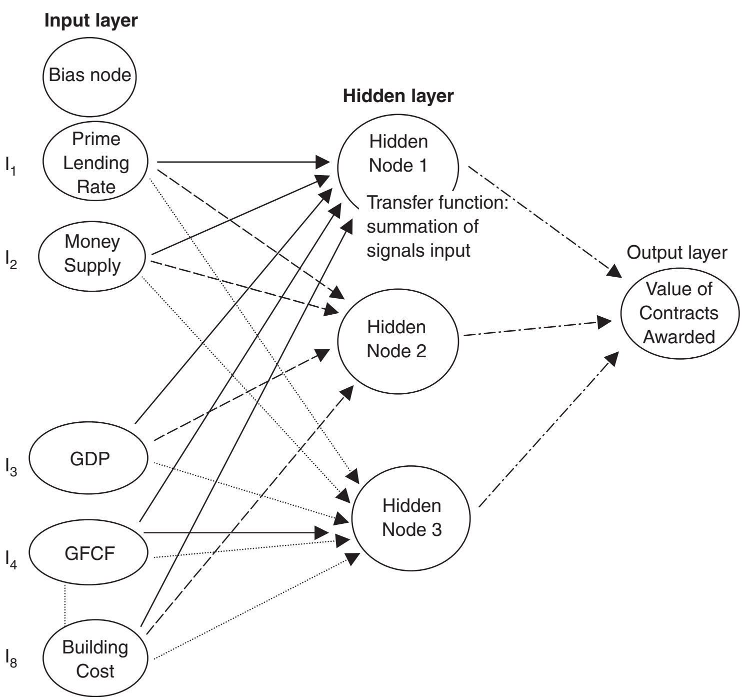

Figure 4.2 A simple architecture of a Back-propagation Neural Network to model construction demand

various indicators as PEs to be investigated upon. As in biological systems, the strength of these connections changes in response to the strength of each input and the use of transfer function by the PEs. All nodes (which are indicators) in the input-layer are fully connected to each of the hidden nodes in the hidden-layer and the process of learning involves all the input nodes and only one of the hidden nodes, H_1. In other words, learning also involves all the other input nodes with each input node connected to every hidden node. The output value from each node of the hidden layer in turn becomes the excitatory input-value for a particular node in the output layer.

There are eight indicators, that is processing elements (PE), and one bias node in the input layer of the neural network model constructed. All the input values are normalised using the MinMax Table. The principle behind this normalisation process is:

Normalised value, $N =$ [Original value less minimum value]/ [Maximum value less minimum value] where:

$$0 \subseteq N \subseteq 1 \tag{1}$$

The module learns the underlying latent function through an error gradient-descent method and the training stops when the root-mean-square-error for output-target values falls below 5 per cent. More iterations in the training of data improve convergence. Each hidden node (that is H_1 to H_3) receives a set of feed-in signals (or values) from which an output value is generated. Finally, all nodes in the hidden-layer are fully connected to the output node.

It is possible to examine the causal impact of the various indicators on the Value of Contracts Awarded, the output node. The shares of influence (equation (2)) for individual input nodes (or indicators) are imputed in the causal analysis. This application of Garson's method (Garson, 1991) helps to explain the 'black box' rules in the hidden-layer.

Share of Influence Input Node, I_i, asserts on the subject Output Node $= S_i\%$

$$= \frac{\sum_{j=1}^{n_j} (|w_{ij}||o_j|)/(\sum_{i=1}^{n_i} |w_{ij}|)}{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (|w_{ij}||o_j|)/(\sum_{i=1}^{n_i} |w_{ij}|)} \times 100\% \tag{2}$$

where n_i = number of input nodes
n_j = number of hidden nodes
w_{ij} = connection-weight from input node I_i to hidden-node H_j
o_j = connection-weight from hidden-node H_j to subject output node S_i

A neural network learns to solve specific problems without the need for problem-specific algorithms. The learning strategy incorporates the minimisation of mean square error across all training patterns, and the network can use supervised training technique with a noise to perturb the network to circumvent the local minima (Hecht-Nielsen, 1989). The user can set a desirable result and compare the network's performance with the target training set. In the next section, the application of the following three types of neural networks from NeuralWare Professional II (1995) is explained: Fast-learning Back-propagation (FBP) Neural Network, Modular Neural Network (MNN), and Reinforcement Neural Network (RNN/DRS) using Directed Random Search as the learning rule.

9 Application of neural networks

In the development of feasible neural network solutions, all the eight selected indicators were used to ascertain the effect and to predict the level of construction demand, thereby to preserve consistency in subsequent comparison on the accuracy of neural network solutions.

The Fast-learning Back-propagation (FBP) Neural Network was chosen as a basic neural network to compare with two other neural networks since Back-propagation is widely accepted. The Modular Neural Network (MNN) offers a new dimension of learning process through the window of gating network and local experts. However, the global error function is still based on back-propagation of errors. Finally, the Reinforcement Neural Network (RNN) uses DRS to adjust the connection weights rather than Back-propagation.

Fast-learning Back-propagation Neural Network is a variation of the traditional Back-propagation algorithm presented by Samad (1988). The aim of the learning process is to minimise the global error *E* of the system by modifying the weights. A gradient descent rule is adopted in the learning across the training set. Suppose a vector i is presented as the input layer of the network and the desired output is D. Let O denote the actual output produced by the network with its current set of weights. Then the measure of the error in achieving that desired output is given by:

$$E = 0.5 \sum_{k} (D_k - O_k)^2 \tag{3}$$

Modular Neural Network consists of a group of networks (referred to as 'local experts') competing to learn the different aspects of the research problem (Jacobs et al., 1991). A gating network controls the competition and learns to assign different regions of the data space to different local expert networks. Both the gating network and the local experts have full connections from the input layer. The gating network has as many output nodes as there are local experts, and the output values of the gating network are normalised to the sum of 1. These normalised output values are used to weight the output vector from the corresponding local expert. The final output vector is the sum of these weighted output vectors. The learning rule is to encourage competition among the local experts so that, for a given input vector, the gating network will tend to choose a single local expert rather than a mixture of them. Training of the local experts and the gating network is achieved using back-propagation of error, that is:

$$E = (d - y)\left(\frac{dy}{dI}\right) \tag{4}$$

where d = desired output vector (for whole network)
y = output vector (for whole network)

Reinforcement Neural Network with Directed Random Search (Matyas, 1965) as the learning rule uses a different learning algorithm compared to traditional Back-propagation Neural Network. Random steps are taken in the weight space and a directed component is added to the random step to provide an impetus to pursue previously successful search directions. The objective of DRS is to choose a set of connection weights that minimise the network prediction error over all the training cases. The prediction error is regarded as the square of the difference between the desired output pattern and the network output pattern for all exemplars in the training set:

$$E = \sum_{j}(D_j - O_j)^2 \tag{5}$$

where D_j = desired output of the network for training exemplar j

O_j = predicted network output for exemplar j

All the neural networks are set to a certain number of iterations. Training stops when convergence obtains at the required root-mean-square-error or when the error across the learning maxim generated by network has become consistently stable. 'Ex post' forecasts are being made over a historical period between 1994Q4 and 1995Q3 and the Run/Test dialog box in the Neuralware programme will help to establish the actual output. We compare the findings of the three selected neural networks, in order to assess the prediction ability of each solution. The share of influence of various selected indicators is usually established using Garson's method (Garson, 1991). In this regard, the Neuralware programme has the Explain/Now dialog box that shows the change in output caused by dithering (its value is actually output divided by input and then multiplied by 100). This mechanism allows us to know which of the indicators has the most effect on the output. Cross-comparison shows the explanatory ability of the neural network solutions. Finally, we predict levels of construction demand for 1997 to 2000.

10 Results and discussion

This section looks at three aspects of model building. Firstly, the predictive abilities of the neural network solutions are compared. Secondly, the explanatory strength of indicators on the output PE, Value of Contracts Awarded, is examined. Lastly, the future level of construction demand between the period 1997 and 2000 is also projected.

(a) Historical forecasts: results and their significance

Prediction tests were run on the historical data between 1994Q4 and 1995Q3. The prediction ability of the three neural network solutions

Table 4.1 Prediction ability of Back-propagation Neural Network (FBP)

Year	Quarter	Actual	Predicted	STDEV
1994	4th	0.9827	1.0144	0.0224
1995	1st	1.1139	1.0344	0.0562
1995	2nd	1.1910	1.0030	0.1329
1995	3rd	0.9693	1.0870	0.0832

Table 4.2 Prediction ability of Modular Neural Network (MNN)

Year	Quarter	Actual	Predicted	STDEV
1994	4th	0.9827	0.9352	0.0335
1995	1st	1.1139	0.9480	0.1173
1995	2nd	1.1910	0.8945	0.2096
1995	3rd	0.9693	0.9659	0.0024

Table 4.3 Prediction ability of Reinforcement Neural Network (RNN/DRS)

Year	Quarter	Actual	Predicted	STDEV
1994	4th	0.9827	0.9770	0.0040
1995	1st	1.1139	1.1473	0.0236
1995	2nd	1.1910	1.1938	0.0020
1995	3rd	0.9693	0.9770	0.0054

is shown in Tables 4.1, 4.2 and 4.3 respectively. The last column indicates standard deviation form the actual (historical) value. The results confirm that the neural network solutions are generally robust and acceptable.

Another parameter used in this comparative study is the mean absolute percentage error, MAPE:

$$\frac{\sum_i \left| \left(\frac{X_i - F_i}{X_i} \right) \right|}{n} \tag{6}$$

where X_i = historical (actual) value
F_i = predicted value
n = number of iterations used in the calculation

The MAPE values of the various neural network solutions are generally below 10 per cent, which implies that the selected indicators may be used as reliable inputs for the modelling of construction demand and this finding provides further justification for the conclusions drawn by Bon (1989) and Tan (1989), that a close relationship exists between building and economic cycles. In this connection, the prediction ability of the RNN/DRS solution is much more accurate than those offered by the FBP and MNN solutions. A comparison of their MAPE values yields 1.15 per cent for the RNN/DRS solution, 9.57 per cent for FBP and 10.04 per cent for MNN.

(b) Classification of significant indicators in terms of explanatory strength

$\Sigma_{j=1}[|w_{ij}||o_j|/\Sigma_{i=1}|w_{ij}|]$ is the sum of signal transfers from input to output, shown in column 3 of Tables 4.4, 4.5 and 4.6. It measures the relationship between the input signals and the output PE in the respective neural network models. Equation (2) is applied to

Table 4.4 Explanatory strength of indicators in FBP model

Indicators	**Node**	$\Sigma_{j=1}[\lvert w_{ij}\rvert\lvert o_j\rvert/\Sigma_{i=1}\lvert w_{ij}\rvert]$	**Share of influence (%)**	**Classification by ranking the strength of indicators**
Prime Lending Rate	2	9.4186	15.35	
Money Supply	3	6.3438	10.34	
Gross Domestic Product	4	5.0684	8.26	
Gross Fixed Capital Formation	5	9.8386	16.03	2
Consumption Expenditure	6	4.5195	7.37	
Increase in Stock	7	1.0516	1.71	
Manufacturing Output	8	15.343	25.00	1
Building Cost	9	9.7773	15.94	3
Sum of signal transfers		61.3608	100.00	

Table 4.5 Explanatory strength of indicators in the MNN model

Indicators	Node	$\Sigma_{j=1}[\|w_{ij}\|\|o_j\|/\Sigma_{i=1}\|w_{ij}\|]$	Share of influence (%)	Classification by ranking the strength of indicators
Prime Lending Rate	2	5.7284	10.07	
Money Supply	3	4.3963	7.73	
Gross Domestic Product	4	5.9837	10.52	
Gross Fixed Capital Formation	5	10.5571	18.56	2
Consumption Expenditure	6	5.5695	9.79	
Increase in Stock	7	0.9795	1.72	
Manufacturing Output	8	14.8225	26.05	1
Building Cost	9	8.8564	15.56	3
Sum of signal transfers		56.8934	100.00	

Table 4.6 Explanatory strength of indicators in the RNN/DRS model

Indicators	Node	$\Sigma_{j=1}[\|w_{ij}\|\|o_j\|/\Sigma_{i=1}\|w_{ij}\|]$	Share of influence (%)	Classification by ranking the strength of indicators
Prime Lending Rate	2	10.8203	9.57	
Money Supply	3	8.5475	7.56	
Gross Domestic Product	4	5.6393	4.99	
Gross Fixed Capital Formation	5	11.1028	9.82	3
Consumption Expenditure	6	0.0354	0.03	
Increase in Stock	7	0.0547	0.05	
Manufacturing Output	8	25.573	22.62	2
Building Cost	9	51.2673	45.36	1
Sum of signal transfers		113.0403	100.00	

calculate the contribution of individual indicators to explain their share of influence towards the output node, the Value of Contracts Awarded. All the three neural network solutions have ranked Building Cost, Manufacturing Output and Gross Fixed Capital Formation (GFCF) as influencing factors towards construction demand. The neural network solutions emphasise these three indicators because they account for a large change in the Value of Contracts Awarded. In addition, they influence the expected mix of construction and pattern of construction demand. This convergence in the ranking of very significant indicators further shows that a few strong variables are of sufficient merit to explain the movement of construction demand and this is supported by the accumulated shares of influence of these three indicators, being represented by 56.94 per cent (FBP), 60.17 per cent (MNN) and 77.80 per cent (RNN/DRS) respectively. Among the three neural network solutions, the RNN/DRS attributes the strongest explanatory power to the three most significant indicators.

A unit change in building cost will change the Value of Contracts Awarded. This explains the proportionate relationship between the change in building cost over previous years and the Value of Contracts Awarded. The boom-and-slump effect of the manufacturing industry is directly experienced by the construction industry because a boom offers opportunities for new development and refurbishment projects. Hence, the performance of manufacturing over time can be used to devise suitable measures to prepare the construction industry to meet any future changes in the size or nature of construction demand. The GFCF variable accounts directly for output from the construction industry. Government's capital spending is normally in the form of capital investment to provide a good network of infrastructure and buildings necessary for economic activities and to meet societal needs. Besides, it can be used to stabilise construction demand.

(c) Forecasts of construction demand

This section attempts to predict the future Value of Contracts Awarded through appropriate assumptions for the eight indicators for 1997 to 2000. Due to the slow-down of the economy during the second half of 1996, the first half of 1997 is assumed to mark a slow recovery. The manufacturing sector is expected to rebound by 1998 and its growth rate is projected at five per cent, which is conservative. Construction activity is believed to slow down in 1997 but government spending on several major projects will introduce more construction activities from 1998 onwards. Hence, the GFCF is expected to increase by 5 per cent in 1997 to at least 8 per cent in the year 2000. The prime lending rate is expected to remain relatively constant during 1997 to 2000 because the policy is to maintain a healthy economic growth. Based on the past trends, the velocity of money supply (M2) is increasing and this is expected to grow by 2.5 per cent per quarter. As far as increase in stock is concerned, the projected increase is 5 per cent to reflect a progressively healthy economy. Consumption expenditure may increase moderately by 1 per cent each year. An increase in building cost is justifiable by the strong demand for building materials and the shortage of manpower faced by the construction industry in Singapore. A 5 per cent increase in building cost over previous years is expected.

These above assumptions on the indicators were fed in the trained network and the network was allowed to test the hypothetical data. The results in Table 4.7 show the future forecast of Value

Table 4.7 Forecasts of construction demand for 1997–2000 ($ million)

Quarter	FBP network	MNN network	RNN/DRS network
1997Q1	5295.740	4751.710	4718.126
1997Q2	5307.249	4758.758	4759.727
1997Q3	5318.654	4765.727	4801.381
1997Q4	5329.950	4772.610	4843.096
1998Q1	5358.116	4788.241	4921.042
1998Q2	5369.039	4794.863	4962.770
1998Q3	5442.241	4852.896	5223.546
1998Q4	5451.911	4858.749	5264.726
1999Q1	5457.942	4857.283	5188.962
1999Q2	5467.908	4863.088	5230.246
1999Q3	5477.773	4868.819	5271.400
1999Q4	5487.548	4874.484	5312.410
2000Q1	5492.400	4872.727	5231.024
2000Q2	5502.031	4878.345	5293.918
2000Q3	5511.566	4883.893	5313.184
2000Q4	5521.010	4889.376	5354.025

of Contracts Awarded ($ millions) by the various neural network solutions.

From the results, the predictions offered by the RNN/DRS network appear to converge towards those of the FBP network after the second quarter of 1998. The MNN network shows that the construction work volume in terms of Value of Contracts Awarded will be relatively constant over the next four years. The predictions suggest that in 1997 and early 1998, there may be a fall in total workload. From mid-1998 onwards, there will be some recovery in construction activity.

11 Conclusion

Neural networks represent a state-of-the-art approach that intelligently searches for underlying relationships among the time series concerned, through adapting or changing the connection weights which represent the array of variables, thereby overcoming the problems associated with sharp corrections, and the paucity and non-stationarity of the data. Unlike the traditional statistical method which needs a priori parametric knowledge of the form of linear or non-linear function to be tested, neural networks do not need such information beforehand to predict the future possible outcomes. They are designed to capture the non-linear relationship between the input and output variables automatically. They are useful for solving complex problems which are too difficult to apply constrained-optimisation algorithms. A creative, flexible solution can be 'invented' through neural networks.

The chapter demonstrates the estimating of construction demand via the use of neural network models to predict the output factor, the Value of Contracts Awarded. The results seem to suggest that the RNN/DRS network has the best trainability network for the period 1981 to 1996. However, the MNN and FBP networks are still able to offer reasonably good explanatory strength towards the prediction of the level and pattern of construction demand. Neural networks offer a realistic measure of construction demand which is necessary if effective effort is to be made to maintain, and improve upon, the capacity of the industry. They can also advise on how to moderate the swings in construction outputs through various measures such as monetary and fiscal policies.

12 Acknowledgement

The authors are grateful to Dr David Ho Kin Hin for his assistance in explaining the finer details of neural networks and their applications, which was of immense help to them in preparing this chapter.

References

Bergstrom, A.R. (1967), *The Construction and Use of Economic Models*, London, English Universities Press.

Bon, R. (1989), 'Direct and Indirect Resource Utilisation by the Construction Sector: The Case of the USA since World War II', *Habitat International*, 12(1), 49–74.

DiPasquale, D. and W.C. Wheaton (1995), *Urban Economics and Real Estate Markets*, Englewood Cliffs, NJ, Prentice Hall.

Garson, G.D. (1991), 'Interpreting Neural-Network Connection Weights', *AI Expert*, 6, 47–51.

Goh, B.H. (1996), 'Residential Construction Demand Forecasting Using Economic Indicators: A Comparative Study of Artificial Neural Networks and Multiple Regression', *Construction Management & Economics*, 14(1), 25–34.

Hebb, D.O. (1949), *The Organisation of Behavior*, New York, Wiley.

Hecht-Nielsen, R. (1989), 'Theory of the Backpropagation Neural Network', *Proceedings of the International Joint Conference on Neural Networks*, 1, 593–611, New York, IEEE Press.

Hillebrandt, P. (1984), *Analysis of the British Construction Industry*, London, Macmillan.

Hirshleifer, J. (1958), 'On the Optimal Investment Decision', *Journal of Political Economy*, 66(4), 329–52.

Hodgkin, A.L. and A.F. Huxley (1952), 'A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve', *Journal of Physiology*, 117, 500–44.

Hopfield, J.J. (1982), 'Neural Networks and Physical Systems with Emergent Collective Computational Abilities', *Proceedings of the National Academy of Sciences*, 79, 2554–58.

Jacobs, R.A., M.I. Jordan, S.J. Nowlan and G.E. Hinton (1991), 'Adaptive Mixtures of Local Experts', *Neural Computation*, 3, 79–87.

Koh, A.M.M. (1987), *An Econometric Model for Forecasting Industrial Space Demand in Singapore*, Unpublished PhD dissertation, University of Georgia, USA.

Matyas, J. (1965), 'Random Optimization', *Automation and Remote Control*, 26, 246–53.

Minsky, M. and S. Papert (1969), *Perceptrons: An Introduction to Computational Geometry*, Cambridge, MA, MIT Press.

Ofori, G. (1988), 'Construction Industry and Economic Growth in Singapore', *Construction Management and Economics*, 6, 57–70.

Ofori, G. (1993), *Managing Construction Industry Development: Lessons from Singapore's Experience*, Singapore, Singapore University Press, NUS.
Rosenblatt, F. (1958), 'The Perceptron: A Probabilistic Model for Information Storage and Organisation in the Brain', *Psychological Review*, 65(6), 386–408.
Samad, T. (1988), 'Backpropagation is Significantly Faster if the Expected Value of the Source Unit is Used for Update', *International Neural Network Society Conference Abstracts*.
Tan, W. (1989), *Subsector Fluctuations in Construction*, Occasional Paper No. 1, Construction Management and Economics Research Unit, School of Building and Estate Management, National University of Singapore, Singapore.
Tang, J.C.S., P. Karasudhi and P. Tachopiyagoon (1990), 'Thai Construction Industry: Demand and Projection', *Construction Management & Economics*, 8, 249–57.
Turin, D.A. (1973), *The Construction Industry: its Economic Significance and its Role in Development*, London, University College London.
Widrow, B. and M.E. Hoff (1960), 'Adaptive Switching Circuits', *IRE WESCON Convention Record*, New York, 96–104.

5
Lump-sum Moving Cost

Gerbert Romijn

1 Introduction

When one sets out to model the market for office space, one of the central variables that requires modelling is the demand for office space or the occupied stock of office space. In one of the first attempts Rosen (1984) models the occupied stock of office space as being a function of the employment level in the key service producing industries and the real rental rate. When one projects this aggregate model to the microeconomic level it implies that office space using firms change their use of office space every time their employment or the rent changes. However, we see in practice that firms only infrequently change their use of office space. To remedy this Wheaton (1987) models the net absorption (that is the change in occupied stock) as a partial-adjustment process: current absorption equals a fraction of the difference between desired office space use and lagged actual office space use. Desired office space use is again modelled as a function of office employment and the real rental rate. Additionally, Wheaton adds office employment growth to account for expectations regarding future office space needs. Hence, Wheaton accommodates the gradual change in occupied office space in two ways, the partial adjustment of net absorption and the appearance of office employment growth in desired office space use. However, both mechanisms are incorporated in an ad hoc fashion at the aggregate level. At the firm level it is not realistic to assume that firms adjust gradually to some desired level of office space use since this involves continual adjustment of office space use, something Wheaton claims hardly occurs. Instead he says: 'It is likely that the long-term leasing structure of the office market reflects a high cost to moving and relocating business' (Wheaton, 1987, p. 285).

In a recent paper, Romijn, Hakfoort and Lie (1996) (henceforth referred to as RHL) introduce adjustment costs in a microeconomic model for the use of office space. They motivate these costs primarily as relocation or moving cost for which it is reasonable to assume that these costs are to a large extent unrelated to the amount of office space in use or the change therein. Therefore they model them as being lump-sum. This adjustment cost structure implies that the individual firm's office space use is governed by a (s, S)-rule.[1] Hence office space using firms only infrequently relocate, just as observed in practice. Their empirical findings are based on a cross section survey of individual firms and indicate that at the firm level the lump-sum nature of relocation costs matters.

This paper uses the set-up of RHL with lump-sum adjustment cost at the firm level, but instead of the heuristic solution given in RHL, here we solve explicitly an intertemporal maximisation programme. Additionally, instead of focusing on the microeconomic implications as did RHL, we investigate the implications the individual firm's behaviour has for aggregate office space use. The data we use concern the Dutch office market for which a consistent dataset has been compiled in Romijn (1997). As main office space using industries we identify the government, banking and insurance, and other commercial services. Output and employment in these sectors are assumed to be a good indicator for office-related output and office employment.

The rest of this chapter is set up as follows. In Section 2 we formulate and solve the model for firm behaviour, which is governed by a so-called control band policy or (s, S)-rule. The mathematical argumentation in this section is heuristic and not entirely rigorous. For more rigour, we refer to the paper by Harrison, Sellke and Taylor (1983) (henceforth referred to as HST).

Obviously when all firms are exactly identical, aggregate behaviour would coincide with individual behaviour. However, we assume that individual firms face stochastic shocks that are only partly shared by other firms. This implies that at every point in time a certain fraction of all firms will relocate whereas others will not. To determine what fraction of firms relocates, we have to concern ourselves with distributional issues. In Section 3 this problem is addressed using the framework of Bertola and Caballero (1994). This results in a relation between aggregate desired office space and aggregate actual office space use with the gap between them depending on the growth rate of desired office space use.

In Section 4, we first calculate aggregate desired office space use which is subsequently used for calculating the aggregate gap and obtaining an estimate for actual office space use. Our estimate for aggregate office space use tracks actual office space use much better than desired office space, implying that observed aggregate behaviour can indeed be accounted for by our relocation-cost-cum-stochastic-aggregation model. Finally, Section 5 summarises and concludes.

2 Optimal demand for office space with lump-sum adjustment costs

Consider a firm that uses office space in the production process. We want to focus on the demand for office space exercised by this firm. Assume that there exists a desired demand for office space that summarises all relevant information about the office space use of the firm. This desired demand evolves stochastically over time. We interpret desired office space use as the minimum cost or maximum efficiency office space use for the firm in the sense that when actual office space use equals the desired office space use, the intensity of use of the office space is optimal. Actual office space use may deviate from its desired level. These deviations result in extra costs or loss of efficiency. This implies that in a frictionless environment a cost minimising firm would like to adjust its demand for office space continually in response to the stochastic fluctuations in desired office space use. However, we assume that in order for the firm to change its demand, it has to move to a new location and that this move is costly. Specifically, we assume that the cost of moving is lump-sum. This implies that the firm has to balance two types of cost. On the one hand, there is the opportunity cost of not adjusting demand for office space to its desired value. On the other hand, the more frequently the firm moves, the higher will be the moving costs.

The above problem can be reformulated as a special case of a more general problem studied by HST. They show that the optimal policy is a so-called 'control band policy' (CBP). This policy entails an optimal demand for office space depending on the state variable (in our case desired office space use) and intervention bands around this optimal demand. As long as actual demand remains within the bands, the firm does not move and hence does not adjust demand. When actual demand is on or outside the band, it is optimal for the firm to move and adjust its demand to the optimal demand.

Denote by $z(t)$ the logarithm of actual office space use and by $z^d(t)$ the logarithm of desired office space use. We model $z^d(t)$ as a $(-\mu, \sigma)$ Brownian motion, that is:

$$dz^d(t) = -\mu dt + \sigma dw(t) \tag{1}$$

with $w(t)$ a standard Brownian motion. For simplicity, we model the above-mentioned loss of efficiency due to deviations from actual office space use from desired office space use to be quadratic in the deviation, that is $\frac{1}{2}[z(t) - z^d(t)]^2$. The firm now faces the problem of deciding whether or not to move and what demand to exercise if it moves. When the firm moves to a new building, it incurs a lump-sum moving cost of magnitude γ. Otherwise, it remains in the old building without changing its demand for office space.

To solve this problem, note first that it is not optimal for the firm to adjust its demand for office space continually as it would then incur the strictly positive moving cost γ at each moment in time making total moving cost infinite. Hence, the firm will only move at discrete intervals. Denote the times at which it moves by $T_n, n \in \{0, 1, \ldots\}, 0 = T_0 < T_1 < \ldots \to \infty$. Furthermore, denote the change in the demand for office space at points in time when the firm moves by ξ_n. Because of the assumption of $T_0 = 0$ we have to allow for $\xi_0 = 0$. Hence, we see that a policy consists of sequences of stopping times $\{T_0, T_1, \ldots\}$ and associated stochastic jumps $\{\xi_0, \xi_1, \ldots\}$. Now define the cost function of moving:

$$\phi(\xi) = \begin{cases} 0 & \text{for } \xi = 0 \\ \gamma & \text{for } \xi \neq 0 \end{cases} \tag{2}$$

Next, define $N(t) = sup\{n \geq 0 : T_n \leq t\}$ and $y(t) = \xi_0 + \cdots + \xi_{N(t)}$ the cumulated change in the demand for office space from time zero to time t, $x(t) = -z^d(t) + z(0)$ a (μ, σ) Brownian motion, and $u(t) = z(t) - z^d(t)$ the gap between actual and frictionless demand. The latter variable follows a process that is the sum of two processes: $u(t) = x(t) + y(t)$. Note that $\xi_n = u(T_n) - u(T_n-)$, with T_n- the moment directly before T_n. We see that without any action by the firm, the instantaneous rate of cost is given by $\frac{1}{2}x(t)^2$. This may however result in large negative rate of profit so occasionally – at the stopping times – it is profitable for the firm to change its demand by an amount x so as to bound losses.

For any feasible policy $\{(T_n, \xi_n)\}$ and initial value $x(0) = x$, define the value function $V(x)$, to be the current value of all expected future cost discounted to the present at rate r:

$$V(x) = E\left[\int_0^\infty \tfrac{1}{2}u(t)^2 e^{-rt}dt + \sum_{n=0}^{\infty} \phi(\xi_n)e^{-rT_n}\right] \tag{3}$$

The firm will be assumed to choose the CBP $\{(T_n, \xi_n)\}$ that minimises this value function.

The value function

It can be shown by arguments similar to those in HST that an optimal CBP is characterised by a set of numbers $(s, S, Q), s < Q < S$, with Q the optimal demand for office space, and s and S the lower bound and upper bound of the CBP, respectively. The CBP parameters are parameters of the value function. Since the value function does not depend on time, we know that the CBP parameters cannot be functions of time. Additionally, the CBP parameters are values of the state variable for which the value function meets certain criteria: Q is the value of the state variable that maximises the value function. s and S are the values of the state variable for which the boundary of the control band is reached and the boundary conditions for the value function apply. This implies that the CBP parameters are constants and not function of time or the state variable.

When $u(T_n-)$ reaches the lower bound s, the firm will move into a new building and change its demand to Q. Hence, for $n \geq 1$, the jump in u_t will be given by $\xi_n = Q - s > 0$. Analogously, when $u(T_n-)$ reaches the upper bound S, the jump is equal to $\xi_n = Q - S < 0$. For time zero, we have to allow for the possibility of a jump of size zero.

Hence, we define:

$$\xi_0 = \begin{cases} 0 & \textit{if } s < x < S \\ Q - x & \textit{otherwise} \end{cases} \tag{4}$$

To find an explicit solution for the value function, note that between stopping times, by definition, no jumps occur, and u will remain between the upper and lower intervention band. Hence, for values of u between the upper and lower band (or analogously for points in time between two adjacent stopping times), the Bellman equation that follows from (3) can be obtained by forgetting about the second term on the right-hand side of (3) and using Ito's Lemma (see for instance Dixit and Pindyck, 1994, chapter 3). This of course also holds for the initial value x when $s \leq x \leq S$ and we obtain:

$$rV(x) = \tfrac{1}{2}x^2 + \tfrac{1}{2}\sigma^2 V''(x) + \mu V'(x) \textit{ for } s \leq x \leq S \tag{5}$$

Additionally, we have boundary conditions given by:

$$\begin{aligned} V(s) &= V(Q) + \gamma \\ V(S) &= V(Q) + \gamma \end{aligned} \tag{6}$$

The solution to differential equation (5) is given by:

$$V(x) = Ae^{\alpha x} + Be^{-\beta x} + v_0 + v_1 x + \tfrac{1}{2}v_2 x^2, s \leq x \leq S \tag{7}$$

with

$$\begin{aligned} \alpha &= [(\mu^2 + 2\rho\sigma^2)^{1/2} - \mu]/\sigma^2 > 0 \\ \beta &= [(\mu^2 + 2\rho\sigma^2)^{1/2} + \mu]/\sigma^2 > 0 \end{aligned} \tag{8}$$

and

$$v_0 = \frac{1}{2}\frac{\sigma^2}{r^2} + \frac{\mu^2}{r^3}, v_1 = \frac{\mu}{r^2}, v_2 = \frac{1}{r} \tag{9}$$

The constants A and B can be found by substituting (7)–(9) into the boundary conditions (6). Define $a(y) = e^{\alpha y} - e^{\alpha Q}$ and $b(y) = e^{-\beta y} - e^{-\beta Q}$. We obtain:

$$A = \frac{[v_1(Q-s) + \frac{1}{2}v_2(Q^2 - s^2) + \gamma]b(S) - [v_1(Q-S) + \frac{1}{2}v_2(Q^2 - S^2) + \gamma]b(s)}{a(s)b(S) - a(S)b(s)} \tag{10}$$

$$B = \frac{[v_1(Q-S) + \frac{1}{2}v_2(Q^2 - S^2) + \gamma]a(s) - [v_1(Q-s) + \frac{1}{2}v_2(Q^2 - s^2) + \gamma]a(S)}{a(s)b(S) - a(S)b(s)} \tag{11}$$

Now we want to extend the value function for values of x outside the control band. To see how, note that when x lies outside the control band, by definition the firm will immediately pay the moving cost γ and jump to Q. Hence:

$$V(x) = V(Q) + \gamma, x \notin [s, S] \tag{12}$$

This completes the characterisation of the value function for CBP $(s, S, Q), s < Q < S$, and starting value x.

Optimal control band parameters

Having obtained the value function, we can now find the optimal CBP consisting of the set of numbers $(s, S, Q), s < Q < S$. If the firm starts outside the interval control band, that is $x \notin [s, S]$, it will immediately jump to Q and follow the CBP (s, S, Q). The total reward from this will be $V(Q) + \gamma$. For Q to be optimal, we should have:

$$V'(Q) = 0, V''(Q) > 0 \tag{13}$$

Additionally, by arguments similar to those in Section 5 of HST, it can be shown that the following conditions hold at the boundaries:

$$\begin{aligned} V'(s) &= V'(Q) \\ V'(S) &= V'(Q) \end{aligned} \tag{14}$$

These can be interpreted as some sort of smooth pasting conditions for problems involving jump processes.

Solving equations (13) and (14) yields the following expressions for parameters of the optimal CBP. The derivative of the value function is given by:

$$V'(x) = \begin{cases} \alpha A(s,S,Q)e^{\alpha x} - \beta B(s,S,Q)e^{-\beta x} + v_1 + v_2 x & \textit{for } s \le x \le S \\ 0 & \textit{otherwise} \end{cases} \tag{15}$$

The roots of (15) yield a system of equations the solutions of which are the optimal control band parameters (s, S, Q). Unfortunately the roots of (15) cannot be determined analytically and hence we have to resort to numerical solutions for reasonable values of the model parameters (μ, σ, γ, r).

3 Aggregate demand for office space

We now turn to the aggregate implications of our lump-sum adjustment cost micromodel. For aggregation of individual units' actions we rely heavily on Bertola and Caballero (1994, pp. 229–34).

First, the markets for real estate are populated by a large number of units which we approximate by continuum indexed by $i \in [0, 1]$. To facilitate the subsequent discussion we introduce some notations. Let $x_i(t)$ denote the value of a variable x for unit i at time t. Additionally, let $\tilde{x}(t)$ denote a random variable with a probability distribution (x, t) identical to that of the cross-section distribution of the $x_i(t)$ (see Caballero and Engel, 1991, p. 1663, for this construct). Note that the following holds:

$$E\tilde{x} = \int_0^1 x_i di \tag{16}$$

Finally, let $x(t)$ denote the associated aggregate.

First consider actual office space use. To aggregate we simply sum over all units, that is $Z = \int_0^1 Z_i di$. Hence, the process followed by the logarithm of aggregate actual office space use $z(t)$ is found as:

$$dz(t) = \int_0^1 h_i(t) dz_i(t) di \tag{17}$$

with $h_i(t)$ unit i's share in aggregate office space use with $\int_0^1 h_i di = 1$.

Individual frictionless demand $z_i^d(t)$ follows a process given by:

$$dz_i^d(t) = -\mu dt + \sigma dw_i(t) \tag{18}$$

with $w_i(t)$ a standard Brownian motion. To aggregate individual desired office space use, we have to make assumptions about how individual uncertainty translates into aggregate uncertainty. We assume that the correlation structure between the individual firms can be described by $E[dw_i(t)dw_j(t)] = \rho^2$, for all $i,j \in [0,1], i \neq j$. This implies that the covariance between an individual shock $dw_i(t)$ and the aggregate shock $dw(t)$, which is given as:

$$\rho dw(t) = \int_0^1 h_i(t)dw_i(t)di \tag{19}$$

equals ρ. Hence we can decompose the individual Brownian motions $w_i(t)$ into an aggregate and a purely idiosyncratic component $w_{Ii}(t)$ according to:

$$dw_i(t) = \rho dw(t) + \sqrt{1-\rho^2}\, dw_{Ii}(t) \tag{20}$$

By construction the idiosyncratic components are uncorrelated among each other and the aggregate shock, and wash out in the aggregate. Using the expression for aggregate uncertainty in (19) and aggregating (18) we obtain:

$$dz^d(t) = -\mu dt + \sigma\rho dw(t) \tag{21}$$

Now define stochastic variables $\tilde{u}$ and $\tilde{h}$ with probability density functions identical to the cross-sectional distribution of the u_i and h_i. Since there is no reason to assume that the u_i and h_i vary systematically with each other, we assume the opposite, that is $\tilde{u}$ and $\tilde{h}$ are independent. Note that the following relations hold:

$$\int_0^1 h_i u_i di = E(\tilde{h}\tilde{u}) = E(\tilde{h})E(\tilde{u}) = E(\tilde{u}) = \int_0^1 u_i di \equiv u \tag{22}$$

Using these we obtain an expression for the process for actual aggregate office space use given as:

$$dz(t) = dz^d(t) + du(t) \tag{23}$$

and we see that the difference between actual and desired aggregate net absorption ratios differ by the change in the average difference of logged actual and desired office space use at the firm level. To obtain the change in this average, we need to track the probability density function $\phi(\tilde{u}, t)$ through time.

First consider the case where no aggregate shocks are present and all shocks are fully idiosyncratic, that is $\rho = 0$, and the cross section density has settled into its steady state. Due to the independence of the different shocks and the fact that the number of units is large, this density $\phi(\tilde{u})$ is identical to the ergodic density of a single $u_i(t)$. The derivation of this density is detailed in the Appendix. It is given as:

$$\phi(\tilde{u}) = \begin{cases} A_1[e^{\theta\tilde{u}} - e^{\theta s}], & s < \tilde{u} < Q \\ A_2[e^{\theta\tilde{u}} - e^{\theta S}], & Q < \tilde{u} < S \end{cases} \tag{24}$$

with $\theta = 2\mu/\sigma^2$, $A_1 = cA_2$, $c = [e^{\theta Q} - e^{\theta S}][e^{\theta Q} - e^{\theta s}]^{-1}$ and $A_2 = -[ce^{\theta s}(Q - s) - e^{\theta S}(Q - S)]^{-1}$. From this, it follows that in steady state with uncorrelated shocks $u(t)$ is given as:

$$\begin{aligned} E(\tilde{u}) = {} & (A_1/\theta)[(Q-1)e^{\theta Q} - (s-1)e^{\theta s}] - (A_1/2)e^{\theta s}(Q^2 - s^2) \\ & + (A_2/\theta)[(S-1)e^{\theta S} - (Q-1)e^{\theta Q}] - (A_2/2)e^{\theta S}(S^2 - Q^2) \end{aligned} \tag{25}$$

Now we want to introduce aggregate shocks. When aggregate shocks are present, the shocks faced by individual firms are correlated and the steady state cross-sectional density can no longer be represented by the ergodic density of a single random walk in a CBP. Instead the cross-sectional density is changing at every point in time and will not settle down into a steady state density as long as new aggregate shocks keep arriving. To model aggregate shocks we use the approach used by Bertola and Caballero (1994). They approximate the ongoing aggregate shocks by discrete changes in the drift rate μ. In other words, we assume that the realisations of aggregate uncertainty are evenly spread within an observation interval. Bertola and Caballero (1994) motivate this as follows: 'accumulation over a finite time period of abnormally positive aggregate shocks has roughly the same effect for the cross-sectional distribution as a larger mean rate of growth' (Bertola and Caballero, 1994, p. 232).[2]

This approximation neglects within period path-dependency and infinite variation of Brownian motions. About this Bertola and Caballero (1994) say: '... any empirical importance of these issues is overshadowed by the substantial simplification of the analytical and estimation problem' (Bertola and Caballero, 1994, p. 242).

In addition to these simplifications, note that data on the Dutch office markets is available on an annual basis only. Since this constitutes relatively low-frequency data, this motivates another simplification. We assume that the length of the time interval relative to the time-scale at which the infinitesimal processes in our micromodel operate is large. Hence, we assume that by the end of a period, the effects of the change in the aggregate growth rate at the beginning of the period have petered out and the cross-section distribution has settled into its steady state distribution associated with a drift rate of μ_t.

4 Empirical implications and evidence

In this section, we assess the aggregate empirical implications and importance of microeconomic lumpy adjustment. In order to do so, we first have to find an estimate for aggregate desired office space use z^d. Subsequently, we calibrate the micromodel parameters (μ, σ, γ, r).

Aggregate desired office space use

Consider again the individual firm. The firm takes input prices and the rate of output as given and minimises cost conditional on input prices and output. We assume that the firm has two inputs, labour L and office space Z (capitals indicate levels whereas lowercase letters indicate logarithms) with wage rate W and rental rate R. Denote output (= value added) by office-using sectors by Y. Additionally, we use a unit of output as numeraire so both rent and wage rate are deflated by the output price index for office-using sectors.

In a frictionless world, cost minimisation yields a cost function as a function of output and input prices alone. Denote this cost function by $C(Y, R, W)$. Frictionless conditional factor demand is then given by $Z^d = \partial C/\partial R$ and $L = \partial C/\partial W$. In the presence of relocation cost, total costs consist of frictionless cost, costs of relocation and costs associated with deviation of actual office space use from desired office space use. We model the latter as in Section 2, that is quadratic in the difference of the logarithm of actual and desired office space use. Hence, the rate of total costs are given by:

$$C(Y, R, W) + \kappa\left[\tfrac{1}{2}(z - z^d)^2 + \phi(\xi)\right] \tag{26}$$

The present value of expected future cost is then given as:

$$E\left[\int_0^\infty C(Y, R, W)e^{-rt}dt\right] + \kappa E\left[\int_0^\infty \tfrac{1}{2}(z - z^d)^2 e^{-rt}dt + \sum_{n=0}^{\infty}\phi(\xi_n)e^{-rT_n}\right] \tag{27}$$

The first term is given to the firm whereas the second term is just κ times the value function of equation (3). Hence, we see that minimising total cost as given in (27) is equivalent to the cost minimisation problem of Section 2.

To find an estimate for z^d we now only have to specify a frictionless cost function $C(Y, R, W)$, take its derivative with respect to R and transform to logarithms. What should the functional form of $C(Y, R, W)$ be? To get a clue, we graphed the share of office space expenditure in output and the logarithm of the rental rate over the wage rate in Figure 5.1 (we standardised both series to have zero mean and unit variance to fit into one graph). Obviously, these two series have a lot in common which leads us to consider a translog functional form for the frictionless cost function. Hence:

$$\begin{aligned}\log C(Y, R, W) &= \alpha_0 + \alpha_1 \log R + (1 - \alpha_1)\log W + \tfrac{1}{2}\alpha_2(\log R)^2 \\ &\quad - a_2(\log R)(\log W) + \tfrac{1}{2}\alpha_2(\log W)^2 + \alpha_3 \log Y \end{aligned} \tag{28}$$

Note that in the absence of frictions, profit maximisation yields $C = (1/\alpha_3)Y$. Usually α_3 is restricted to unity so that $C = Y$, that is no profit is made. In our case this cannot be imposed because in addition to frictionless cost C, we also have costs associated with the friction. Setting α_3 to unity would then imply that the firm would

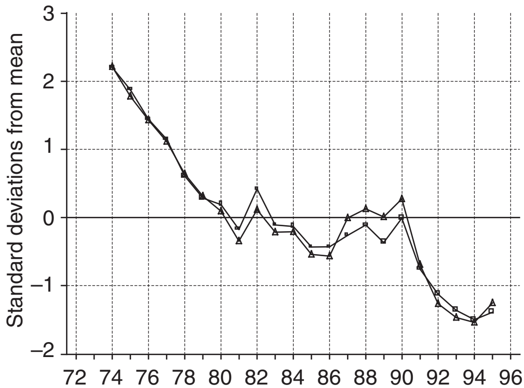

Figure 5.1 The office space expenditure share and the rent–wage ratio

The share of office space expenditure in output (☐) and the logarithm of the rental rate over the wage ratio (△). Both series have been standardised to have zero mean and unit variance over the sample for scaling.

never make a positive profit and make a loss some of the time. This is clearly inconsistent. Instead we impose that the firm should make a strictly positive profit when it is at its desired demand, that is $\alpha_3 > 1$.

Conditional desired demand for office space can be found by differentiating frictionless cost with respect to the office space rental rate. Using equation (27) we find:

$$S^d = \alpha_3\alpha_1 + \alpha_3\alpha_2[\log R - \log W] \tag{29}$$

with S^d the desired share of office space expenditure in output, that is $S^d = RZ^d/Y$. Using equation (23), we see that $S^d = S/U$ with S the actual share. From the way in which U is constructed, we know that it must be stationary. Figure 5.1 indicates, however, that the office expenditure share is not stationary over the sample period. Linearising the relation between S^d, S and U, we see that a co-integrating relation exists between S and S^d. Hence, to obtain an estimate of S^d, we estimate a co-integrating relation between S and $r - w$. The results are reported in Table 5.1. The Vector Error Correction Model (VECM) that was ultimately used to calculate the co-integrating relation is of order 2. The deterministic term has an unrestricted constant

Table 5.1 Co-integration between office space and the rent–wage ratio

This table contains the results of the Johansen's trace test for co-integration. The associated vector error correction model (VECM) is of order 2 with unrestricted constant and linear trend restricted in the co-integration relation, that is:

$$\Delta X_t = \mu_0 + \alpha(\mu_1 t + \beta' X_{t-1}) + \Gamma_1 \Delta X_{t-1} + \Gamma_2 \Delta X_{t-2} + u_t \quad u_t \sim IIN(0, \Omega)$$

with $X_t = (S, r - w)'_t$. For a detailed description of the tests and the issues involved see Johansen (1995).

Johansen co-integration likelihood ratio test statistics		
eigen values	0.561	0.357
trace test	24.0*	8.40

Normalised co-integrating relation β (standard error)

S	$r - w$	trend (1966 = 1)	constant
1	−0.0609	0.000551	−0.426
	(0.00368)	(6.6E-5)	

* Significant at 10 per cent using critical values reported in Johansen (1995, section 15.3, Table 15.4).

and a trend component that is restricted in the co-integrating relation. Hence, the estimated co-integrating relation between S and $r - w$ includes a time trend. At a significance level of 10 per cent, we find one co-integrating relation. This co-integrating relation is then used to calculate desired frictionless share of office space expenditures S^d. The growth rate of desired frictionless office space use can then be found as $dz^d = ds^d + dy - dr$.

Figure 5.2 contains the growth rates for actual aggregate office space use and desired office space use. We see that desired office space use is a lot more volatile than actual demand. This is also indicated by the summary statistics in Table 5.2. The growth rate standard deviation of actual office space use is 1.43 per cent per annum, whereas the growth rate standard deviation of desired office space use equals 2.28 per cent per annum.

The cross-correlation reported in Table 5.2 show that the contemporaneous correlation between the growth rates of actual and

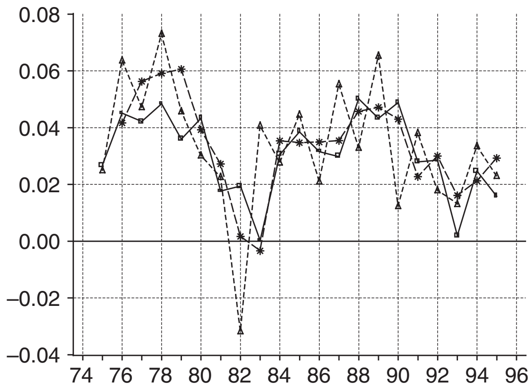

Figure 5.2 Growth rate of office space use in the Netherlands.

Actual office space use (□, solid), desired office space use (Δ) and fitted office space use (*, dashed).

Table 5.2 Summary statistics

	Δz	Δz^d	$\Delta \hat{z}$
Sample	75–95	75–95	76–95
Number of obs.	21	21	20
Mean	0.0310	0.0334	0.0339
Standard error	0.0143	0.0228	0.0169
Skewness	−0.591	−0.691	−0.466
Kurtosis	2.54	4.31	2.80
Jarque-Bera	1.41	3.18	0.760
Probability	0.495	0.0204	0.684
AR(1) parameter	0.397	0.0338	0.640
t-value	1.77	0.143	3.46
Cross-correlations			
k	$\text{corr}(\Delta z_t, \Delta z^d_{t+k})$	$\text{corr}(\Delta z_t, \Delta \hat{z}_{t+k})$	$\text{corr}(\Delta \hat{z}_t, \Delta z^d_{t+k})$
−3	0.049	−0.194	0.097
−2	0.231	0.187	0.237
−1	0.700	0.589	0.833
0	0.435	0.821	0.590
1	0.315	0.517	0.285
2	−0.118	0.206	−0.039
3	−0.081	−0.168	−0.278

desired aggregate office space use is only 0.435 leaving ample space for improvement. Additionally, from the cross correlation pattern, we see that actual office space use lags desired office space by approximately one year. This can also be seen in Figure 5.2.

Calibration of CBP model parameters

In Section 2, we assumed that desired office space use to be a $(-\mu, \sigma)$ Brownian motion. This implies that the increments for desired office space use are identically independently normally distributed. The Jarque-Bera test for normality and the AR(1) parameter and its t-value that are also reported in Table 5.2, do not indicate important deviations from these assumptions. Hence, using the figures in Table 5.2, we set $\mu = -0.0334$ and $\rho\sigma = 0.0228$. Note that this choice of parameters implicitly sets the unit of time to a year.

Next, we need to find a value for the discount rate r. The discount rate only affects the boundaries of the inaction interval (s, S) and the optimal value Q. Moreover, these parameters are not very sensitive to the actual choice of r, so that the precise choice of r is not very critical. We set the discount rate at 5 per cent per annum. This is a reasonable choice that is also frequently employed in the real business cycle literature.

Finally, we need to find values for the correlation between individual and aggregate shocks ρ, and for the lump-sum moving cost γ. This is, however, a bit of a problem since we do not have any information regarding their magnitude. We estimate γ and ρ so that the growth rates of calculated and actual demand for office space resemble each other as much as possible. We do this by minimising the variance of the difference between calculated and actual demand using a grid search. This yields $\rho = 0.15$ and $\gamma = 0.025$. This choice implies that individual shocks correlate relatively weakly with aggregate shocks and that a large share of the shocks faced by individual firms is purely idiosyncratic.

Implied aggregate office space use

Having found values for our micromodel parameters, we can now set out to calculate aggregate office space use as implied by our model. Denote its logarithm by $\hat{z}$ and hence its growth rate by $\Delta\hat{z}$. The micromodel parameters are used to calculate values for the CBP bounds (s, S, Q). Next we calculate a value for u_t using equation (25) with

θ substituted by $\theta_t = 2\mu_t/\sigma^2 = -2\Delta z_t^d/\sigma^2$. Adding the change in u_t to the growth rate of desired aggregate office space use, we obtain the growth rate of fitted aggregate office space use as implied by our model, that is:

$$\Delta\hat{z}_t = \Delta z_t^d + \Delta u_t \tag{30}$$

We plotted the growth rate of fitted aggregate office space use in Figure 5.2 together with actual and desired aggregate office space use. We see clearly that fitted aggregate office space use tracks actual office space use much better than desired office space use. This is confirmed by the contemporaneous correlations which equal 0.821 for the growth rates of actual and fitted office space and only 0.435 between actual and desired office space use. Additionally, the cross-correlation pattern shows that the time series patterns of the growth rates of actual and fitted office space use coincide since the contemporaneous correlation is the largest cross-correlation and the cross-correlations taper off symmetrically in both directions. Also, just like actual office space use, fitted office space lags desired office space use by about one year.

Let us take a closer look at the relation between $\Delta\hat{z}$ and Δz^d. u_t is calculated as a function f of Δz_t^d with $f(\Delta z_t^d)$ given by equation (25) with θ substituted by $\theta_t = -2\Delta z_t^d/\sigma^2$. Hence, we have:

$$\Delta\hat{z}_t = \Delta z_t^d + \Delta f(\Delta z_t^d) \tag{31}$$

Now define:

$$\omega(\Delta z_t^d, \Delta z_{t-1}^d) = -\frac{\Delta f(\Delta z_t^d)}{\Delta z_t^d - \Delta z_{t-1}^d} \tag{32}$$

and we can rewrite (31) as:

$$\Delta\hat{z}_t = [1 - \omega(\Delta z_t^d, \Delta z_{t-1}^d)]\Delta z_t^d + [\omega(\Delta z_t^d, \Delta z_{t-1}^d)]\Delta z_{t-1}^d \tag{33}$$

We see that the growth rate of fitted aggregate office space use is a weighted average of Δz_t^d and Δz_{t-1}^d with time-varying weights that depend on the current and lagged state of the economy (that is aggregate desired office space use). At any point in time, only a fraction of the firms will actually relocate so that only a fraction of actual office space use is determined by current market conditions. Also the fraction of firms that relocates in a certain period will depend on market conditions in that period. If, for instance, during a time period market conditions remain relatively stable, only few firms will relocate, making the fraction of actual office space use that depends on current market conditions small. Highly volatile market conditions will induce a lot of firms to relocate implying that a large fraction of actual demand depends on current market conditions.

But suppose that we can approximate it well by a linear function $g(\Delta z^d) = f(-\mu) + f'(-\mu)(\Delta z^d + \mu)$. In that case $\omega(\Delta z_t^d, \Delta z_{t-1}^d) = -f'(-\mu)$ and equation (33) reduces to a simple weighted average of the current and lagged state. Figure 5.3 contains the graph of the function f for our choice of micromodel parameters ($\mu = -0.0334$, $\rho = 0.15$, $\rho\sigma = 0.0228$, $\gamma = 0.025$). We see that this function is highly non-linear so that generally a linear approximation will not yield satisfactory results. However, the figure also contains the actually observed values for Δz^d and the associated values for u (indicated as ◯). We see that these observed values all lie in a relatively narrow margin for which a linear approximation may well be adequate. From the graph

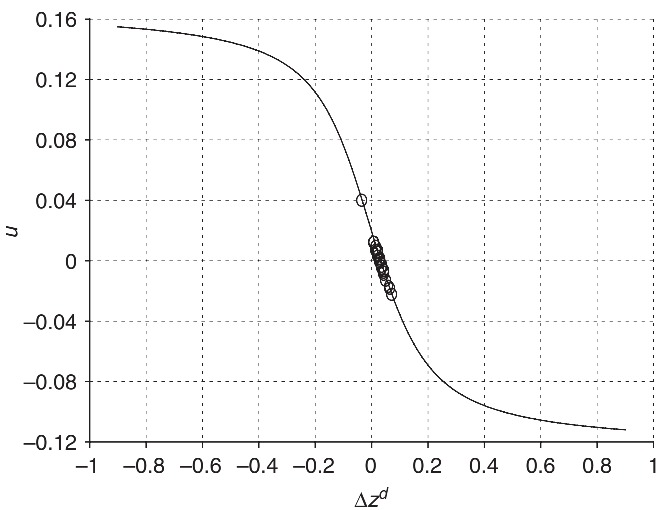

Figure 5.3 Functional relation $u = f(\Delta z^d)$

Functional relation $u = f(\Delta z^d)$ (—) and observed values (◯) for Dutch office space use. Micro model parameters: $\mu = -0.0334$, $\rho = 0.15$, $\rho\sigma = 0.0228$, $\gamma = 0.025$.

we conclude that the non-linearity of f only becomes important for values of the growth rate of desired office space use below −10 per cent and above 15 per cent. Our dataset does not include values of those magnitudes.

The above merits an investigation whether we cannot simply explain the growth rate of actual office space use by a fix-weight average of the current and lagged growth rates of desired office space use. To investigate this, we run three simple regressions of the growth rate of actual office space use on (1) the growth rate of desired office space use, (2) the growth rate of desired office space use and lagged growth rate of desired office space use, and (3) the growth rate of fitted office space use. The results are reported in Table 5.3.

The results indicate that regression (1) performs poorly compared to the other two and should be discarded as a model for the use of office space. The statistics of regressions (2) and (3), however, do not differ very much although regression (3) performs slightly better on all statistics.[3] We interpret the results in Table 5.3 as implying that the apparent dependence of the actual aggregate net absorption rate on the current and lagged state as indicated by regression (2) can be accounted for by our relocation-cost-cum-stochastic-aggregation model. Apparently, the restrictions that are imposed by our stochastic

Table 5.3 Office space use growth rate regressions

Independent (x)	**(1) Δz^d**	**(2) Δz^d**	**(3) $\Delta \hat{z}$**
constant	0.0220	0.00768	0.00713
	(4.03)	(1.66)	(1.62)
x_t	0.272	0.257	0.710
	(2.03)	(2.90)	(6.09)
x_{t-1}	–	0.436	–
		(4.92)	
R^2	0.186	0.664	0.673
$\bar{R}^2$	0.141	0.625	0.655
AIC	−8.50	−9.29	−9.42
SC	−8.40	−9.14	−9.32

Dependent variable: Δz
Sample: 1976–95
$\bar{R}^2$: adjusted R^2
AIC: Akaike Information Criterion
SC: Schwartz Criterion

aggregation model on the relative importance of the current versus the lagged state variable constitute an improvement over a simple model with constant weights, although the degree of improvement is not dramatic. It also confirms our suspicions that the non-linearity of the function f in Figure 5.3 is not very important for the dataset at our disposal.

5 Summary and conclusion

This chapter sets up and solves a model for the demand for office space by individual firms with lumpy adjustment costs and studies its implications for aggregate office space use when both idiosyncratic and aggregate uncertainty are present. Additionally, it provides some empirical evidence for the model using aggregate time series data for the Dutch office market over the period 1974–95.

The most distinguishing feature of the micromodel is the lump-sum adjustment cost. These are motivated by noting that firms in order to adjust their demand for office space in many instances have to relocate and that this entails moving costs that are – at least to a large extent – independent of the amount of office space rented or the change therein. The resulting behaviour by firms is a so-called control band policy or (s, S)-rule in which a firm only adjusts its demand for office space at discrete points in time when the deviation of actual from desired office space use exceeds a certain threshold. When the deviation is smaller than this threshold, the deviation is said to fall within the inaction interval and the firm will not relocate. This obviously describes an important feature of actual behaviour as firm relocations are generally infrequent whereas business conditions change frequently and significantly.

We then go on to investigate the aggregate implications of this lumpy individual behaviour. The rate of relocation is determined by the measure of firms that are in the immediate neighbourhood of their relocation threshold. Hence, we have to find the distribution of firms over the inaction interval. Specifically, we need the cross-section distribution of the deviation of actual from desired office space use at the firm level. When no aggregate uncertainty is present, this distribution will settle into a steady state that equals the ergodic distribution of the process followed by the deviation for a single firm. However, in the presence of aggregate uncertainty, this convenient relation breaks down since then individual shocks are correlated among each other. Although theoretically it is possible to track the cross-sectional distribution over time, no closed-form solution exists. We prefer to follow the approach of Bertola and Caballero (1994) who approximate the infinite variation of the aggregate shock by a discrete variation in the mean growth rate of individual shocks. This approach lends itself readily to further analytical and empirical work while preserving the most important features of the model. We find that the logarithm of actual aggregate office space use equals the logarithm of desired aggregate office space use plus the cross sectional mean of the log-deviations at the firm level. The latter depends on the growth rate of desired aggregate office space use, implying that actual office space use is a weighted average of current and lagged desired office space, with weights that are time-varying and dependent on current and lagged desired office space.

We apply the above lumpy-adjustment-cum-stochastic-aggregation model to aggregate time series for the Dutch office market. We find that aggregate desired office space use is much more volatile than actual aggregate office space and does not track actual office space use very well. This implies that for the Dutch office market a simple static model for the demand for office space, which could be compared to the approach taken by Rosen (1984), is not adequate. Calculated aggregate office space use as implied by our model tracks actual office space use much better. This indicates that deviations of desired office space use and actual office space use can be accounted for by lumpy adjustment at the individual unit's level. Remarkably, we find that, for the Dutch office market data, a fix-weight weighted average of current and lagged desired office space use constitutes a good approximation of the time-varying weighted average. This is due to the fact that the variation in the data is too small to make the weights vary very much over time. Hence, we see that the office space use in the Netherlands can be described nearly equally well by some form of the fix-weight partial-adjustment approach as taken by Wheaton (1987). However, the partial-adjustment model does not apply at the individual firm level and hence it is not clear what economic principles lie at the heart of the partial-adjustment model. The model proposed in this chapter explicitly looks at microeconomic behaviour and in fact rationalises the ad hoc partial-adjustment assumption at the aggregate level from microeconomic principles. Additionally, it shows the limitations of the partial-adjustment approach since the fix-weight partial-adjustment approach is an adequate approximation only when the variation in the data is not too large. When the data are more volatile, the fix-weight partial-adjustment model no longer constitutes an adequate approximation to our model and the advantages of our model should become more apparent.

One way to look into this is to look at a more localised market. The aggregate shocks in our data cover all of the Netherlands and it is likely that the shocks observed at the national level smooth out the shocks at regional or local levels. Hence, we expect that at the regional or local level, the office markets exhibit much more volatility so that the time-variation of the weights as implied by our model become much more pronounced. Hence, it is interesting to take a look at the office market of Amsterdam for which regional accounts exist and for which the office market is relatively well documented. This is however a topic for future research.

6 Appendix

In this appendix we derive the steady-state distribution of a (μ, σ) Brownian motion u in a (s, S, Q) control band. Approximate the continuous time process for u by a discrete time–discrete state process with time jumps Δt and state jumps Δh. At any point in time, this approximating process can jump from u at time t_0 to $u + \Delta h$ or $u - \Delta h$ at time $t_0 + \Delta t$ with probabilities p and q respectively. To make sure that this discrete process converges to the actual continuous time process, we impose the following:

$$\begin{aligned} &\Delta h = \sigma\sqrt{\Delta t} \\ &p = \tfrac{1}{2}\left[1 + \tfrac{\mu}{\sigma}\sqrt{\Delta t}\right], q = \tfrac{1}{2}\left[1 - \tfrac{\mu}{\sigma}\sqrt{\Delta t}\right] \end{aligned} \tag{34}$$

The density function ϕ_u can be found by solving the following difference equation (using steady-state occupancy rates):

$$\phi(u) = p\phi(u - \Delta h) + q\phi(u + \Delta h), u \neq Q \tag{35}$$

Expanding this expression around u, dividing by Δh^2 and letting $\Delta h \to 0$, we obtain (see Dixit and Pindyck, 1994, pp. 83–4):

$$\phi''(u) = \theta\phi'(u), \theta = \frac{2\mu}{\sigma^2}, u \neq 0 \tag{36}$$

Equation (36) is an ordinary differential equation with general solution:

$$\phi(u) = \begin{cases} A_1 e^{\theta u} + B_1, & s < u < Q \\ A_2 e^{\theta u} + B_2, & Q < u < S \end{cases} \tag{37}$$

The full solution to (36) can be found by substituting equation (37) into the boundary conditions. The boundary conditions are given by:

$$\lim_{u \uparrow Q} \phi(u) = \lim_{u \downarrow Q} \phi(u) \equiv \phi(Q) \tag{38}$$

$$\phi(s) = \phi(S) = 0 \tag{39}$$

Note that ϕ is continuous at Q but not (necessarily) differentiable. The boundary conditions in (38) imply $B_1 = -A_1 e^{\theta s}$ and $B_2 = -A_2 e^{\theta S}$ which yields:

$$\phi(u) = \begin{cases} A_1 [e^{\theta u} - e^{\theta s}], & s < u < Q \\ A_2 [e^{\theta u} - e^{\theta S}], & Q < u < S \end{cases} \tag{40}$$

Note that, as $s < Q < S$ and a proper density should be nonnegative, it follows that $A_1 \geq 0$ and $A_2 \leq 0$ when $\theta > 0$ and vice versa when $\theta < 0$. Substituting equation (24), (40) we obtain:

$$\frac{A_1}{A_2} = \frac{e^{\theta Q} - e^{\theta S}}{e^{\theta Q} - e^{\theta s}} < 0 \tag{41}$$

which we rewrite as $A_1 = cA_2$ with c defined as the right-hand side of (41).

The final condition we use to determine the constants of the differential equation is that the integral of any proper density function over its support should equal unity. This yields:

$$\begin{aligned} &\int_s^S \phi(u) du = \int_s^Q cA_2 (e^{\theta u} - e^{\theta s}) du + \int_Q^S A_2 (e^{\theta u} - e^{\theta S}) du = 1 \\ &\Rightarrow \; A_2 = -[ce^{\theta s}(Q - s) + e^{\theta S}(S - Q)]^{-1} \end{aligned} \tag{42}$$

which completes the characterisation of the steady-state density function of $u(t)$.

Finally, Figure 5.1 contains a simulated density function for 100 000 replications, with $\mu = 0.1$, $\sigma = 1$, $s = -10$, $Q = 5$, $S = 5$. We see that the density consists of two exponential distributions with most probability mass to the right of Q due to the positive drift term.

Notes

1. For more on (s, S)-rules and lump-sum adjustment costs, see for instance Blanchard and Fischer (1989, chapter 8), or Caballero and Engel (1991).

2. Note that normally the drift rate affects the parameters of the CBP. However, in this case the changing drift rate is an approximation to an aggregate process. To the individual firms the growth rate remains constant and hence does not affect the CBP parameters. So we let the drift rate vary with unchanging CBP parameters.
3. We have to bear in mind, however, that the application of the stochastic aggregation model involves the 'estimation' of two additional parameters, that is ρ and γ, the sampling variability of which have not been taken into account when comparing models (2) and (3).

References

Bertola, G. and R.J. Caballero (1994), 'Irreversibility and Aggregate Investment', *Review of Economic Studies*, 61, 223–46.

Blanchard, O.J. and S. Fischer (1989), *Lectures on Macroeconomics*, Cambridge, MA, MIT Press.

Caballero, R.J. and E.M.R.A. Engel (1991), 'Dynamic (S,s) Economies', *Econometrica*, 59(6), 1659–86.

Dixit, A.K. and R.S. Pindyck (1994), *Investment Under Uncertainty*, Princeton, NJ, Princeton University Press.

Harrison, J.M., T.M. Sellke and A.J. Taylor (1983), 'Impulse Control of Brownian Motion', *Mathematics of Operations Research*, 8(3), 454–66.

Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford, UK, Oxford University Press.

Romijn, G. (1997), 'Office Space in the Netherlands, 1974–1995', *Tinbergen Institute Discussion Paper Series*, 97-112/3.

Romijn, G., J. Hakfoort and R. Lie (1996), 'A Model of the Demand for Office Space per Worker', *Tinbergen Institute Discussion Paper Series*, 96-85/5.

Rosen, K.T. (1984), 'Toward a Model of the Office Building Sector', *AREUEA Journal*, 12(3), 261–9.

Wheaton, W.C. (1987), 'The Cyclic of Behaviour of the National Office Market', *AREUEA Journal*, 15(4), 281–99.

6
Rent Growth Control and the Transition of Land to Urban Use

Alastair McFarlane

1 Introduction

The notion that a rigid price-ceiling discourages investment in rental housing has led to the introduction of more flexible rent controls over the past few decades. What distinguishes this 'second generation' of rent regulations is that the upward adjustment of rents is allowed to account for inflation and other costs of holding residential real estate. Another characteristic that is common among these dynamically oriented rent controls is that the initial contract rent is not subject to regulation. It is only the growth of rents that is regulated. Yet, in much of economic theory, the static rent ceiling is still used as a benchmark case. Arnott (1995) has called for additional conceptual work to fill this void. It is the purpose of this chapter to make a small contribution to the understanding of how the construction of urban housing may be affected by a flexible rent ceiling.

There are two points that must be made at the outset of this chapter. The first is that although this work was inspired by the case of Switzerland, it is not meant to be a model of the Swiss housing market. Rather, it is intended as a model of an 'endogenous rent-ceiling', which is only one aspect of Swiss rent regulations. By 'endogenous', I refer to the property that the initial rent on rental housing is the result of a negotiation between the tenant and landlord. By choosing when to lease the apartment, a landlord is explicitly choosing the base from which future rents will evolve. Fortunately, the generality of the model presented in this chapter is not a loss as there are many other countries in which rent regulations of this nature can be found. For example, Austria, Belgium, Finland, France, the Netherlands, Portugal, and Sweden are all reported to have rent controls such that the initial setting of rents on private rental housing is unrestricted but increases are governed by regulations (European Economic Commission, 1991, pp. 31–2).

The second remark to be made is that this chapter ignores issues that are very much the centre of other work on housing investment and rent regulations.[1] I refer especially to maintenance and housing quality. Although most work emphasizes the negative effects of these regulations on investment, there is a literature that points to the hidden incentive effects that different designs may have on maintenance. Kutty (1996) and Olsen (1988) both find ambiguous effects of rent controls on maintenance activity. Instead, in this essay, I discuss the construction of new housing. One motivation for this focus was to provide a simple model of the conversion of vacant to urban land. Another motivation is that when rent controls are flexible new construction is a significant portion of housing investment. Most work by Swiss economists point to the adverse effects of rent growth control on housing investment (for a classic example, see Lambelet and Zimmermann, 1991). An exception is the work of Raess and von Ungern-Sternberg (1999), who suggest that a policy of indexing rents may be socially preferable to an unregulated market when search costs are present.

2 Model of urban land market

In order to analyse the likely effect of rent growth control on the supply of built urban land, I start with a model of the land market free of regulations. To this purpose, I choose the Capozza-Helsley (1989) model of land development and durable housing in an open and monocentric city. The following notation will be used in our discussion:

$R(t,z)$	rent at time t and distance z from the centre, with $R_t(t,z)$ the first derivative with respect to time and $R_{tt}(t,z)$ the second derivative
A	rent on agricultural (undeveloped) land
r	real interest rate
C	cost of building one unit of housing

t	current time
t^*	optimal time of conversion in an unregulated market
$\tilde{t}$	optimal time of conversion in a regulated market
T	per distance transportation cost
$y(t)$	income at time t
x	consumption of the non-housing good

2.1 Market rents

Land upon which no structure is built is assumed to earn an agricultural rent, A, that is invariable across space and time. Development of land yields an annual urban rent, $R(t,z)$, that varies over time, t, and distance, z. Following the exposition of Capozza and Helsley (1990), I assume that the residents consume one unit of housing each so that their budget constraint at each point in time is $y(t) = R + x + Tz$. Income, $y(t)$, is divided between urban land rent, R, a consumption good, x, and transportation costs. Transportation costs are assumed to be linear so that expenditures are equal to the distance from the Central Business District (CBD), z, times the per distance transport cost, T.

The implication of the open city assumption is that given perfect mobility of consumers, any advantages of living in a particular city will be capitalized into the rent of urban land. Given the possibility of arbitrage by consumers, rents will adjust so that net revenue (income less rent and transport cost) is equivalent in all cities and at every location within a city. Substituting this condition into the budget constraint and rearranging yields a rent function:

$$R(t,z) = y(t) - \bar{x}(t) - Tz \tag{1}$$

where $\bar{x}(t)$ is the consumer's 'reference' level of consumption. If by living at a particular site, the consumer cannot attain this common reference level of consumption, the consumer will move to another location or to another city. Rents are assumed to be growing over time, which requires that $y'(t) > \bar{x}'(t)$. $R(t,0)$ is assumed to be a deterministic and non-decreasing, unbounded function of time such that $R(t,0) \leq \bar{R} + R(0,0)e^{gt}$ for constants $\bar{R}$ and g such that $g < r$ and $\bar{R} \geq 0$. This condition guarantees that the present value of future rents is finite.

2.2 Rents under rent growth control

Switzerland's current rent regulation has been in place since 1972. The policy was strengthened even further when a federal law for renter protection came into effect on July 1, 1990. All residential buildings are subject to the regulation. The spirit of the rent growth control is that landlords may increase rents only if the cost of providing housing increases. The tenant has the right to object to any increase that is not directly related to an increase in maintenance costs, the interest rate[2] or inflation rate. However, landlords are allowed to charge the prevailing market rent on new apartments. The legal limit imposed upon the initial rent is that it must be below what is defined as an 'excessive' rent. A rent is deemed to be excessive either if it gives the landlord a return that is more than one half of a percentage point greater than the mortgage rate (Lachat and Micheli, 1992, p. 210) or if it is greater than what is normal for the neighbourhood (Lachat and Micheli, 1992, p. 220). However, it is the responsibility of the tenant to oppose the initial rent or any rent increase. In this chapter I make a number of assumptions concerning the rent regulation. They are as follows:

1. An excessive rent is one that is above the average rent of the neighbourhood in question.
2. A neighbourhood is defined as the set of locations sharing the same distance from the central business district.
3. There are no transaction costs to contesting an excessive rent.
4. Tenants have perfect information concerning rent regulations and the normal rent of their neighbourhood.
5. The market for housing is perfectly competitive.

Assumption 1 defines the legal limit of a rent. Assumption 2 is a stylized definition of a neighbourhood. Assumptions 3 and 4 ensure that tenants will register a complaint against a contract that could be considered illegal. It should be noted that this model possesses none of the information asymmetries found in models of tenant mobility (for example, see Basu and Emerson, 1998). These assumptions will result in the prevailing market rent being the average neighbourhood rent. Why? As I will show the city expands from the centre (see Section 3). Thus, all houses in the same neighbourhood will be developed simultaneously. If the market is perfectly competitive, developers will compete amongst one another for tenants. This competition will make it impossible for a landlord to charge anything above the market rent. Second, as development of all of the housing in the same neighbourhood occurs at the same time, the rent on every apartment will equal the average rent for the neighbourhood. The outcome is that while the rents of housing near the central city will not reflect market demand, newly built housing will.

2.3 The objective function of a developer in an unregulated market

The price of developed land, the output from construction, is given by the present value of discounted future rents:

$$P^d(t,z) = \int_t^{\infty} R(s,z) \cdot e^{-r(s-t)} ds$$

where $P^d(t,z)$ is the price of urban land at time t and location z. The present value is found by discounting the urban rents $R(s,z)$ from now, time t, to infinity by the interest rate, r. The rent received on urban land is a function of the time, s, and distance from the CBD, z.

The economic problem of the developer is to choose the time to convert agricultural land to urban land in order to maximize the value of vacant agricultural land, $P^v(t,z)$. Even if the current owner of the land is not willing or able to make the development decision, the market value of vacant land will reflect its best possible use. Thus, the price of vacant land is given by the optimal solution to the developer's problem. The objective function is:

$$\max_{t^*} P^v(t,z) = \int_t^{t^*} Ae^{-r(s-t)} ds + \int_{t^*}^{\infty} R(s,z)e^{-r(s-t)} ds - Ce^{-r(t^*-t)} \quad (2)$$

The value of agricultural land is equal to the present value of the agricultural rents, A, earned from the current time, t, to the time of development, t^*, plus the present value of urban rents earned after development less the cost of construction, C, paid at the time of development. Development occurs at a fixed density. Capital is durable and does not depreciate.

The optimal timing condition can be found by differentiating the objective function with respect to t^*. Applying Leibnitz's rule (Spiegel, 1990, p. 163) gives us the first-order condition with respect to the time of development:

$$R(t^*, z) = A + rC \tag{3}$$

This first-order condition has a straightforward interpretation. For development to be a profitable exercise, the urban rent at location z must be at least as great as the benefits of leaving this land vacant. From the right-hand side of equation (3) we see that the opportunity cost of development is the sum of the agricultural rent that is foregone plus the annualized cost of construction. Development occurs the moment that the rent reaches the 'hurdle rent' given by the optimal timing condition. Thus, it is apparent that higher agricultural rents, interest rates and construction costs will delay development. Comparative static results can be arrived at more formally by applying the implicit function theorem (Spiegel, 1990, p. 107) to equation (3). For example, consider the effect of a change on the interest rate on development timing:

$$\frac{\partial t^*}{\partial r} = \frac{C}{R_t(t^*, z)} > 0 \tag{4}$$

As long as rents are growing, $R_t > 0$, development will be delayed. This delay will be greater, as construction costs are larger and as rent growth is lower. Consider the role that location plays in affecting when a parcel of land is converted to urban use. What we find is that land farther away from the centre is developed later:

$$\frac{\partial t^*}{\partial z} = -\frac{R_z(t^*, z)}{R_t(t^*, z)} > 0 \tag{5}$$

Since rents near the CBD are greater, they will pass the hurdle rent before rents on land far from the CBD. From equation (5), we see that development activity will begin at the centre and work its way outwards. There is no incentive for 'leapfrog' development, which is not always the case in dynamic models.[3]

The second-order condition for a local maximum is:

$$R_t(t^*, z) < 0$$

For development to be an optimal strategy, rents must be increasing over time.

In the remainder of this chapter, I will provide examples of solutions and comparative static results assuming specific functional forms for the rent function. In these examples, I assume that there is growth of income but no change in the level of consumption relative to the world level. As long as rents are growing, development occurs sooner in cities where growth is higher:

$$\frac{\partial t^*}{\partial g} = -\frac{R_g(t^*, z)}{R_t(t^*, z)} < 0 \tag{6}$$

This expression gives us the change in the time of conversion if rents had been growing at a higher rate for t^* periods. For the examples of additive growth of income where $R(t, 0) = R(0, 0) + gt$ and geometric growth where $R(t, 0) = R(0, 0)e^{gt}$, the above derivative is equal to $-t^*/g$.

2.4 Development under rent growth control

Under rent growth control, the goal of the land developer is to maximize the value of a plot of vacant land, P^v, by choosing the optimal time of development, $\tilde{t}$, subject to the restriction that the rent be set at the market rent at the time of development. The 'market' rent at the time of development is defined as the rent at which the tenant attains a level of utility such that they are indifferent to moving (see equation (1)). One would expect the endogeneity between the rent and the development decision to lead to a different kind of optimal timing decision than when the market is unregulated.

The maximization problem at time t for a plot of land at distance z from the CBD can be represented by:

$$\max_{\tilde{t}} P^v(t, z) = \int_t^{\tilde{t}} Ae^{-r(s-t)}ds + \int_{\tilde{t}}^{\infty} R(\tilde{t}, z) \cdot e^{-r(s-t)}ds - Ce^{-r(\tilde{t}-t)} \tag{7}$$

By choosing when to build, a developer chooses the rent the building will receive. Assuming that the rent is growing, rent growth control fixes the rent at $R(\tilde{t}, z)$. As before, the optimal timing condition can be found by using Leibnitz's rule to find the first-order condition of the objective function of developer:

$$R(\tilde{t}, z) = A + rC + R_t(\tilde{t}, z)/r \tag{8}$$

In equation (8) we see that the opportunity cost of development under rent growth control is the sum of the agricultural rent that is foregone, the annualized cost of construction and the lost growth of urban rents that would have been gained by waiting to develop. It is the 'withholding premium' that makes a regulated city different from an unregulated one. The faster that rents are growing, the more advantageous it will be to wait before developing. Rent growth control creates an incentive to withhold.

When there is rent growth control and rents are growing, land will be developed later than when the market is unregulated. An informal proof of this proposition can be achieved by comparing equations (3) and (8). If $R_t(t, z) > 0$, then $R(\tilde{t}, z) > R(t^*, z)$. All other things equal, this is possible only if $\tilde{t} > t^*$.

That a rent regulation would discourage investment is not a surprising conclusion. This is a standard result and a standard criticism. However, it is possible to arrive at more subtle conclusions concerning the differences between the two regimes.

To better understand the optimal development strategy, one must consider the second-order condition as well:

$$R_{tt}(\tilde{t}, z)/r - R_t(\tilde{t}, z) < 0 \tag{9}$$

The interpretation is that for development to be an optimal strategy, rents must be growing faster than the withholding premium. As long as rents are growing at a rate less than the real interest rate (as assumed), this second-order condition will be satisfied.

The change in development timing from a change in the interest rate under rent growth control is found by applying the implicit function theorem to equation (8):

$$\frac{\partial \tilde{t}}{\partial r} = \frac{C - R_t(\tilde{t}, z)/r^2}{R_t(\tilde{t}, z) - R_{tt}(\tilde{t}, z)/r}$$

The direction of the effect will depend upon the numerator because, by equation (9), the denominator of the above term is positive. However, the sign of the numerator is ambiguous because under rent growth control, a change of the interest rate has two effects. As before, a higher interest rate will raise the cost of construction in the form of higher annualized interest payments and thus delay conversion of vacant land. The second effect is that an increase of the interest rate reduces the value of the withholding premium. The net effect depends on the difference between construction and the rent growth. Thus, under rent growth control, the effect on timing of a change in the interest rate is ambiguous. As long as $C > R_t/r$, there will be a delay from an increase of the interest rate.

Although the effect of the interest rate under rent growth control is ambiguous, it is possible to compare the magnitude of the effect under rent growth control with that of the unregulated case. Doing so necessitates the evaluation of the denominator of the above comparative static equation. Thus, assumptions concerning the acceleration of rents, R_{tt}, play a role in determining the difference between the two scenarios. Consider the simplest case, where rent growth is linear. It follows that acceleration of rents is zero ($R_{tt} = 0$) and that the denominator is constant. In comparing, we see that $\partial\tilde{t}/\partial r = \partial t^*/\partial r - 1/r^2$. Thus, in the linear case, an increase of the interest rate leads to less of a delay under rent growth control than for the unregulated case. Geometric growth is not as straightforward because the denominator is not constant.[4] It turns out that the delay from an increase in the interest rate could be smaller or larger than the delay in the unregulated case. The larger the growth rate, the likelier that a change in the interest rate will lead to a greater delay under the regulated case.

Next, we can examine the effect of location on timing:

$$\frac{\partial\tilde{t}}{\partial z} = -\frac{R_z(\tilde{t}, z)}{R_t(\tilde{t}, z) - R_{tt}(\tilde{t}, z)/r} > 0 \tag{10}$$

Land that is farther from the CBD is developed later, just as before. From equation (9), we know the denominator to be positive and from equation (1), the numerator is negative. We see that the magnitude of this effect will be different between rent growth control and the unregulated market and will depend on $R_{tt}(t, z)$. If $R_{tt}(t, z) > 0$, then location may be relatively more important under rent growth control.

Comparing equations (3) and (8), it is easy to see that the developer's response to a change in agricultural rents or construction costs will be identical under the two regimes. By allowing initial rent levels to be flexible, the policy of rent growth control does not prevent construction even when the costs of development are high. This would not be the case with a strict-rent ceiling. If we put a ceiling on the level of the rent, then development would proceed up until the point where $A + rC$ equals the rent ceiling and then cease.

Table 6.1 Comparative static results for optimal timing

Case	*A*	*C*	*r*	*z*	*T*	*g*
Unrestricted	+	+	+	+	+	−
Rent control	+	+	?	+	+	?

Finally, we can look at the effect of a change in growth:

$$\frac{\partial \tilde{t}}{\partial g} = -\frac{R_g(\tilde{t}, z) - R_{tg}(\tilde{t}, z)/r}{R_t(\tilde{t}, z) - R_{tt}(\tilde{t}, z)/r} \tag{11}$$

This result cannot be signed even when the denominator is assumed to be positive, as the sign of the numerator is ambiguous. Raising growth has two effects: rents will increase faster but at the same time, the withholding premium will also increase. Thus, a change in the growth of rents has an ambiguous effect on the optimal timing.

Table 6.1 summarizes the comparative static results.

2.5 Example: linear income growth

Consider the linear form of income growth, $y(t) = y(0) + gt$, where income grows by a lump-sum amount, g, every period.[5] If income growth takes this form, then markets rents will follow $R(t, z) = R_0 + gt - Tz$ where $R(0, 0) = y(0) - x$. The expositional advantage of examining the case of linear growth is that the first and second derivatives with respect to time are constant, that is $R_t = g$ and $R_{tt} = 0$.

Given linear growth, the opportunity cost of not withholding for a higher rent is equal to the present value of that lost growth, $R_t/r = g/r$, so that $R(\tilde{t}, z) = A + rC + g/r$. Then, the solutions for the optimal time of development for the regulated and unregulated case respectively are $t^* = [A + rC - (R_0 - z)]/g$ and $\tilde{t} = [A + rC - (R_0 - z)]/g + 1/r$.

Optimal development under the two different regimes is represented in Figure 6.1. The dark line is the time path of rent under rent growth control on a plot of land at distance z, developed at time $\tilde{t}$. Before development, it receives the agricultural rent, A, and after development it receives the market rent at the time of development. In this case, the hurdle rent that gives the optimal time of development is equal to $A + rC + g/r$.

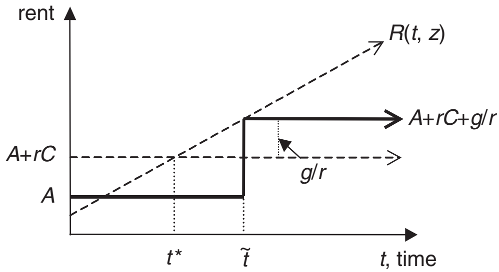

Figure 6.1 Optimal development of land under rent growth control

The optimal solution in an unregulated market as characterized by equation (3) is at t^*. Developers build earlier and at a lower rent than under rent growth control. The hurdle rent in the unregulated case is $A + rC$, and after development, the rent follows $R(t, z)$. What is important to take from Figure 6.1 is that the rent a building earns is not exogenous, rather it is determined by the development decision. We also see that there is a widening between the market rent and the regulated rent over time.

3 Application: the urban–rural boundary

A question that is interesting for cities where the majority of housing is under rental growth control is how the regulation will affect urban growth.

3.1 Urban–rural boundary and the unregulated market

Following Capozza and Helsley (1989), one can rewrite the optimal timing condition, equations (3) and (8), to get an explicit solution for the urban–rural boundary. We can use the fact that at the urban–rural boundary, $\bar{z}(t)$, the optimal timing condition is met:

$$R(t, \bar{z}(t)) = A + rC \tag{12}$$

As there is a unique boundary, z can be replaced by $\bar{z}(t)$. To understand how the city size changes with the radius, the city is assumed to be a solid disk. The area is equal to $\pi\bar{z}^2$. When the equilibrium boundary expands, construction is called for. When the equilibrium boundary shrinks, there is no new construction.

To make these insights more explicit, I rewrite equation (12) in terms of the urban–rural boundary:

$$\bar{z}(t) = \frac{R(t,0) - A - rC}{T} \tag{13}$$

The city will be smaller in an area with higher transport costs, agricultural rents, interest rates and construction costs. Cities will be larger with a larger base rent. As rents grow, development is encouraged and the city expands.

3.2 Urban–rural boundary and rent growth control

The first-order condition for the case of rent growth control, equation (8), in terms of the urban–rural boundary under rent growth control, $\tilde{z}(t)$, is:

$$R(t, \tilde{z}(t)) = A + rC + R_t(t, \tilde{z}(t))/r \tag{14}$$

As with the case of the unregulated market, this can be rewritten explicitly in terms of the urban fringe:

$$\tilde{z}(t) = \frac{R(t,0) - A - rC - R_t(t,0)/r}{T} \tag{15}$$

The city will be smaller in an area with high transport costs, agricultural rents, and construction costs. As with timing, the effect of a change in the interest rate or growth rate is ambiguous.

3.3 Comparative statics and geometric growth

Instead of modelling growth as linear, suppose that it is based on geometric growth. If the time-path of income is given by $y(t) = y(0)e^{gt}$, then the market rent would be given by $R(t,z) = y(0)e^{gt} - \bar{x} - Tz$. Inserting this expression into equation (13) and taking the derivative with respect to time, we get:

$$\frac{\partial \bar{z}(t)}{\partial t} = g \cdot \frac{y(t)}{T} \tag{16}$$

The radius of the city grows over time, always accelerating. Growth is greater when transport costs are lower. Concerning the dynamic path of the regulated city, one sees that the growth of the city is not stopped by rent growth control:

$$\frac{\partial \tilde{z}(t)}{\partial t} = (1 - g/r) \cdot g \cdot \frac{y(t)}{T} \tag{17}$$

But it will not grow as quickly as in an unregulated market. The extent of this slowdown depends upon the ratio between the growth and interest rates.[6] Thus the difference between the two city sizes grows over time with geometric growth. For example, the growth under rent growth control is three-fourths of what it would be in an unregulated market when the growth rate is 1 per cent and the real interest rate is 4 per cent. In low-growth, high-interest rate areas, we should not expect rent growth control to have as adverse an effect on the expansion of the boundary. Rent growth control is most harmful to the supply of urban land precisely when demand is increasing rapidly.

From the above discussions, we see that rent growth control allows the city to grow steadily with demand growth. But what happens when the rate of growth changes unexpectedly? I compare equations (18) and (19):

$$\frac{\partial \bar{z}(t)}{\partial g} = t \cdot \frac{y(t)}{T} \tag{18}$$

Not surprisingly, a city with a higher growth rate will be larger. The response of the boundary under rent growth control to a different growth rate is as follows:

$$\frac{\partial \tilde{z}(t)}{\partial g} = \frac{y(t)}{T} \cdot \left(\left(1 - \frac{g}{r}\right) \cdot t - \frac{1}{r} \right) \tag{19}$$

The direction of the effect is ambiguous because there are two contradictory effects. The first, $y(t)(1-g/r)t/T$, represents what would have been the city size if it had been expanding to meet the faster growth of demand since time zero. The second effect is the speculative withholding effect and captures the fact that the withholding of land is more profitable as the gains to be had by waiting are increased, as can be seen in equation (11).

To understand how a city will react to an 'immediate' change in the growth rate requires careful interpretation. Both equations (18) and (19) depict the effects of having had a different growth rate since $t = 0$. Since this growth has not actually occurred, the city size will not increase by the full amount. However, the withholding effect is realized immediately and thus will dominate in the short-run:

$$\Delta\tilde{z}(t) = \left(-\frac{1}{r} + \left(1 - \frac{g}{r}\right)(t - t_0)\right) \cdot \Delta g \cdot \frac{y(t)}{T}$$

The immediate effect of an increase in the growth rates will be to lower the equilibrium size of the city. By putting the two effects together, we see that an increase in the growth rate will halt construction activity initially. This kind of speculative withholding will be especially evident during periods of high growth. After a period of withholding, building will resume at a later period and at a faster pace.

Under rent growth control, an increase in the growth rate leads to a short-run decrease in the equilibrium city size, followed by a long-run increase.

3.4 An illustrative example

In this section, I present a simulation of city size based on equations (13) and (15). The purpose is to illustrate the impact of rent growth control using empirically plausible values. To start, I assume that the city is built upon a featureless plain. The area of the city is then $\pi\bar{z}^2$ where $\bar{z}$ is the distance of urban–rural boundary from the CBD. The density of development is fixed at one floor. I assume the following magnitudes per 100 square metres: the agricultural rent is 150 francs/year; the cost of construction is 120 000 francs; and the annual rent is 13 000 francs/year, which grows by 150 francs/year. The real interest rate is 2 per cent and the transport cost is 300 francs/km/year. Given these values, the unregulated hurdle rent would be 1325 francs. With rent growth control, the hurdle rent would be 8825 francs. The difference in hurdle rents leads to a significant difference in city size: the unregulated city covers 237 778 hectares and the regulated city has an area of only 30 407 hectares. Thus, it appears that under fairly reasonable assumptions, the rent growth control policy can substantially diminish the supply of land.

4 Application: average rents and prices

After discussing the effects of rent growth control on city size in Section (3), we can use these results (equations (13) and (15)) to derive rents and prices for the two cases.

4.1 Average rents and prices in the unregulated market

The rent at any location can be expressed as a function of the size of the city. The rent at the urban boundary is equal to the hurdle rent. Following Capozza-Helsley (1989), the rent on developed land can be expressed as:

$$R(t, z) = A + rC + T(\bar{z}(t) - z) \tag{20}$$

As the city grows, so does rent on urban land at all locations. Equation (20) is illustrated in Figure 6.2 (taken from Capozza-Helsley, 1989). Rent on urban land consists of location rent, rent on the capital used in construction, and agricultural rent. Outside of the urban boundary, land earns the agricultural rent. Performing the integration over time of the rent as defined in equation (20), and substituting in $\bar{z}'(t) = R_t(s, z)/T$, we get an expression for the price of developed land:

$$P^d(t, z) = \frac{A}{r} + C + \frac{T}{r} \cdot (\bar{z}(t) - z) + \frac{1}{r} \cdot \int_t^{\infty} R_t(s, z) \cdot e^{-r(s-t)} ds \tag{21}$$

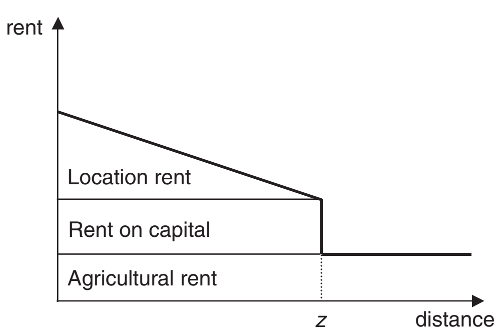

Figure 6.2 Rent in an unregulated market

The price of urban land is equal to the sum of the present value of the agricultural rent, the cost of conversion, the present value of location rent and the value of expected future growth.

The price of agricultural land is defined in equation (2) and by substituting the solution to the optimal timing problem as expressed in equation (3), the price of agricultural land expressed in terms of location is:

$$P^{v}(t,z) = \frac{A}{r} + \frac{1}{r} \cdot \int_{t^*(z)}^{\infty} R_t(s,z) \cdot e^{-r(s-t)} ds \qquad (22)$$

The price of agricultural land is equal to the present value of agricultural rents plus the value of expected growth. The value of expected growth declines with distance from the CBD. This relation is illustrated in Figure 6.3 (taken from Capozza-Helsley, 1989).

The average rents and prices of developed land are found by integrating over z and dividing this sum by the city size. Integrating the rent function as given in equation (20) over z and dividing by $\pi\bar{z}^2$ yields the following expression for the average rent of urban housing:

$$AR(t) = A + rC + \frac{1}{3} T\bar{z}(t) \qquad (23)$$

The average rent in an unregulated city increases with city size.

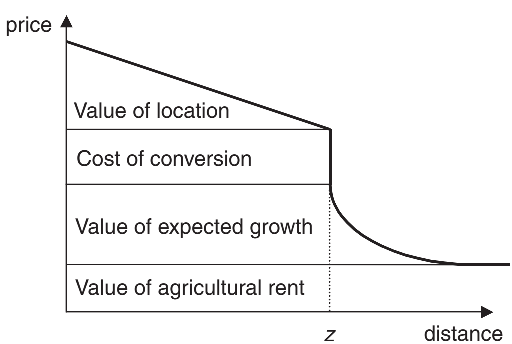

Figure 6.3 Land price in an unregulated market

The average price of urban land is equal to the present value of the average rent plus a growth premium:

$$AP^d(t) = \frac{A}{r} + C + \frac{T}{r} \cdot \frac{\bar{z}(t)}{3} + \frac{1}{r} \cdot \int_t^{\infty} R_t(s, 0) \cdot e^{-r(s-t)} ds \tag{24}$$

The average price of developed land also increases with city size.

4.2 Average rents and prices under rent growth control

The rent on urban land under rent growth control is dependent on location to the extent that location influences when the land is developed and thus the withholding premium. Rewriting equation (8) in terms of z, we have that rents are a function of distance and only incidentally of time:

$$R(t, z) = A + rC + R_t(\tilde{t}(z), z) \tag{25}$$

In Figure 6.4, a rent gradient is illustrated for the case of linear growth in a regulated city. When the first derivative of rents with respect to time is constant, the rent gradient is flat. With linear growth, developers at every location face the same hurdle rent because the withholding premium with linear rent is constant over time. However, inspecting the diagram it appears as if there is a rent ceiling. In a dynamic model, the rent under rent growth control remains at the level of the hurdle rent at the time of development. If the second derivative with respect to time were positive, that is if growth were

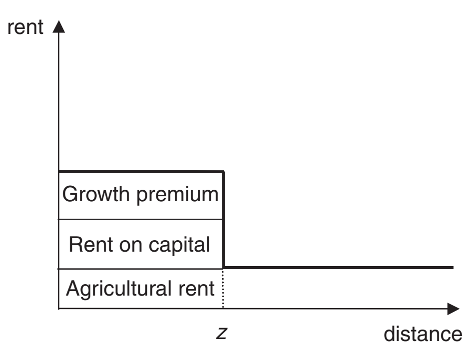

Figure 6.4 Rent under rent growth control (linear growth)

accelerating as in the geometric case, then the rent gradient would be upward sloping.

In the long run, rents under rent growth control are higher at the urban–rural boundary and lower at the city than if the market is unregulated. Compare equations (12) and (14). We know that rents at the boundary must be larger under rent growth control than when they are unregulated because of the withholding premium. Over time, as rents are allowed to grow, the location rent in the unregulated economy will overtake the withholding premium. This confirms the intuition that while such a policy aids established households, it makes housing more expensive for new arrivals.

The price of developed land under the regime of rent growth control is the present value of all future rents and is:

$$\tilde{P}^d(t,z) = \frac{A}{r} + C + \frac{R_t(\tilde{t}(z),z)}{r^2} \tag{26}$$

In Figure 6.5, the case of linear growth[7] is illustrated. If the growth of rents were geometric the price gradient would rise towards the edge and then gently decline.

Following the same steps as above, the price of agricultural land under rent growth control is:

$$\tilde{P}^v(t,z) = \frac{A}{r} + \frac{R_t(\tilde{t}(z),z)}{r} \cdot e^{-r(\tilde{t}(z)-t)} \tag{27}$$

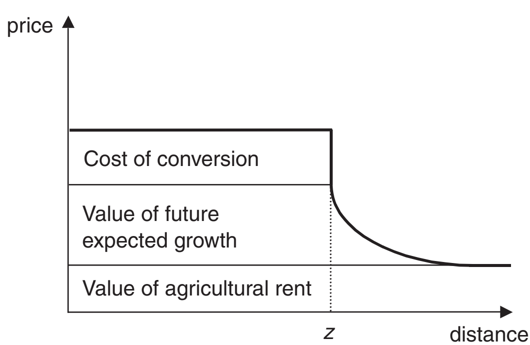

Figure 6.5 Land price under growth control (linear growth)

It is only in comparing the price of vacant land under the two regimes, equations (22) and (27), that we see the source of loss resulting from rent growth control. For land that is already developed, the policy redistributes the value of rent growth directly from landlords to tenants. Thus, there is no net loss. However, for land that has not yet been developed, there is a deadweight loss resulting from the lost rents due to the delay of conversion.

Finding an expression for the average rent in a city involves integrating the expression for rent found in equation (25). Unlike the market rent in equation (20), rents do not increase after development:

$$A\tilde{R}(t) = A + rC + \frac{2}{r\bar{z}(t)^2} \cdot \int_0^{\bar{z}(t)} z \cdot R_t(\tilde{t}(z), z)dz$$

The last term represents the average withholding premium. In the case of additive growth, this expression simplifies to g/r. In the case of geometric growth, the average withholding premium increases over time so that the average rent rises over time. The comparison of rents in an unregulated city with those in a regulated city depends on time, that is on size. In the case of small cities, the regulated one is likely to have higher average rents. The reason being that, in the regulated city, the opportunity cost of development is higher. As time passes and location rents surpass the withholding premium, then average rents will be higher in the unregulated city.

The average price is found by integrating the expression found in equation (25) over distance:

$$A\tilde{P}^d(t) = \frac{A}{r} + C + \frac{2}{r^2\bar{z}(t)^2} \cdot \int_t^{\infty} z \cdot R_t(\tilde{t}(z), z)dz$$

The average value of developed land in a regulated city is equal to the present value of agricultural rents, the cost of construction, and the average value of rent growth that is 'lost' from developing.

5 Fiscal incentives

There is a perception in Switzerland that developers withhold land for speculative purposes (Favarger, 1996). Here, it has been shown that this behaviour may be the result of unexpected growth and rent growth control. In Favarger and McFarlane (1996), a variety of taxes were proposed to regulate land development in Switzerland. One of these is the property tax on vacant land to punish the withholding of land from non-urban use. Here, I would like to explore this proposal and ask whether such a tax could achieve the desired result of hastening development. To do so, I draw on the tax and development literature,[8] especially the work of Anderson (1986).

5.1 The differential property tax

Suppose that the government taxes the value of urban and vacant agricultural land at different rates. Up until the time of development, when land is vacant, taxes are paid on P^v at the rate τ^v. After development, taxes are paid on P^d at the rate τ^d. The present value of the flow of these property tax payments enter the price of land as a cost:

$$P^v(t,z) = \int_t^{\tilde{t}} A \cdot e^{-r(s-t)} ds + \int_{\tilde{t}}^{\infty} R(\tilde{t},z) \cdot e^{-r(s-t)} ds - Ce^{-r(\tilde{t}-t)}$$
$$- \int_t^{\tilde{t}} \tau^v \cdot P^v(s,z) \cdot e^{-r(s-t)} ds - \int_{\tilde{t}}^{\infty} \tau^d \cdot P^d(s,z) \cdot e^{-r(s-t)} ds$$

Following the work of Anderson (1986), the above term can be simplified by taking the time derivative and then re-integrating. Thus, I can express the objective function of the developer as follows:

$$\max_{\tilde{t}} P^v(t,z) = \int_t^{\tilde{t}} A \cdot e^{-(r+\tau^v)(s-t)} ds$$
$$+ \left[\int_{\tilde{t}}^{\infty} R(\tilde{t},z) \cdot e^{-(r+\tau^d)(s-\tilde{t})} ds - C \right] \cdot e^{-(r+\tau^v)(\tilde{t}-t)}$$

Maximizing with respect to the time of development gives us the following first-order condition:

$$R(\tilde{t},z) = A + (r+\tau^v)C + \frac{R_t(\tilde{t},z)}{r+\tau^d} + (\tau^d - \tau^v) \cdot \frac{R(\tilde{t},z)}{r+\tau^d} \qquad (28)$$

We have the condition that development will be optimal when the urban rent is equal to the lost agricultural rent, the rent on capital, the benefit of withholding, and a term that represents the increase in the property tax bill due to conversion. The second-order condition is:

$$R_t(\tilde{t},z) > \frac{R_{tt}(\tilde{t},z)}{r+\tau^d} + (\tau^d - \tau^v) \cdot \frac{R_t(\tilde{t},z)}{r+\tau^d}$$

By applying the implicit function theorem to the first-order condition, equation (28), and using the above second-order condition to show that the denominator is positive, we arrive at the following conclusion:

$$\frac{\partial \tilde{t}}{\partial \tau^v} = -\frac{(r+\tau^d)A + R_t(\tilde{t}, z)}{(r+\tau^v)^2 R_t(\tilde{t}, z) - (r+\tau^v) R_{tt}(\tilde{t}, z)} < 0$$

Given positive rent growth, development can be hastened by raising the property tax on vacant land. Whether or not it is desirable to do so is another question, especially when the distortion that we are trying to correct is one that is the result of government regulation. However, I am not going to explore the normative aspects of the situation. I will merely ask whether tax rates of a reasonable level can achieve the goal of counteracting the withholding incentive.

I start by asking whether the tax can counteract the withholding effect for a specific plot of land. The condition for no extra delay due to regulation at location z is:

$$R(\tilde{t}, z) = R(t^*, z)$$

where the hurdle rent on the left-hand side is that under rent regulation and incentive taxation and that on the right-hand side is in the absence of both taxes and regulation. Substitute equations (3) and (28) into the above condition. For the ease of exposition, suppose that the tax rate on developed land is zero,[9] and rearrange to find a solution for the necessary tax on land value to overcome the delay due to rent regulation. Expressing the timing of development as a function of location, the appropriate tax must be of the following magnitude:

$$\tau^v = \frac{R_t(\tilde{t}(z), z)}{A}$$

Not surprisingly, the incentive tax must be larger with the growth of rents. However, it is not obvious that the tax should vary inversely with agricultural rents. To understand this finding, observe that if agricultural rents were zero, then the tax would have to be infinite. Since the only difference between the two hurdle rents would be the withholding premium, the tax has to reduce the value of the withholding premium to an infinitely small amount. This result casts doubt on the practicability of such an incentive tax. First, it is unrealistic to imagine taxes greater than 100 per cent. However, one can easily imagine that the increase in urban rent from one year to the next would be greater than the level of agricultural rent on the same plot of land. Second, we see that if growth varies over time, then the tax as calculated above does not exactly counteract the withholding effect for all locations. In other words, a constant tax rate is not general unless rent growth is linear. Consider the case of geometric growth. For land closer to the city than at location z above, the tax will lead to earlier development than in the unregulated case. Here, the correct tax must be a function of time. Thus, finding the tax that achieves the goal of counterbalancing the rent regulation for every site is technically quite difficult.

5.2 The construction subsidy

A policy common to countries with housing shortages, and an alternative to the punitive tax on owners of undeveloped land, would be a construction subsidy. Here, I model the subsidy as a percentage reduction, at rate s, of construction costs. The first-order condition with a subsidy and rent growth control is:

$$R(\tilde{t}, z) = A + (1 - s)rC + R_t(\tilde{t}, z)/r$$

Increasing subsidies to construction hastens development activity. As with the land tax, we ask whether a subsidy can be used to counterbalance the delay introduced by rent growth control at a specific location z. As before, we assume no fiscal instruments for the unregulated case. Then, the subsidy rate where $R(\tilde{t}, z) = R(t^*, z)$ is as follows:

$$s = \frac{R_t(\tilde{t}(z), z)}{r^2 C}$$

A subsidy given to the developer at the time of development eliminates withholding because it is equal to the lost growth premium. As an illustrative example, if the cost of construction were 500 000 francs, the real interest rate 4 per cent, and the growth of urban rents 50 francs/year, then the subsidy would have to be 6¼ per cent. If growth were linear, then this subsidy rate would have the desired effect for all locations in the city. If rent growth were accelerating, then the subsidy rate must increase as well. However, when the subsidy rate is not constant over time, the problem becomes more complex. If the subsidy increases over time, then the after-tax physical cost of construction, $(1-s(\tilde{t}))C$, decreases. Although the purpose of the subsidy is to hasten development, a time-varying rate introduces an incentive to delay when growth accelerates. Accounting for this unwanted distortion requires fairly specific knowledge of the growth process of urban rents and solving a nonhomogeneous first-order differential equation.

6 Negotiated rents

If the starting rent of a rental contract were not mandated by statute but instead negotiated between the landlord and prospective tenant, then landlords would rent to the tenant willing to bid up the initial rent until the surplus from the rent control is entirely capitalized into the initial rent. When the initial rent is competitively determined, the net present value of rental income over the leasing period would be the same with the rent control as without. To formalize this insight, suppose that the consumer desires a lease that minimizes the net present value of the stream of rent payments during the period of the lease, from $\tilde{t}$ over an infinite horizon. Since the evolution of rents is governed by regulation, renters and landlords can choose only the base rent. Arbitrage is possible if there are unregulated sectors in the housing market or if the tenant can move to another city. Individuals will choose the contract that minimizes the net present value of rent paid to the landlord. Thus, the initial negotiated rent must satisfy the condition that the present value of the stream of rents under rent regulation must not be larger than the present value of the stream of future market rents:

$$\int_{\tilde{t}}^{\infty} \tilde{R}(\tilde{t}) \cdot e^{-r(s-\tilde{t})} ds \leq \int_{\tilde{t}}^{\infty} R(s) \cdot e^{-r(s-\tilde{t})} ds$$

where $\tilde{R}(\tilde{t})$ is the regulated rent of housing built at time $\tilde{t}$, $R(t)$ is an unregulated rent at time t, and r is the interest rate.[10] If this condition were not met, then tenants would be willing to pay market rents.

The initial rent of the regulated housing would be bid up to the point where consumers are indifferent between rent-controlled housing and its alternative. Integrating the above expression by parts and rearranging yields an expression for the initial rent:

$$\tilde{R}(\tilde{t}) = R(\tilde{t}) + \int_{\tilde{t}}^{\infty} R'(s) \cdot e^{-r(s-\tilde{t})} ds \tag{29}$$

The negotiated initial rent under rent stabilization equals the market rent at the time the contract was made plus a premium reflecting the present value of the future rent growth prevented by the control. Initial rents under regulation overshoot the market rent. Strong evidence of the rent premium that tenants are willing to pay is presented in the empirical work of Nagy (1997) who assesses New York City's rent stabilization programme.

The movement of initial contract rents is given by the following equation:

$$\frac{\partial \tilde{R}(\tilde{t})}{\partial \tilde{t}} = r \cdot \int_{T}^{\infty} R'(s) \cdot e^{-r(s-T)} ds \tag{30}$$

The above expression is positive, which implies that the base rent increases over time. By waiting, the base rent will change for three reasons. First, there is an increase of the market rent by $R'(\tilde{t})$. Second by waiting, the growth restriction premium falls by $R'(\tilde{t})$. These two effects cancel out. The third change, shown in equation (30), is the increase of the discounted value of the rent control premium by developing later.

We can then re-examine the optimal timing decision when rents are regulated as specified by the objective function in equation (7). The first-order condition for the optimal time of development, equation (8), can be re-expressed using the notation of this section:

$$\tilde{R}(\tilde{t}) = A + rC + \tilde{R}_t(\tilde{t})/r$$

Insert the definition of the initial negotiated rent, equation (29), and change in the initial rent over time, equation (30), into the above first-order condition to obtain:

$$R(\tilde{t}) = A + rC$$

This is equivalent to equation (12), the first-order condition for the development of unregulated rents. The hurdle rent when rent growth is regulated but the starting rent is negotiated will be the same as the hurdle rent when rents are unregulated. Thus, the timing of development is not affected by regulation when there is no limit set on the initial rent. The explanation of this result is that the net present value of rents collected by the landlord is not reduced by this form of regulation. Such a rent growth control is neutral to both the welfare of producers and consumers.

7 Conclusion

In this essay I have attempted to describe why rent regulations may lead to seemingly irrational behaviour by landlords. One can set forth the hypothesis that construction will not respond as negatively to an increase in interest rates under rent growth control and may even respond positively. Also, when there is an unexpected increase in economic growth, it is my hypothesis that we will observe a period of withholding vacant land from development. The second purpose of this essay is to present some hypotheses concerning the market for developed land under rent growth control. First, the size of a city will be smaller under rent growth control and may grow more slowly. Second, the price of land is always less under rent growth control. Third, despite positive transport costs, the rent gradient may be flat or upward sloping under rent growth control. Fourth, rent on newly developed land will always be greater under rent growth control than in an unregulated market. Fifth, in 'small' cities, average rents will be larger under rent growth control than in the unregulated case. Finally, it is possible to counterbalance the withholding due to rent regulations with incentive taxes. However, doing so correctly has been shown to be considerably difficult.

This work can be extended in a number of directions. First, one may ask whether considering rent processes governed by different assumptions would change our theoretical results. For example, the incorporation of a stochastic process into the rent function would more closely approximate the fluctuations we observe in the market and give us an idea as to how uncertainty interacts with rent growth control. A second addition could be to explore in greater depth the modelling of fiscal incentives designed to encourage development in Switzerland (see Favarger and McFarlane, 1996). Third, one of the most important issues in the economics of rent regulation is understanding the effect of a regulation upon the renovation and redevelopment of housing. Considering this same policy in a model of sequential development would give greater insight than we have now (see McFarlane, forthcoming). Fourth, it was found that allowing flexibility in the determination of the initial rent will nullify the impact of a rent growth control. It would be worthwhile to explore this type of regulation in more detail (see McFarlane, forthcoming). Finally, a number of testable hypotheses were put forward in this essay. Nonetheless, in actuality, the withholding premium may be a very small part of total construction costs. An empirical examination of these hypotheses would serve the useful purpose of establishing a framework for explaining development behaviour under rent growth control and to compare Swiss research with recent empirical work elsewhere (Capozza and Li, 2001; Mayer and Somerville, 2000).

Notes

1. The reader should see Arnott (1995) for a brief review of some of the possible theoretical empirical approaches to studying rent regulations.
2. The work of von Ungern-Sternberg (1997) points at the costs to renters of indexing rents to the nominal and not the real interest rate.
3. For a description of models where outwards-in development is the outcome of profit maximizing decisions, see Brueckner (2000).
4. Under geometric growth, velocity and acceleration increase over time. For this reason, we must state specifically whether we are comparing the same city at different times or different cities at the same time.
5. For more on rent stabilization and the linear growth case, see McFarlane (1997).
6. When growth is linear there will not be a slowdown of the expansion of the urban–rural boundary with rent stabilization.
7. Under linear growth this would be $\tilde{P}^d(t, z) = A/r + C + g/r^2$.
8. For further references, see McFarlane (forthcoming).
9. The hurdle rent under regulation without the tax on developed land is $R(\tilde{t}, z) = (r/r + \tau^v) \cdot A + rC + (R_t(\tilde{t}, z)/r + \tau^v)$.
10. Since we are comparing regulated and unregulated rents in this section, a '~' is placed over the regulated rent for clarity. In addition, to simplify the exposition, we drop location from the rent function.

References

Anderson, J.E. (1986), 'Property Taxes and the Timing of Urban Land Development', *Regional Science and Urban Economics*, 16, 438–92.

Arnott, R. (1995), 'Time for Revisionism in Rent Control?', *Journal of Economic Perspectives*, 9, 99–120.

Basu, K. and P. Emerson (2000), 'The Economics of Tenancy Rent Control', *Economic Journal*, 110, 939–62.
Brueckner, J. (2000), 'Urban Growth Models with Durable Housing: An Overview', in J.M. Huriot and J.-P. Thisse (eds), *Economies of Cities: Theoretical Perspectives*, Cambridge University Press, 263–89.
Capozza, D. and R. Helsley (1989), 'The Fundamentals of Land Prices and Urban Growth', *Journal of Urban Economics*, 26, 295–306.
Capozza, D. and R. Helsley (1990), 'The Stochastic City', *Journal of Urban Economics*, 28, 187–203.
Capozza, D. and Y. Li (2001), 'Residential Investment and Interest Rates: An Empirical Test of Development as a Real Option', *Real Estate Economics*, 29, 503–19.
European Economic Commission (1991), *Politique des loyers dans les pays de la CEE*, Geneva, United Nations.
Favarger, P. (1996), *Logement, économie de marché et politique publique*, Doctoral dissertation, University of Geneva.
Favarger, P. and A. McFarlane (1996), 'Aménagement du territoire et construction: l'outil fiscal', in C. Jaccoud, M. Schuler and M. Bassand (eds), *Raisons et déraisons de la ville: Approches du champ urbain*, Lausanne, Presses polytechniques et universitaires romandes.
Kutty, N. (1996), 'The Impact of Rent Control on Housing Maintenance: A Dynamic Analysis Incorporating European and North American Regulations', *Journal of Housing Studies*, 11, 69–88.
Lachat, D. and J. Micheli (1992), *Le Nouveau Droit du Bail*, ASLOCA, Lausanne.
Lambelet, J.-C. and C. Zimmermann (1991), *Droit au logement ou économie de marché? Une analyse de l'immobilier Suisse*, Lausanne, Editions Payot.
Mayer, C. and T. Sommerville (2000), 'Residential Construction: Using the Urban Growth Model to Estimate Housing Supply', *Journal of Urban Economics*, 48(1), 85–109.
McFarlane, A. (1997), 'Timing of Land Development in the Presence of Rent Growth Control', *Final Proceedings of the 55th International Conference of the Applied Econometric Association*, University of Neuchâtel, 83–8.
McFarlane, A. (1999), 'Taxes, Fees and Development', *Journal of Urban Economics*, 46, 416–36.
McFarlane, A. (forthcoming), 'Rent Stabilization and the Long-run Supply of Housing', *Regional Science and Urban Economics*.
Nagy, J. (1997), 'Do Vacancy Decontrol Provisions Undo Rent Control?', *Journal of Urban Economics*, 42, 64–78.
Nilles, D. (1996), *Faut-il libérer le loyers?*, unpublished working paper, Department of Economics, University of Lausanne (December).
Office fédéral du logement, *Rapport de la commission d'étude loyer libre*, Rapports du travail sur le logement, Cahier 28, Bern, 1993.
Olsen, E. (1988), 'What do Economists Know about the Effect of Rent Control on Housing Maintenance?', *Journal of Real Estate Economics and Finance*, 1, 295–307.

Raess, P. and T. von Ungern-Sternberg (1999), 'A Model of Regulation in the Housing Market', *Cahiers de recherches économiques* 9903, Département d'économétrie et d'économie politique, University of Lausanne, February.
Spiegel, M.R. (1990), *Theory and Problems of Advanced Calculus*, *Schaum's Outline Series*, McGraw-Hill.
von Ungern-Sternberg, T. (1997), 'Bases économiques pour une loi sur la protection des locataires', *Cahiers de recherches économiques*, 9717, Département d'économétrie et d'économie politique, University of Lausanne, December.

Author and Subject Index