

PLATFORM FOR INTEGRATING ECONOMIC ANALYSIS WITH URBAN FORM ASSESSMENT

Éva-Terézia Vesely

Landcare Research (Manaaki Whenua) New Zealand Ltd
Private Bag 92170
Auckland

Landcare Research Contract Report No. 0506/120

Prepared for:

OPUS International Consultants Limited
PO Box 30845
Lower Hutt

Date: June 2006

Reviewed by:

Approved for release by:

Michael Krausse
Scientist
Landcare Research

Charles Eason
Science Manager
Landcare Research

© Landcare Research New Zealand Ltd 2006

No part of this work covered by copyright may be reproduced or copied in any form or by any means (graphic, electronic or mechanical, including photocopying, recording, taping, information retrieval systems, or otherwise) without the written permission of the publisher.

Disclaimer

The findings in this report are specific to this project. Landcare Research accepts no responsibility where information in the report is used for any other purpose, and will not be liable for any loss or damage suffered as a result of such other use.

Table of Contents

1	SUMMARY	7
2	BACKGROUND	10
3	OBJECTIVES	10
4	METHODOLOGY	10
5	THE PROPOSED PLATFORM	13
5.1	Methodological issues for urban economics	18
6	ECONOMIC DETERMINANTS OF URBAN FORM	21
6.1	Agglomerative forces	22
6.1.1	Sources of agglomerative forces	23
6.1.2	Estimating agglomeration forces	25
6.1.3	Agglomeration at smaller spatial scale	27
6.2	Dispersing Forces	27
7	APPROACHES TO MODELLING THE URBAN FORM	29
7.1	The Bid-Rent Model	29
7.2	Extensions to the Bid-Rent Model	30
7.2.1	Polycentric Structure	30
7.2.2	Housing	30
7.2.3	Transportation	31
7.2.4	Non-central employment	31
7.2.5	Neighbourhood characteristics	32
7.3	New Economic Geography Models	33
7.3.1	Localisation economies, price competition and transport costs	34
7.3.2	Industrial agglomeration under monopolistic competition	35
7.3.3	Provision of public services	37
7.3.4	Communication externalities	37
7.4	Summary	38
8	ECONOMIC CONSEQUENCES OF THE URBAN FORM	40
8.1	Effects on Production	41
8.1.1	City Size	41
8.1.2	Concentration (Primacy)	42
8.1.3	Industry concentration	42
8.2	Effects on Consumption	43
8.3	Summary and the New Zealand Experience	43

		4
9	COMPACT CITY VERSUS SPRAWL	46
9.1	Definition of the Concepts	46
9.2	Costs and Benefits of Sprawl versus Compact Development.....	48
9.2.1	Land-use effects.....	49
9.2.2	Infrastructure, property development and transportation costs	49
9.2.3	Benefits of sprawl.....	51
9.3	Summary	54
10	WHERE TO FROM HERE?	56
11	REFERENCES	57

Figures

Figure 1 Framework for the integration of economic analysis with urban form assessment 15

Tables

Table 1 Extract from ‘Summary of findings about key urban design elements’ (McIndoe et al. 2005). Each finding is characterized in terms of quality of the evidence it offers, using an asterisk system: *** conclusive, ** strong, * suggestive.	40
Table 2 Land saved (in acres) under compact growth scenario, United States, 2000–2025 (Burchell et al. 2005; Centre for Urban Policy Research, Rutgers University)	49
Table 3 Savings under compact growth scenario, United States, 2000–2025 (Burchell et al. 2005; Centre for Urban Policy Research, Rutgers University)	49
Table 4 Testing whether the alleged benefits of sprawl are true benefits to society as a whole (Source: Downs (2002) cited in Burchell et al. (2005))	53
Table 5 Definition of sprawl based on eight distinct dimensions of land use patterns: Sprawl is a condition of land use represented by low values on one or more of these dimensions (Galster et al. 2001).	64

1 SUMMARY

This report is prepared for the *Learning Sustainability* research programme commissioned by the Foundation for Research, Science and Technology (FRST). The research programme asks how we might manage urban form to reduce environmental degradation and optimize 'liveability'. The objectives of the programme's economic component are to assess the plausibility and economic implications of alternative settlement forms. The objectives of this report are to develop a framework for the integration of economic analysis and urban form assessment, to outline the economic determinants and impact of urban form, and to discuss the economic elements of the urban sprawl debate.

Urban form assessment involves measuring a series of descriptors covering different dimensions of urban form, such as density, mixed uses, continuity, clustering, nuclearity, and the accessibility of the street network¹ (Galster et al. 2001; Ewing et al. 2002). The associated economic analysis has two main focuses: the economic determinants and consequences of urban form.

The framework proposed in this report is built on the choices households and firms make about where to locate. While present form was determined by past choices, new choices are slowly changing this. Such locational choices are linked in a circular fashion – people attract jobs and jobs attract people. These choices impact on a range of dimensions of urban form, from size and density to clustering and centrality. The framework is split in two: one side focuses on the economic determinants of the locational choices, while the other side explores the economic consequences of the choices made.

The economic determinants of urban form are explored using the theory that spatial structure is determined by the balancing of centripetal² (also referred to as agglomerative) forces and centrifugal³ (also referred to as dispersing) forces. The agglomerative forces arise from combinations of factors such as interregional differences in resource endowments; scale economies in production, transport technologies, and the provision of public services and

¹ See Appendix I for definitions, units of analysis and operationalisations proposed by Galster et al. (2001).

² Centripetal: proceeding or acting in a direction towards a centre or axis.

³ Centrifugal: proceeding or acting in a direction away from a centre or axis.

facilities; indivisibility of commodities, input-output links, and public or quasi-public goods; speed of communication; and social interaction. Dispersing forces are generated by limited land supply, traffic congestion, crowding, and cultural externalities.

Central and local government intervention modifies the interplay of these forces. The proposed framework highlights the impact of infrastructure investment and controls on housing and property development and their combined influence on locational choice. These impacts include the direct effect of Government investment in built infrastructure, and the indirect effects created by investment in transport infrastructure or by constraining land and property development.

A series of economic models of urban form have been built based on the interaction of the different agglomerative and dispersing forces. In the traditional bid-rent model, individual households optimize their location by trading off accessibility to the urban centre and land rents, which are bid up for locations closer to the centre. Extensions to the bid-rent model explicitly acknowledge the durability of the housing stock, provide a more realistic specification of the transport production function, introduce non-central employment, and build in neighbourhood characteristics. More recently, new economic geography models explain the formation of urban spatial structure as an endogenous⁴ process resulting from interactions among individual economic agents distributed in space. Such interactions arise from a variety of factors, e.g., demand and supply links between customers and firms, knowledge spillover among firms, or congestion effects between residential land uses.

The urban form that results from the locational choices of households and firms has a series of economic consequences. Size, primacy, density and structure impact on production by influencing the availability of skills and specialisations, innovation, production factor prices, and, ultimately, productivity and economic growth. Large cities also impact on consumption by providing private, quasi-public and public goods and services that are not available in smaller cities. They offer different aesthetic characteristics and allow high-speed interaction.

Urban growth by sprawl has different infrastructure, property development and transportation costs, and results in a different mix of land use, housing prices and consumer lifestyles compared with growth in a more compact form. The relative magnitude and distribution of these costs and benefits between individuals, communities and society as a whole are contested and form the basis of the debate on urban sprawl vs compact development.

⁴ Arising from within the generating structure.

The report concludes by recommending a series of actions to progress the economic component of the 'Learning Sustainability' programme:

- Analysis of the New Zealand research space by mapping NZ research efforts around the elements of the proposed framework
- Identification of research gaps by detecting those elements of the proposed framework that have not been researched in New Zealand
- Prioritisation of the identified research gaps based on their relevance to the objectives of the Learning Sustainability project
- Selection of research project(s) to fill the prioritised gap(s)

2 Background

This report is prepared for the *Learning Sustainability* research programme, which asks how we can manage urban form to reduce environmental degradation and optimize 'liveability'. In 2005, while reviewing this FRST programme, Landcare Research and the University of Auckland identified the integration of the economic thinking with environmental, social and liveability assessments as a priority. The acceleration of the economic component of the project was seen to provide a reality check for different urban forms, inform the plausibility of urban futures and help evaluate the economic implications of alternative future settlement form scenarios.

3 Objectives

In order to accelerate the economic component of the 'Learning Sustainability' research programme, the main objective pursued in this report is the development of a platform for the integration of economic analysis and urban form assessments. Urban form assessment is carried out by measuring a series of descriptors covering different dimensions of the urban form such as density, mixed uses, continuity, clustering, nuclearity, and the accessibility of the street network⁵ (Galster et al. 2001; Ewing et al. 2002). The associated economic analysis has two main focuses: the economic determinants of urban forms, and the economic consequences of urban forms. Understanding the organisation of cities from an economic point of view is expected to yield insights into economy-wide growth processes and also shed light on economic concepts of long-standing interest, such as returns to scale, monopolistic competition, vertical integration, technological innovation, innovation diffusion, and international specialisation (Anas et al. 1997).

4 Methodology

To develop a platform for integrating economic analysis with urban form assessment, an international and New Zealand literature review has been initiated. The methodology followed was to look for writings that have specifically urban content and contain specifically

⁵ See Appendix I for definitions, units of analysis and operationalisations proposed by Galster et al. (2001).

economic analysis, which is Mills⁶ (2000) definition of urban economics. The main journal of this field is the *Journal of Urban Economics*, published for the first time in 1974. Due to the blurring of the boundaries between urban economics and other specialties, journals devoted to housing, transportation, local government, human resources, labour relations, poverty, racial problems, and general economics also often contain urban economics articles. And, of course, urban economics research appears in books, monographs and reports. Consequently, the literature relevant to the objectives of this report is extensive and ranges from the early work of von Thünen (1826) and Marshall (1920) to the recent work of new economic geographers such as Krugman, Thisse and Fujita.

Given the wide breadth of the relevant literature, the effort has been directed toward a selection that would inform the development and description of a platform for the integration of economic analysis and urban form assessment. This report is based extensively on articles published in the *Journal of Urban Economics*. The different volumes⁷ of *The Handbook of Regional and Urban Economics* and the *Brookings-Wharton Papers on Urban Affairs* also proved to be very useful sources of information. In addition, a series of books such as Burchell et al.'s (2005) *Sprawl Costs*, Fujita and Thisse's (2002) *Economics of Agglomeration*, and Williams' (2005) *Spatial Planning, Urban Form and Sustainable Transport* helped shape the thinking.

In a New Zealand context, the literature review identified working papers, information resources, reviews and studies that cover topics related to this report. Working papers by Riley (2004) and Arbury (2004), produced previously in the 'Learning Sustainability' programme, have provided the basis for exploring concepts like urban form and compact city. Resource packages on urban sustainability by Hargreaves and Davies (2003) and O'Connell (2003) have urban economics⁸ as one dimension, along with urban forms and transport.

There are a series of reviews that are directly or indirectly relevant. Box (2000) reviews the key concepts of economic geography and explores how such concepts may apply to New Zealand. Crawford (2006) assesses the empirical evidence on mechanisms and magnitudes of

⁶ See Mills (2000) for a thematic history of urban economic analysis.

⁷ Volume II: *Urban Economics*, Volume III: *Applied Urban Economics*, and Volume IV: *Cities and Geography*

⁸ The urban economics dimension has a strong business management focus (e.g., cleaner production information, good environmental practices, and sustainable development reporting practices).

agglomeration on economic activity. McIndoe et al. (2005) look at the range of benefits and costs associated with urban design, while Gow (2000) investigates the urban growth management in the United States and draws lessons for New Zealand.

Finally, there are a series of finished and forthcoming studies with a specific New Zealand focus. Grimes et al. (2004) investigate the bi-directional impacts of economic, social and environmental changes and the New Zealand housing market and identify a series of research questions for future work. Maré (2005) assesses the concentration, specialisation and agglomeration of firms in New Zealand, with future work to examine in greater detail the relationship between firm productivity and patterns of own-industry and between-industry agglomeration.

The insights gained from the reviewed literature have been organised into a platform that integrates economic analysis with urban form assessment. The proposed platform and its elements are described in the following sections:

- section 5 provides a general overview of the platform and sub-section 5.1 discusses some of the methodological issues associated with urban economics;
- section 6 is devoted to the analysis of the economic determinants of the urban form, the two sub-sections covering the agglomerative and dispersing forces;
- section 7 builds on the forces identified in the previous section and gives a selective review of some of the approaches to build economic models to explain and predict urban form;
- section 8 looks at the economic consequences of the urban form;
- section 9 further targets the exposure of the costs and benefits of different urban forms by tapping into the literature on the debate around compact city versus sprawl;
- section 10 outlines proposed future research steps.

5 The Proposed Platform

Urban form is assessed by using a range of descriptors. Despite on-going issues with the sensitivity of urban form to the location of urban boundary and scale, a number of descriptors have been put forward to describe urban form and methodological advances have been made in measuring these. Galster et al. (2001), for example, propose descriptors that cover eight dimensions: density; continuity; concentration; clustering; centrality; nuclearity; mixed uses, and proximity⁹. Individual measures are obtained for each aspect; these are then weighted and combined to form an overall index. In comparison, Ewing et al. (2002) characterise urban form by measuring only four factors: residential density; neighbourhood mix of homes, jobs and services; strengths of activity centres and downtowns; and the accessibility of the street network.

Because they fall short of explaining the underlying processes at work, descriptors can only provide a limited understanding of the urban form (Hall 2001). Hall (2001) argues that urban forms are the visible outcome of a series of complex economic, political, social and cultural processes and the urban landscape must be read within this context. In addition, the city not only reflects the nature of the processes of urbanisation, but is also active in affecting them. As Hall (2001) explains, changes in economics, politics, society and culture are not simply imprinted on the city, but rather are frequently opposed, resisted, misread or encouraged.

While recognising that political, social and cultural processes interact with economic processes to shape cities, this report focuses on economic processes. Understanding the organisation of cities from an economic point of view yields insights about economy-wide growth processes and sheds light on economic concepts of long-standing interest, such as returns to scale, monopolistic competition, vertical integration, technological innovation, innovation diffusion, and international specialisation (Anas et al. 1997).

The platform (see Figure 1) for integrating urban form assessment with economic analysis is built on the choices households and firms make about where to locate. The locational choices impact on the different dimensions of the urban form such as density, nuclearity, continuity, clustering and mix of land use. Based on the values scored for these dimensions, a given urban form can be mapped on a continuum ranging from compact urban form to dispersed

⁹ See Appendix I for definitions, units of analysis and operationalisations.

urban form (sprawl). The existing urban form reflects past locational choices and is shaped by present ones. The two entities – households and firms – influence each other's decisions: people attract jobs and jobs attract people (Mills 2000).

Existing households change location as their life-cycle stages change (Grimes et al. 2004). As a result of internal growth and migration from rural areas, other urban areas and other countries new households are added to a city (Mills 2000). Climate, topography, demographic composition, current industrial and institutional base, general economic conditions and government policies are all potential factors influencing migration choices (Downs 1999, Simmie 2004). New firms become established and the circumstances of the existing ones change, dictating a change in location that is conditional on the mobility level of the industry.

With the locational choices of households and firms in the centre of the proposed framework, one side investigates the economic determinants of the locational choices (see upper part of Figure 1 and Sections 6 and 7), while the other side explores the associated economic consequences (see lower part of Figure 1 and Sections 8 and 9). Here as well there is a circular causation link. The economic consequences of a given urban form become the economic determinants that shape that particular urban form. For example, the size of the city influences productivity, productivity impacts on competitiveness, competitiveness is related to growth, and growth causes changes in size that impact on productivity. Consequently, history matters and any analysis will be strongly influenced by which elements are considered exogenous¹⁰ and which ones endogenous¹¹.

¹⁰ An exogenous variable affects the operation of an economic model but is not itself affected by any of the relationships depicted in the model.

¹¹ An endogenous variable affects the operation of an economic model and is itself affected by the relationships depicted in the model.

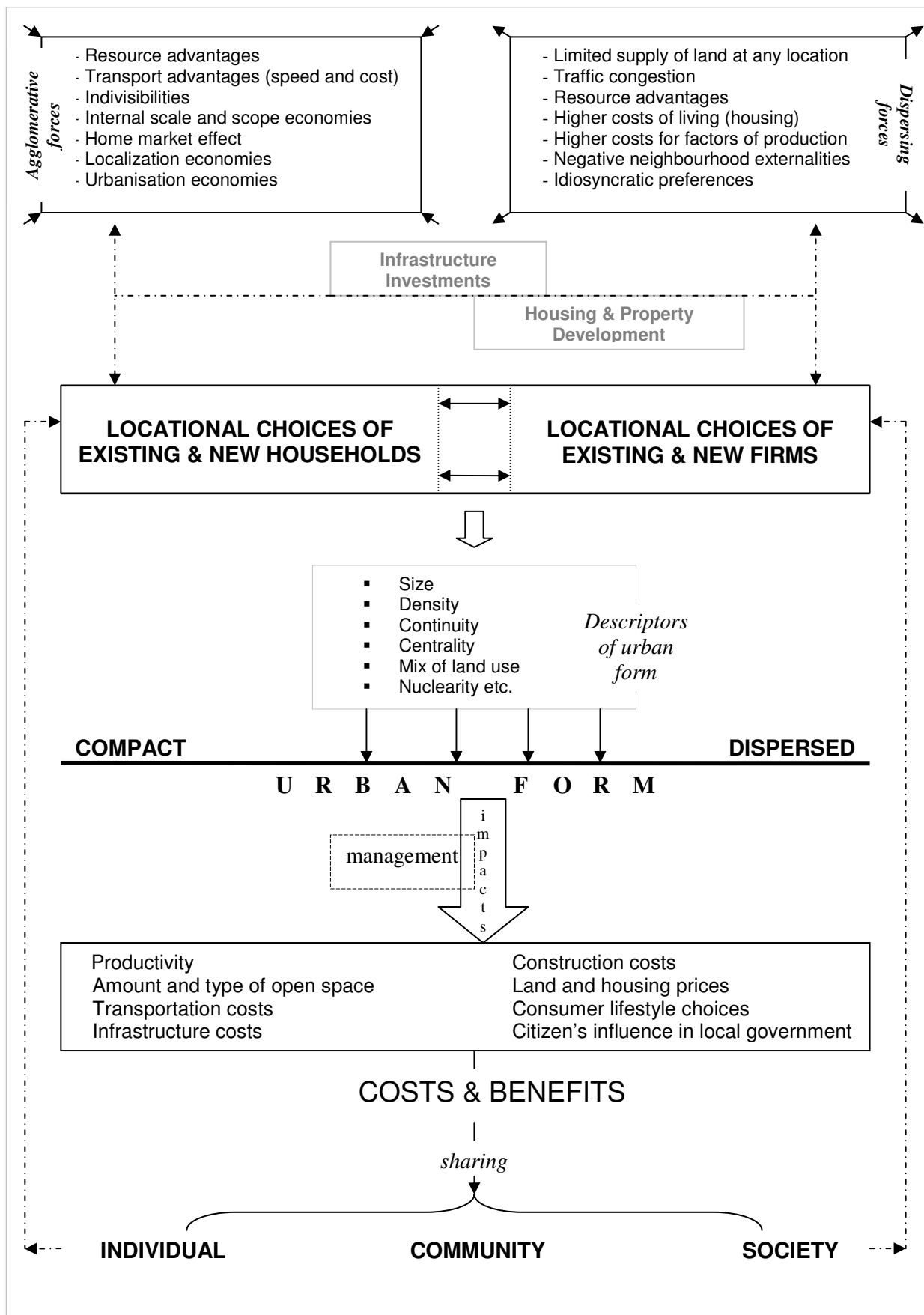


Figure 1 Framework for the integration of economic analysis with urban form assessment

The economic determinants of urban form are explored by using the theory that spatial structure is determined by the balancing of centripetal (or agglomerative) forces (upper left box in Figure 1) and centrifugal (or dispersing) forces (upper right box in Figure 1) (Anas et al. 1997; Verhoef & Nijkamp 2002; Fujita & Thisse 2002; Duranton & Puga 2004).

The agglomerative forces arise from the appropriate combinations of factors such as interregional differences in resource endowments, transport advantages, scale economies in transport technologies and public services and facilities, economies of scale in production, indivisibility of some commodities, input-output linkages, public and quasi-public goods, communication, and social interaction (Fujita 1989). For many agglomerative forces the underlying economic processes are sharing, matching and learning (Duranton & Puga 2004). Agglomeration makes possible the sharing of indivisible facilities, the sharing of gains from the wider variety of input suppliers that can be sustained with larger production, and the sharing of risks. Agglomeration improves either the expected quality of matches or the probability of matching and alleviates hold-up problems. Finally, agglomeration results in the generation, diffusion and accumulation of knowledge (Duranton & Puga 2004).

The most fundamental dispersing force comes from the limitation of geography, land at any location being in limited supply (Anas et al. 1997). Dispersing forces also have their sources in crowding and racial (ethnic) externalities as well as traffic congestion (Fujita 1989). Both the agglomerative and dispersing forces imply a series of externalities opening the way for policy intervention¹² (Hall 2001, Button 2002).

The proposed framework highlights the role of infrastructure investments in shaping the urban form. Transportation investments, for example, are considered to largely determine where growth goes (Frank 2004). As Frank (2004) explains, these investments create or limit access and establish a set of economic tradeoffs that help determine if an area will be compact or dispersed, and mixed or single use. Gow (2000, p. 92), in an attempt to draw lessons for New Zealand from urban growth management in the United States¹³, argues that if he “took anything from the US experience, it is that roads are a key shaper of urban form. [...] Where roads go, people go – assuming roads lead to desirable locations.”

¹² This report does not cover the available policy tools and their implementation.

¹³ Gow (2000) visited and explored in more depth Minneapolis St Paul (Minnesota), Portland (Oregon) and Seattle (Washington).

In a New Zealand context, Coman (1989) studied the role of transport in shaping the Auckland Metropolitan Area. Many factors were recognised as contributing to Auckland's form, such as the surface topography of the area, the isthmus site of the metropolitan area, soil types, historical trends and inertia, town planning consequences, central government policies such as immigration and state housing, building industry trends, and so on. Transport and its innovations, however, are argued as having the greatest impact. The proposed study by the Parliamentary Commissioner for the Environment, 'Transport and Mobility' (by September 2006), will examine the sustainability of the current transport activities focusing on New Zealand's three largest cities.

Along with transportation investments, other infrastructure investments (such as water, storm-water, sewer and energy) also have a strong influence (Burchell et al. 2005). Eberts and McMillen (1999) point out that public infrastructure is an important input into the production process¹⁴ (and in the household utility function) as well as a key factor in facilitating the benefits of agglomeration and in easing the congestion associated with large cities.

Housing and property development, strongly dependent on the infrastructure investments, is another factor of great interest as this sector plays a significant role in shaping the supply that interacts with the demand from new and moving households and firms (Meen & Andrew 2004). Housing and property development is influenced by local government land controls (e.g., limits on business and residential density, prohibition of multi-family dwellings, minimum lot size requirements, height limitations, urban boundaries, etc.) and fiscal policies (e.g., deductibility of mortgage interests and property taxes, and agricultural subsidies that presumably get capitalised in farm land prices) (Brueckner 2001; Mills 2001; Grimes et al. 2004).

The blurring boundaries between urban economics and other specialties such as those analysing housing, transportation and local government, are a reflection of the strong interconnections.

Once locational choices are made, they result in a series of economic consequences. The second part of the proposed framework focuses on the economic implications of different urban forms (see the bottom half of Figure 1). The economic consequences are investigated in two ways. First, the empirical evidence of the impact of urban form on production and

¹⁴ Timmins (2005) evaluates the effect of specific infrastructure projects on firm productivity within New Zealand.

consumption is reviewed. Second, the debate on compact city versus urban sprawl is tapped into. Burchell et al.'s (2005) study that quantifies potential savings from a more compact growth for the United States is reviewed in detail. This study is used to highlight the costs and benefits associated with two scenarios that have different infrastructure, property development and transportation costs, and which result in different land and housing prices and consumer lifestyle choices.

The materialisation of the economic consequences of the urban form is the result of the interplay of many factors and depends to a great extent on local circumstances and management interventions. Furthermore, the economic consequences impact on individuals, communities and the society in different manner. For example, the benefits of sprawl – such as more housing for less cost – tend to accrue to individual households, while sprawl's cost in infrastructure building, energy generation, pollution mitigation, tends to be borne by society overall (Krieger 2004). The way in which the costs and benefits associated with a given urban form are shared between the individual, community and society becomes an important determinant of future locational choices.

5.1 Methodological issues for urban economics

This section provides a short review of some of the methodological issues to be addressed when integrating economic analysis with urban form assessment. Jacobs (1961) stresses that the inherent nature of the subject itself will determine which avenues of thinking are likely to be useful. In her opinion, cities happen to be problems in organised complexity, “situations in which a half-dozen or even several dozen quantities are all varying simultaneously and in subtly interconnected ways” (p. 433). She argues that treating cities as problems of simplicity or of disorganised complexity is mistaken because

City processes in real life are too complex to be routine, too particularised for application as abstractions. They are always made up of interactions among unique combinations of particulars, and there is no substitute for knowing the particulars (p. 441).

To understand cities effectively, Jacobs (1961) identifies the following: (1) think about processes, (2) work inductively, reasoning from particulars to the general rather than the reverse, and (3) seek for ‘unaverage’ clues involving very small quantities that reveal the way larger and more ‘average’ quantities are operating.

Fujita (1989) observes that much of traditional economic theory could not be readily applied to the study of cities. Among the impediments are the strong non-convexity¹⁵ in consumers' preferences and production technologies. In addition, externalities are a common feature since the essence of cities is the presence of many people and firms in close quarters. Many interactions such as information exchange through face-to-face communication are non-price interactions. Oligopolistic or monopolistic competition is a common feature of urban markets, for example, the existence of distance between cities implies that the producers of local public and private goods can enjoy a monopoly. The same is true for producers of neighbourhood services within each city. Finally, buildings and other urban infrastructures are among the most durable of all human products, and this limits the usefulness of classical static theory. Many spatial phenomena such as urban sprawl and renewal can be treated in a satisfactory way only within a dynamic framework.

Omitted variables, imperfect data sets and endogenous changes¹⁶ cause difficulties with the econometric analysis of the processes involved in shaping the urban form (Eberts & McMillen 1999; Mills 2000; Rosenthal & Strange 2004). A major issue is accounting for those unobserved characteristics of cities, industries and individual workers that are systematically related both to economic outcomes and to the features of agglomeration under study. Generalisations are hindered by differences in economic structure, institutional arrangements affecting factor mobility¹⁷, and economic and product cycles¹⁸ (Crawford 2006).

Despite the impressive record of progress in formal econometric work, there is much to be learned from less formal research (Rosenthal & Strange 2004). Case studies are an important part of the entire empirical story, both confirming and placing in context the formal empirical work and identifying important details that the formal work misses. Referring to the link between urban economic development and the state of the local environment, Button (2002)

¹⁵ Non-convexity is explained by the incapacity of the *additivity* and *divisibility* hypotheses on production/consumption to hold. The additivity assumption says that if two production plans are technologically feasible, a new production plan consisting of the sum of these two will also be possible. Divisibility, on the other hand, states that if a production plan is feasible, then any production plan consisting of a reduction in scale will also be feasible.

¹⁶ Endogenous change is one that comes from inside the model and is explained by the model itself.

¹⁷ Factor mobility refers to the degree to which a factor of production (i.e. land, capital, labour) is willing or able to move between different locations or uses.

¹⁸ Product cycles refer to the stages through which a product passes from development to being withdrawn from the market: development, launch, growth, maturity, and decline.

also argues for a case study approach. Given the importance of particular local characteristics of a city, he considers generalisations across urban areas to be often inappropriate and sees the case study approach to provide more insights than the more general statistical examinations of links.

In summary, when developing methodologies for integrating economic analysis with urban form assessment, cities should be portrayed as problems of organised complexity. Statistical examination of links will help with generalisations but is hindered by imperfect data sets, omitted variables, and endogeneity issues. The economic analysis is further complicated by externalities, imperfect competition and non-market interactions. A case study approach is seen to complement the statistical examination of linkages confirming and placing in context the empirical work.

6 Economic Determinants of Urban Form

The economic determinants of urban form are explored by using the theory that spatial structure is determined by the balancing of centripetal agglomerative forces and centrifugal dispersing forces. Fujita and Thisse (2002) see the spatial configuration of economic activities to result from a complicated balance of forces that push and pull consumers and firms. Duranton and Puga (2004) describe the efficient city size as the outcome of a trade-off between urban agglomeration economies (or localised aggregate increasing returns) and urban crowding (or the costs of urban congestion). Verhoef and Nijkamp (2002) also emphasise the governing role of centripetal and centrifugal forces:

Cities exhibit a wide array of attractive and unattractive features, and are hence governed by centripetal and centrifugal forces. Whereas, for instance, a rich cultural life or proximity to a wide variety of goods, services and jobs may make a city an attractive location for living, countervailing negative consequences of an urban way of life may include noise and smell, higher criminality rates, smaller and more expensive housing, and a generally lower environmental quality than many rural locations would offer (p.?).

This dichotomy of forces shaping the urban form has a long history. Thünen ([1826] 1966) was one of the earliest to explain the centripetal forces that stand behind industrial agglomerations. He wrote:

1. Only in large-scale industrial plants is it profitable to install labour-saving machinery and equipment, which economise on manual labour and make for cheaper and more efficient production. 2. The scale of an industrial plant depends on the demand for its products ... 4. For all these reasons, large scale plants are viable only in the capital in the many branches of industry. But the division of labour (and Adam Smith has shown the immense influence this has on the size of the labour product and on economies of production) is closely connected with the scale of an industrial plant. This explains why, quite regardless of economies of machine-production, the labour product per head is far higher in large than in small factories. ... 7. Since it takes machines to produce machines, and these are themselves the product of many different factories and workshops, machinery is produced efficiently only in a place where factories and workshops are close enough together to help each other work in unison, i.e. in large towns (pp. 287–290 of the English translation).

When asking whether industrial firms are better off located in major cities (especially in the capital), Thünen ([1826] 1966) started by describing the main centrifugal forces at work:

1. Raw materials are more expensive than in the country towns on account of the higher cost of transport. 2. Manufactured articles incur the cost of haulage to the provincial towns when they are distributed to rural consumers. 3. All necessities, especially firewood, are much more expensive in the large town. So is rent for flats and houses, for two reasons (1) construction costs are higher because raw materials have to be

brought from a distance and are consequently more expensive, and (2) sites that may be bought for a few thalers in a small town are very dear. Since food, as well as fuel and housing, cost so much more in the large town, the wage expressed in money, must be much higher than in the small one. This adds appreciably to production costs (pp. 186–287 of the English translation).

After this historical perspective, the next two sections explore in more detail the agglomerative and dispersing forces that shape the urban form. The balance between the agglomerating and dispersing forces will differ between locations and across industries. From an efficiency point of view, it is probably important to let the forces work naturally so that an agglomeration is neither stopped before it reaches its prime nor encouraged to grow beyond its natural abilities (Box 2000).

6.1 Agglomerative forces

The existence of cities is explained by a range of agglomerative forces. The agglomerative forces are exogenous or endogenous (Box 2000). Exogenous forces are classed as those things that firms or individuals take as given in a location and that influence their location decision, e.g., climate, harbours, the regulatory framework, and tariffs. Endogenous forces are those things that firms or individuals can actually influence, and in doing so they influence the decisions of other firms and individuals on where to locate, e.g., lower transport costs and economies of scale and scope in dense areas (Box 2000).

An important aspect of the localisation and urbanisation externalities is that the effects of agglomeration extend over at least three different dimensions (Rosenthal & Strange 2004). These are the geographic, industrial, and temporal scope of economic agglomeration economies. Geographic distance is often measured in terms of politically defined discrete areas within which all establishments are modelled as experiencing the same effects. Industrial distance has generally been measured crudely as a binary variable that differentiates between own industry and other industries. For all three dimensions the literature suggests that agglomeration economies attenuate with distance (Rosenthal & Strange 2004).

The analysis of agglomeration traditionally focused on the ability of cities to enhance productivity due externalities and non-price interactions between firms. Recent work on agglomeration has emphasised the consumption possibilities of cities (Waldfogel 2003; George & Waldfogel 2003).

6.1.1 Sources of agglomerative forces

Using the sources of agglomerative forces as a basis for differentiation, and combining insights from Fujita (1989), Eberts and McMillen (1999), Rosenthal and Strange (2004) and Duranton and Puga (2004), the following categorisation can be made:

1. Resource advantages

Natural advantages are also referred to as ‘first nature’ (Krugman 1993) and ‘locational fundamentals’ (Davis & Weinstein 2002). The geographic distribution of production (consumption) is explained by the exogenous geographic distribution of productive resources.

While space is not homogeneous, it is futile to try to justify the marked unevenness of development solely on the basis of space being naturally heterogeneous (Duranton & Puga 2004). For example, Cronon (1991) argues that the land on which Chicago has been built is not all that different from other places on the shore of Lake Michigan that have been sparsely developed.

2. Transport advantages (cost and speed)

The spatial impossibility theorem states that once the heterogeneity of the underlying space is abstracted, without indivisibilities or increasing returns, any competitive equilibrium in the presence of transport costs will feature only fully self-sufficient locations (Starrett 1978). However, due to indivisibilities and increasing returns – discussed in the following paragraphs – the existence of urban areas can be attributed to savings on the cost of moving goods, people and messages due to proximity. In Mills’ (2000) view, this characterisation covers early commercial urban areas, medieval fortress urban areas, medieval cathedral-based centres, and the largest metropolitan complexes in the late twentieth century.

3. Indivisibilities

In the view of Duranton and Puga (2004), perhaps the simplest argument to justify the existence of cities is to invoke the existence of indivisibilities in the provision of certain goods or facilities. They take an ice hockey rink as an example and argue that while having a community of 1000 people to share a rink is feasible, building a rink for each of those people at 1/1000th of the usual scale is not. An ice hockey rink is therefore an indivisible facility that can be shared by many users. Examples of shared facilities that are subject to indivisibilities range from parks, museums, opera houses, rugby stadiums, and schools to airports, train stations and even power plants. Public infrastructure that helps achieve agglomeration

economies and is itself an important input into the production process and household utility function is also subject to indivisibilities (Eberts & McMillen 1999).

4. Internal scale and scope economies

Internal-scale economies exist when expanding production at some given site lowers a firm's unit costs. These can be attributed to the existence of indivisible inputs, division of labour, ability to take advantage of bulk purchases and more efficient use of specialised machinery. Internal-scale economies might lead to the formation of 'company towns': steel in Gary, rubber in Akron, glass in Toledo, chocolate in Hershey, or aircraft manufacture in Seattle (Eberts & McMillen 1999).

Goldstein and Gronberg (1984) argue that agglomerative economies exist when it is less costly to produce a variety of products within one firm than to produce each product in separate firms pointing toward scope economies¹⁹.

5. Home-market effect

Suppose that increasing returns lead to the concentration of employment into a large factory. This in turn, creates a large market, which, in the presence of transportation costs, induces other firms to choose the same location. The idea here is that the interaction between internal-scale economies in production and transport costs leads to a "magnification", where home market size expands in a self-reinforcing process of agglomeration (Rosenthal & Strange 2004).

6. Localisation and urbanisation economies

These refer to the economies from proximate locations or urban sizes that are external to individual firms (Mills 2000). As Anas et al. (1997) explain, at the urban scale, factor mobility is much greater and interactions are more spatially intensive. Firms interact with suppliers (backward linkages), customers (forward linkages), and each other (sideways linkages). Links cause external economies between firms within or across industries. The resulting economies are called economies of localisation in the former case and economies of urbanisation in the latter. The conclusion to the debate over localisation versus urbanisation

¹⁹ Scope economies refer to the long-run reduction in average (or unit) costs that occurs as the scope of the firm's activities increases.

economies is an empirical one and, for the time being, the outcome is not clear (Fujita & Thisse 2002).

Jacobs (1969) stresses the importance of urban diversity as an agglomerative force as diversity fosters cross-fertilisation of ideas. Urban areas function as unified markets that facilitate the idiosyncratic matching of firms and workers, or of firms and customers (Eberts & McMillen 1999). Mills and Hamilton (1994) argue that the most important aspect of agglomerative economies is an application of the law of large numbers. Sales of outputs and purchases of inputs fluctuate in many firms and industries for random, seasonal, cyclical, and secular reasons. To the extent that fluctuations are imperfectly correlated among employers, an urban area with many employers can provide more nearly full employment of its labour force than can an urban area with few employers. Helsley and Strange (1991) consider agglomeration economies arise because the expected second best use of an immobile asset is worth more in a large city than in a small one. This means that resource productivity rises with city size in two ways: used assets are better matched, and risk is reduced. Eberts and McMillen (1999) identify the city's role as an 'urban warehouse' as another source of agglomeration economies that allows firms to carry lower inventories. Fujita et al. (1999) highlight a specific type of agglomeration economies that concerns the increase in the number of product variants on offer.

6.1.2 Estimating agglomeration forces

The most natural way to look for internal scale and agglomeration economies is to directly estimate the production function (Rosenthal & Strange 2004). The usual approach is to include as an argument in the production function one of the following arguments: (1) output of the firm, to represent internal scale economies; (2) output of the industry, to represent localisation economies; and (3) population or output of the urban area, to represent agglomeration economies. The approach taken depends on the type of firm that is being modelled, but for the urban area as a whole there is no reason to focus on just one of the three approaches, as all may exist simultaneously (Rosenthal & Strange 2004).

To estimate the production function it is necessary to have measures of various elements, including employment, land, capital, and materials. Thus a fundamental challenge that must be faced in estimating a production function is to find data on inputs. Omitted variables, measurement errors and endogeneity problems cause difficulties with the production function approach. In addition, wages and other factor prices vary with city size and affect net productivity advantages. This resulted in the use of cost functions rather than production functions in some later studies (Eberts & McMillen 1999).

Estimating the production or cost functions directly is not the only way to look for evidence of agglomeration economies. Rosenthal and Strange (2004) review four indirect approaches that have been used because of the challenges associated with the production function approach. The first indirect approach is to consider growth, e.g., Glaeser et al. (1992) and Henderson et al. (1995). The idea here is that agglomeration economies enhance productivity and productive regions grow more rapidly as a result. The endogeneity problem, however, remains. Existing employers are constrained by prior choices, most importantly the level and kind of capital previously installed. Those fixed factors affect how the employer values the marginal worker, and consequently how s/he changes the employment level in response to a change in the operational environment. Consequently not only is the growth of total employment in a given area sensitive to the composition of employment in an area, but growth affects the level and composition of employment.

A second approach to studying the scope and effect of agglomeration on productivity has been to focus on the births of new establishments and their employment, e.g., Carlton (1983) and Rosenthal and Strange (2003). The idea here is that entrepreneurs seek profit-maximizing locations and are disproportionately drawn to the most productive regions. The principal drawback with this approach is that many locations do not receive any birth in a given period, which can lead to technical challenges on the econometric side. In addition, births are more likely to occur in areas where there is already an existing concentration of industrial activity as spin-offs.

The third approach is to study wages, e.g., Glaeser and Maré (2001) and Wheaton and Lewis (2002). This approach assumes that in competitive markets labour is paid the value of its marginal product. Even without perfect competition, in more productive locations, wages will therefore be higher. As with the production function approach, here too omitted variables and endogenous regressors cause difficulties.

The fourth approach is to use rents. The idea here is that if firms are paying higher rents in a particular location, all else being equal, then the location must have some compensating productivity differential. Dekle and Eaton (1999) use this approach to measure agglomeration economies in Japan. Finding reasonably refined data on rents is problematic.

The empirical evidence for agglomeration economies is reviewed by Rosenthal and Strange (2004). They conclude that while there is evidence of urbanization economies in several industries, there is evidence of localization economies in more, while some industries exhibit no evidence of external economies at all.

6.1.3 Agglomeration at smaller spatial scale

Although agglomeration economies are the *raison d'être* of most cities, their exact nature is in flux and only partially understood (Anas et al. 1997). The sources of agglomeration economies at city or regional level are better known than those at smaller spatial scales. Research into the foundations of the micro-spatial aspects of agglomeration has been identified as a research priority by Duranton and Puga (2004) due to the public interest in urban issues being related to local aspects of land use that affect the immediate environment in which people live and work.

Marshall (1920) identified as micro-foundations of external economies of scale labour market pooling, input sharing and knowledge spillover. Labour market pooling allows a better match between an employer's needs and a worker's skills and reduces risk for both. Input sharing is expected to lead to increasing returns to scale during production. Knowledge spillover takes place when an industry is localised allowing workers to learn from each other. Evidence to date supports the presence of all three of these forces (Rosenthal and Strange 2004).

Duranton and Puga's (2004) propose the sharing, matching and learning processes as micro-foundations of urban agglomeration economies. These apply more generally than Marshall's micro-foundations. Agglomeration makes possible the sharing of indivisible facilities, the sharing of gains from the wider variety of input suppliers that can be sustained with larger production, and the sharing of risks. It improves either the expected quality of matches or the probability of matching and alleviates hold-up problems. Finally, agglomeration results in the generation, diffusion and accumulation of knowledge. Because many of the micro-foundations generate final outcomes that are observationally equivalent in most respects, empirically separating these mechanisms becomes difficult (Rosenthal and Strange 2004).

6.2 Dispersing Forces

Anas et al. (1997) identify as the most fundamental centrifugal tendency the limitation of geography, land at any location being in limited supply. Other centrifugal tendencies recognised by them are created by congestion, by disamenities²⁰ associated with urban activities such as pollution, and by idiosyncratic preferences for different locations.

²⁰ Disamenity is a feature of real property that decreases its attractiveness and decreases the occupant's or user's satisfaction although the feature is not essential to the property's use.

Fujita (1989) identifies three types of negative externalities that act as dispersing forces: crowding and racial (ethnic) externalities, and traffic congestion. Crowding externalities are due to closeness of residences and represent neighbourhood externalities. They refer to an increase in noise, littering, and crimes, and a decrease in open spaces and green areas. Racial (ethnic) externalities are caused by prejudices among different groups. Finally, traffic congestion is due to each car on a highway adding to the congestion and hence increasing travel time for all others. In Fujita's view probably traffic congestion represents the most important type of negative externality²¹ in cities.

Box's (2000) list of dispersion forces includes (1) higher costs of living, particularly housing costs; (2) higher costs for factors of production; (3) pressure on essential infrastructure such as sewerage and roads; (4) pollution; and (5) social problems such as crime.

²¹ Externality: external cost or benefit affecting those not directly involved in production or consumption.

7 APPROACHES TO MODELLING THE URBAN FORM

Building on the agglomerative and dispersive forces described in Section 6, a number of models have been proposed to explain the spatial organisation of cities. Traditional economic models that describe urban spatial patterns of land use can be broadly classified as either microeconomic models that describe equilibrium land use patterns within an urban area or regional economic models that describe the equilibrium flows of population, employment, or other economic factors across regions (Irwin & Geoghegan 2001). More recent urban economic models have focused on explaining the formation of the urban spatial structure as an endogenous process that is the result of ‘interactions’ among individual economic agents distributed in space (Fujita et al. 1999; Krugman 1991; Anas & Kim 1996). This section describes a selection of approaches to build economic models²² of the urban form starting from the traditional bid-rent model and its extensions through to the new economic geography models based on interactions among individual economic agents distributed in space.

7.1 The Bid-Rent Model

Economists understood first why economic activity spreads out, not why it becomes concentrated – and thus the central model of spatial economics became one that deals only with the way competition for land drives economic activities away from a central market (Krugman 1995).

The traditional urban economic model of land-use pattern is the bid-rent model (or monocentric model), which presumes the location of an exogenous central business district to which households commute (Alonso 1964; Mills 1967). All other features of the landscape are ignored, so that distance to the centre is the underlying determinant of land-use change. Individual households optimize their location by trading off accessibility to the urban centre against land rents, which are bid up higher for locations closer to the centre. In its simplest form, the resulting equilibrium pattern of land use is described by concentric rings of residential development around the urban centre and decreasing residential density as distance from the urban centre increases.

Land use in the simple monocentric city is efficient – that is, the equilibrium density pattern is Pareto optimal (Fujita 1989). As Fujita (1989) explains, this is basically because there are no

²² For a mathematical specification of these models refer to the cited references.

externalities; land-use decisions are based entirely on tradeoffs between desire for space and recognition of commuting costs, both of which are purely private. The need for commuting is exogenous in the model, so no agglomerative effects are present. Other considerations ignored are the costs of relocation, imperfect information and uncertainty (Straszheim 1987). When changes occur in income, family size, place of work, preferences, or neighbourhoods, a household's valuations of its current and alternative possible residences are affected. Yet, search for a different residence is costly, moving entails monetary outlays and time and psychological costs, and there are inevitably uncertainties about the future course of other neighbourhoods, the quality of units being contemplated, and the household's future income and housing preferences.

Despite its shortcomings, the monocentric city model provides a useful starting point in explaining the observed spatial structure of cities. The model provides key insights into the two most pervasive facts about urban structure, namely that densities decline, albeit non-monotonically, with distance from the centre, and that most cities have been steadily decentralising for a century or more (Anas et al. 1997).

7.2 Extensions to the Bid-Rent Model

7.2.1 Polycentric Structure

Three functional forms have been suggested as appropriate to generalise monocentric formulations to a polycentric structure (Heikkila et al. 1989). Each is based on a different assumption about how the occupant of a given land parcel interacts with multiple centres. The first assumes that centres are viewed as perfect substitutes; each centre therefore generates its own declining bid-rent function for surrounding land, and land-use density at any point is determined by the highest of these bid-rent functions. What matters at any location is only the centre with the largest influence at that point, and space is divided into strictly separate zones of influence. An alternative assumption is that centres are complements. The occupant of a given location then requires access to every centre in the area. An intermediate case is when every centre has an influence, but a centre's influence becomes negligible at large distances.

7.2.2 Housing

The bid-rent model treats housing implicitly since the household can combine its residential lot with some of its other goods to produce housing. Brueckner (1987) extends the model to make this treatment explicit. In the extended model, land rent, housing rent, household density, and housing density all decline monotonically with distance from the central business district (CBD). A rise in income or a fall in marginal transport cost causes the household and

housing density functions to flatten, whereas a rise in agricultural rent or in population causes them to become steeper.

The monocentric model does not account for the durability of housing. Harrison and Kain (1974) observe that cities tend to grow outwards by adding rings of housing at a density that reflects contemporaneous economic conditions, with the density of earlier rings remaining unchanged due to housing durability. Dynamic versions of the monocentric model with durable housing have been constructed, leading to results that have conflicting implications for the value of density gradients compared with those predicted from the basic model. In spatial models with durable housing, the density gradient depends not only on the past time path of income and transport costs, but also on developers' expectation over time and the prospects for redevelopment. Explanations for observed density gradients are correspondingly complex (Anas et al. 1997).

7.2.3 Transportation

Urban models are considerably complicated when a realistic specification of the production function for transportation is included. Several characteristics of production functions for providing transport service are responsible. One of the necessary inputs in the provision of transportation services is the household's time, incurred by each commuting household. Different technologies have widely varying factor proportions between land, time costs, and other inputs. The choice of the least cost technology depends on the density of trip ends, corridor volumes, and on the value of travellers' time. Another influential characteristic is that urban highway systems are often characterised by congestion. When traffic volumes affect travel time, transport costs become endogenous, jointly determined with location choices and land rents. Congestion costs represent a divergence between average and marginal costs of highway use which distorts location decisions in the absence of appropriate Pigovian taxes or tolls. Similarly, in the presence of automobile pollution externalities, location choices will be non-optimal in the absence of taxes that reflect these externalities (Irwin & Geoghegan 2001).

7.2.4 Non-central employment

The simplest approach to introducing non-central employment into household location models assumes work-site location is exogenous and treats space by distance gradients (Irwin & Geoghegan 2001). If a household is employed at a single non-central work site, with both place of work and earnings exogenous, only a minor modification is necessary to the standard location model, namely the definition of transportation costs with respect to place of work rather than the city centre. With absent neighbourhood effects, the household's bid-rent function must peak at its place of work and decline with distance from the work site in all

directions. If market rents decline with distance from the city centre, households will choose a residential location more distance from the city centre than place of work.

More complex models permit work-site location, wage income, transportation costs, and residential location to be jointly determined (Irwin & Geoghegan 2001). Because bid rent and wage offer gradients are interdependent, complexity arises when decentralised employment opportunities and endogenous work place locations are introduced into a general equilibrium model of urban spatial structure. Wheaton (2004) models commuting, congestion, and employment dispersal in cities with mixed land use. The paper first presents empirical evidence that in US cities actual employment turns out to be almost as dispersed as residences. This range of urban forms is generated with analytic ease in a model that assumes land can have “mixed” rather than exclusive use at any location. The model has firms trading off a central-agglomeration force against the lower wages that accompany shorter commuting distances in peripheral locations. At one extreme, with very high agglomerative forces, employment is approximately centred, with long commute distances and high congestion levels. At the other extreme, lower agglomerative forces lead to employment that is completely dispersed, commute distances of zero, and the absence of congestion.

7.2.5 Neighbourhood characteristics

The interpretation of actual location patterns must consider factors besides those in the simple bid-rent model (Straszheim 1987). The standard monocentric model focusing on land consumption and one-dimensional distance gradients ignores the role of neighbourhood characteristics such as socioeconomic composition, population density, air quality, or public services in location decisions. Superior environmental amenities, public services, and better quality housing are the most frequently cited factors for high income households’ willingness to bid more for distant sites. Whatever the basis for preferring more distant sites, these outweigh the disutility to higher income households of longer commutes.

Hedonic pricing models were the first to illustrate the importance of neighbourhood characteristics. These models try to explain land values by including variables that measure the distance to urban centre(s) as well as specific locational features of the land parcel. They combine the conventional location models with Ricardian models. The Ricardian tradition explains differences in land rents by differences in land qualities that arise from a heterogeneous landscape.

By defining a neighbourhood as a small area centred at any given distance to the city centre, the monocentric model can be readily extended to represent environmental or other neighbourhood characteristics that continuously vary with distance. This approach preserves

all the analytic advantages of defining space by one dimensional distance gradients. However, when there are fixed characteristics of housing units or neighbourhoods and neighbourhood boundaries assume importance at particular locations, discontinuities arise in the choice set and the conventional location models are less useful. Econometric and simulation techniques are used to represent choices when discontinuities exist.

7.3 New Economic Geography Models

More recent urban economic models have focused on explaining the formation of the urban spatial structure as an endogenous process that is the result of 'interactions' among individual economic agents distributed in space (Fujita et al. 1999; Krugman 1991; Anas & Kim 1996). These models, which are part of the new economic geography literature, hypothesize an interdependence among individual households and/or firms that leads to the location decisions of one individual affecting the location decisions of others. Such interdependence can arise due to a variety of factors, e.g. demand and supply linkages between customers and firms, knowledge spillovers among firms, or congestion effects among residential land uses. Depending on the type and magnitude of these interactions, a monocentric, polycentric, or fully dispersed land use pattern may result. Because these models explain the emergence of agglomerations and urban spatial structure, they are much more robust than the traditional bid-rent models. However, in order to solve the model for an equilibrium solution that describes the urban spatial structure, much of the actual heterogeneity of the landscape is ignored. As such, these models offer a fairly abstract description of land-use pattern based on equilibrium conditions.

There are now at least three monographs written by combinations of leading theorists in the field – Fujita et al. (1999), Fujita and Thisse (2002), and Baldwin et al. (2003) – that provide thorough analyses of the theoretical aspects of the new economic geography literature (Head & Meyer 2004). Head and Meyer (2004) identify five essential ingredients that distinguish new economic geography models from other approaches to understanding the geography of economic activity: increasing returns to scale that are internal to the firm; imperfect competition; trade costs; endogenous firm locations; and endogenous location of demand with two mechanisms for the mobility of demand – mobile workers who consume where they work and firms that require the outputs of their sector as intermediate inputs. NEG theories are characterized by magnification, bifurcation, multiple equilibria, and the possibility of catastrophe.

In the new economic geographers' view, the space-economy has to be understood as the outcome of the interplay between agglomeration and dispersion forces within a general

equilibrium framework accounting explicitly for market failures (Ottaviano & Thisse 2004). Because clusters appear at different geographical scales and involve various degrees of sectoral details, it would be futile to look for the model explaining different types of economic agglomerations (Papageorgiou 1983). As Anas et al. (1998) explain, “[i]t may be that the patterns that occur at different distance scales are influenced by different types of agglomeration economies, each based on interaction mechanisms with particular requirements for spatial proximity” (p. 1440).

Consequently, most of the contributions to location theory by industrial organisation deal with partial equilibrium models. Although a comprehensive general equilibrium model with imperfect competition has so far been out of reach, specific models have been developed that, taken together, have significantly improved the understanding of how the spatial economy works (Fujita & Thisse 2002). The new economic geographers’ approach is illustrated in the following sub-sections.

7.3.1 Localisation economies, price competition and transport costs

Fujita and Thisse (2002) examine a microeconomic setting in which firms compete in price while being able to benefit from Marshallian externalities in their location choices. The authors critique the standard general equilibrium analysis for failing to capture the essential impact of transport and land when one comes to study the spatial distribution of economic activities. The reason for failure is seen in abstaining from the consideration of indivisibilities or increasing returns to scale. Their model investigates the interplay between the fall in trade costs, the cost reductions associated with the implicit cooperation arising among firms located in the same area, and the intensity of competition between firms in the domestic and foreign product markets.

Localization economies are an agglomeration force, but geographical proximity renders price competition on the product market fiercer. Whereas firms enjoy low cost when they concur in their locational choices, they can sell their products at higher prices when they are dispersed. Even if price competition is relaxed through product differentiation, it is still true that firms want to be far apart when transport costs are high. Because the spatial distribution of demand is supposed to be unaffected by the locations and sizes of clusters, the cost reduction associated with the agglomeration may be more than offset by the fall in exports. In contrast, firms could enjoy higher profits by being local monopolists. Consequently, transport costs have to be low for firms to congregate. Therefore, the formation of an industrial cluster appears to depend on the relative strengths of three different forces: the magnitude of localisation economies, the intensity of price competition, and the level of transport costs.

When trade costs keep falling, an asymmetric distribution of firms gradually emerges from the interplay between these three forces. Fujita and Thisse (2002) observe that the recent fall in trade costs seems to allow for a great deal of flexibility in where particular activities can locate, but once spatial differences have developed, they tend to become rigid ('putty-clay economic geography'). Hence, regions that were once similar may end up having very different production structures. A small initial advantage may lead to the emergence of a strongly polarised space once the existence of localised production externalities, the existence of natural amenities or both are explicitly accounted for. This effect is magnified when the mobility of factors or the transportability of products are high, or both. Either of the first two possibilities is seen to allow the localised externalities to display their full impact. In addition, although they are active at the local level only, localization economies potentially appear to be a strong agglomeration force in shaping the interregional landscape. This is especially true when transport costs between regions are low, varieties are highly differentiated, or both. In these cases, agglomeration occurs because firms are able to enjoy a higher level of localisation economies while they are still able to sell a substantial fraction of their output on distant markets. More product differentiation fosters more agglomeration because a higher degree of product differentiation allows firms to relax price competition, thus permitting any existing agglomeration force to dominate the dispersion force. Thus, the same causes lead to the same effects, although the geographical scales, as well as the forces at work, are very different.

7.3.2 Industrial agglomeration under monopolistic competition

Krugman (1991a) investigates the effect of the pecuniary externalities associated with market interactions. He observes that manufacture production will tend to concentrate where there is a large market, but the market will be large where manufacture production is concentrated. The market outcome is likely to depend on the initial conditions, so history matters. The process of circular causation, then, looks like a snowball effect that leads industrial firms to be locked in the same region for long periods. As Myrdal (1957) explains:

Within broad limits, the power of attraction of a centre has its origin mainly in the historical accident that something was once started there, and not in a number of other places where it could equally well have been started, and that start met success (pp. 26–27).

In such a context, small initial differences between territories, minor changes in the socio-economic environment, or both, may eventually result in vastly different economic configurations.

The core-periphery model can be used to study the growing geographical concentration of business services in large metropolitan areas (Kolk 1999). In this context the manufacturing sector is to be reinterpreted as the service industry in which firms not only supply consumers and manufacturing firms but also serve each other. This circularity in demand gives the corresponding firms incentives to agglomerate in a fashion similar to that described by Krugman (1991a).

So far agglomeration has been considered as the outcome of a circular causation process fed by the mobility of workers. Yet, in the international marketplace, it is reasonable to expect this mobility to be low, and thus the core-periphery model can not be used to explain the agglomeration of industries in the world economy. Dealing with the intermediate sector, however, allows the explanation of the possible emergence of a core-periphery structure at the international level (Venables 1996). According to Venables (1996), the core-periphery structure may emerge not only as a result of the migration of workers and the imperfectly competitive nature of the final sector but also as a result of the existence of an imperfectly competitive intermediate sector when workers are immobile.

The concentration of economic activity in large metropolitan regions is also explored by Calthorpe (2001). He argues that, because of the increasing globalisation of the economy, metropolitan regions instead of nations have emerged as cohesive economic units that operate as important players in the global economy. Calthorpe (2001) gives two reasons why the global economy operates best at the regional scale. First, despite the advances in telecommunications technology, proximity still matters a great deal. Second, because of the decentralised nature of the economy, networking among a large number of highly specialised people and businesses matters more than ever. He argues that due to volatile and unpredictable economic activity, the single most important component of economic success, either for a business or a worker, is access to networks of all kinds: job networks, money networks, idea networks, and networks of vendors and services. And the only sure way to operate successfully in the networks economy is to be physically located in what might be called a 'network metropolis' – a region where all these networks are located in close enough proximity that they can remain lively and active without a heavy investment in travel or long distance telecommunications. The recognition that it is the network that matters and not any individual business resulted in a 'cluster approach' to economic development. This approach revolves around analysing and understanding business and industrial clusters – geographically based groups of companies, entrepreneurial networks, and labour skills that permit any region to find and keep its place in the global economy year after year.

7.3.3 Provision of public services

Fujita and Thisse (2002), considering the provision of public services as the reason for the existence of cities, observe that most public services are not public goods²³ as they suffer from congestion; furthermore, consuming the public good often involves travelling. If a public good is located in space, there is competition for the limited land close to the public good. Hence, the social cost increases with the number of users, because higher transportation costs are required to use the public good. In the context of local public goods, it is generally not desirable to increase the size of the city population indefinitely even though the per capita cost of the public good is decreasing with the number of users. Indeed, even when the public good is pure, the marginal social cost of a consumer, which is identical to the additional commuting cost, increases. Therefore, there is a trade off between transport costs and the cost of supplying the public good. The model concludes that when the population is such that the residents' common utility level is maximised, the cost of the public good is equal to the aggregate differential land rent within each city.

7.3.4 Communication externalities

Because personal relations are the essence of societies and information is a basic input in firms' activities, Fujita and Thisse (2001) investigate the spatial structure of cities under communication externalities. The information on which their model is built is difficult to codify because it is tacit, and can typically be collected only through face-to-face communications that require travel by high-skilled people whose time is valuable. The authors argue that the transmission of knowledge and ideas is not a routine activity that can be performed through standardised procedures; it is a cognitive process that is made easier when the individuals involved are close to each other. Furthermore, face-to-face communications often stimulate new ideas, combining insights from each party that are crucial for innovations. Although telecommunications maybe a substitute for face-to-face meetings, these two forms of communication may also be complementary. The contrast between the results obtained

²³ A pure public good's consumption is nonrivalrous in the sense that each individual's consumption does not subtract from any other individual's consumption of that good. A pure public good's benefits are also nonexcludable because, once the good is provided, it is virtually impossible to exclude any member of the community from the benefits. Hence, pure public goods can serve unlimited numbers of consumers without having quantity or quality degraded through congestion or increased costs.

under linear and exponential distance-decay functions²⁴ suggests the form taken by the social process of interaction in the transmission of knowledge and information is crucial for the type of urban configuration that may emerge.

7.4 Summary

A range of models have been suggested to explain the urban form. The traditional bid-rent model provides a starting point in explaining the observed spatial structure of cities. In this model, distance to the centre is the underlying determinant of land-use change. All land users benefit from increased accessibility (some more than others, e.g., in general, commercial and industrial users benefit more than households) and thus bid to be at, or close, to the city centre. Because land is more expensive close to the centre, it tends to be used more intensively. There are no externalities such as congestion, air pollution, economies of agglomeration in the model. Land-use decisions are based entirely on tradeoffs between desire for space and recognition of commuting costs, both of which are purely private. Other considerations ignored are the costs of relocation, imperfect information and uncertainty.

The monocentric bid-rent city model is readily modified to accommodate more than one centre, differential patterns of accessibility, non-central employment, durable housing, and a range of neighbourhood characteristics.

The new economic geographers account explicitly for market failures (externalities) and consider the space-economy as the outcome of the interplay between agglomeration and dispersion forces within a general equilibrium framework. Because clusters appear at different geographical scales and involve various degrees of sectoral details, specific partial equilibrium models have been developed based on localisation and urbanisation economies that, taken together, have significantly improved the understanding of how the spatial economy works. A comprehensive general equilibrium model with imperfect competition has so far been out of reach.

The legacy of location theory can be summarised in the following points (Ottaviano & Thisse 2004):

- the economic space is the outcome of a trade-off between various forms of increasing returns and different types of mobility costs;

²⁴ The distance-decay function describes the gradual reduction (decrease) of a response variable with increased distance from the source of the impact.

- price competition, high transport costs and land use foster the dispersion of production and consumption;
- cities provide a wide array of final goods and specialized labour markets that make them attractive to consumers/workers; and
- agglomerations are the outcome of cumulative processes involving both the supply and demand sides.

In the view of Ottaviano and Thisse (2004), future new economic geography studies are expected to consider the welfare implications of agglomeration and, instead of a two-region setting, to account for the rich and complex hierarchy that characterises the space-economy. Future models also need to investigate the geography of employment, labour mobility between sectors, and the impacts of institutions on the spatial organisation of economic activities. Progress is expected from taming increasing returns and imperfect competition within more general models than those used so far.

8 ECONOMIC CONSEQUENCES OF THE URBAN FORM

This section continues the description of the elements of the framework proposed for the integration of economic analysis with urban form assessment. After investigating in the previous two sections, the economic determinants of the urban form and the different approaches to build economic models, attention is now focused on the economic consequences of the urban form.

An initial insight is offered by the McIndoe et al. (2005) study. This study was motivated by interest in what kinds of value urban design offers and how New Zealand's towns and cities might benefit. McIndoe et al. (2005) identify a series of values, the extract in Table 1 provides examples of design elements associated with urban form and their economic consequences.

Table 1 Extract from 'Summary of findings about key urban design elements' (McIndoe et al. 2005). Each finding is characterized in terms of quality of the evidence it offers, using an asterisk system: * conclusive, ** strong, * suggestive.**

Urban Design Elements	Economic Value Findings
Density	Provides land savings***
	Provides infrastructure and energy savings**
	Reduces the economic cost of time allocated to mobility**
	Is associated with concentration of knowledge and innovative activity in urban cores*
Connectivity	Increases viability of local service shops and facilities**
	Increases a site or area's accessibility, thereby enhancing land value**
Mixed Use	Enhances value for those preferring a mixed-use neighbourhood***
	Utilises parking and transport infrastructure more efficiently***
	Increases viability of local service shops and facilities**
	Significantly lowers household expenditure on transportation**

For more insights, the following sub-sections review the empirical evidence on the effects of urban form on production and consumption, and then the literature on compact city versus sprawl is examined to highlight the arguments of both sides.

8.1 Effects on Production

8.1.1 City Size

Although data and conceptual approaches vary greatly, conclusions are remarkably uniform: each doubling of the size of an urban area increases total factor productivity by 5 to 10% (Mills 2000). Empirical work from the United States suggests that doubling employment density can in some cases lead to a 6% increase in average labour productivity (Box 2000).

The benefits associated with city size are only potential; they are contingent on the quality of management (Alonso 1971). City size would therefore define an efficiency frontier, with effective efficiency often significantly below this frontier. The distance between a particular point (that is, a city) and the frontier is a measure of the quality of its management. Mechanisms by which management can influence productivity include efficiency of the transport system (speed) and the relative location of jobs and homes (sprawl), which are the output of transport policies and urban policies respectively (Alonso 1971).

Prud'homme and Lee (1999) test the hypothesis that the efficiency of cities, adjusted for differences in industry-mix, is a function of the 'effective size' of their labour market. The efficiency of cities is defined as labour productivity, while the 'effective size' of the labour market is defined as the average number of jobs available in less than 't' minutes to workers in the city. The effective size of the labour market is further explained by three factors: the size of the city; the average potential job-home distance (sprawl); and the average speed at which journey to work takes place. A sample of 23 French cities is used to test the hypothesis and determine the elasticities of the effective size of the labour market with respect to each of these factors. Results show that when the labour market size increases by 10% productivity increases by slightly less than 2%; when the city size increases by 100 persons, the labour market increases by about 20 jobs and 18 workers within 25 minutes; when the average potential job-home distance increases by 1 km, the size of the labour market at 25 minutes is reduced by about 16 000 jobs, all other things equal; and a 10% increase in average speed, all other things constant, leads to a 15–18% increase in the labour market size.

Large cities are also considered to concentrate skill. Costa and Kahn (2000) show that power couples – where both partners have at least a bachelor's degree – are much more likely to be located in large cities now than 50 years ago, and that roughly a third of this change seems to be because such cities facilitate dual careers. Duranton and Jayet (2005) provide evidence for France that scarce occupations are over-represented in large cities.

8.1.2 Concentration (Primacy)

Henderson (2003) considers the impact of primacy – the percentage of the urban population in the largest city – on productivity growth. In his preferred results, a one standard deviation (.15) increase in primacy from the best level leads to a seven percentage point reduction in productivity growth over five years. For example, for Argentina to be one standard deviation above its best level of primacy (around 0.25) would cost 1.3% growth. Henderson's (2003) interpretation of the findings is that anything that leads to excessive urban concentration can have substantial negative effects on economic performance. In general, optimal primacy reduces as countries become more developed, and is lower for large countries.

8.1.3 Industry concentration

Spatial concentration has been found to impact on the diffusion of knowledge, at least at the early stages of the diffusion process. Controlling for the geographical concentration of sectors affecting the location of patent use, Jaffe et al. (1993) found that in the United States, citation of patents are more likely to be domestic and to come from the same states and metropolitan statistical areas. Feldman and Audretsch (1999) link the innovative output of product categories within a specific city to the extent to which the economic activity of that city is concentrated in that industry, or diversified in terms of complementary industries sharing a common science base. They find that diversity across a common science base is more conducive to innovation than is specialisation.

Fogarty and Garofalo (1988), using total factor productivity and a variable elasticity of substitution production function model, identify and test for the contribution of spatial structure to productivity growth. They also investigate the potential contribution of declining central densities and flattening density gradients to productivity growth in the manufacturing sector. The estimated production function supports the importance of agglomeration economies to productivity. The decomposition of the growth rate in output indicates that declining central densities and flattening of density gradients may have reduced productivity growth over the period 1957–1977. One interpretation is that changes in urban spatial structure have reduced the benefits of agglomeration economies associated with density, thereby producing a loss that subtracts more than 1% per year from the growth rate in real manufacturing output (reducing the overall growth rate by about one-third). Results suggest that the efficiency of cities depends not only on optimal population size but also on the spatial arrangement of economic activity.

Henderson (2003), using panel data, estimates plant level production functions for machinery and high-tech industries that allow for scale externalities from other plants in the same

industry (localization effects) and from the scale or diversity of local economic activity outside the own industry (urbanization effects). He uses the count of other own industry plants as a representation of a count of local information spillover sources. As Henderson (2003) explains, each plant engages in a set of experiments about contemporaneous choices of suppliers, of specific fixed and variable inputs, and of methods for dealing with local regulators; all local plants then benefit from learning the outcome of such experiments, and the spillovers are proportional to the number of experiments, or plants. The count of other own industry plants is found to have strong productivity effects in high-tech but not machinery industries. There is little evidence of economies from the diversity or scale of local economic activity outside the own industry.

Rosenthal and Strange (2004), reviewing the evidence on localisation and urbanisation economies, conclude that while there is evidence of urbanization economies – linked to city size – in several industries, there is evidence of localization economies – linked to own industry size and concentration – in more, while some industries exhibit no evidence of external economies at all.

8.2 Effects on Consumption

The size of cities is argued to influence not only the production but also the consumption side. Glaeser et al. (2001) argue that large cities enhance consumption in four fundamental ways. First, there may be goods and services available in large cities that are not available elsewhere (i.e. opera or restaurants). Second, large cities may offer various aesthetic charms (i.e. the Los Angeles' climate or Paris's architecture). Third, large cities may allow the provision of public and quasi-public goods that would not be possible in a smaller place (i.e. specialized schools). Fourth, the relatively dense settlement of a large city allows speed of interaction that would not be possible in a smaller city (i.e. social interactions).

Waldfogel (2003) and George and Waldfogel (2003) argue that a larger market may allow goods to be more closely tailored to individual consumers' tastes. As population rises, the proportion listening to the radio increases more than proportionately, suggesting provision is more closely tailored to individual tastes (Waldfogel 2003). There are similar effects in internet usage (Sinai & Waldfogel 2001) and newspaper purchases (George & Waldfogel 2003).

8.3 Summary and the New Zealand Experience

Urban form has been found to impact on both production and consumption. The dimensions of the urban form associated with impacts on production are size, primacy, industry

concentration, and the relative location of employers and employees. The main effects manifest in the form of changes in productivity, concentration of skills and diffusion of knowledge. On the consumption side, the city size is the main explanatory variable and this is linked to a larger market that allows for enhanced consumption. Consumption is improved through a closer tailoring of goods and services to individual consumers' taste, higher speed of interaction, the supply of specialised services, and the provision of specific public and quasi-public goods.

In New Zealand, the impact of urban form on production and consumption is not well researched. Maré's (2004) paper is the first to assess the degree of geographic concentration of New Zealand industries based on the Statistics New Zealand business demography micro-data, with plans to examine in greater detail the relationship between firm productivity and patterns of own-industry and between-industry agglomeration in later work. Maré found that, overall, around 30% of the FTE employment is in highly concentrated industry groups. These groups contain mainly Concentrated Manufacturing (6%), Wholesaling (6%) and Business Services (17%). All three groups of industries are disproportionately located in larger cities: the Concentrated Manufacturing, and to a lesser extent Wholesaling industries are most over-represented in South Auckland, whereas Business Services are disproportionately concentrated in Auckland and Wellington. The degree of concentration is similar for the three groups. In concentrated industries the degree of localization is greatest at relatively short distances, between zero and 50 km. In contrast, around 60% of employment is in Local Services (43%) and Local Manufacturing (18%) groups, where industry employment is spread fairly evenly in proportion to total employment. A resource-based group of industries account for a further 7% of employment, and are distributed more evenly across labour market areas than is total employment.

Rates of job creation, job destruction, and net employment growth are higher for industries that are more geographically concentrated, but the relationship disappears when area and industry fixed effects are controlled for. Areas with a more diverse mix of local industries appear to have lower job flow rates and lower employment growth but this relationship is reversed and becomes insignificant once area and industry fixed effects are controlled for. Industry competition and labour market thickness are the only variables to have a significant positive impact on net employment growth, operating through both higher job creation rates, and lower job destruction rates.

Establishing that agglomeration economies exist does not mean that more agglomeration at the margins will necessarily lead to better outcomes in terms of productivity or welfare (Crawford 2006). It is possible to have too much (Crawford 2006) and too little (Mills 2000)

agglomeration. Judgements on policy will require good empirical evidence on the marginal effects on economic and welfare outcomes of agglomeration in the range relevant to the policy focus (Crawford 2006). In addition, different micro-foundations have very different welfare and policy implications (Duranton & Puga 2004). While different assumptions regarding individual behaviour and technology may support similar aggregate outcomes, the normative implications of alternative micro-foundations can differ substantially, and the micro-foundations of urban agglomeration economies interact with other building blocks of urban models in ways that can not be recognised unless they are explicitly stated.

Furthermore, what may be ineffective from an efficiency view may be justified from an equity view (Box 2000). For example, from an efficiency perspective a drift northwards in New Zealand is reasonable. It is likely to increase productivity and wages in Auckland and create a larger denser area in which knowledge and innovation are more likely to flourish. From an equity perspective, however, the drift may be seen as undesirable as other regions may face decline (Box 2000).

9 COMPACT CITY VERSUS SPRAWL

Further insights on the costs and benefits of different urban forms can be gained by tapping into the literature on the debate about compact city versus sprawl. This literature is growing (Krieger 2003) and is divided between those critiquing sprawl and proposing ‘smart growth’ and ‘new urbanism’ (Duany et al. 2000; Calthorpe & Fulton 2001) and those critiquing the critics’ of sprawl (Gordon & Richardson 1998, 2000).

9.1 Definition of the Concepts

First, what constitutes compact city and sprawl? Generally, three aspects of the compact city are identified: a high-density city, a mixed-use city and an intensified city (Burton 2002). There is general agreement that the ‘compact city model’ is based on an increase in density from current levels. Getting the right mix of densities so that the infrastructure can handle the increased pressure but without having a negative impact on the existing urban environment is an extremely difficult task. It has further been argued that the method of urban intensification is often of more importance for both acceptability and sustainability, than the absolute densities. Thus the compact city appears to be a highly complex concept not only related to an increase in density but also to a variety of densities across the urban landscape in order to achieve ‘its benefits’. To further add to this complexity, a mix of urban uses has also been identified as an important aspect of the compact city.

Galster et al.’s (2001), in an attempt to ‘wrestle to the ground’ the other elusive concept, defined sprawl as a condition of land use represented by low values on one or more of eight distinct dimensions of land use patterns: density, continuity, concentration, clustering, centrality, nuclearity, mixed uses, and proximity. In the literature, however, the term sprawl is a metaphor rich in ambiguity. As Galster et al. (2001) explain, sprawl has been attached to patterns of residential and non-residential land use, to the process of extending the reach of urbanised areas, to the causes of particular practices of land use, and to the consequences of those practices.

Burchell (1998), for example, defines sprawl by looking inductively at all the criticism of sprawl in the literature and deriving traits that would cause them. Sprawl then is a particular form of metropolitan growth and development dominated by the following traits: unlimited outward extension of development; low-density residential and commercial settlements; leapfrog development; fragmentation of power over land use among many small localities; dominance of transportation by private automotive vehicles; lack of centralised planning or

control of land uses; widespread strip commercial development; great fiscal disparities among localities; segregation of types of land use in different zones; and reliance mainly on the trickle-down or filtering process to provide housing to low income households.

Two examples of sprawl are given by Nivola (1998): the spatial size of the Chicago metropolitan area grew by 46% between 1970 and 1990, while the area's population grew by only 4%; in the Cleveland metropolitan area, spatial growth of 33% occurred over this period even though population declined by 8%.

Downs (1999) considers sprawl to cause, or contribute to, two sets of economic and social problems:

The first occurs mainly in fast-growing areas, but it spreads to others as well. It includes traffic congestion, air pollution, large-scale absorption of open space, extensive use of energy for movement, inability to provide adequate infrastructure, inability to locate region serving facilities that produce negative local impacts (such as airports), shortages of affordable housing near where jobs are being created, and suburban labour shortages. These problems mainly harm people who benefit from other aspects of sprawl.

The second set of problems occurs mainly in big cities, inner-ring suburbs, and a few outer-ring suburbs. These problems arise because our development process, not sprawl per se, concentrates poor households – especially poor minority households – in certain high-poverty neighbourhoods that become sites for high crime rates, poor quality public schools, dysfunctional big-city bureaucracies, and lack of fiscal resources. These poverty-related problems soon spread to inner-ring suburbs. Also, many outer-ring suburbs with low commercial tax bases but much moderate-cost housing do not have enough taxable resources to pay for decent schools and other services (pp. 956–957).

Brueckner's (2001) review of criticism against sprawl identifies similar problems:

Critics of sprawl argue that urban expansion encroaches excessively on agricultural land, leading to a loss of amenity benefits from open space as well as the depletion of scarce farmland resources. The critics also argue that the long commutes generated by urban expansion create excessive traffic congestion and air pollution. In addition, growth at the urban fringe is thought to depress the incentive for redevelopment of land closer to the city centres, leading to decay of downtown areas. Finally, some commentators claim that, by spreading people out, low-density suburban development reduces social interaction, weakening the bonds that underpin a healthy society (p. 65).

The validity of some of these criticisms has been questioned. Mills (2001) argues that there is no intrinsic relationship between sub-urbanisation and traffic congestion. He points to a scenario in which business and population suburbanisation occurred that reduced commuting distances dramatically compared with a monocentric business location pattern. Gordon and Richardson (1998, 2000) refer to sub-urbanisation as 'the traffic safety valve' arguing that increasingly footloose industry has followed workers into the suburbs and exurban areas. The invasion of farmland is dismissed as an issue because excess not deficient agricultural output

was the characteristic scenario for the US for most of the 20th Century. Regarding open space, it is argued that federal, state and local governments can buy as much land for parks, forests, and so on as their constituents are willing to pay for. On the other hand, these authors (Gordon & Richardson 1998, 2000; Mills 2001) recognise a series of market and government failures – the most important being the under-pricing of transportation – that might lead to excessive urban expansion.

The actual costs and benefits of sprawl have remained elusive (Burchell et al. 2005).

9.2 Costs and Benefits of Sprawl versus Compact Development

Burchell et al. (2005) provide credible evidence in a United States context of the costs of sprawl as well as the potential savings as a function of more compact growth. While reviewing their study the point is taken from Downs (1999) who emphasises that “it is the basic traits of our growth and development process themselves that produce our most serious urban problems, not sprawl[; e]ven compact growth would produce the same problems” (p. 961). Consequently, the attention is focused not on the costs associated with growth per se but on the potential savings from shifting from one urban growth form to another.

Burchell et al. (2005) base their book, *Sprawl Costs – Economic Impacts of Unchecked Development*, on the findings of *The Costs of Sprawl – 2000* study by the National Research Council, published in 2002. This research document meant to update and improve on the seminal *Costs of Sprawl* study from 1974. *The Costs of Sprawl – 2000* set out two different futures for the United States with regard to the accommodation of growth²⁵ – one of low-density sprawl development at the outer reaches of the metropolitan area, the other of more compact, smarter development – and compared their differing impacts on land, infrastructure, housing costs and public services. They quantified the potential costs of sprawl over the next 25 years while also acknowledging its benefits.

Both development alternatives involve growth that produces 53 million residential and non-residential development units nationwide over the period 2000 to 2025. Under the compact growth scenario, 9.2% of the development units are shifted to the more urban and suburban locations. An additional 2 million development units are relocated within their county to more developed areas. These development units are projected to be built at 20% higher density. In the undeveloped areas about 20% of the residential units are projected to be developed in

²⁵ The study does not account for changes to land use patterns that affect current residents.

cluster developments where density is twice as high as the prevailing density of undeveloped areas. In addition, 25% more housing units are developed as town-homes or multi-family buildings rather than as single-family detached or mobile home units. The sprawl costs are assigned in the following manner: infrastructure and transportation costs to the community; housing and quality of life costs to the individual; and energy and environmental costs, both to the community and to society as a whole.

9.2.1 Land-use effects

Under the compact growth scenario an estimated 4 million acres of land (see Table 2) is saved from development. Urban growth boundaries that redirect growth to more developed counties account for 60% of the savings, while urban service areas that focus growth within a county account for 40% of the savings. While this acreage savings is significant, it still is only 0.3% of all non-federal rural land. Further, the agricultural land savings (1.5 million acres) is only 0.5% of the cropland and pastures in use.

Table 2 Land saved (in acres) under compact growth scenario, United States, 2000–2025 (Burchell et al. 2005; Centre for Urban Policy Research, Rutgers University)

Agricultural Lands	Environmentally Fragile Lands	Other Lands	Total Land
1 499 636	1 505 434	997 156	4 002 231

9.2.2 Infrastructure, property development and transportation costs

Table 3 summarises the savings under the compact growth scenario associated with infrastructure, property development and transportation.

Table 3 Savings under compact growth scenario, United States, 2000–2025 (Burchell et al. 2005; Centre for Urban Policy Research, Rutgers University)

	Savings (US\$(2000))	Percent Savings
Water and sewer infrastructure costs	12,606,000,000	6.6
Road construction costs	109,700,000,000	11.8
Aggregate property development costs	420,300,000,000	6.6

Transportation costs	24,071,000	2.4
----------------------	------------	-----

The model considers water and sewer costs to arise from three factors: the demand for water and sewer services, the number of hook ups to water mains, and the type of water system used. All these factors are influenced by the way a community develops and the types of houses that are built.

Future road demand is projected based on the historical relationship between population density and the density of the road network. The model takes the projected population density in 2025 and determines an ideal road density for each county. The need for new road construction is then predicted by comparing the ideal level of required lane-miles with the existing lane-miles found in a county. The cost to construct these required lane-miles is calculated based on per-mile concrete roadway construction costs for urban and rural development environments. The costing takes into account the applicable development standards for roads but excludes both required structures and land acquisition. The cost of repairing and maintaining an ever-expanding road network is not dealt with by the model.

To determine the effects of sprawl and compact growth development futures on housing costs, the model calculates the prices of various types of new housing for each of the counties. Although aggregate property development costs are lower under a compact development scenario, it is highlighted that the long-term benefits of this type of development are obscured by several immediate considerations for developers and the banks that finance them. Such considerations refer to the compact growth being unfamiliar, mixed-use elements being risky, and the ‘stitching’ together of urban parcels, some of which may have legal encumbrances, as problematic.

When calculating the transportation costs of the two scenarios, the premise is that people who live in more compact areas are more likely to drive shorter distances and to take public transport instead. The model distinguishes between personal vehicle and public transport use only. While the model is not able to account for potential increases in walking trips, it is noted that walking is the lowest cost form of personal transport. Transportation costs include the fixed and variable costs of driving and taking public transport, as well as the monetary value of time spent travelling.

The cost of driving includes the costs incurred by vehicle owners, governments and society. Travel costs per passenger mile for public transport are calculated based on two basic types of

public transport: bus and train travel (bus includes motor bus and trolleybus; rail includes commuter rail, heavy rail, and light rail) and include operating costs, capital costs and social costs.

9.2.3 Benefits of sprawl

Considering both the costs and benefits of sprawl is important because, as Gordon and Richardson (1998) point out, most people are not cost minimisers but rather trade off costs for perceived benefits. Burchell et al. (2005), based on Downs (2002), evaluate thirteen benefits of sprawl by subjecting them to the following tests:

- Is this condition perceived by a large number of people to be a benefit?
- Is this condition actually caused by sprawl or by certain traits that are part of the basic definition of sprawl?
- Is this condition widespread enough in the United States to be socially significant?
- Does this condition have serious negative side-effects or consequences that offset its benefits when considering its influence on society as a whole?
- Is this condition perceived by a large number of people to be a disadvantage?
- In sum, is this condition unequivocally a net benefit to society as a whole?

The results (see Table 3) show that five of the thirteen benefits of sprawl are ‘true’ net benefits to society as a whole. These five benefits (with bold in Table 4) are: lower land and housing costs on farther-out sites; larger average lot sizes on such sites; meeting widespread consumer preferences for low-density living; providing households with wider choices of combinations of tax levels and social services than would occur under non-sprawl development; and permitting greater individual citizen participation and influence over local government affairs than would occur under fewer and larger governments²⁶.

²⁶ Although the benefit of stronger participation in local government as a result of smaller government units is considered a social benefit of sprawl, the down side of such fragmentation is highlighted in the form of difficulties associated with influencing decisions that have regional impact.

Due to the lack of readily available data on variations in housing unit size by distance from the centre of each metropolitan region, the benefit associated with potentially larger home and room sizes under the sprawl scenario could not be estimated.

The benefits associated with the superiority of schools and low crime rate in relatively remote, low-density neighbourhoods are contested because these are caused not by the distant locations or lower density of those neighbourhoods but by the particular groups of people who live there. The perceived benefit of more homogeneous communities is also criticised because it is based on very exclusionary behaviour.

The argument that sprawl provides lower traffic congestion, on average, than more compact and therefore higher density settlement patterns, is taken as valid. However, due to limited data, the precise quantification of this benefit is not considered possible. The claim that commuting times are shorter for residents of sprawl than for people living in more dense areas is examined based on a Los Angeles case study. This suggests that where jobs are highly scattered regionally, average commuting times are not closely related to residential distance from the regional centre, and residents of highly sprawled areas are at a slight average commuting time disadvantage compared with those in denser areas. The argument that communities with low densities have lower total transportation costs than higher density ones because public transport costs per mile are higher than private vehicle costs per mile is dismissed. The counter-argument is that residents in low density areas typically travel more miles each day than those living in higher density areas. However, it is recognised that travel cost savings from adopting the compact growth scenario, compared with the sprawl growth scenario are extremely small compared with the total national costs of travel.

A final alleged benefit of sprawl is that leapfrog development makes the use of infill more efficient. It is argued that as a large metropolitan area expands outward and sites located at its edges become more central, the sites that have been left vacant can then be developed at much higher densities without incurring the substantial cost of redevelopment. How important this sprawl benefit will be depends on the percentage of once-peripheral land in a region that is initially leapt over by developers, the rate at which the region expands outward and thereby raises the densities at which it is most appropriate to develop skipped-over vacant sites, the costs of demolition, and the connection between skipped-over sites and the region's major transportation arteries. These factors all vary enormously from one region to another and data shortage prevents the estimation of the associated savings. Leapfrog development also creates a potential for inner-city/suburban open space as a result of skipped-over lands (not covered in the summary table). In practice, however, Burchell et al. (2005) find that the new open space rarely materialises in either developed or developing neighbourhoods.

Table 4 Testing whether the alleged benefits of sprawl are true benefits to society as a whole (Source: Downs (2002) cited in Burchell et al. (2005))

Benefits of sprawl	Perceived as a benefit by many people	Actually caused by sprawl	Appears wide-spread in regions of the US	Has serious negative side effects	Perceived as a disadvantage by many people	Unequivocally a net benefit to society as a whole
Lower land and housing cost	Yes	Yes	Probably	No	Partly	Probably
Larger average lot size²⁷	Yes	Yes	Yes	No	No	Yes
Larger home and room sizes	Yes	Not clear	Not clear	No	No	No, because actual extent of occurrence is not clear
Reflects low density preferences	Yes	Yes	Yes	No	Unclear; some say not enough other choices are available	Yes
Neighbourhoods with lower crime rates	Yes	Partly	Yes	Yes, partly caused by exclusionary behaviour	Yes	No, because partly caused by exclusionary behaviour
Better quality public schools	Yes	Partly	Yes	Yes, partly caused by exclusionary behaviour	Yes	No, because partly caused by exclusionary behaviour
Greater consumer lifestyle choices	Yes	Yes	Yes	Yes; helps perpetuate exclusionary behaviour	Yes	Yes
More homogeneous communities	Yes	Partly	Yes	Yes; based directly on very exclusionary behaviour	Yes	No, because based directly on very exclusionary behaviour
Stronger citizen participation and influence in local governments	Yes	Yes	Yes	Yes; helps perpetuate exclusionary behaviour	No	Probably

²⁷ These potentially can provide open space in form of private yards.

Less-intensive traffic congestion	Only by a few people	Not clear	Not clear	No	Yes	No, because actual extent of occurrence is not clear
Lower overall transport costs	No	No	No	No	Yes	No
More efficient use of infill sites	Only by a few people	Yes	Not clear	No	Yes	No, because actual extent of occurrence is not clear

Burchell et al. (2005) argue that a strong economic and moral argument for changing sprawl is that many households moving into distant subdivisions are not being required to pay the full social cost of gaining the benefits of living there, including infrastructure trunk lines and increased social costs in inner city areas. The most persuasive part of this cost-based argument is that merely continuing current sprawl densities for future growth will require massive additional spending on infrastructure that is not consistent with adequately maintaining existing roads, sewer and water systems, and other infrastructure. Whether charging households more fully for the costs they generate will encourage more of them to locate in closer in areas depends on their trade-offs of costs against benefits.

9.3 Summary

The stakes in the sprawl debate are substantial, as the following reasoning from Brueckner (2001) demonstrates:

Ultimately, an attack on urban sprawl would lead to denser cities containing smaller dwellings. The reason is that antisprawl policies would limit the supply of land for residential development, so that the price of housing, measured on a per-square-foot basis, would rise. In response to this price escalation, consumers would reduce their consumption of housing space, making new homes smaller that they would have been otherwise. [...] If the criticisms of sprawl are correct, then public policies should be altered to restrict the spatial expansion of cities. The resulting losses from lower housing consumption would be offset by gains such as improved access to open space and lower traffic congestion, and consumers on balance would be better off. But if the attack on sprawl is misguided, with few benefits arising from restricted city sizes, consumers would be worse off in the end. People would be packed into denser cities for no good reason, leading to a reduction in the American standard of living. The same conclusion would arise if some limitation of city sizes is desirable, but policy makers are overzealous. If only mild measures are needed to restrict urban growth that is slightly excessive, but draconian measures are used instead, consumers are likely to end up worse off (pp. 66–67).

Looking into the debate on sprawl and the compact city, the following general observations can be made:

- There is a shared recognition of recent increases in the spatial extent of cities (for examples of increases in the spatial extent of cities see Nivola (1998)).
- Advances have been made in defining both concepts (Burchell 1998; Brueckner 2001; Galster et al. 2001; Burton 2002).
- There is a relatively consistent line-up of issues with sprawl – anti-sprawl arguments include excessively costly municipal infrastructures, inefficient transportation patterns, unnecessary environmental damage, alienation and communal breakdown, and suboptimal economic development (Downs 1999; Brueckner 2001).
- The anti-sprawl arguments are contested (Mills 2001; Gordon & Richardson 1998, 2000).
- There is a call to separate the effects of growth per se from the effects of the spatial pattern of growth (Downs 1999; Burchell et al. 2005).
- There is a call to recognise both the costs and the benefits of alternative spatial patterns (Burchell et al. 2005).
- There is recognition of the significance of how the costs and benefits of alternative spatial growth patterns are shared between the individual, community and society (Krieger 2004; Burchell et al. 2005).

10 WHERE TO FROM HERE?

Integrating economic thinking with environmental, social and liveability assessments of urban form is a priority for the ‘Learning Sustainability’ programme. In an effort to accelerate the economic component of the project, this report has developed a framework for the integration of urban form assessment with economic analysis. This framework is built on the choices households and firms make about where to locate, and has two parts: one investigates the economic determinants of these choices, while the other explores the associated economic consequences.

To further develop the economic component of the ‘Learning Sustainability’ programme the following steps are recommended:

- Analysis of the New Zealand research space by mapping NZ research efforts around the elements of the proposed framework
- Identification of research gaps by detecting those elements of the proposed framework that have not been researched in New Zealand
- Prioritisation of the identified research gaps based on their relevance to the objectives of the Learning Sustainability project
- Selection of research project(s) to fill the prioritised gap(s)

11 REFERENCES

- Alonso, W 1964. Location and land use. Harvard University Press, Cambridge MA.
- Anas A, Arnott R, Small KA 1997. Urban spatial structure. Working Paper UCTC No. 357, The University of California Transportation Centre, University of California at Berkeley.
- Anas A, Arnott R, Small KA 1998. Urban spatial structure. *Journal of Economic Literature* 36: 1426–1464.
- Arbury J 2004. What contribution has the ‘Compact City’ made to the search for urban sustainability? Working Paper, School of Geography and Environmental Science, The University of Auckland, Auckland, NZ.
- Baldwin R, Forslid R, Martin P, Ottaviano G, Robert-Nicoud F 2003. The core-periphery model: key features and effects. In: *Public Policies and Economic Geography*, Chapter 1, Princeton University Press, Princeton.
- Box S 2000. Economic geography: key concepts. Treasury Working Paper 00/12, Treasury, Wellington, NZ.
- Brueckner JK 2001. Urban sprawl: lessons from urban economics. In: Gale WG, Rothenberg Pack J eds *Brookings-Wharton papers on urban affairs*. Brookings Institution Press, Washington DC. Pp. 65–89.
- Burchell RW 1998. The costs of sprawl: revisited. Transportation Cooperative Research Program Report 39. National Academy Press, Washington DC.
- Burchell RW, Downs A, McCann B, Mukherji S 2005. *Sprawl costs: economic impacts of unchecked development*. Island Press, Washington DC.
- Burton E 2002. Measuring urban compactness in UK towns and cities. *Environment and Planning B: Planning and Design* 30: 503–522.
- Button K 2002. City management and urban environmental indicators. *Ecological Economics* 40: 217–233.
- Calthorpe P 2001. *Regional city: new urbanism and the end of sprawl*. Island Press, Covelo, CA, USA.
- Carlton DW 1983. The location and employment choices of new firms: an econometric model with discrete and continuous endogenous variables. *Review of Economics and Statistics* 65: 440–449.

Coman PJ 1989. Transport and the urban form: a case study of the Auckland Metropolitan Region. Unpublished Bachelor of Property Administration research project, School of Architecture, The University of Auckland, Auckland, NZ.

Costa DL, Kahn ME 2000. Power couples: changes in the locational choice of the college educated. *Quarterly Journal of Economics* 115(4): 1287–1315.

Crawford R. January 2006. The effects of agglomeration on economic activity: the empirical evidence on mechanisms and magnitudes. Ministry of Economic Development, Wellington, NZ.

Cronon W 1991. *Nature's metropolis: Chicago and the Great West*. Norton, New York.

Dekle R, Eaton J 1999. Agglomeration and land rent: evidence from the prefectures. *Journal of Urban Economics* 46: 200–214.

Downs A 1999. Some realities about sprawl and urban decline. *Housing Policy Debate* 10(4): 955–89.

Downs A 2002. Have housing prices risen faster in Portland than elsewhere? *Housing Policy Debate* 13(1): 7–31.

Duany A, Plater-Zyberk E, Speck J 2000. *Suburban nation: The rise of sprawl and the decline of the American dream*. North Point Press, New York.

Duranton G, Jayet H 2005. Is the division of labour limited by the extent of the market? Evidence from French cities. University of Toronto, URL: <http://individual.utoronto.ca/gilles/research.html>

Duranton G, Puga D 2004. Micro-foundations of urban agglomeration economies. In: Henderson JV, Thisse JF eds *Handbook of regional and urban economics, Volume 4 (Cities and geography)*. Elsevier B. V., Amsterdam, Holland. Pp. 2063–2117.

Duranton G, Puga D 2005. From sectoral to functional urban specialisation. *Journal of Urban Economics* 57: 343–370.

Eberts RW, McMillen DP 1999. Agglomeration economies and urban public infrastructure. In: Milles ES, Cheshire P. eds *Handbook of regional and urban economics, Volume 3 (Applied urban economics)*. Elsevier Science B.V., Amsterdam, Holland.

Feldman MP, Audretsch DB 1999. Innovation in cities: science-based diversity, specialisation and localised competition. *European Economic Review* 43(2): 409–429.

Fogarty MS, Garofalo GA 1988. Urban spatial structure and productivity growth in the manufacturing sector of cities. *Journal of Urban Economics* 23: 60–70.

Frank LD 2004. Economic determinants of urban form: resulting trade-offs between active and sedentary forms of travel. *American Journal of Preventive Medicine* 27(3S): 146–153.

Fujita M 1989. *Urban economic theory: land use and city size*. Cambridge University Press, Cambridge.

Fujita M, Krugman P, Venables A 1999. *The spatial economy: cities, regions, and international trade*. MIT Press, Cambridge.

Fujita M, Thisse J-F, 1991. Spatial duopoly and residential structure. *Journal of Urban Economics* 30: 27–47.

Fujita M, Thisse J-F 2001. Agglomeration and growth with migration and knowledge externalities. Discussion Paper No. 531, Institute of Economic Research, Kyoto University, Kyoto.

Fujita M, Thisse J-F 2002. *Economics of agglomeration: cities, industrial location and regional growth*. Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

Galster G, Hanson R, Ratcliffe MR, Wolman H, Coleman S, Freihage J 2001. Wrestling sprawl to the ground: defining and measuring an elusive concept. *Housing Policy Debate* 12(4): 681–717.

George L, Waldfogel J 2003. Who benefits whom in daily newspaper markets? *Journal of Political Economy* 111(4): 765–785.

Glaeser E, Kahn ME 2004. Sprawl and urban growth. In: Henderson JV, Thisse JF eds *Handbook of regional and urban economics*, Vol 4 (cities and geography). Elsevier B.V., Amsterdam, Holland

Glaeser E, Kallal HD, Scheinkman JA, Shleifer A 1992. Growth in cities. *Journal of Political Economy* 100: 1126–1152.

Glaeser E, Kolko J, Saiz A 2001. Consumer city. *Journal of Economic Geography* 1: 27–50.

Glaeser E, Maré DC 2001. Cities and skills. *Journal of Labour Economics* 19(2): 316–342.

Goldstein GS, Gronberg TJ 1984. Economies of scope and economies of agglomeration. *Journal of Urban Economics* 16: 91–104.

Gordon P, Richardson HW 1998. Prove it: the costs and benefits of sprawl. *The Brookings Review* 16(4): 23–25.

Gordon P, Richardson HW 2000. Critiquing sprawl's critics. Policy Analysis No. 365, Cato Institute, Washington DC.

Gow L 2000. *Curbing the sprawl: urban growth management in the United States – lessons for New Zealand*. ME no. 342, Ministry for the Environment, Wellington, NZ.

Grimes A, Kerr S, Aitken A 2004. Bi-directional impacts of economic, social and environmental changes and the New Zealand housing market. Motu Working Paper 04-09. Motu Economic and Public Policy Research, Wellington, NZ.

Hall T 2001. *Urban geography*. Routledge, Florence, KY, USA.

Hargreaves R, Davies K 2003. *Urban sustainability in New Zealand: an information resource for urban practitioners*. Ministry for the Environment, Wellington, NZ.

Haring JE, Slobko T, Chapman J 1976. The impact of alternative transportation systems on urban structure. *Journal of Urban Economics* 3: 14–30.

Harrison D, Kain JF 1974. Cumulative urban growth and urban density functions. *Journal of Urban Economics* 1: 61–98.

Head K, Mayer T 2004. The empirics of agglomeration. In: Henderson JV, Thisse JF eds *Handbook of regional and urban economics*, Vol 4 (cities and geography). Elsevier B. V., Amsterdam, Holland. Pp. 2609–2669.

Heikkila E, Gordon P, Kim JI, Peiser RB, Richardson HW, Dale-Johnson D 1989. What happened to the CBD-distance gradient?: land values in a polycentric city. *Environment and Planning A* 21: 221–232.

Helsley RW, Strange WC 1990. Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics* 20: 189–212.

Helsley RW, Strange WC 2004. Knowledge barter in cities. *Journal of Urban Economics* 56: 327–345.

Henderson JV 2003. Marshall's scale economies. *Journal of Urban Economics* 53: 1–28.

Henderson JV, Kuncoro A, Turner M 1995. Industrial development in cities. *Journal of Political Economy* 103: 1067–1085.

Hobson PAR 1987. Optimum product variety in urban areas. *Journal of Urban Economics* 22: 190–197.

Imi A 2005. Decentralisation and economic growth revisited: an empirical note. *Journal of Urban Economics* 57: 449–461.

Irwin EG, Geoghegan J 2001. Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems and Environment* 85: 7-23.

Jacobs J 1961. *The death and life of great American cities*. Random House, New York.

Jaffe AB, Trajtenberg M, Henderson R 1993. Geographic localisation of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* 108(3): 577–598.

Kolko J 1999. Can I get some service here? Information technology, service industries, and the future of cities. Department of Economics, Harvard University.

Krieger A 2004. The costs-or have there been benefits, too?-of sprawl. Harvard Design Magazine 19 (2003-2004 Fall-Winter)

Krugman PR 1991a. Increasing returns and economic geography. *Journal of Political Economy* 99: 483–499.

Krugman PR 1991b. *Geography and trade*. MIT Press, Cambridge.

Krugman PR 1993. First nature, second nature, and metropolitan location. *Journal of Regional Science* 33: 129–144.

Krugman PR 1995. *Development, geography, and economic theory*. MIT Press, Cambridge, MA.

Krugman PR, Venables AJ 1995. Globalisation and the inequality of nations. *Quarterly Journal of Economics* 110: 857–880.

Macrae CD 1982. Urban housing with discrete structures. *Journal of Urban Economics* 11: 131–147.

Maré DC 2005. Concentration, specialisation and agglomeration of firms in New Zealand. Motu Working Paper 05-12. Motu Economic and Public Policy Research, Wellington, NZ.

Marshall A 1920. *Principles of economics*. Macmillan, London.

McGrath DT (revised January 2005, in press June 2005). More evidence on the spatial scale of cities. *Journal of Urban Economics* Any possibility of a date for this?.

McIndoe G, Chapman R, McDonald C, Holden G, Howden-Chapman P, Sharpin AB 2005. The value of urban design: the economic, environmental and social benefits of urban design. Ministry for the Environment Manatū Mō Te Taiao, Wellington, NZ.

McMillen DP, Smith SC 2003. The number of sub-centres in large urban areas. *Journal of Urban Economics* 53: 321–338.

Meen G, Andrew M 2004. The role of housing in city economic performance. In: Boddy M, Parkinson M eds *City matters: competitiveness, cohesion and urban governance*. The Policy Press, Bristol, UK. Pp. 199–215.

Mills ES 1967. An aggregative model of resource allocation in a metropolitan area. *American Economic Review* 57: 197–210.

Mills ES 2000. A thematic history of urban economic analysis. In: Gale WG, Rothenberg Pack J eds *Brookings-Wharton papers on urban affairs 2000*. Brookings Institution Press, Washington DC. Pp. 1–52.

Mills ES 2001. Comments (to Urban sprawl: lessons from urban economics by JK Brueckner). In: Gale WG, Rothenberg Pack J eds *Brookings-Wharton papers on urban affairs*. Brookings Institution Press, Washington DC. Pp. 90-93.

Mills ES, Hamilton BW 1994. *Urban economics*, 5th ed. Harper Collins, New York.

Mills ES, Nijkamp P 1987. Advances in urban economics. In: Mills ES ed. *Handbook of regional and urban economics*, Vol 2 (urban economics). Elsevier Science B.V., Amsterdam, Holland. Pp. 703–714.

Moore B, Begg I 2004. Urban growth and competitiveness in Britain: a long-run perspective. In: Boddy M, Parkinson M eds *City matters: competitiveness, cohesion and urban governance*. The Policy Press, Bristol, UK. Pp. 93–109.

Myrdal G 1957. *Economic theory and underdeveloped regions*. Duckworth, London.

Nivola PS 1998. Fat city: understanding American urban form from a transatlantic perspective. *Brookings Review* 16(4):17–19.

O’Connell MJ 2003. *Urban sustainability worldwide: an information resource for urban practitioners*. Ministry for the Environment, Wellington, NZ.

Ottaviano G, Thisse J-F 2004. Agglomeration and economic geography. In: Henderson JV, Thisse JF eds *Handbook of regional and urban economics*, Vol 4 (cities and geography). Elsevier B.V., Amsterdam, Holland. Pp. 2563–2608.

Papageorgiou YY 1983. Models of agglomeration. *Sistemi Urbani* 3: 391–410.

Prud’homme R, Lee C-W 1999. Size, sprawl, speed and the efficiency of cities. *Urban Studies* 36(11): 1849–1859.

Riley D 2004. Review of settlement form descriptors: Working Paper. Central Laboratories Report 520951.02, Opus International Consultants Ltd, New Zealand.

Rosenthal SS, Strange WC 2003. Geography, industrial organisation, and agglomeration. *Review of Economics and Statistics* 85(2): 377–393.

Rosenthal SS, Strange WC 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson JV, Thisse JF eds *Handbook of regional and urban economics*, Vol 4 (cities and geography). Elsevier B.V., Amsterdam, Holland. Pp. ??–??

Simmie J 2004. Innovation clusters and competitive cities in the UK and Europe. In: Boddy M, Parkinson M eds *City matters: competitiveness, cohesion and urban governance*. The Policy Press, Bristol, UK. Pp. 171-196.

Sinai T, Waldfoegel J 2004. Geography and the internet: is the internet a substitute or a complement of cities? *Journal of Urban Economics* 56: 1–24.

Stansel D 2005. Local decentralisation and local economic growth: a cross-sectional examination of US metropolitan areas. *Journal of Urban Economics* 57: 55–72.

Starrett DA 1978. Market allocations of location choice in a model with free mobility. *Journal of Economic Theory* 17: 21–37.

Straszheim M 1987. The theory of urban residential location. In: Mills ES ed. *Handbook of regional and urban economics*, Vol II. Elsevier Science B.V., Amsterdam, Holland. Pp. 717–757.

Thünen JH von (1826). *Der isolierte staat in beziehung ant landwirtschaft und nationalökonomie*. Perthes, Hamburg. (English trans: 1966) *The Isolated State*, Pergammon Press, Oxford.

Timmins J 2005. Is infrastructure productive? Evaluating the effect of specific infrastructure projects on firm productivity within New Zealand. Mimeo, Ministry of Economic Development, Wellington, NZ.

Venables AJ 1996. Equilibrium locations of vertically linked industries. *International Economic Review* 37: 341–359.

Waldfogel J 2003. Preference for externalities: an empirical study of who benefits whom in differentiated product markets. *RAND Journal of Economics* 34(3): 557–568.

Wheaton WC 2004. Commuting, congestion, and employment dispersal in cities with mixed land use. *Journal of Urban Economics* 55: 417–438.

Wheaton WC, Lewis MJ 2002. Urban wages and labour market agglomeration. *Journal of Urban Economics* 51: 542–562.

Williams K ed.. 2005. *Spatial planning, urban form and sustainable transport*. Ashgate Publishing, Hampshire, England.

APPENDIX 1

Table 5 Definition of sprawl based on eight distinct dimensions of land use patterns: Sprawl is a condition of land use represented by low values on one or more of these dimensions (Galster et al. 2001).

Dimension	Definition	Unit of analysis	Operationalisation
Density	The average number of residential units or the average number of employees per square mile of developable land in a UA.	One-mile-square grids	Total number of housing units (or employees) in a UA/area of a UA.
Continuity	The degree to which developable land has been developed in an unbroken fashion throughout the UA.	One-half-mile-square grids	A one-half-mile-square grid is considered developed if it contains 10 or more housing units or 50 or more employees. The proportion of all such grids in the UA that are so developed is a measure of continuity.
Concentration	The degree to which housing units or jobs are disproportionately located in a relatively few areas or spread evenly in the UA.	One-mile-square grids	Three potential measures: (1) Very high density grids (for housing units or employees) as a percentage of all grids with developable land within the UA. Very high density grids are two standard deviations or more above the mean of the density of all grids in the 100 largest UAs (or in a sample of the 100 largest UAs). (2) The coefficient of variation (standard deviation divided by the mean) of the density of housing units or employees among the grids of scale m in a UA. (3) Delta index. This is analogous to the dissimilarity index and can be interpreted as the share of land use i (e.g., housing units) that would need to shift areal units of scale m to achieve a uniform distribution across the UA. Higher values of Delta indicate more concentration of a use in certain subareas, thus less sprawl.
Clustering	The degree to which development within any one-mile-square area is clustered within one of the four one-half-mile-squares contained within (as opposed to spread evenly throughout).	One-half-mile- and one-mile-square grids	The average for all one-mile squares of the standard deviations of the density of a particular land use (e.g., housing units or employees) among the four squares of each one-mile grid with developable land, standardised by the average density across m -scale grids.

Centrality	The degree to which observations of a given land use are located near the CBD of a UA.	One-mile-square grids	Two measures (in both, the CBD is defined as the address of city hall): (1) The average distance of a land use (e.g., housing units) from the CBD. This is measured as the inverse of the average of the sum of the distance from the centre of the CBD grid to the centre of each one-mile-square grid weighted by the number of observations of land use (e.g., housing units) in the grid, with the resulting average standardised by the square root of the area of the UA. Lower values therefore reflect more sprawl. (2) A centralisation index that measures how rapidly a given land use accumulates relative to land areas as one moves progressively outward in concentric rings from the CBD.
Nuclearity	The extent to which a UA is characterised by a mononuclear pattern of development.	One-mile-square grids	Nuclearity involves the identification of nodes or nuclei.
Mixed uses	The degree to which substantial numbers of two different land uses (e.g., housing units and employees) exist within the same area and this pattern is typical throughout the UA.	One-mile-square grids	The intuitive interpretation of this index is the average density of a particular land use (e.g., housing units) in another land use (e.g., non-residential or employees) area.
Proximity	The degree to which a particular land use or pair of land uses are close to each other across the UA.	One-mile-square grids	This measure can be defined not only for a given use (average distance between households, between jobs, etc.) but, potentially more interesting, between uses. For example, one can define proximity between households and jobs as the measure of sprawl most closely associated with spatial mismatch.